

INSTITUTO TECNOLÓGICO DE AERONÁUTICA



Marcelo Gonçalves Concli

**CARACTERÍSTICAS SOCIOECONÔMICAS E
COMPORTAMENTAIS DE PASSAGEIROS:
INFLUÊNCIAS E CORRELAÇÕES COM O ASSENTO
OCUPADO A BORDO DE AERONAVES.**

Trabalho de Graduação
2021

Curso de Engenharia Civil-Aeronáutica

Marcelo Gonçalves Concli

**CARACTERÍSTICAS SOCIOECONÔMICAS E
COMPORTAMENTAIS DE PASSAGEIROS:
INFLUÊNCIAS E CORRELAÇÕES COM O ASSENTO
OCUPADO A BORDO DE AERONAVES.**

Orientador

Prof. Dr. Alessandro Vinícius Marques de Oliveira (ITA)

ENGENHARIA CIVIL-AERONÁUTICA

**SÃO JOSÉ DOS CAMPOS
INSTITUTO TECNOLÓGICO DE AERONÁUTICA**

Dados Internacionais de Catalogação-na-Publicação (CIP)
Divisão de Informação e Documentação

Concli, Marcelo Gonçalves

Características socioeconômicas e comportamentais de passageiros: influências e correlações com o assento ocupado a bordo de aeronaves. / Marcelo Gonçalves Concli.

São José dos Campos, 2021.

148f.

Trabalho de Graduação – Curso de Engenharia Civil-Aeronáutica– Instituto Tecnológico de Aeronáutica, 2021. Orientador: Prof. Dr. Alessandro Vinícius Marques de Oliveira.

1. Clusterização. 2. Layout e Disposição de assentos. 3. Perfil socioeconômico. 4. Passageiros.
I. Instituto Tecnológico de Aeronáutica. II. Título.

REFERÊNCIA BIBLIOGRÁFICA

CONCLI, Marcelo Gonçalves. **Características socioeconômicas e comportamentais de passageiros: influências e correlações com o assento ocupado a bordo de aeronaves..** 2021. 148f. Trabalho de Conclusão de Curso (Graduação) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSÃO DE DIREITOS

NOME DO AUTOR: Marcelo Gonçalves Concli

TÍTULO DO TRABALHO: Características socioeconômicas e comportamentais de passageiros: influências e correlações com o assento ocupado a bordo de aeronaves..

TIPO DO TRABALHO/ANO: Trabalho de Conclusão de Curso (Graduação) / 2021

É concedida ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias deste trabalho de graduação e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho de graduação pode ser reproduzida sem a autorização do autor.

Verified by PDFFiller

Marcelo G. Concli

11/24/2021

Marcelo Gonçalves Concli

R. H8A, 127

12.228-460 – São José dos Campos–SP

CARACTERÍSTICAS SOCIOECONÔMICAS E COMPORTAMENTAIS DE PASSAGEIROS: INFLUÊNCIAS E CORRELAÇÕES COM O ASSENTO OCUPADO A BORDO DE AERONAVES.

Essa publicação foi aceita como Relatório Final de Trabalho de Graduação

Verified by PDFFiller

Marcelo G. Concli

11/24/2021

Marcelo Gonçalves Concli

Autor



Alessandro Vinícius Marques de Oliveira (ITA)

Orientador



Prof. Dr. João Claudio Bassan de Moraes
Coordenador do Curso de Engenharia Civil-Aeronáutica

São José dos Campos, 19 de novembro de 2021.

Este trabalho é o encerramento de um dos ciclos mais importantes da minha vida, dedico este feito à minha família, à minha companheira e a todos os parceiros de jornada. Todos vocês tiveram papel essencial e foram fundamentais para que eu conseguisse chegar até aqui. Espero que estejamos juntos daqui pra frente e possamos compartilhar muitas mais vitórias!

Agradecimentos

Primeiramente, gostaria de agradecer profundamente aos meus pais Paulo e Cleide por terem cultivado este sonho junto a mim e terem me apoiado de maneira amorosa ao longo de todos os percalços e conquistas. Da mesma maneira, deixo registrado também meu agradecimento à minha companheira Taís pelo apoio incondicional e suporte emocional necessário nesta jornada.

Segundamente, gostaria de agradecer ao Prof. Dr. Alessandro Vinícius Marques de Oliveira, por ter apoiado o desenvolvimento deste trabalho e fornecido ideias e contexto para a pesquisa.

Terceiramente, gostaria de deixar o agradecimento mais sincero aos colegas da Turma 20 e 21 por todos os momentos vividos ao longo desta jornada, desde aqueles que foram os mais desafiadores até os de maior glória. Fazer parte deste instituto é uma honra e foi um privilégio dividir este tempo com vocês.

Por último, deixo o agradecimento geral ao instituto e todos seus colaboradores e professores, bem como aos responsáveis pelo H8 e seu bom funcionamento.

"Persistence is the path to success."

— CHARLES CHAPLIN

Resumo

Desde seu princípio, a indústria da aviação experimenta de maneira constante evolução e modernização em seus diversos pilares. Um dos assuntos que entra nessa jornada evolutiva é o layout ou disposição dos assentos, bem como a modernização dos materiais utilizados na construção deles.

No aspecto da utilização de materiais, a busca sempre foi direcionada à redução de peso da aeronave. No aspecto de posicionamento, a otimização proposta e atualmente percebida na prática é o aumento da capacidade de transporte da aeronave, de maneira de que coloquem maior quantidade de assentos em menor área, para que se haja melhor diluição de custos da viagem pelo número de passageiros.

Os recentes avanços dessa estrutura de pensamento estão sendo levados por uma geração que idealiza novos tipos de configuração. Como destaques e objeto de estudo, estão as configurações Skyrider e Checkerboard. A primeira traz uma proposta que tem por base alterar o assento para uma estrutura similar a um banco de bicicleta, na qual o encosto e assentos se alinham numa posição quase vertical, possibilitando o aumento de capacidade das aeronaves. A segunda proposta de configuração pode ser comparada a um tabuleiro de damas, na qual se obtém liberdade e flexibilidade para modelar a configuração de acordo com a demanda do vôo, de maneira que seja possível transitar entre assentos mais espaçados (considerados premium) e assentos mais juntos (considerados econômicos), de maneira que se ganhem possibilidades de combinação entre áreas premium e áreas econômicas.

Nessa linha de pensamento, a pesquisa "Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial" em seu escopo global tem o objetivo de analisar a viabilidade econômica das configurações citadas de layout e posicionamento de assentos dentro de aeronaves. Como parte fundamental dessa análise, reside o estudo direcionado ao longo deste trabalho de graduação sobre a segmentação de perfis dos passageiros que ocupam os assentos, de maneira que se identifiquem dentro de um conjunto de características socioeconômicas e comportamentais possíveis agrupamentos e correlações com o posicionamento daquele passageiro.

A base de dados que orienta o trabalho é resultado de um questionário aplicado entre

os meses de Fevereiro e Setembro no ano de 2014 em 66 aeroportos, com 122.039 respondentes em um total de 3.793 voos. O questionário é composto por 107 perguntas. Dentre as respostas coletadas, obtêm-se informações que variam desde informações sobre trecho, gastos envolvidos durante os trajetos e estadias, perfil do consumidor (faixa salarial, profissão e outros) até um mapeamento sobre a opinião do respondente acerca do segmento de infraestrutura aeroportuária.

Através de análise na base de dados proveniente do questionário, será possível a identificação de “clusters” ou grupos de passageiros e relações claras de influência / correlação destes nas posições ocupadas por cada um dos passageiros.

O método de clusterização a ser utilizado deve se aplicar tanto à variáveis categóricas (como profissão, motivo da viagem, grupo salarial, meios de transporte utilizados nos trajetos e assento ocupado) quanto à variáveis não-categóricas (como gastos no interior dos aeroportos, nos trajetos, idades e outros). Como proposta inicial, serão exploradas as metodologias “K-Modes” e “K-Prototypes”, derivadas da “K-Means” (aplicável para variáveis não-categóricas) com as devidas modificações para incluir variáveis categóricas.

Abstract

Since its inception, the aviation industry has constantly experienced evolution and modernization in its various pillars. One of the subjects that enters this evolutionary journey is the layout or arrangement of the seats, as well as the modernization of the materials used in their construction.

As far as the use of materials is concerned, the quest has always been directed at reducing the aircraft's weight. In the aspect of positioning, the optimization proposed and currently perceived in practice is to increase the carrying capacity of the aircraft, in order to fit more seats in a smaller area, in contemplation of a better dilution of costs of the trip by the number of passengers.

The recent advances in this thought structure are being carried by a generation that idealizes new types of configuration. As centerpiece and study objects are the Skyrider and Checkerboard configurations. The former brings a proposal based on changing the seat to a structure similar to a bicycle seat, in which the backrest and seats are aligned in an almost vertical position, making it possible to increase the capacity of the aircraft. The latter configuration proposal can be compared to a checkerboard, in which one gets freedom and flexibility to model the configuration according to the flight demand, in a manner which is possible to move between more widely spaced seats (considered premium) and closer together seats (considered economy), in order to gain combination possibilities between premium and economy areas.

In this line of thought, the research "Study of cabin densification and segmentation by introduction of innovative seating concepts in commercial aviation aircraft" in its global scope has the objective of analyzing the economic feasibility of the mentioned configurations of layout and positioning of seats inside aircraft. As a fundamental part of this analysis, lies the study directed throughout this final paper on the segmentation of passenger profiles that occupy the seats, in a way that possible groupings and correlations with the positioning of that passenger can be identified within a set of socioeconomic and behavioral characteristics.

The database that guides the paper is the result of a questionnaire applied between the months of February and September in the year 2014 in 66 airports, with 122,039

respondents in a total of 3,793 flights. The questionnaire is composed of 107 questions. Among the answers collected, information is obtained that ranges from information about the route, expenses involved during trips and stays, consumer profile (salary range, profession, and others) to a mapping of the respondent's opinion about the airport infrastructure segment.

Through analysis of the database from the questionnaire, it will be possible to identify clusters or groups of passengers and clear relations of influence / correlation of these in the positions occupied by each passenger. The clustering method to be used should apply both to categorical variables (such as occupation, purpose of the trip, wage group, means of transport used on the routes and seat occupied) and to non-categorical variables (such as spending inside the airports, on the routes, age and others). As an initial proposal, the "K-Modes" and "K-Prototypes" methodologies will be explored. "K-Prototypes" methodologies, derived from "K-Means" (applicable for non-categorical variables) will be explored with appropriate modifications to include categorical variables.

Lista de Figuras

FIGURA 1.1 – Modelo dos assentos - SkyRider 2.0 - Fonte: (AVIOINTERIORS, 2021)	25
FIGURA 1.2 – Configuração dos assentos - Checkerboard - Fonte: (BUTTERFLY, 2021)	25
FIGURA 1.3 – Modelo dos assentos premium - Checkerboard - Fonte: (BUTTERFLY, 2021)	26
FIGURA 1.4 – Modelo dos assentos econômicos - Checkerboard - Fonte: (BUTTERFLY, 2021)	26
FIGURA 3.1 – Número de respostas por estado - 20 mais representativos	32
FIGURA 3.2 – Número de respostas por aeroporto - 20 mais representativos	33
FIGURA 3.3 – Distribuição percentual de passageiros por fileira de assentos	38
FIGURA 3.4 – Custo de clusterização de acordo com o número de clusters - K-Prototypes	47
FIGURA 3.5 – Custo de clusterização de acordo com o número de clusters - K-Modes	48
FIGURA 4.1 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in	50
FIGURA 4.2 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem	51
FIGURA 4.3 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método	52
FIGURA 4.4 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável	52
FIGURA 4.5 – Número de passageiros por cluster dentro de cada segmentação - Profissão	53

FIGURA 4.6 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta	53
FIGURA 4.7 – Número de passageiros por cluster dentro de cada segmentação - Idade	54
FIGURA 4.8 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano	55
FIGURA 4.9 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto	56
FIGURA 4.10 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias	56
FIGURA 4.11 – Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos	57
FIGURA 4.12 – Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos	58
FIGURA 4.13 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in	65
FIGURA 4.14 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem	66
FIGURA 4.15 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método	67
FIGURA 4.16 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável	67
FIGURA 4.17 – Número de passageiros por cluster dentro de cada segmentação - Profissão	68
FIGURA 4.18 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta	69
FIGURA 4.19 – Número de passageiros por cluster dentro de cada segmentação - Idade	69
FIGURA 4.20 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano	70
FIGURA 4.21 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto	71
FIGURA 4.22 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias	72

FIGURA 4.23 –Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos	73
FIGURA 4.24 –Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos	73
FIGURA 4.25 –Número de passageiros por cluster dentro de cada segmentação - Método de Check-in	82
FIGURA 4.26 –Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem	83
FIGURA 4.27 –Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método	84
FIGURA 4.28 –Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável	84
FIGURA 4.29 –Número de passageiros por cluster dentro de cada segmentação - Profissão	85
FIGURA 4.30 –Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta	86
FIGURA 4.31 –Número de passageiros por cluster dentro de cada segmentação - Idade	86
FIGURA 4.32 –Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano	87
FIGURA 4.33 –Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto	88
FIGURA 4.34 –Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias	89
FIGURA 4.35 –Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos	90
FIGURA 4.36 –Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos	90
FIGURA 4.37 –Número de passageiros por cluster dentro de cada segmentação - Método de Check-in	100
FIGURA 4.38 –Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem	101
FIGURA 4.39 –Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método	102

FIGURA 4.40 –Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável	103
FIGURA 4.41 –Número de passageiros por cluster dentro de cada segmentação - Profissão	104
FIGURA 4.42 –Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta	104
FIGURA 4.43 –Número de passageiros por cluster dentro de cada segmentação - Idade	105
FIGURA 4.44 –Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano	106
FIGURA 4.45 –Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto	106
FIGURA 4.46 –Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias	107
FIGURA 4.47 –Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos	108
FIGURA 4.48 –Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos	109
FIGURA 5.1 – Distribuição espacial dos clusters encontrados	121
FIGURA 5.2 – Distribuição espacial dos clusters encontrados após revisão e agru- pamento de resultados	123
FIGURA 5.3 – Legenda - Distribuição espacial dos clusters encontrados - variável renda	141
FIGURA 5.4 – Distribuição espacial dos clusters encontrados - variável renda . . .	141
FIGURA 5.5 – Legenda - Distribuição espacial dos clusters encontrados - variável profissão	142
FIGURA 5.6 – Distribuição espacial dos clusters encontrados - variável profissão . .	143

Lista de Tabelas

TABELA 3.1 – Tabela resumitiva - Método de check-in	34
TABELA 3.2 – Tabela resumitiva - Motivo de viagem	34
TABELA 3.3 – Tabela resumitiva - Meio de compra da passagem	35
TABELA 3.4 – Tabela resumitiva - Responsável pela compra da passagem	35
TABELA 3.5 – Tabela resumitiva - Profissão	36
TABELA 3.6 – Tabela resumitiva - Renda familiar mensal	37
TABELA 3.7 – Contagem de passageiros e distribuição percentual por coluna do assento ocupado	38
TABELA 3.8 – Contagem e distribuição percentual de métodos de check-in por co- luna do assento	39
TABELA 3.9 – Contagem e distribuição percentual de motivos de viagem por coluna do assento	39
TABELA 3.10 – Contagem e distribuição percentual de método de compra da passa- gem por coluna do assento	40
TABELA 3.11 – Contagem e distribuição percentual de responsável pela compra da passagem por coluna do assento	40
TABELA 3.12 – Contagem e distribuição percentual de profissão por coluna do assento	40
TABELA 3.13 – Contagem e distribuição percentual de faixa de renda familiar men- sal bruta por coluna do assento	41
TABELA 3.14 – Gastos dos passageiros dentro do aeroporto - mínimo, média, máximo	41
TABELA 3.15 – N° de vezes que realizou o mesmo trecho no último ano - mínimo, média, máximo	42
TABELA 3.16 – N° de vezes que utilizou transporte aéreo no último ano - mínimo, média, máximo	42

TABELA 3.17 –Idade - mínimo, média, máximo	42
TABELA 3.18 –Nº de dias antecedência compra - mínimo, média, máximo	43
TABELA 4.1 –Número de passageiros por cluster - N=5 - K-modes	49
TABELA 4.2 –Perfil identificado - 1º Cluster - K-Modes	59
TABELA 4.3 –Perfil identificado - 2º Cluster - K-Modes	59
TABELA 4.4 –Perfil identificado - 3º Cluster - K-Modes	60
TABELA 4.5 –Perfil identificado - 4º Cluster - K-Modes	60
TABELA 4.6 –Perfil identificado - 5º Cluster - K-Modes	61
TABELA 4.7 –Perfil identificado - 1º Cluster - K-Prototypes	62
TABELA 4.8 –Perfil identificado - 2º Cluster - K-Prototypes	62
TABELA 4.9 –Perfil identificado - 3º Cluster - K-Prototypes	63
TABELA 4.10 –Perfil identificado - 4º Cluster - K-Prototypes	63
TABELA 4.11 –Perfil identificado - 5º Cluster - K-Prototypes	64
TABELA 4.12 –Número de passageiros por cluster - N=7 - K-modes	64
TABELA 4.13 –Perfil identificado - 1º Cluster - K-Modes	74
TABELA 4.14 –Perfil identificado - 2º Cluster - K-Modes	75
TABELA 4.15 –Perfil identificado - 3º Cluster - K-Modes	75
TABELA 4.16 –Perfil identificado - 4º Cluster - K-Modes	76
TABELA 4.17 –Perfil identificado - 5º Cluster - K-Modes	76
TABELA 4.18 –Perfil identificado - 6º Cluster - K-Modes	77
TABELA 4.19 –Perfil identificado - 7º Cluster - K-Modes	77
TABELA 4.20 –Perfil identificado - 1º Cluster - K-Prototypes	78
TABELA 4.21 –Perfil identificado - 2º Cluster - K-Prototypes	78
TABELA 4.22 –Perfil identificado - 3º Cluster - K-Prototypes	79
TABELA 4.23 –Perfil identificado - 4º Cluster - K-Prototypes	79
TABELA 4.24 –Perfil identificado - 5º Cluster - K-Prototypes	80
TABELA 4.25 –Perfil identificado - 6º Cluster - K-Prototypes	80
TABELA 4.26 –Perfil identificado - 7º Cluster - K-Prototypes	81
TABELA 4.27 –Número de passageiros por cluster - N=8 - K-modes	81

TABELA 4.28 –Perfil identificado - 1° Cluster - K-Modes	91
TABELA 4.29 –Perfil identificado - 2° Cluster - K-Modes	92
TABELA 4.30 –Perfil identificado - 3° Cluster - K-Modes	92
TABELA 4.31 –Perfil identificado - 4° Cluster - K-Modes	93
TABELA 4.32 –Perfil identificado - 5° Cluster - K-Modes	93
TABELA 4.33 –Perfil identificado - 6° Cluster - K-Modes	94
TABELA 4.34 –Perfil identificado - 7° Cluster - K-Modes	94
TABELA 4.35 –Perfil identificado - 8° Cluster - K-Modes	95
TABELA 4.36 –Perfil identificado - 1° Cluster - K-Prototypes	96
TABELA 4.37 –Perfil identificado - 2° Cluster - K-Prototypes	96
TABELA 4.38 –Perfil identificado - 3° Cluster - K-Prototypes	97
TABELA 4.39 –Perfil identificado - 4° Cluster - K-Prototypes	97
TABELA 4.40 –Perfil identificado - 5° Cluster - K-Prototypes	98
TABELA 4.41 –Perfil identificado - 6° Cluster - K-Prototypes	98
TABELA 4.42 –Perfil identificado - 7° Cluster - K-Prototypes	99
TABELA 4.43 –Perfil identificado - 8° Cluster - K-Prototypes	99
TABELA 4.44 –Número de passageiros por cluster - N=10 - K-modes	100
TABELA 4.45 –Perfil identificado - 1° Cluster - K-Modes	110
TABELA 4.46 –Perfil identificado - 2° Cluster - K-Modes	110
TABELA 4.47 –Perfil identificado - 3° Cluster - K-Modes	111
TABELA 4.48 –Perfil identificado - 4° Cluster - K-Modes	111
TABELA 4.49 –Perfil identificado - 5° Cluster - K-Modes	112
TABELA 4.50 –Perfil identificado - 6° Cluster - K-Modes	112
TABELA 4.51 –Perfil identificado - 7° Cluster - K-Modes	113
TABELA 4.52 –Perfil identificado - 8° Cluster - K-Modes	113
TABELA 4.53 –Perfil identificado - 9° Cluster - K-Modes	114
TABELA 4.54 –Perfil identificado - 10° Cluster - K-Modes	114
TABELA 4.55 –Perfil identificado - 1° Cluster - K-Prototypes	115
TABELA 4.56 –Perfil identificado - 2° Cluster - K-Prototypes	115

TABELA 4.57 –Perfil identificado - 3° Cluster - K-Prototypes	116
TABELA 4.58 –Perfil identificado - 4° Cluster - K-Prototypes	116
TABELA 4.59 –Perfil identificado - 5° Cluster - K-Prototypes	117
TABELA 4.60 –Perfil identificado - 6° Cluster - K-Prototypes	117
TABELA 4.61 –Perfil identificado - 7° Cluster - K-Prototypes	118
TABELA 4.62 –Perfil identificado - 8° Cluster - K-Prototypes	118
TABELA 4.63 –Perfil identificado - 9° Cluster - K-Prototypes	119
TABELA 4.64 –Perfil identificado - 10° Cluster - K-Prototypes	119
TABELA 5.1 – Perfis identificados - eq1km, eq2km, eq3km	124
TABELA 5.2 – Perfis identificados - eq4km, eq5km, eq9km	124
TABELA 5.3 – Perfis identificados - 7km7, 1km10	125
TABELA 5.4 – Perfis identificados - eq1kp, eq7kp, 9kp10	125
TABELA 5.5 – Perfis identificados - eq1km, eq2km, eq3km	126
TABELA 5.6 – Perfis identificados - eq4km, eq5km, 7km7	126
TABELA 5.7 – Perfis identificados - 4km8, 4km10	127
TABELA 5.8 – Perfil identificado - eq7kp	127
TABELA 5.9 – Perfis identificados - eq4km, eq7km, eq9km	128
TABELA 5.10 –Perfis identificados - 7km7, 4km8, 4km10	128
TABELA 5.11 –Perfil identificado - eq3kp	129
TABELA 5.12 –Perfil identificado - 4km7	129
TABELA 5.13 –Perfis identificados - eq3km, 3km8, 3km10	130
TABELA 5.14 –Perfil identificado - eq3km	130
TABELA 5.15 –Perfil identificado - eq8kp	131
TABELA 5.16 –Perfil identificado - 7km7	131
TABELA 5.17 –Perfis identificados - eq5kp, 4kp5	132
TABELA 5.18 –Perfil identificado - 7km7	132
TABELA 5.19 –Perfil identificado - eq7kp	133
TABELA 5.20 –Perfis identificados - eq2km, eq3km, eq6km	133
TABELA 5.21 –Perfil identificado - 3km8	134

TABELA 5.22 –Perfis identificados - eq4kp, 6kp8	134
TABELA 5.23 –Perfis identificados - eq2km, eq5km, eq6km	135
TABELA 5.24 –Perfis identificados - eq8km, 3km8, 5km10	135
TABELA 5.25 –Perfil identificado - eq2kp	136
TABELA 5.26 –Perfis identificados - eq2km, eq6km	136
TABELA 5.27 –Perfis identificados - eq6kp, 5kp8	137
TABELA 5.28 –Perfis identificados - eq9km, 7km7	137
TABELA 5.29 –Perfil identificado - 8kp10	138
TABELA 5.30 –Perfil identificado - 7km7	138
TABELA 5.31 –Perfil identificado - 6kp10	139
TABELA 5.32 –Perfis identificados - eq7km, eq8km, eq9km	139
TABELA 5.33 –Perfil identificado - eq10km	140
TABELA 5.34 –Perfis identificados - eq3kp, 8kp8	140

Sumário

1	INTRODUÇÃO	23
1.1	Relevância do tema	24
1.2	Definição do problema	27
1.3	Objetivo principal da pesquisa	28
1.4	Objetivos intermediários da pesquisa	28
1.5	Estrutura do trabalho	29
2	REVISÃO DE LITERATURA	30
2.1	Mensuração do conforto percebido pelo passageiro	30
2.2	Inovações e cenário de práticas adotadas pelas companhias aéreas	31
3	METODOLOGIA	32
3.1	Análise do Escopo da pesquisa	32
3.2	Definição das variáveis de interesse	33
3.3	Análise dos assentos de passageiros e agrupamentos nas variáveis de interesse	37
3.4	Definição do método de clusterização	43
3.5	Implementação do algoritmo e tratamento prévio dos dados	45
4	RESULTADOS	49
4.1	5 Clusters - K-Modes	49
4.1.1	Método de Check-in	50
4.1.2	Motivo da viagem	51
4.1.3	Compra de passagem e Responsável	52
4.1.4	Profissão, Renda familiar mensal bruta e Idade	53

4.1.5	Uso do transporte aéreo	55
4.1.6	Gasto no aeroporto	56
4.1.7	Antecedência de compra da passagem em dias	56
4.1.8	Fileira e coluna dos assentos	57
4.1.9	5 Clusters - K-Modes - Resumo	59
4.2	5 Clusters - K-Prototypes	62
4.3	7 Clusters - K-Modes	64
4.3.1	Método de Check-in	65
4.3.2	Motivo da viagem	66
4.3.3	Compra de passagem e Responsável	67
4.3.4	Profissão, Renda familiar mensal bruta e Idade	68
4.3.5	Uso do transporte aéreo	70
4.3.6	Gasto no aeroporto	71
4.3.7	Antecedência de compra da passagem em dias	72
4.3.8	Fileira e coluna dos assentos	73
4.3.9	7 Clusters - K-Modes - Resumo	74
4.4	7 Clusters - K-Prototypes	78
4.5	8 Clusters - K-Modes	81
4.5.1	Método de Check-in	82
4.5.2	Motivo da viagem	83
4.5.3	Compra de passagem e Responsável	84
4.5.4	Profissão, Renda familiar mensal bruta e Idade	85
4.5.5	Uso do transporte aéreo	87
4.5.6	Gasto no aeroporto	88
4.5.7	Antecedência de compra da passagem em dias	89
4.5.8	Fileira e coluna dos assentos	90
4.5.9	8 Clusters - K-Modes - Resumo	91
4.6	8 Clusters - K-Prototypes	96
4.7	10 Clusters - K-Modes	100

4.7.1	Método de Check-in	100
4.7.2	Motivo da viagem	101
4.7.3	Compra de passagem e Responsável	102
4.7.4	Profissão, Renda familiar mensal bruta e Idade	104
4.7.5	Uso do transporte aéreo	106
4.7.6	Gasto no aeroporto	106
4.7.7	Antecedência de compra da passagem em dias	107
4.7.8	Fileira e coluna dos assentos	108
4.7.9	10 Clusters - K-Modes - Resumo	110
4.8	10 Clusters - K-Prototypes	115
5	ANÁLISE ESPACIAL E DE ROBUSTEZ DOS RESULTADOS	120
5.1	Distribuição espacial dos resultados	120
5.2	Análise de equivalências entre resultados	121
5.3	Distribuição espacial dos resultados iterada	123
5.3.1	Coluna A	124
5.3.2	Coluna B	129
5.3.3	Coluna C	131
5.3.4	Coluna D	133
5.3.5	Coluna F	137
5.4	Distribuição espacial dos resultados - variável renda	140
5.5	Distribuição espacial dos resultados - variável profissão	142
6	CONCLUSÕES	145
6.1	Proposta de integração entre trabalhos	145
6.2	Limitações do trabalho de graduação	146
	REFERÊNCIAS	148

1 Introdução

Desde o pontapé inicial da aviação, tal indústria experimentou um processo constante de renovação de pensamentos, desenvolvimento de novas tecnologias e validações. Tais renovações e novas tecnologias não se restringiram aos aspectos técnicos da engenharia aeronáutica, mas sim foram avançadas as discussões nos mais diversos campos, tais como os aspectos operacionais da aviação comercial.

Neste contexto operacional, a capacidade de uma aeronave e a dinâmica de transporte de passageiros foram assuntos relevantes que se modificaram de maneira expressiva ao longo dos anos diante das realidades da indústria. Ainda neste tópico, o assento ocupado pelo passageiro, suas características físicas e o posicionamento ou disposição são variáveis de alto impacto em diversos aspectos operacionais.

Ainda diante de regulamentação e enrijecimento de diretrizes, os fabricantes e responsáveis pelo desenvolvimento de assentos conseguiram entregar cada vez assentos mais leves e compactos (GARSTEN, 2019). Nessa linha de pensamento, se considera em primeiro plano a redução de peso da aeronave, que acaba por otimizar o consumo de combustível, reduzindo de maneira expressiva os custos das operadoras. De maneira complementar, o layout ou disposição dos assentos na cabine também é capaz de gerar diluição de custos, sendo alvo também de melhorias e otimizações. Através da utilização de estratégias de redução de espaço entre fileiras de assentos, é possível propiciar o aumento da densidade de passageiros por aeronave. Apesar da controvérsia gerada pelo nível de percepção de conforto do passageiro nesta estratégia, esse foi o passo dado pela maior parte das operadoras, visando a redução de custos da operação e focados em serem competitivos. Esta decisão repercutiu na discussão acerca do layout diferenciado para passageiros que buscassem um maior nível de conforto e acabou por fazer com que isso pudesse ser adicionado às estratégias de "up-selling" nos anos subsequentes, acrescentando ao preço da passagem além das várias opcionalidades de bagagens e serviços a opção de adquirir um assento mais espaçado dos demais.

Dado o cenário de relevância do tema na otimização de custos da aviação comercial, o layout de assentos foi a motivação de muitos novos registros de patente, criação de novas empresas e start-ups focadas na proposição de conceitos que influenciassem de maneira

significativa a melhoria ou otimização de alguma das partes da cadeia. Dentre as proposições de layouts, a pesquisa que orienta este trabalho de graduação, intitulada "Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial"(OLIVEIRA, 2021) destaca e aprofunda dois principais: "Skyrider", proposto pela Aviointeriors (AVIOINTERIORS, 2021) e "Checkerboard Convertible Seating System", proposto pela Butterfly (BUTTERFLY, 2021).

O objetivo global do projeto reside em modelar econometricamente tais conceitos de layout e configurações da cabine das aeronaves, analisando a viabilidade econômica dessas e desenvolvendo o estudo do mix de passageiros nas aeronaves. Deste modo, o direcionamento deste trabalho de graduação está definido em categorizar características socioeconômicas de passageiros e propor correlações das variáveis categorizadas entre os assentos ocupados, de maneira que se encontrem, minimamente grupos de passageiros e características principais que possam se acoplar ao modelo econométrico conduzido no objetivo global do projeto.

1.1 Relevância do tema

Dado o cenário de relevância do tema na otimização de custos da aviação comercial, o layout de assentos foi a motivação de muitos novos registros de patente, criação de novas empresas e start-ups focadas na proposição de conceitos que influenciassem de maneira significativa a melhoria ou otimização de alguma das partes da cadeia. Dentre as proposições de layouts, a pesquisa que orienta este trabalho de graduação destaca e aprofunda dois principais: "Skyrider", proposto pela Aviointeriors (AVIOINTERIORS, 2021) e "Checkerboard Convertible Seating System", proposto pela Butterfly (BUTTERFLY, 2021).

A primeira traz uma proposta que tem por base alterar o assento para uma estrutura similar a um banco de bicicleta, na qual o encosto e assentos se alinham numa posição quase vertical, possibilitando o aumento de capacidade das aeronaves. Tal configuração é ilustrada abaixo:



FIGURA 1.1 – Modelo dos assentos - SkyRider 2.0 - Fonte: (AVIOINTERIORS, 2021)

A segunda proposta de configuração pode ser comparada a um tabuleiro de damas, na qual se obtém liberdade e flexibilidade para modelar a configuração de acordo com a demanda do voo, de maneira que seja possível transitar entre assentos mais espaçados (considerados premium) e assentos mais juntos (considerados econômicos), de maneira que se ganhem possibilidades de combinação entre áreas premium e áreas econômicas.

A configuração checkerboard pode ser exemplificada pela distribuição entre assentos organizada de acordo com a imagem abaixo:

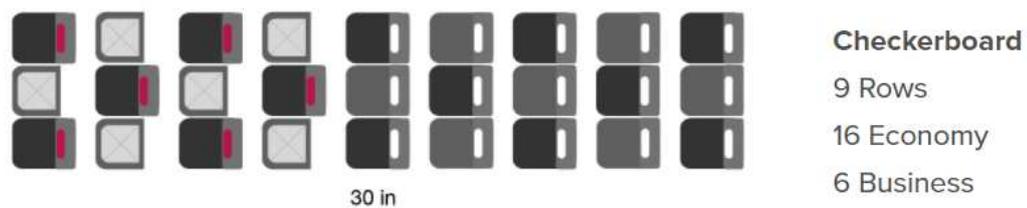


FIGURA 1.2 – Configuração dos assentos - Checkerboard - Fonte: (BUTTERFLY, 2021)

Na qual os assentos assinalados em rosa tem a configuração ilustrada por:



FIGURA 1.3 – Modelo dos assentos premium - Checkerboard - Fonte: (BUTTERFLY, 2021)

Enquanto que os assentos completamente cinzas, econômicos, são construídos da seguinte maneira:



FIGURA 1.4 – Modelo dos assentos econômicos - Checkerboard - Fonte: (BUTTERFLY, 2021)

Diante do cenários de inovações e do pool de opções de layouts e disposições, faz-se necessária a exploração e simulação dos efeitos hipotéticos sobre o mercado diante da adoção de tais inovações. O estudo que orienta este trabalho, então, tem por fim analisar quais seriam os efeitos sobre o grupo de variáveis: custos operacionais, preços das

passagens, segmentação de mercado e receitas operacionais das companhias aéreas. Dentro do grande objetivo, reside o mapeamento de perfil socioeconômico dos ocupantes de cada assento em cada região da aeronave, de maneira que seja possível inputar a variável de perfil e mix de passageiros no estudo econométrico.

1.2 Definição do problema

Diante do escopo definido, o problema principal a ser resolvido ao longo deste trabalho é a proposição de uma maneira analítica para a segmentação de perfis socioeconômicos de passageiros a bordo de aeronaves, de maneira que, ao fim do trabalho, saibam-se com maior precisão e acurácia quais são os principais grupos de passageiros, suas características principais e como estão dispersos espacialmente nas aeronaves.

Nesta linha, encaixa-se o problema desta pesquisa.

A base de dados que orienta este trabalho é resultado de uma pesquisa realizada em 2014 pelo Instituto Olhar intitulada "Pesquisa de Origem/Destino do Transporte Aéreo de Passageiros"(LOGÍSTICA, 2014), cujo escopo se concentra em analisar um conjunto de informações amplo acerca de cada passageiro dentro de um aeroporto, tal conjunto pode ser resumido por:

- 1) Identificação de Unidade Federativa e aeroporto em que a resposta foi coletada;
- 2) Seção dedicada a coletar turno, origem, país, aeroporto de origem e destino do passageiro;
- 3) Identificação de companhia aérea e número de vôo;
- 4) Local de partida para ida ao aeroporto, meio de transporte utilizado, duração do trajeto, custo do trajeto, uso de estacionamento e tempo de permanência, antecedência de chegada, método de check-in utilizado, utilização de serviço de despacho de bagagem, detalhes sobre conexões e o mesmo escopo de informação sobre o destino final;
- 5) Motivação de viagem, forma de pagamento e custo da passagem, motivação de decisão sobre aeroporto de origem/destino, detalhamento sobre uso de transporte aéreo e preferências;
- 6) Valor gasto em compras realizadas no aeroporto, idade do passageiro, renda familiar mensal bruta, condição profissional, entre outros.
- 7) Fotografia do cartão de embarque com informações mais detalhas, tal qual o assento ocupado pelo passageiro.

Dentro dessas grandes áreas, estão dispersas, ao todo, 107 perguntas. Dada a abundância de respostas, há grande volume de informação, principalmente das etapas do processo mais frescas/gravadas na memória do passageiro, como é o caso do valor gasto dentro do aeroporto. É importante ressaltar que uma das etapas que antecedem este trabalho de graduação foi o tratamento das imagens de cartões de embarque e coleta de informações mais detalhadas sobre o assento ocupado pelo passageiro.

Diante desse objeto de análise, pretende-se obter alguma direção sobre perfis de clientes e sua influência sobre os assuntos ocupados pelos mesmos.

1.3 Objetivo principal da pesquisa

O intuito desta pesquisa é explorar diferentes métodos computacionais de agrupamento de dados para chegar em um modelo mínimo que garanta confiabilidade para definição de perfis socioeconômicos de passageiros.

Além da definição, será tratado também o aspecto espacial de distribuição desses grupos identificados no interior de aeronaves, a fim de servir como variável adicional na modelagem econométrica proposta pela pesquisa "Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial"(OLIVEIRA, 2021).

1.4 Objetivos intermediários da pesquisa

O objetivo principal deste trabalho de graduação pode ser subdividido na realização de quatro objetivos intermediários:

- 1) Análise exploratória da base de dados e insights possíveis para conduzir as clusterizações;
- 2) Limpeza e pré-tratamento da base de dados de acordo com as variáveis de interesse;
- 3) Construção de perfis socioeconômicos gerais de passageiros, buscando o número ótimo de grupos e a devida segmentação, incluindo variáveis de assento ocupado (por linha e coluna); e
- 4) Análise aprofundada da distribuição espacial dos grupos e conclusões sobre os resultados gerados.

1.5 Estrutura do trabalho

O capítulo 1 contém a introdução do trabalho, onde são expostos o objetivo, a motivação do mesmo, a descrição da pesquisa e interesses de análise.

O capítulo 2 contém a revisão de literatura do trabalho, na qual trazem-se referências sobre o tema de assentos em aeronaves.

O capítulo 3 contém a análise exploratória inicial, a metodologia de definição das variáveis de interesse, a análise exploratória das mesmas variáveis, as definições acerca de métodos de clusterização, a implementação do algoritmo e as limpezas de dados realizadas na base da pesquisa.

O capítulo 4 apresenta os resultados obtidos utilizando tanto o método K-Prototypes como K-Modes com todos os números de clusters definidos na seção anterior.

O capítulo 5 introduz a análise aprofundada feita com o objetivo de mapear os perfis de passageiros e traçar correlações entre o assento ou região do assento ocupado pelo passageiro e o perfil mapeado nas clusterizações.

Por fim, o capítulo 6 encerra com as conclusões sobre o trabalho e os resultados obtidos, de maneira a fornecer direcionamentos para o uso de resultados e esclarecer as limitações do trabalho, discutindo também seus possíveis desdobramentos.

2 Revisão de literatura

No que tange à literatura que se relaciona e contextualiza o tema proposto, traz-se uma breve revisão de referências que abordam de alguma forma o tema de assentos. Quanto ao objeto principal de aprofundamento desta tese, agrupamento de passageiros por características socioeconômicas, não foram encontrados trabalhos com este fim. As referências principais estão, dessa maneira, relacionadas diretamente com as referências propostas ao longo da proposta submetida ao CNPq citada na introdução (OLIVEIRA, 2021).

A literatura relacionada aos efeitos da disposição de assentos sobre os passageiros se subdivide em dois principais pilares:

- 1) Mensuração do conforto percebido pelo passageiro;
- 2) Inovações e cenário de práticas adotadas pelas companhias aéreas.

Exploram-se, nestes itens, alguns destaques:

2.1 Mensuração do conforto percebido pelo passageiro

Neste subtópico, existe literatura que trata sobre a relação entre a distância dos assentos e o nível de conforto percebido pelos passageiros, dentre os autores, vale citar (ANJANI, 2020) e (KREMSER, 2012). Em suma, a proposta desses autores pode ser resumida por estudar de que maneira a disposição entre fileiras (considerando o "pitch" do assento) influencia o conforto / desconforto do passageiro. Os estudos encontram relações significativas entre tais objetos e apontam maiores desconforto gerado pelo pitch reduzido em detrimento ao posicionamento do passageiro (janela, meio ou corredor).

Já quando se tratam de estudos nacionais, destacam-se (SILVA; MONTEIRO, 2009), (SOUZA, 2010) e (ROSSI, 2011).

2.2 Inovações e cenário de práticas adotadas pelas companhias aéreas

Neste subtópico, a grande parte dos estudos se concentra na observação da questão de precificação da marcação de assentos em vôos comerciais e eventuais consequências. Sob essa ótica, vale citar os autores (KYPARISIS; KOULAMAS, 2018), que estudam preços e alocações ideais de passageiros em termos de otimização do gerenciamento de receitas de companhias aéreas, propondo uma formulação matemática que considera a relação preço-demanda.

Já os autores (MUMBOWER *et al.*, 2015) estudam a adesão de passageiros à compra de seleção de assentos nas classes econômicas premium. Tal estudo traz contribuições significativas no que tange às preferências dos passageiros em situações diversas de marcação de assentos, como, por exemplo, a conclusão de que passageiros preferem comprar a marcação de seu assento em cenários que não existem mais assentos de janela / corredor possíveis para marcação gratuita. Ainda nesta linha, o estudo também traz conclusões sobre a sensibilidade de passageiros ao preço de acordo com a antecedência de compra da passagem. Ainda nesse estudo, concluem sobre aumentos de receita dada a adoção de práticas de reserva de fileiras para classes econômicas premium pelas companhias aéreas.

Os autores (ROUNCIVELL *et al.*, 2018) estudam a disposição de passageiros ao pagamento da marcação de assentos. Nesse estudo, os autores inferem sobre a correlação negativa entre a sensibilidade dos passageiros ao preço das passagens e a disposição dos mesmos a pagar pela marcação do assento. Há também inferências sobre a percepção do cliente em relação à companhia aérea e a probabilidade de adquirir o serviço de marcação. Além disso, também é apontada a forte dependência encontrada entre passageiros que já realizaram compras de marcação de assentos e sua probabilidade de adquirir o serviço novamente.

Por fim, os autores (SHAO *et al.*, 2020) estudam sobre o processo decisório de adquirir seleção de assentos e, diante da tomada de decisão, quais são as preferências dos passageiros no que tange ao assento escolhido e ao tempo prévio de escolha. Dentre as conclusões, os autores sugerem comportamentos no sentido de evitar assentos do meio e ter preferências por assentos dianteiros no processo decisório. Além disso, o estudo também infere sobre a possível dependência entre a probabilidade de pagar pela marcação de assentos e o conjunto de variáveis: relação entre valor para marcação do assento e preço da passagem, do canal de aquisição da passagem, da antecedência de compra e de efeitos sazonais presentes na aviação comercial.

3 Metodologia

3.1 Análise do Escopo da pesquisa

Dada a introdução feita acerca do questionário elaborado para ser aplicado, essa seção tem por interesse analisar a abrangência da coleta efetuada.

A base de dados em seu estado inicial conta com 122.039 observações, num total de 115 variáveis atribuídas a cada observação. Tais variáveis incluem identificadores únicos de respondentes, características do vôo/etapa e todas as respostas para cada pergunta do questionário.

A pesquisa em questão abordou 3.793 vôos distintos em 66 aeroportos espalhados em todos os estados brasileiros, além de no Distrito Federal.

Na Figura 2.1, ilustram-se os 20 estados brasileiros mais representativos em termos de número total de respostas. O Estado de São Paulo é o que soma maior número de respostas, totalizando 20.624, mais que o dobro do segundo colocado.

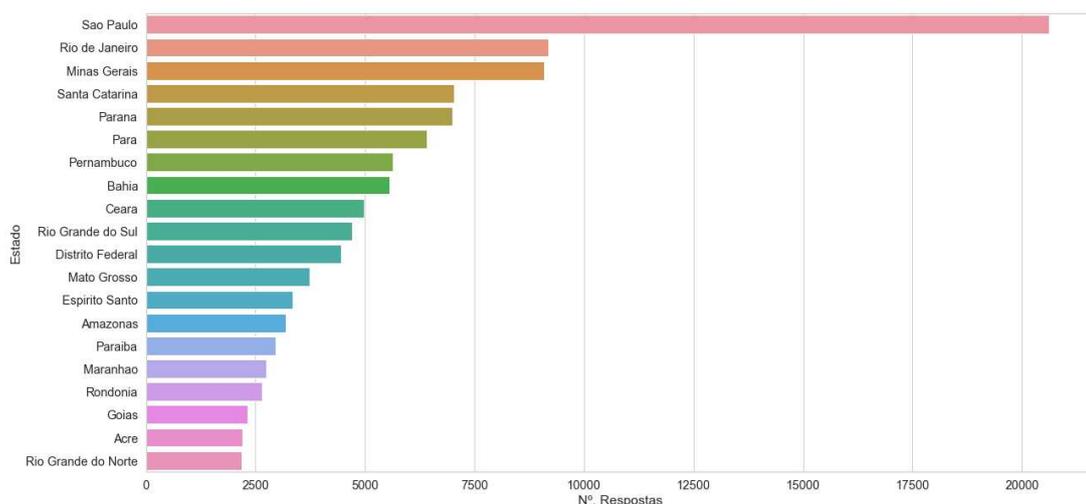


FIGURA 3.1 – Número de respostas por estado - 20 mais representativos

Já na Figura 2.2, são ilustrados os aeroportos da mesma maneira. Sob essa perspectiva, o Aeroporto Internacional de Guarulhos (SBGR) conta com 7.698 respostas e acaba por

ser o mais representativo, acompanhado pelo Aeroporto de Congonhas (SBSP).

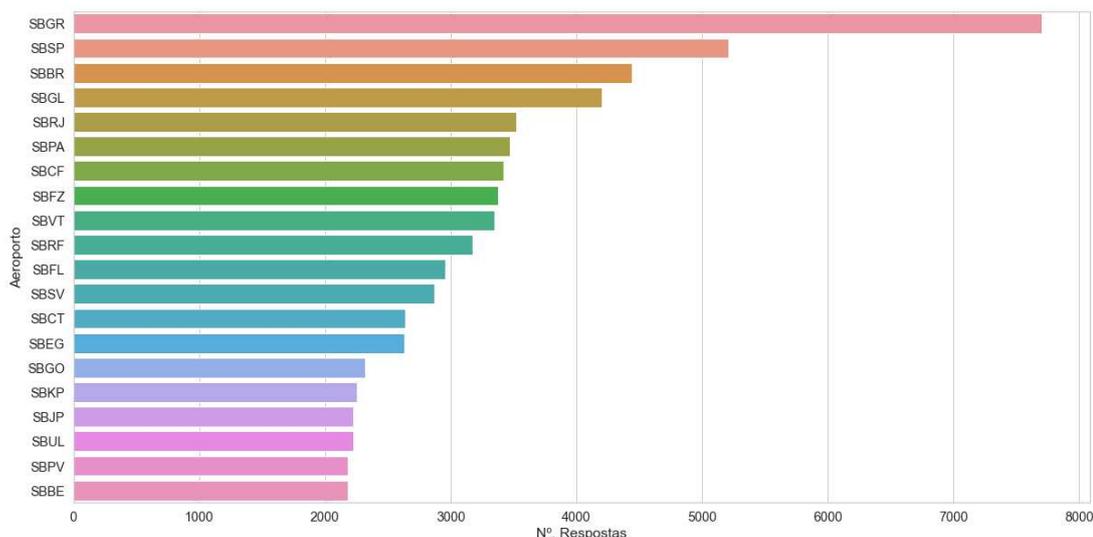


FIGURA 3.2 – Número de respostas por aeroporto - 20 mais representativos

3.2 Definição das variáveis de interesse

Para realizar a definição de variáveis de interesse, foi realizado o seguinte procedimento, que pode ser subdividido em três etapas:

- 1) Definição de tipos de variáveis: classificação entre variáveis categóricas, numéricas contínuas, numéricas discretas;
- 2) Definição de valores únicos em cada uma das variáveis categóricas e intervalos de variáveis numéricas;
- 3) Classificação de colunas entre úteis para visualização e úteis para clusterização de acordo com as categorias de variáveis categóricas e intervalos de variáveis numéricas.

Resultantes desse processo, foram escolhidas as colunas que trariam maior volume de informação útil e de possível diversificação de perfis de consumidor.

Abaixo das variáveis categóricas, há uma tabela resumindo a contagem de ocorrências e seu respectivo percentual representativo na base como todo.

As perguntas / variáveis escolhidas para a realização das clusterizações foram:

- 1) Pergunta 33 - Como você fez o check-in?
 - Balcão

- Totem
- Internet
- Celular
- Despachante
- Não soube responder

TABELA 3.1 – Tabela resumitiva - Método de check-in

Método utilizado para fazer check-in	Contagem	Representatividade percentual
Balcão	26.049	58,67%
Internet	9.738	21,93%
Totem	7.059	15,90%
Celular	1.399	3,15%
Despachante	89	0,20%
Não soube responder	63	0,14%

- 2) Pergunta 54 - Qual o principal motivo dessa viagem?

- Trabalho, negócios, evento profissional
- Lazer
- Visitar amigos ou parentes
- Estudo
- Eventos sociais culturais ou esportivos
- Saúde
- Outro

TABELA 3.2 – Tabela resumitiva - Motivo de viagem

Motivo de viagem	Contagem	Representatividade percentual
Trabalho, negócios, evento profissional	21.188	47,72%
Lazer	11.230	25,29%
Visitar amigos ou parentes	7.119	16,03%
Estudo	1.695	3,82%
Outro	1.255	2,83%
Saúde	1.146	2,58%
Eventos sociais culturais ou esportivos	764	1,72%

- 3) Pergunta 55 - A passagem para o trecho foi comprada com quantos meses e dias de antecedência? - Numérica
- 4) Pergunta 57 - Esta passagem foi comprada ou trocada por milhagem/sistema de milhas?
 - Comprada
 - Trocada por milhagem/sistema de milhas
 - Milhas + dinheiro

TABELA 3.3 – Tabela resumitiva - Meio de compra da passagem

Meio de compra da passagem	Contagem	Representatividade percentual
Comprada	40.693	90,11%
Trocada por milhagem/sistema de milhas	2.994	6,52%
Milhas + dinheiro	1520	3,36%

- 5) Pergunta 59 - Quem pagou pela sua passagem aérea deste trecho da viagem?
 - Eu mesmo
 - Empresa ou órgão que trabalha
 - Familiar ou amigo
 - Outro
 - Não se aplica

TABELA 3.4 – Tabela resumitiva - Responsável pela compra da passagem

Responsável pela compra da passagem	Contagem	Representatividade percentual
Eu mesmo	20.170	45,43%
Empresa ou órgão que trabalha	16.548	37,27%
Familiar ou amigo	4.249	9,57%
Não se aplica	2.944	6,63%
Outro	486	1,09%

- 6) Pergunta 62 - Quanto gastou no aeroporto? - Numérica
- 7) Pergunta 75 - Quantas vezes, nos últimos 12 meses, fez este trecho, utilizando avião (incluindo o trecho de hoje)? - Numérica
- 8) Pergunta 100 - Qual é a sua idade? - Numérica

- 9) Pergunta 101 - Qual é a sua principal condição profissional atual?
 - Funcionário de Empresa Privada
 - Funcionário Público Federal
 - Funcionário Estadual
 - Funcionário Municipal
 - Profissional liberal ou autônomo
 - Empresário
 - Estudante, estagiário ou bolsista
 - Aposentado ou pensionista
 - Atividades do lar
 - Desempregado
 - Outra

TABELA 3.5 – Tabela resumitiva - Profissão

Profissão	Contagem	Representatividade percentual
Funcionário de Empresa Privada	19.669	44,30%
Profissional liberal ou autônomo	5.839	13,15%
Empresário	4.523	9,58%
Funcionário Público Federal	3.828	8,62%
Estudante, estagiário ou bolsista	2.381	5,36%
Funcionário Estadual	2.234	5,03%
Outra	1.825	4,11%
Aposentado ou pensionista	1.679	3,78%
Funcionário Municipal	1.176	2,65%
Atividades do lar	1.033	2,33%
Desempregado	480	1,08%

- 10) Pergunta 102 - Qual é a sua renda FAMILIAR mensal (bruta)?
 - Menor que R\$ 1.448 (Até 2 sm)
 - Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
 - Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
 - Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
 - Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)

- Entre R\$ 14.481 e R\$ 21.720 (Entre 20 e 30 sm)
- Entre R\$ 21.721 e R\$ 36.200 (Entre 30 e 50 sm)
- Maior que R\$ 36.200 (Maior que 50 sm)
- Não quis informar
- Não há renda familiar

TABELA 3.6 – Tabela resumitiva - Renda familiar mensal

Renda familiar mensal	Contagem	Representatividade percentual
Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)	10.023	22,58%
Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)	7.013	15,80%
Não quis informar	6.958	15,67%
Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)	6.394	14,40%
Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)	4.400	9,91%
Entre R\$ 14.481 e R\$ 21.720 (Entre 20 e 30 sm)	3.583	8,07%
Menor que R\$ 1.448 (Até 2 sm)	2.189	4,93%
Entre R\$ 21.721 e R\$ 36.200 (Entre 30 e 50 sm)	2.101	4,73%
Maior que R\$ 36.200 (Maior que 50 sm)	1.154	2,60%
Não há renda familiar	582	1,31%

- 11) Foto do cartão de embarque -> Assento -> Coluna, fileira.

3.3 Análise dos assentos de passageiros e agrupamentos nas variáveis de interesse

Dado o interesse claro deste trabalho de graduação em obter informações socioeconômicas do passageiro e de alguma menira relacionar ao posicionamento do assento, faz-se uma exploração e análise direcionada à identificação de características comuns aos passageiros numa mesma coluna ou numa mesma fileira.

Para isso, é necessário realizar o primeiro tratamento na base de dados, responsável por excluir da base as observações não relevantes para o layout / posicionamento (colunas de assentos definidas entre A e F, fileiras dispersas da primeira à trigésima-segunda). Feita a limpeza, separam-se os valores de coluna e fileira do assento em duas variáveis. Essa etapa faz com que o número total de observações restantes na base seja de 77.381.

Num primeiro momento, analisam-se puramente as distribuições de passageiros em nível de coluna ocupada, de maneira tabulada:

TABELA 3.7 – Contagem de passageiros e distribuição percentual por coluna do assento ocupado

Coluna	Contagem	Percentual representativo
A	20.369	26,32%
D	17.689	22,85%
C	13.837	17,88%
F	10.486	13,55%
B	9.901	12,79%
E	5.099	6,59%

Na Tabela 2.1, pode-se perceber que a distribuição de passageiros nas linhas está em suma maioria nas colunas de A a E, com concentração maior nas colunas A, D e C e menor nas colunas B, E e F. A coluna A tem valor absoluto de passageiros cerca de 4 vezes maior que a coluna E.

De maneira semelhante, distribuem-se os passageiros nas fileiras e ilustra-se como gráfico:

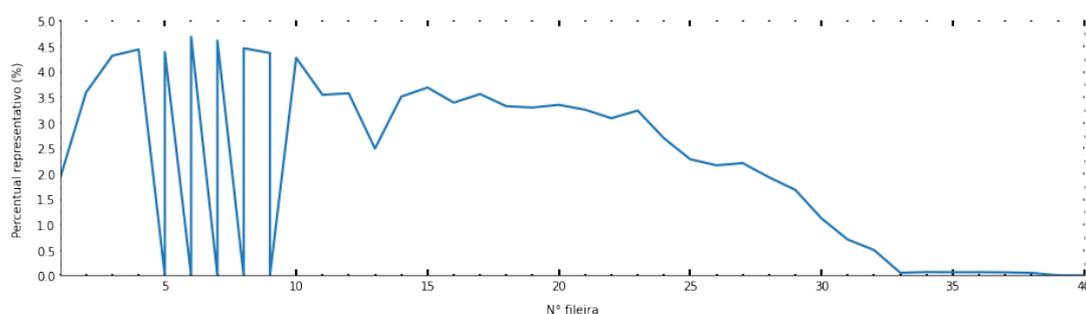


FIGURA 3.3 – Distribuição percentual de passageiros por fileira de assentos

De maneira similar ao que se verifica na etapa anterior, há maior concentração de passageiros dispersos ao longo das fileiras de 1 a 25 e menor concentração no intervalo de fileiras de 25 a 32+.

A fim de trazer clareza para uma possível definição de perfis socioeconômicos de acordo com o posicionamento do passageiro a bordo da aeronave, faz sentido que se analisem as representatividades de cada uma das categorias internas às variáveis categóricas socioeconômicas escolhidas para o desenvolvimento das clusterizações.

Como primeiro passo nesse sentido, analisam-se variáveis categóricas e suas distribuições em cada coluna de interesse:

TABELA 3.8 – Contagem e distribuição percentual de métodos de check-in por coluna do assento

Método utilizado para fazer check-in	A	B	C	D	E	F
Balcão	12099 - 59,4%	6587 - 66,53%	8375 - 60,53%	10590 - 59,87%	3062 - 60,05%	5339 - 50,92%
Celular	580 - 2,85%	171 - 1,73%	410 - 2,96%	479 - 2,71%	107 - 2,1%	365 - 3,48%
Despachante	43 - 0,21%	24 - 0,24%	36 - 0,26%	27 - 0,15%	10 - 0,2%	19 - 0,18%
Internet	5359 - 26,31%	1953 - 19,73%	3163 - 22,86%	4461 - 25,22%	1008 - 19,77%	3010 - 28,7%
Não sabe/ Não respondeu	31 - 0,15%	11 - 0,11%	19 - 0,14%	21 - 0,12%	11 - 0,22%	8 - 0,08%
Totem	2257 - 11,08%	1155 - 11,67%	1834 - 13,25%	2111 - 11,93%	901 - 17,67%	1745 - 16,64%

Na Tabela 2.2, é possível visualizar que há um padrão claro de distribuição concentrada nas categorias presentes na variável "Método utilizado para fazer o check-in": em suma, a maior parcela de passageiros, seja qual for a coluna, utilizou-se de Balcão, Internet e Totem para a realização do check-in. Ainda que as distribuições não estejam dissimilares, é possível inferir que já se identifica uma segmentação de grupos quanto ao método de check-in utilizado.

TABELA 3.9 – Contagem e distribuição percentual de motivos de viagem por coluna do assento

Motivo da viagem	A	B	C	D	E	F
Estudo	934 - 4,59%	317 - 3,2%	445 - 3,22%	713 - 4,03%	173 - 3,39%	518 - 4,94%
Eventos sociais culturais ou esportivos	325 - 1,6%	172 - 1,74%	224 - 1,62%	264 - 1,49%	126 - 2,47%	207 - 1,97%
Lazer	4392 - 21,56%	2549 - 25,74%	2438 - 17,62%	3070 - 17,36%	1794 - 35,18%	2723 - 25,97%
Outro	671 - 3,29%	287 - 2,9%	377 - 2,72%	478 - 2,7%	165 - 3,24%	275 - 2,62%
Saúde	579 - 2,84%	373 - 3,77%	392 - 2,83%	519 - 2,93%	173 - 3,39%	255 - 2,43%
Trabalho, negócios, evento profissional	9788 - 48,05%	4642 - 46,88%	7835 - 56,62%	9889 - 55,9%	1740 - 34,12%	4547 - 43,36%
Visitar amigos ou parentes	3680 - 18,07%	1561 - 15,77%	2126 - 15,36%	2756 - 15,58%	928 - 18,2%	1961 - 18,7%

Na Tabela 2.3, nota-se maior dissimilaridade na distribuição obtida para cada coluna e a motivação da viagem dos passageiros. Ainda assim, há concentração clara de três categorias prevaescentes em todas as colunas, são elas: Trabalho, negócios, eventos profissionais / Lazer / Visitar amigos ou parentes. De todo modo, é possível verificar também que tal variável poderia ser utilizada a fim de segmentar passageiros.

TABELA 3.10 – Contagem e distribuição percentual de método de compra da passagem por coluna do assento

Método - compra de passagem	A	B	C	D	E	F
Comprada	18751 - 92,06%	9240 - 93,32%	12878 - 93,07%	16575 - 93,7%	4606 - 90,33%	9347 - 89,14%
Milhas + dinheiro	351 - 1,72%	154 - 1,56%	207 - 1,5%	253 - 1,43%	84 - 1,65%	187 - 1,78%
Troca - Sistema de milhas	1267 - 6,22%	507 - 5,12%	752 - 5,43%	861 - 4,87%	409 - 8,02%	952 - 9,08%

Na Tabela 2.4, é possível notar que grande parte do mix, quaisquer sejam as colunas, paga a passagem sem o auxílio de sistemas de milhagens, no entanto, ainda há parcelas significativas, como no caso específico da coluna F.

TABELA 3.11 – Contagem e distribuição percentual de responsável pela compra da passagem por coluna do assento

Responsável - pagamento da passagem	A	B	C	D	E	F
Empresa ou órgão que trabalha	7380 - 36,23%	3747 - 37,84%	6438 - 46,53%	7892 - 44,62%	1258 - 24,67%	3231 - 30,81%
Eu mesmo	9498 - 46,63%	4453 - 44,98%	5392 - 38,97%	7245 - 40,96%	2723 - 53,4%	5167 - 49,28%
Familiar ou amigo	2023 - 9,93%	1043 - 10,53%	1096 - 7,92%	1492 - 8,43%	632 - 12,39%	1037 - 9,89%
Não se aplica	1267 - 6,22%	507 - 5,12%	752 - 5,43%	861 - 4,87%	409 - 8,02%	952 - 9,08%
Outro	201 - 0,99%	151 - 1,53%	159 - 1,15%	199 - 1,12%	77 - 1,51%	99 - 0,94%

Em geral, como ilustra a Tabela 2.5, em quaisquer colunas, os responsáveis pelo pagamento da passagem são o próprio passageiro ou a empresa / órgão do qual faz parte. Ainda assim, é um critério a mais de agrupamento que pode ser explorado.

TABELA 3.12 – Contagem e distribuição percentual de profissão por coluna do assento

Profissão	A	B	C	D	E	F
Aposentado ou pensionista	786 - 3,86%	475 - 4,8%	483 - 3,49%	620 - 3,51%	296 - 5,81%	437 - 4,17%
Atividades do lar	485 - 2,38%	311 - 3,14%	295 - 2,13%	367 - 2,07%	176 - 3,45%	221 - 2,11%
Desempregado	275 - 1,35%	119 - 1,2%	136 - 0,98%	198 - 1,12%	91 - 1,78%	150 - 1,43%
Empresário	1747 - 8,58%	857 - 8,66%	1253 - 9,06%	1627 - 9,2%	378 - 7,41%	811 - 7,73%
Estudante, estagiário ou bolsista	1258 - 6,18%	491 - 4,96%	610 - 4,41%	841 - 4,75%	319 - 6,26%	671 - 6,4%
Funcionário Estadual	1010 - 4,96%	510 - 5,15%	688 - 4,97%	845 - 4,78%	303 - 5,94%	585 - 5,58%
Funcionário Municipal	533 - 2,62%	261 - 2,64%	311 - 2,25%	430 - 2,43%	183 - 3,59%	289 - 2,76%
Funcionário Público Federal	1682 - 8,26%	701 - 7,08%	1194 - 8,63%	1532 - 8,66%	443 - 8,69%	972 - 9,27%
Funcionário de Empresa Privada	9039 - 44,38%	4438 - 44,82%	6498 - 46,96%	8261 - 46,7%	1948 - 38,2%	4419 - 42,14%
Outra	927 - 4,55%	458 - 4,63%	628 - 4,54%	726 - 4,1%	266 - 5,22%	530 - 5,05%
Profissional liberal ou autônomo	2627 - 12,9%	1280 - 12,93%	1741 - 12,58%	2242 - 12,67%	696 - 13,65%	1401 - 13,36%

Quanto à profissão, o que se observa na Tabela 2.6 é que há alta concentração de funcionários de empresas privadas e profissionais autônomos ou liberais. Quanto às outras categorias, há distribuição equalizada e pouco concentrada em todas.

TABELA 3.13 – Contagem e distribuição percentual de faixa de renda familiar mensal bruta por coluna do assento

Renda familiar mensal bruta	A	B	C	D	E	F
1) Até 2 sm	1317 - 6,47%	702 - 7,09%	716 - 5,17%	989 - 5,59%	461 - 9,04%	672 - 6,41%
2) Entre 2 e 5 sm	3682 - 18,08%	1789 - 18,07%	2066 - 14,93%	2938 - 16,61%	1067 - 20,93%	1958 - 18,67%
3) Entre 5 e 10 sm	4867 - 23,89%	2166 - 21,88%	2924 - 21,13%	3915 - 22,13%	1164 - 22,83%	2546 - 24,28%
4) Entre 10 e 15 sm	2834 - 13,91%	1315 - 13,28%	2008 - 14,51%	2570 - 14,53%	602 - 11,81%	1468 - 14%
5) Entre 15 e 20 sm	1938 - 9,51%	895 - 9,04%	1441 - 10,41%	1767 - 9,99%	420 - 8,24%	960 - 9,16%
6) Entre 20 e 30 sm	1421 - 6,98%	753 - 7,61%	1221 - 8,82%	1472 - 8,32%	297 - 5,82%	675 - 6,44%
7) Entre 30 e 50 sm	752 - 3,69%	461 - 4,66%	773 - 5,59%	891 - 5,04%	163 - 3,2%	395 - 3,77%
8) Maior que 50 sm	416 - 2,04%	213 - 2,15%	459 - 3,32%	455 - 2,57%	78 - 1,53%	186 - 1,77%
9) Não há renda familiar	328 - 1,61%	142 - 1,43%	173 - 1,25%	217 - 1,23%	87 - 1,71%	162 - 1,54%
Não quis informar	2814 - 13,82%	1465 - 14,8%	2056 - 14,86%	2475 - 13,99%	760 - 14,9%	1464 - 13,96%

Quanto a renda familiar mensal bruta, a distribuição de pessoas nas colunas é muito parecida em todas as colunas utilizadas, a maior parcela dos passageiros analisados se encontram numa faixa entre 2-15 salários mínimos. Não são perceptíveis influências da coluna ocupada para a distribuição de passageiros por renda.

Com a Tabela 2.7, encerram-se as variáveis categóricas que trazem valor diferenciativo para o perfil socioeconômico de passageiros. Entra-se, então, no detalhamento das variáveis numéricas presentes na base de dados.

TABELA 3.14 – Gastos dos passageiros dentro do aeroporto - mínimo, média, máximo

Gasto - No aeroporto (R\$)	Mínimo	Média	Máximo
A	1	30,10	4350
B	1	33,44	2000
C	1	31,20	5000
D	1	30,40	1500
E	1	32,64	2360
F	1	32,84	2490

Como ilustra a Tabela 2.8, em média, todas as colunas concentram passageiros que gastaram cerca de R\$ 30,00 - R\$ 34,00.

TABELA 3.15 – N° de vezes que realizou o mesmo trecho no último ano - mínimo, média, máximo

N° de vezes que realizou o mesmo trecho no último ano	Mínimo	Média	Máximo
A	1	5.82	150
B	1	5.18	190
C	1	7.09	199
D	1	6.59	199
E	1	3.92	120
F	1	5.58	186

Sob a ótica de quantia de vezes que o passageiro utilizou o mesmo trecho no último ano (Tabela 2.10), percebe-se que há diferenciação dos passageiros que se sentam nas colunas C e D, permitindo a inferência de uma relação entre viajar mais e preferir tais colunas.

TABELA 3.16 – N° de vezes que utilizou transporte aéreo no último ano - mínimo, média, máximo

N° de vezes que utilizou transporte aéreo no último ano	Mínimo	Média	Máximo
A	1	18.34	199
B	1	17.34	199
C	1	22.55	199
D	1	21.31	199
E	1	11.25	199
F	1	17.43	199

Para a idade (Tabela 2.11), percebe-se que as médias passam a convergir para o mesmo valor, ainda assim, a distribuição de idades é dissimilar ao longo das colunas, dados os máximos.

TABELA 3.17 – Idade - mínimo, média, máximo

Idade	Mínimo	Média	Máximo
A	14	37.65	85
B	14	38.91	90
C	14	39.82	86
D	14	38.91	95
E	14	38.08	97
F	14	37.27	84

Em termos de antecedência de compra da passagem, nota-se que a média varia no

intervalo de 26 a 43 dias, enquanto que mínimos e máximos são muito próximos nas colunas.

TABELA 3.18 – N^o de dias antecedência compra - mínimo, média, máximo

N ^o de dias antecedência compra	Mínimo	Média	Máximo
A	1	28	361
B	1	33	360
C	1	26	360
D	1	26	360
E	1	43	389
F	1	36	360

3.4 Definição do método de clusterização

Como definido inicialmente, o conjunto de dados que orienta a análise é composto por variáveis numéricas e categóricas.

De acordo com (AHMAD; KHAN, 2019), há uma série de maneiras para se realizar clusterização de bases de dados mistas. Tais métodos ou algoritmos de clusterização podem ser subdivididos da seguinte maneira:

1. Particionais;
2. Hierárquicos;
3. Model-based;
4. Neural network-based;
5. Outros.

Dentro dessas segmentações, as clusterizações que se utilizam de algoritmos particionais são muito adotadas e largamente difundidas no mundo open-source. Hoje, com facilidade, se encontram diversas bibliotecas que fornecem ferramental suficiente para aplicação rápida.

Um dos métodos mais conhecidos e utilizados quando se trata de variáveis numéricas é o K-Means. Quando o foco é analisar variáveis categóricas, o método a se utilizar pode ser o K-modes. Para a abordagem de casos com ambos os tipos de variáveis, surgiram alguns algoritmos que são baseados nos métodos acima citados, como é o caso do K-prototypes.

Neste trabalho de graduação, serão escolhidas os algoritmos K-Prototypes (utilizando a base com variáveis categóricas e não categóricas) e K-Modes (utilizando encoding das variáveis numéricas e transformando-as em categóricas)

Em ambos os métodos, a ideia geral dos algoritmos pode ser subdividida em três etapas, são elas:

1. Definição de centros de clusters que possam representar ambas variáveis numéricas e categóricas;
2. Definição de métrica de distância que seja capaz de combinar ambas variáveis numéricas e categóricas;
3. Definição de uma função de custo de clusterização, que será minimizada iterativamente.

Em linguagem matemática, (HUANG, 1998) define:

A dissimilaridade entre dois objetos de natureza distinta (variável categórica/numérica) X e Y , que são descritos pelos atributos $A_1^\gamma, A_2^\gamma, \dots, A_p^\gamma, A_{p+1}^c, \dots, A_m^c$, pode ser medida por

$$d(X, Y) = \sum_{j=1}^p (x_j - y_j)^2 + \gamma \sum_{j=1}^p \delta(x_j, y_j) \quad (3.1)$$

sendo que o primeiro termo da soma representa o quadrado da distância euclidiana nas variáveis numéricas, enquanto o segundo termo representa a medida de dissimilaridade nas variáveis categóricas. A função do peso γ é evitar o favorecimento de algum dos tipos de atributos.

No algoritmo K-means, dado um conjunto de objetos numéricos X e um número $k (\leq n)$, procura-se por uma partição de X em k clusters que minimiza a soma dos erros quadráticos dentro de cada cluster. A função objetivo a ser minimizada é representada por

$$P(W, Q) = \sum_{l=1}^k \sum_{i=1}^n w_{i,l} d(X_i, Q_l) \quad (3.2)$$

Tal função deve ser minimizada obedecendo o conjunto de condições seguinte:

$$\sum_{l=1}^k w_{i,l} = 1, 1 \leq i \leq n \quad (3.3)$$

$$w_{i,l} \in [0, 1], 1 \leq i \leq n, 1 \leq l \leq k \quad (3.4)$$

em que W é uma matriz que define a partição de dimensão $n \times k$, $Q = [Q_1, Q_2, \dots, Q_k]$ é um conjunto de objetos no mesmo domínio e $d(x, y)$ é a medida de distância euclidiana entre dois objetos.

Adaptando a função de custo do algoritmo K-means para o algoritmo K-prototypes, tem-se:

$$P(W, Q) = \sum_{l=1}^k \left(\sum_{i=1}^n w_{i,l} \sum_{j=1}^p (x_{i,j} - q_{l,j})^2 + \gamma \sum_{i=1}^n w_{i,l} \sum_{j=p+1}^m \delta(x_{i,j}, q_{l,j}) \right) \quad (3.5)$$

Renomeando os elementos, temos:

$$P_l^\gamma = \sum_{i=1}^n w_{i,l} \sum_{j=1}^p (x_{i,j} - q_{l,j})^2 \quad (3.6)$$

$$P_l^c = \gamma \sum_{i=1}^n w_{i,l} \sum_{j=p+1}^m \delta(x_{i,j}, q_{l,j}) \quad (3.7)$$

Finalmente, pode-se reescrever a Eq 3.5 da seguinte maneira:

$$P(W, Q) = \sum_{l=1}^k (P_l^\gamma + P_l^c) \quad (3.8)$$

Com a equação escrita dessa maneira, é possível notar que minimizar o lado esquerdo da equação é equivalente a minimizar cada termo do lado direito individualmente. Desse modo, recorrem-se às estratégias de fixação de uma partição e resolução do problema dada a fixação, de maneira iterativa.

3.5 Implementação do algoritmo e tratamento prévio dos dados

Definido o método de clusterização a ser utilizado, parte-se para a adequação da base de dados.

A implementação do algoritmo foi realizada em Python, utilizando a biblioteca "kmodes", que já tem um de seus módulos dedicados a implementação do algoritmo K-Prototypes, bem como outro dedicado especificamente a aplicação do algoritmo K-Modes. A base de código (para o caso de número de clusters igual a 8) utilizada para a realização deste trabalho de graduação está disponibilizada em ambiente aberto no "Google Colab" (CONCLI, 2021), a fim de que possa ser consultada por quaisquer interessados.

Para as colunas categóricas da base, a tabela abaixo mostra o número de categorias únicas por variável:

Var. Categórica	Qtd. de valores únicos
Método utilizado para fazer check-in	6
Motivo da viagem	7
Método utilizado para comprar a passagem	4
Responsável pelo pagamento da passagem	5
Profissão	11
Renda familiar mensal bruta	10

Para a boa performance do algoritmo, faz-se necessária a continuação da limpeza / tratamento da base de dados.

Previamente, removeram-se as observações da base de dados cujo assento não estava de acordo com o padrão de interesse. O padrão de interesse definido para o escopo do estudo é no formato "XXY", em que "XX" representa a fileira do assento (como número) e "Y" representa a coluna do assento (como letra de A a F).

Como segundo passo, filtram-se as observações cujas variáveis numéricas são positivas e não-nulas.

O número de observações cai drasticamente nesta etapa, totalizando 979 observações. Em tese, essa queda se explica pela qualidade dos dados obtidos após tratamento das imagens de cartão de embarque quando combinados aos dados que representam o perfil socioeconômico dos passageiros em questão. Apenas essas 979 observações contém um assento de interesse e contém informações relevantes para as variáveis socioeconômicas.

Como segundo passo, exclusivamente para a clusterização utilizando o método K-Modes, faz-se a normalização das variáveis numéricas e transforma-se a estrutura de dataframe para matriz.

Com a base já útil em mãos, faz sentido a busca pela definição de um número de clusters a serem calculados. Para isso, utiliza-se a função de custo já implementada para ambos os métodos de clusterização presentes na biblioteca escolhida.

Realiza-se, então, a análise de custo de clusterização do algoritmo K-Prototypes para um número de clusters variando de 1 a 20. O resultado das iterações leva ao gráfico abaixo:

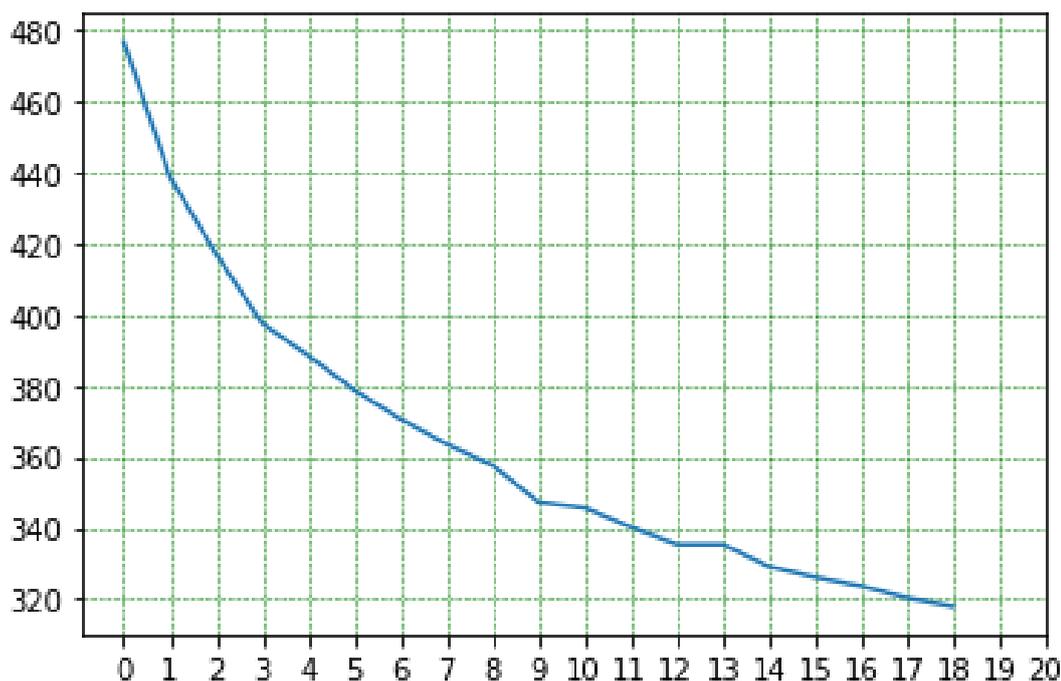


FIGURA 3.4 – Custo de clusterização de acordo com o número de clusters - K-Prototypes

Pode-se notar que a medida que se aumenta o número de clusters, há redução progressivamente atenuada no custo computacional. Num primeiro intervalo, de 0 a 5 clusters, percebe-se a maior redução de custo devido à maior distinção entre grupos encontrada. Após esse número, nota-se que a curva passa a decrescer de maneira mais suave, evidenciando que as novas classes não são tão importantes e de valor diferenciativo quanto as das iterações anteriores.

Da mesma maneira, para efeito de comparação, escolheu-se realizar o método de clusterização K-Modes. Tal escolha se justifica pelo maior número de variáveis categóricas em comparação com numéricas. A transformação feita na base de dados, então, deve ser no sentido de agrupar os valores numéricos em intervalos, criando assim categorias nos mesmos.

Para analisar o custo da clusterização K-modes, plota-se de maneira análoga a anterior o custo de clusterizar de 1 a 25 grupos pelo método:

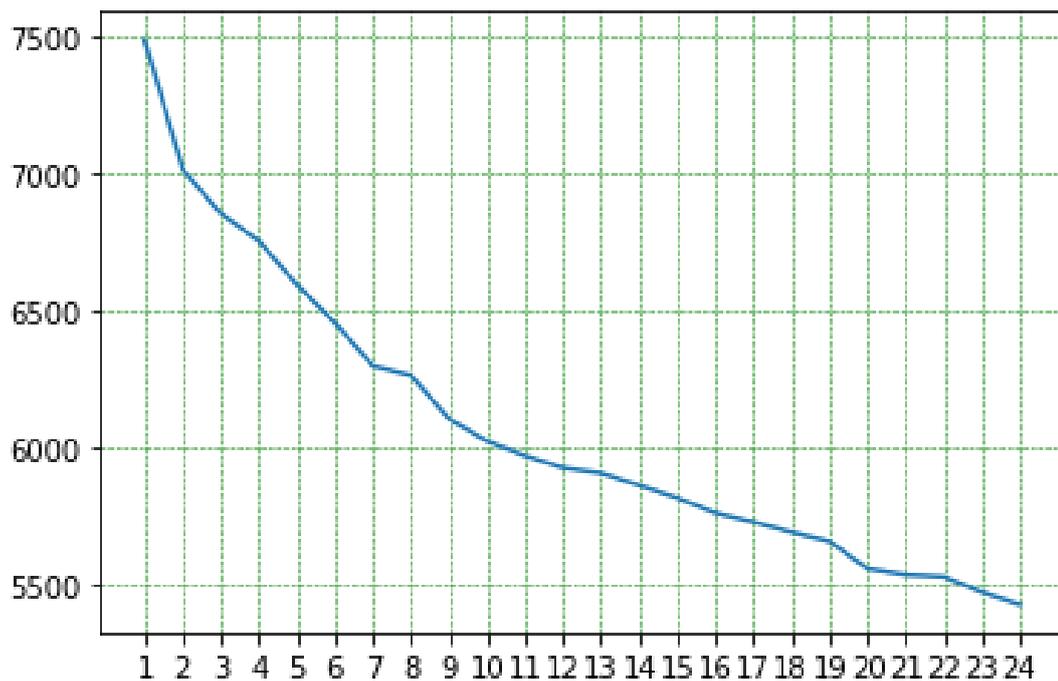


FIGURA 3.5 – Custo de clusterização de acordo com o número de clusters - K-Modes

Quanto aos resultados das análises de custo, esperava-se um padrão de gráfico encontrado em análises como a "Elbow Method". Tal método trata-se de uma otimização que tenta encontrar o menor número de clusters que contabilizam a maior variabilidade nos dados analisados. Como não há clareza em relação ao ponto "Elbow" e na força de sua representatividade, decidiu-se fazer uma série de clusterizações com $N = [5, 7, 8, 10]$ pelos dois métodos distintos citados ao longo do desenvolvimento e então consolidar / comparar o nível de informação adquirida em cada uma das clusterizações.

4 Resultados

4.1 5 Clusters - K-Modes

Utilizando cinco clusters para encaixar os passageiros via K-Modes, os números de passageiros obtidos são mostrados na Tabela 3.7.

TABELA 4.1 – Número de passageiros por cluster - N=5 - K-modes

Cluster	Contagem de passageiros
1 ^o	494
2 ^o	137
3 ^o	135
4 ^o	106
5 ^o	106

Para que se realize uma inspeção sobre o comportamento da clusterização e os possíveis perfis identificados, reúnem-se os gráficos de interesse nas variáveis utilizadas. A seguir, verificam-se os números de passageiros dentro de um dos clusters definidos que, em determinada variável, assumem um valor específico.

4.1.1 Método de Check-in

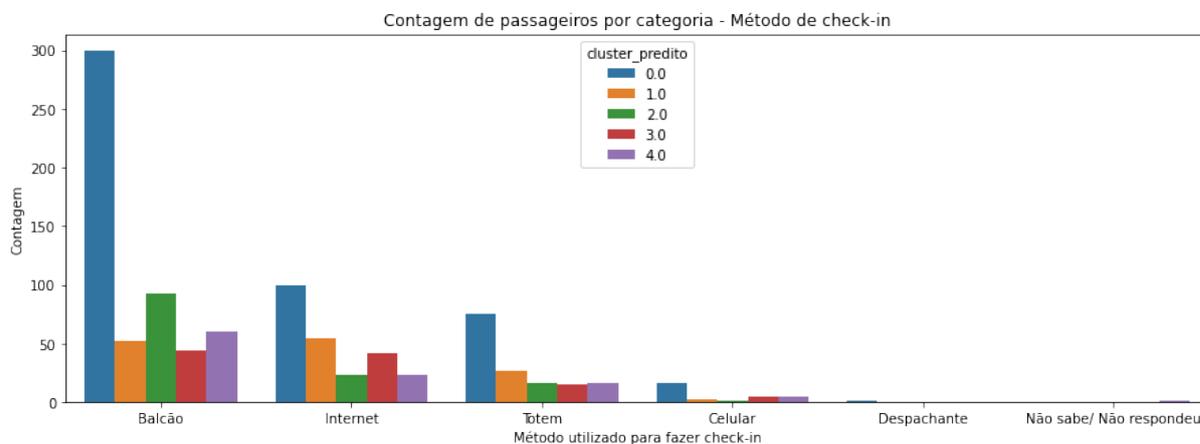


FIGURA 4.1 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in

Para o método de check-in utilizado, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 2º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 3º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 4º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 5º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.

4.1.2 Motivo da viagem

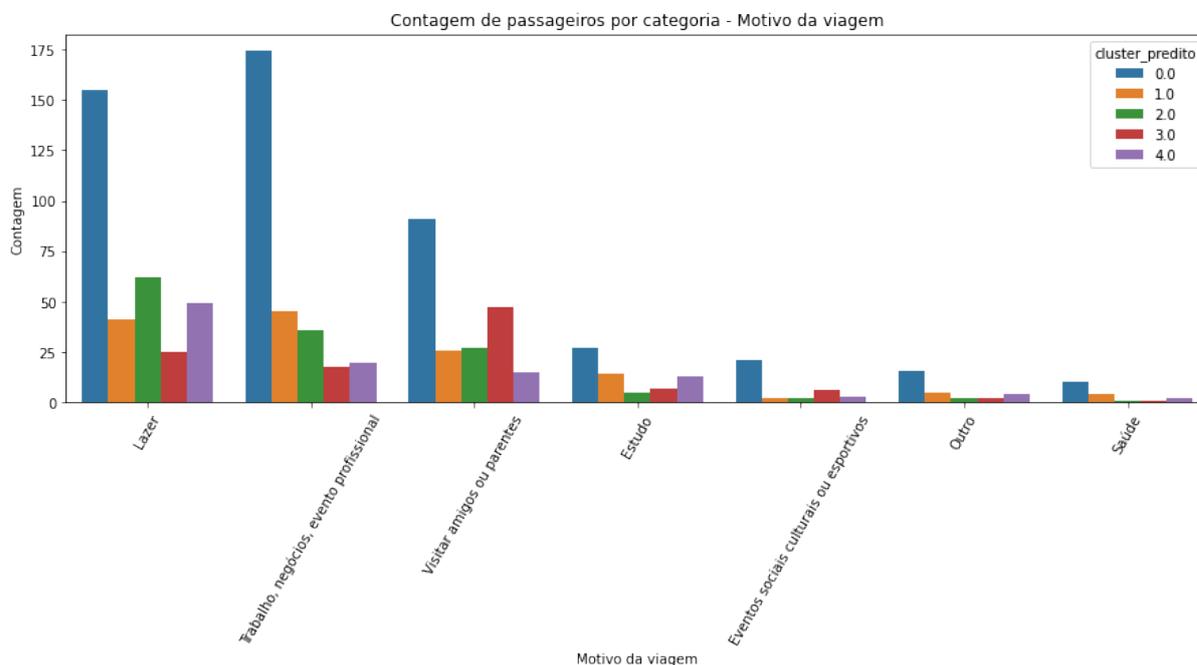


FIGURA 4.2 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem

Para a motivação da viagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais/Lazer.
- 2º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais/Lazer.
- 3º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 4º Cluster: reúne passageiros cuja motivação para viajar é Visitar amigos e parentes.
- 5º Cluster: reúne passageiros cuja motivação para viajar é Lazer.

4.1.3 Compra de passagem e Responsável

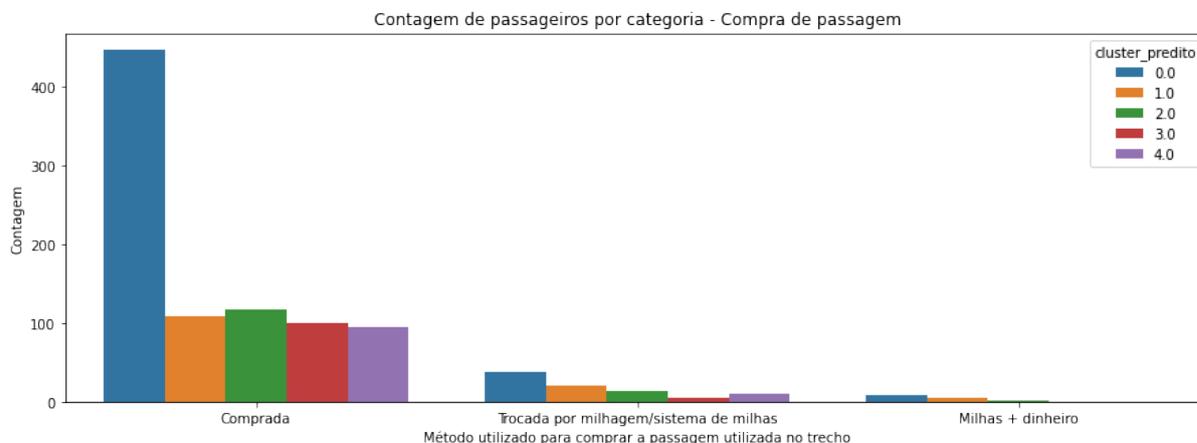


FIGURA 4.3 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método

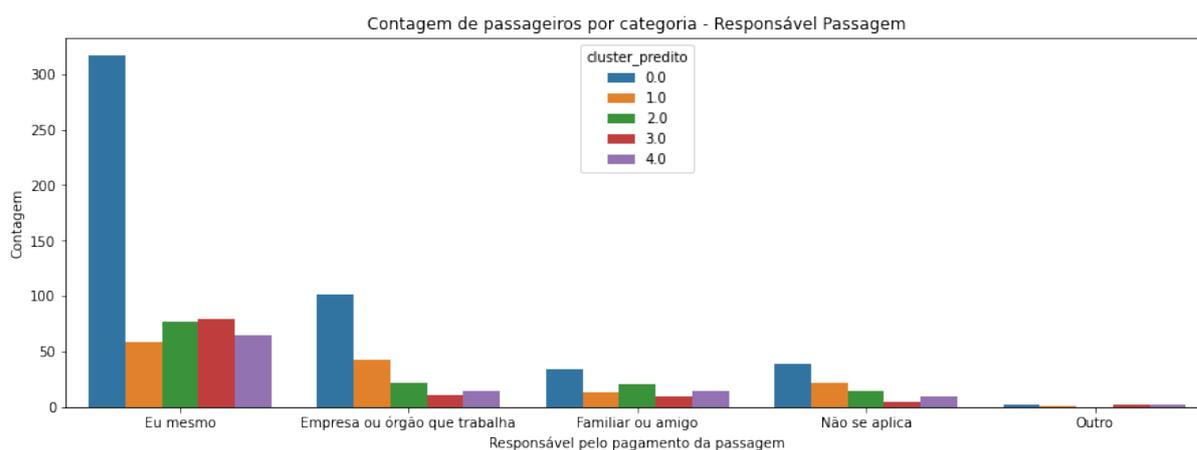


FIGURA 4.4 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável

Para a compra de passagem e seu responsável, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 2º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 3º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 4º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

- 5º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

4.1.4 Profissão, Renda familiar mensal bruta e Idade

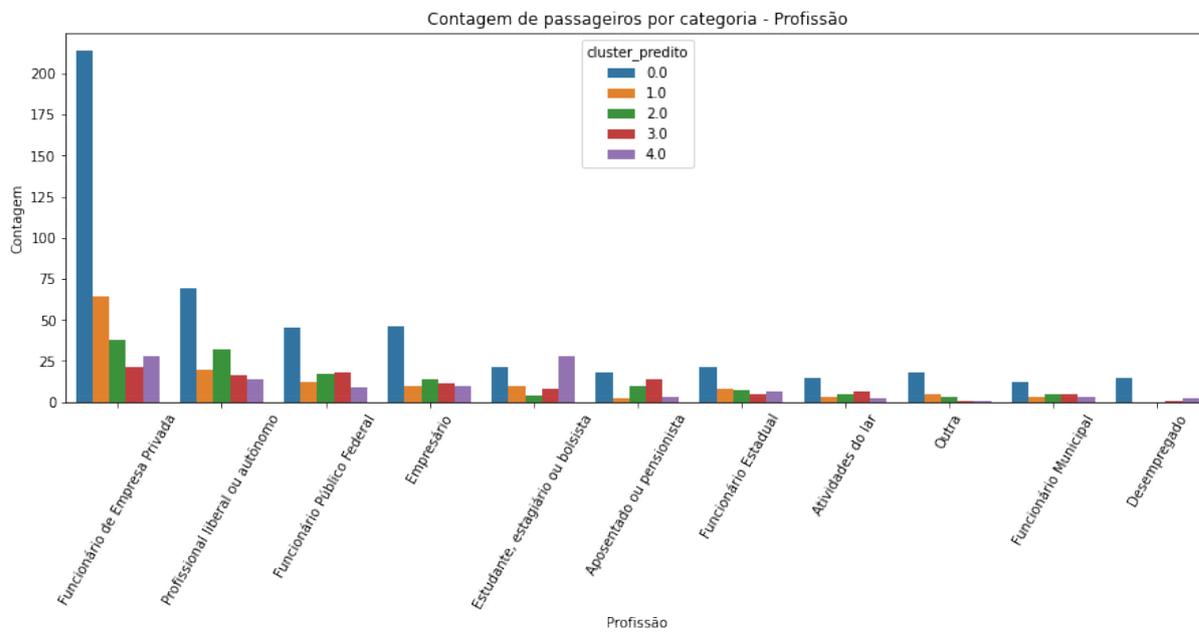


FIGURA 4.5 – Número de passageiros por cluster dentro de cada segmentação - Profissão

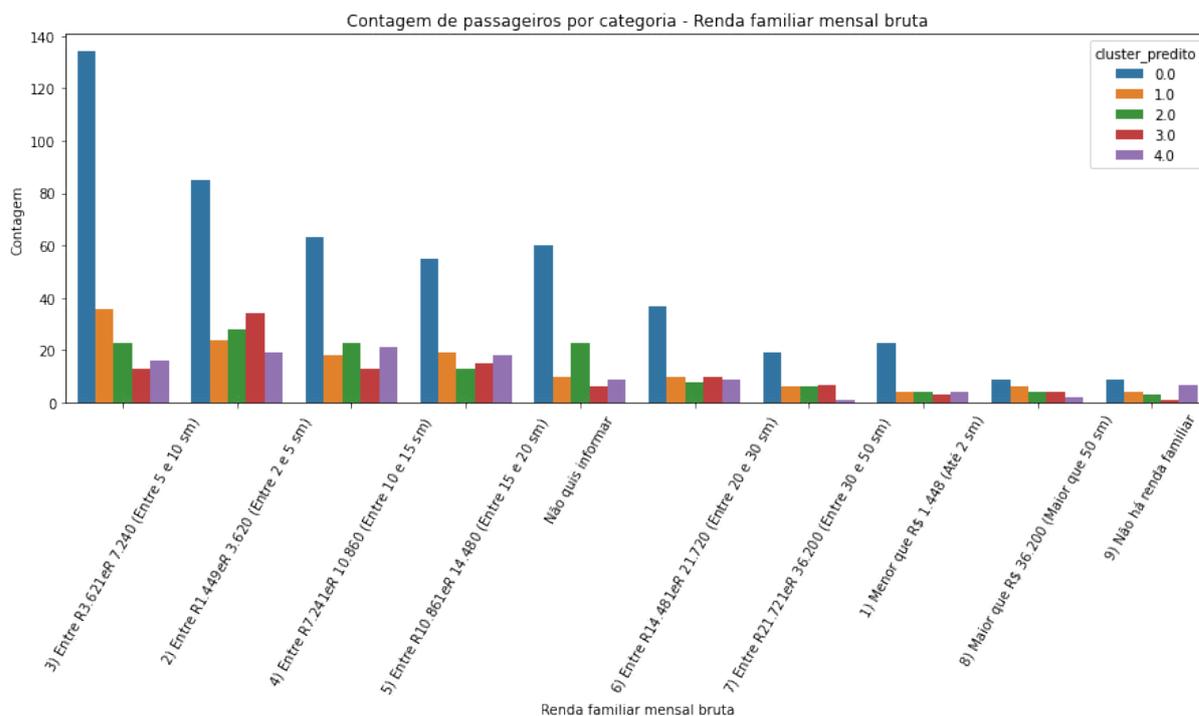


FIGURA 4.6 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta

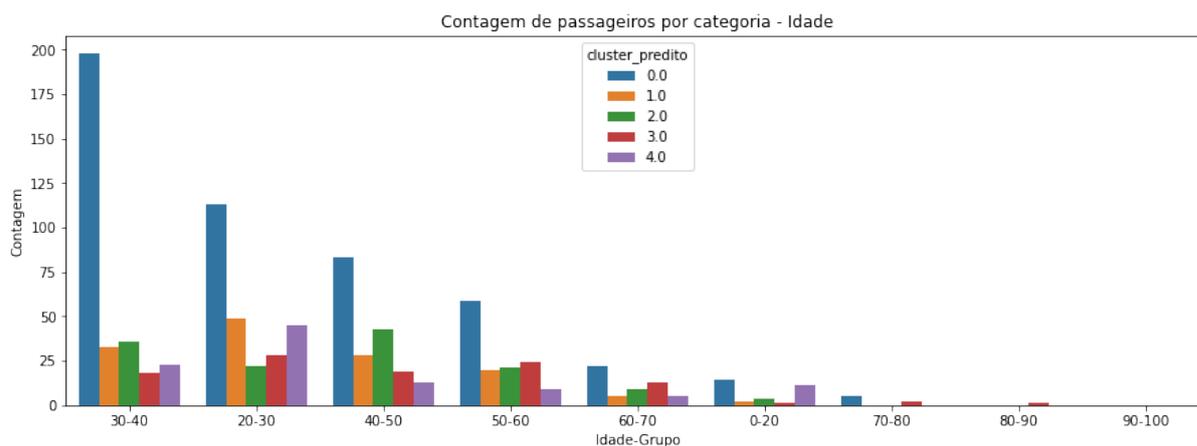


FIGURA 4.7 – Número de passageiros por cluster dentro de cada segmentação - Idade

Para os aspectos de renda, profissão e idade, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 2^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta não foi informada e na faixa de idade de 20 a 30 anos.
- 3^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 40 a 50 anos.
- 4^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 5^o Cluster: reúne passageiros que são estudantes, estagiários ou bolsistas, cuja renda familiar mensal bruta está entre 10 e 15 salários mínimos e na faixa de idade de 20 a 30 anos.

4.1.5 Uso do transporte aéreo



FIGURA 4.8 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano

Para os aspectos de uso do transporte aéreo, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.
- 2º Cluster: reúne passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.
- 3º Cluster: reúne passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.
- 4º Cluster: reúne passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.
- 5º Cluster: reúne passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.

4.1.6 Gasto no aeroporto

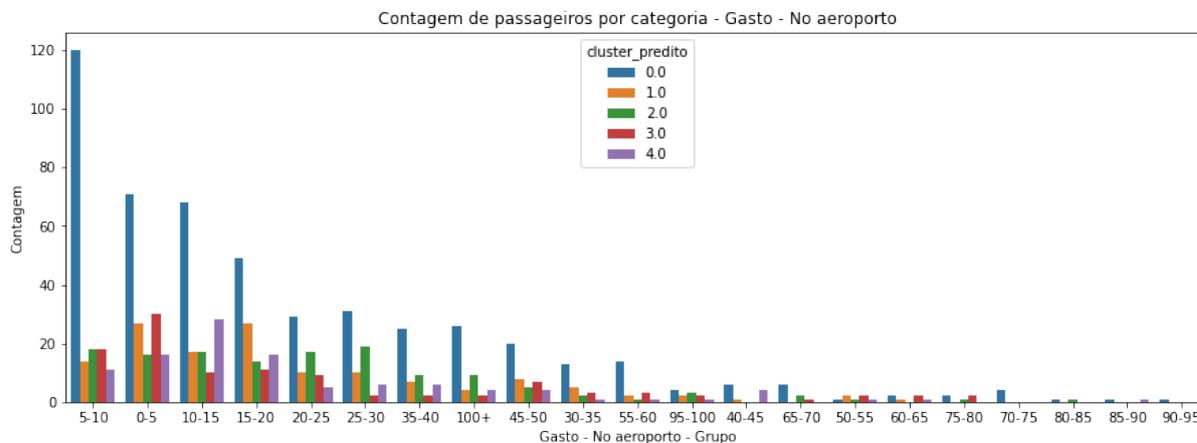


FIGURA 4.9 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto

Para o gasto realizado no aeroporto, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que gastaram entre R\$5 e R\$10.
- 2º Cluster: reúne passageiros que gastaram entre R\$15 e R\$20.
- 3º Cluster: reúne passageiros que gastaram entre R\$25 e R\$30.
- 4º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 5º Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.

4.1.7 Antecedência de compra da passagem em dias

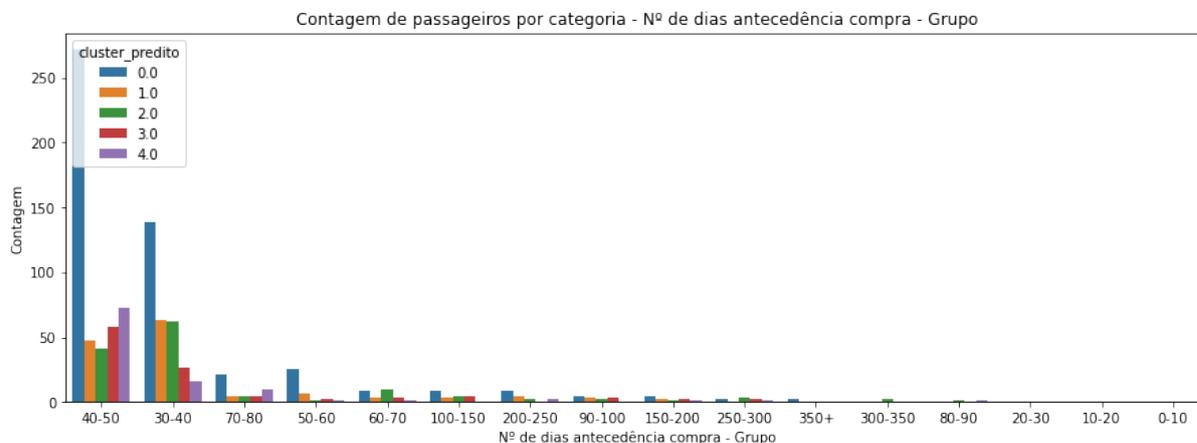


FIGURA 4.10 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias

Para a antecedência de compra da passagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do voo.
- 2º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do voo.
- 3º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do voo.
- 4º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do voo.
- 5º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do voo.

4.1.8 Fileira e coluna dos assentos

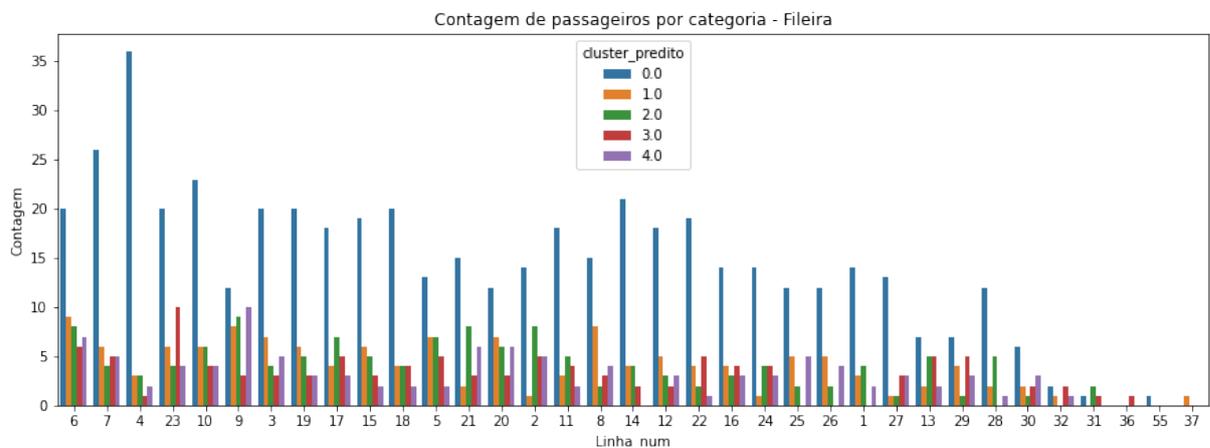


FIGURA 4.11 – Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos

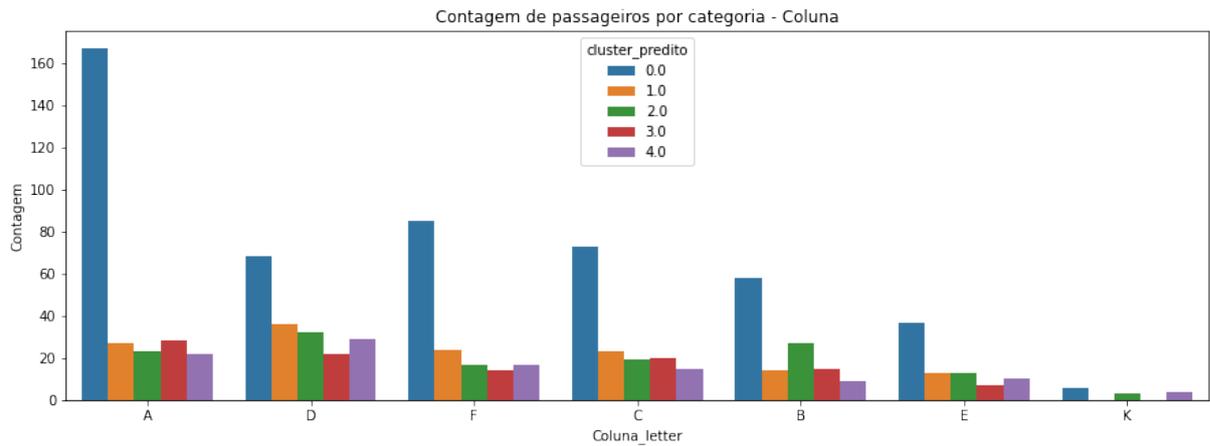


FIGURA 4.12 – Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos

Para a fileira e coluna dos assentos, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileiras 4, 7 e 10.
- 2º Cluster: reúne passageiros cujo assento era majoritariamente distribuído nas colunas D e A e fileiras 6, 9 e 8.
- 3º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D, B e A e fileiras 2, 9 e 21.
- 4º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileiras 23 e 6.
- 5º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A, C e D e fileiras 9 e 6.

4.1.9 5 Clusters - K-Modes - Resumo

TABELA 4.2 – Perfil identificado - 1° Cluster - K-Modes

Variável	1° Cluster - 494 pessoas
Método utilizado para fazer check-in	61,00% - Balcão
Motivo da viagem	35,00% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	90,00% - Comprada
Responsável pelo pagamento da passagem	64,00% - Eu mesmo
Profissão	43,00% - Funcionário de Empresa Privada
Renda familiar mensal bruta	27,00% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	7,00% - 4
Coluna do assento	34,00% - A
Idade-Grupo	40,00% - 30-40
Gasto - No aeroporto - Grupo	24,00% - 5-10
Número de vezes que realizou o mesmo trecho no último ano - Grupo	88,00% - 0-10
Nº de dias antecedência compra - Grupo	55,00% - 40-50

TABELA 4.3 – Perfil identificado - 2° Cluster - K-Modes

Variável	2° Cluster - 137 pessoas
Método utilizado para fazer check-in	40,00% - Internet
Motivo da viagem	33,00% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	80,00% - Comprada
Responsável pelo pagamento da passagem	42,00% - Eu mesmo
Profissão	47,00% - Funcionário de Empresa Privada
Renda familiar mensal bruta	26,00% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	7,00% - 6
Coluna do assento	26,00% - D
Idade-Grupo	36,00% - 20-30
Gasto - No aeroporto - Grupo	20,00% - 15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	83,00% - 0-10
Nº de dias antecedência compra - Grupo	46,00% - 30-40

TABELA 4.4 – Perfil identificado - 3º Cluster - K-Modes

Variável	3º Cluster - 135 pessoas
Método utilizado para fazer check-in	69,00% - Balcão
Motivo da viagem	46,00% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,00% - Comprada
Responsável pelo pagamento da passagem	57,00% - Eu mesmo
Profissão	28,00% - Funcionário de Empresa Privada
Renda familiar mensal bruta	21,00% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	7,00% - 9
Coluna do assento	24,00% - D
Idade-Grupo	32,00% - 40-50
Gasto - No aeroporto - Grupo	14,00% - 25-30
Número de vezes que realizou o mesmo trecho no último ano - Grupo	90,00% - 0-10
Nº de dias antecedência compra - Grupo	46,00% - 30-40

TABELA 4.5 – Perfil identificado - 4º Cluster - K-Modes

Variável	4º Cluster - 106 pessoas
Método utilizado para fazer check-in	42,00% - Balcão
Motivo da viagem	44,00% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	94,00% - Comprada
Responsável pelo pagamento da passagem	75,00% - Eu mesmo
Profissão	20,00% - Funcionário de Empresa Privada
Renda familiar mensal bruta	32,00% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	9,00% - 23
Coluna do assento	26,00% - A
Idade-Grupo	26,00% - 20-30
Gasto - No aeroporto - Grupo	28,00% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,00% - 0-10
Nº de dias antecedência compra - Grupo	55,00% - 40-50

TABELA 4.6 – Perfil identificado - 5° Cluster - K-Modes

Variável	5° Cluster - 106 pessoas
Método utilizado para fazer check-in	57,00% - Balcão
Motivo da viagem	46,00% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	90,00% - Comprada
Responsável pelo pagamento da passagem	61,00% - Eu mesmo
Profissão	26,00% - Estudante, estagiário ou bolsista
Renda familiar mensal bruta	20,00% - 4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	9,00% - 9
Coluna do assento	27,00% - D
Idade-Grupo	42,00% - 20-30
Gasto - No aeroporto - Grupo	26,00% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	95,00% - 0-10
Nº de dias antecedência compra - Grupo	69,00% - 40-50

4.2 5 Clusters - K-Prototypes

Como resultado, apresentam-se os centróides obtidos:

TABELA 4.7 – Perfil identificado - 1° Cluster - K-Prototypes

Variável	1° Cluster - 333 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	6
Coluna do assento	A
Idade	32
Gasto - No aeroporto	36.20
Número de vezes que realizou o mesmo trecho no último ano	4
Nº de dias antecedência compra	54

TABELA 4.8 – Perfil identificado - 2° Cluster - K-Prototypes

Variável	2° Cluster - 119 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	12
Coluna do assento	D
Idade	34
Gasto - No aeroporto	35.74
Número de vezes que realizou o mesmo trecho no último ano	15
Nº de dias antecedência compra	57

TABELA 4.9 – Perfil identificado - 3° Cluster - K-Prototypes

Variável	3° Cluster - 172 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	23
Coluna do assento	F
Idade	38
Gasto - No aeroporto	28.12
Número de vezes que realizou o mesmo trecho no último ano	4
Nº de dias antecedência compra	59

TABELA 4.10 – Perfil identificado - 4° Cluster - K-Prototypes

Variável	4° Cluster - 158 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	3
Coluna do assento	C
Idade	55
Gasto - No aeroporto	32.05
Número de vezes que realizou o mesmo trecho no último ano	4
Nº de dias antecedência compra	54

TABELA 4.11 – Perfil identificado - 5° Cluster - K-Prototypes

Variável	5° Cluster - 196 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	6
Coluna do assento	D
Idade	40
Gasto - No aeroporto	34.30
Número de vezes que realizou o mesmo trecho no último ano	7
Nº de dias antecedência compra	56

4.3 7 Clusters - K-Modes

TABELA 4.12 – Número de passageiros por cluster - N=7 - K-modes

Cluster	Contagem de passageiros
1º	410
2º	127
3º	116
4º	74
5º	98
6º	90
7º	63

4.3.1 Método de Check-in

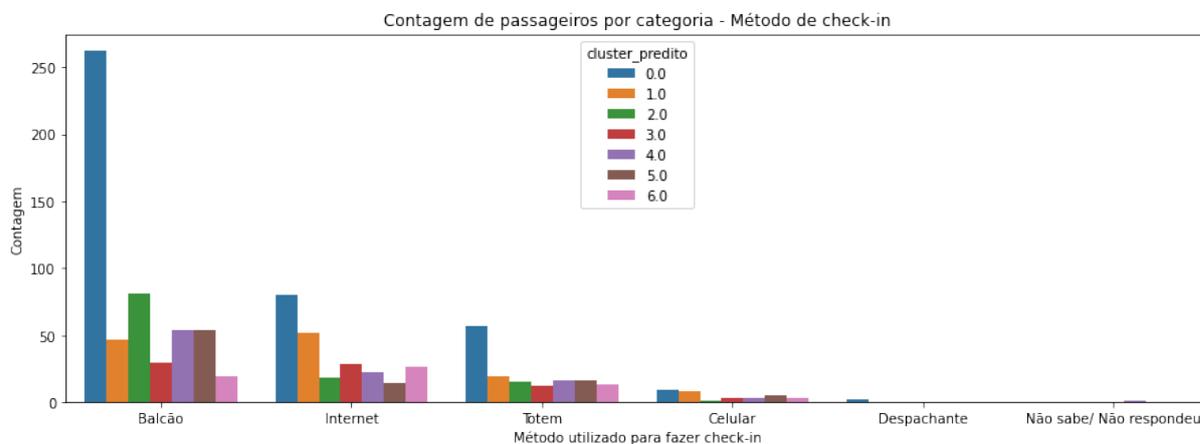


FIGURA 4.13 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in

Para o método de check-in utilizado, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 2º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 3º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 4º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 5º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 6º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 7º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet para realizar o Check-in.

4.3.2 Motivo da viagem

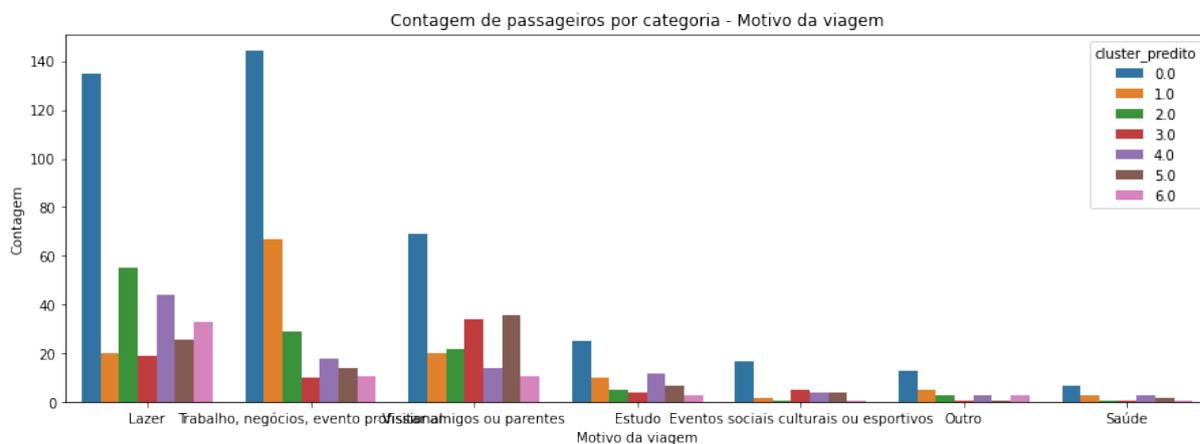


FIGURA 4.14 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem

Para a motivação da viagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais/Lazer.
- 2º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais.
- 3º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 4º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes.
- 5º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 6º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes.
- 7º Cluster: reúne passageiros cuja motivação para viajar é Lazer.

4.3.3 Compra de passagem e Responsável

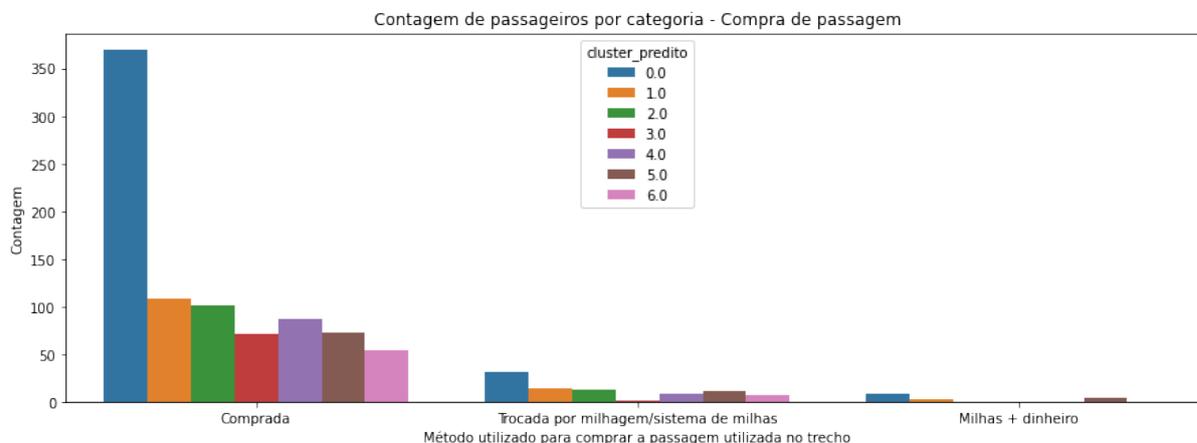


FIGURA 4.15 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método

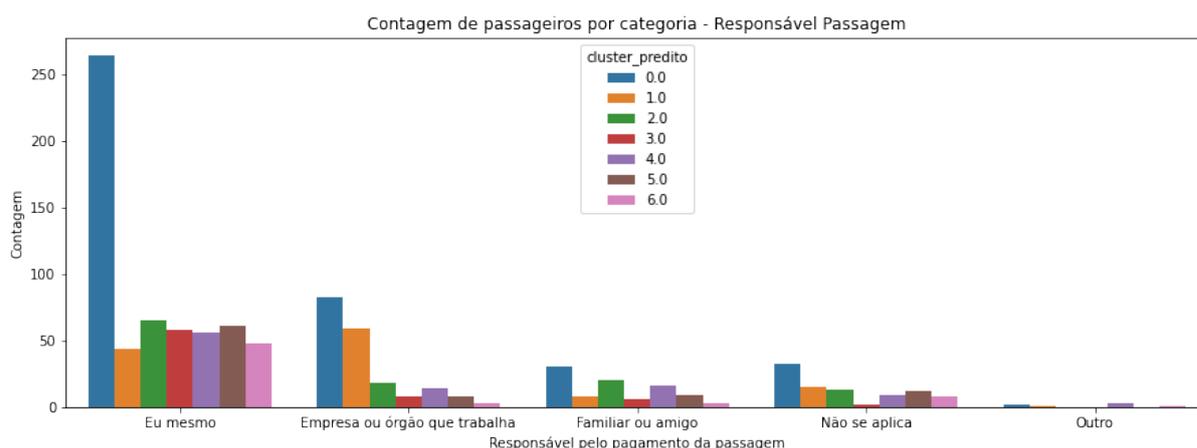


FIGURA 4.16 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável

Para a compra de passagem e seu responsável, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 2º Cluster: reúne passageiros que tiveram a passagem comprada pelo órgão ou empresa na qual trabalham sem auxílio de milhas.
- 3º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 4º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

- 5º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 6º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 7º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

4.3.4 Profissão, Renda familiar mensal bruta e Idade

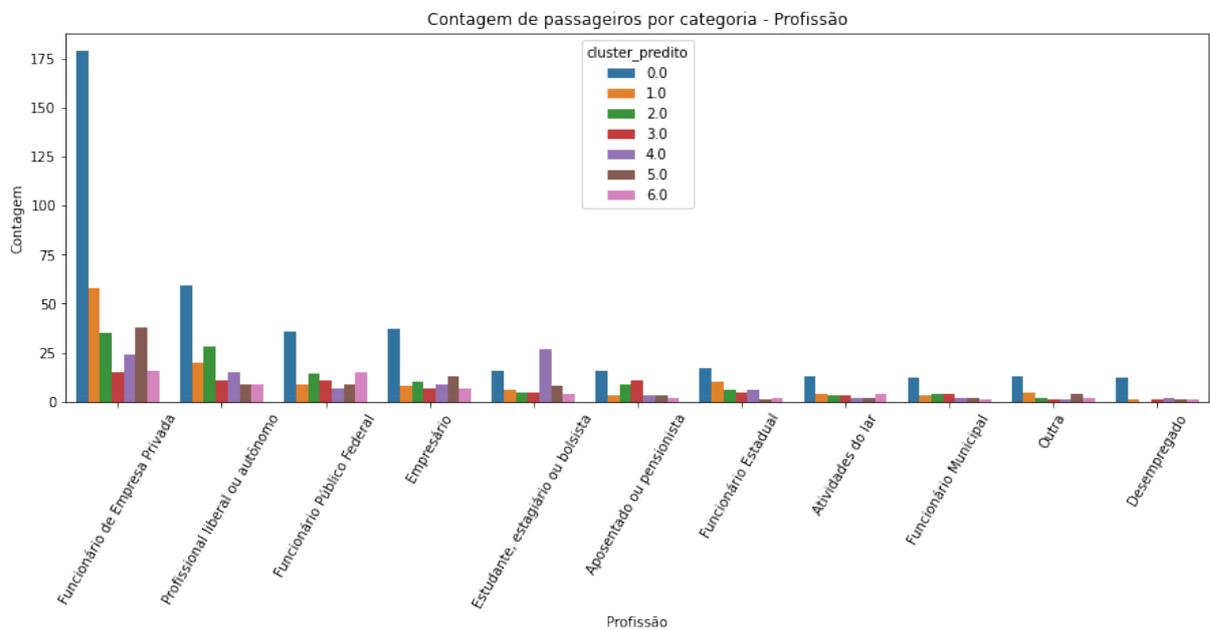


FIGURA 4.17 – Número de passageiros por cluster dentro de cada segmentação - Profissão

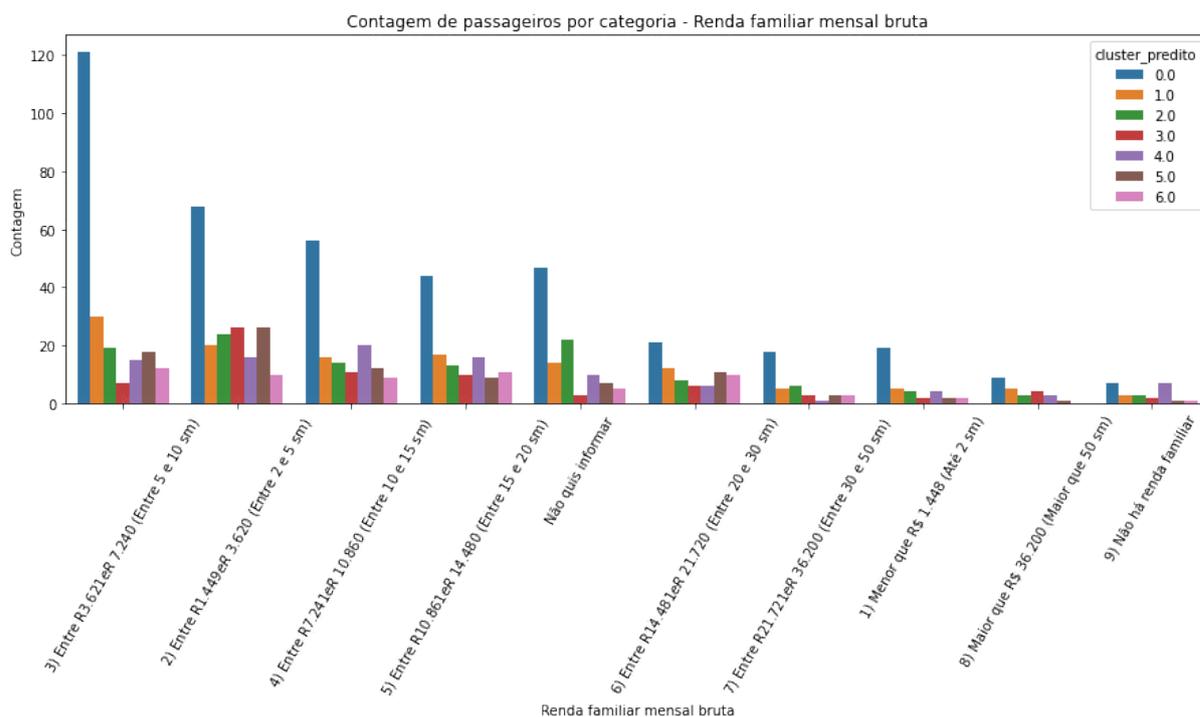


FIGURA 4.18 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta

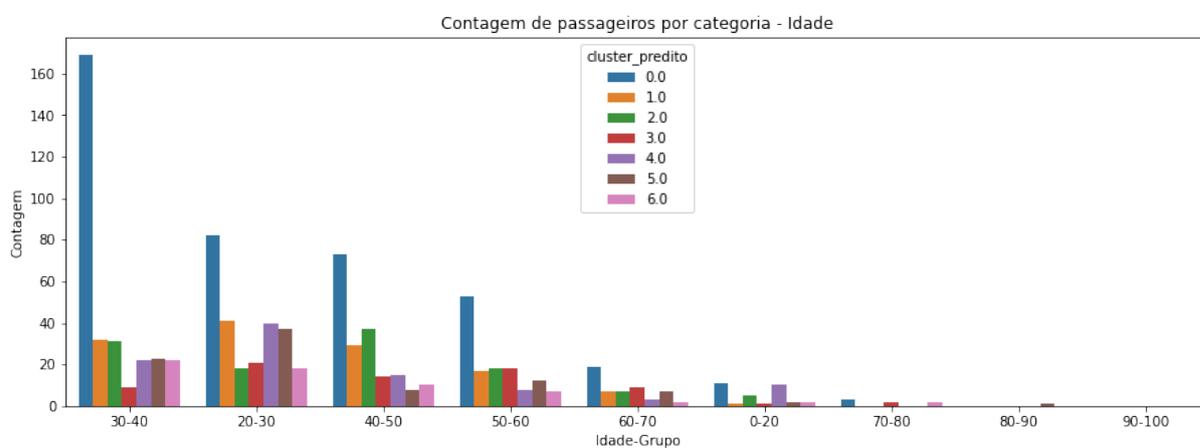


FIGURA 4.19 – Número de passageiros por cluster dentro de cada segmentação - Idade

Para os aspectos de renda, profissão e idade, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 2º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 20 a 30 anos.
- 3º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 40 a 50 anos.

- 4º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 5º Cluster: reúne passageiros que são estudantes, estagiários ou bolsistas, cuja renda familiar mensal bruta está entre 10 e 15 salários mínimos e na faixa de idade de 20 a 30 anos.
- 6º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 7º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.

4.3.5 Uso do transporte aéreo



FIGURA 4.20 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano

Para os aspectos de uso do transporte aéreo, tem-se a seguinte observação:

- Todos os clusters: reúnem passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.

4.3.6 Gasto no aeroporto

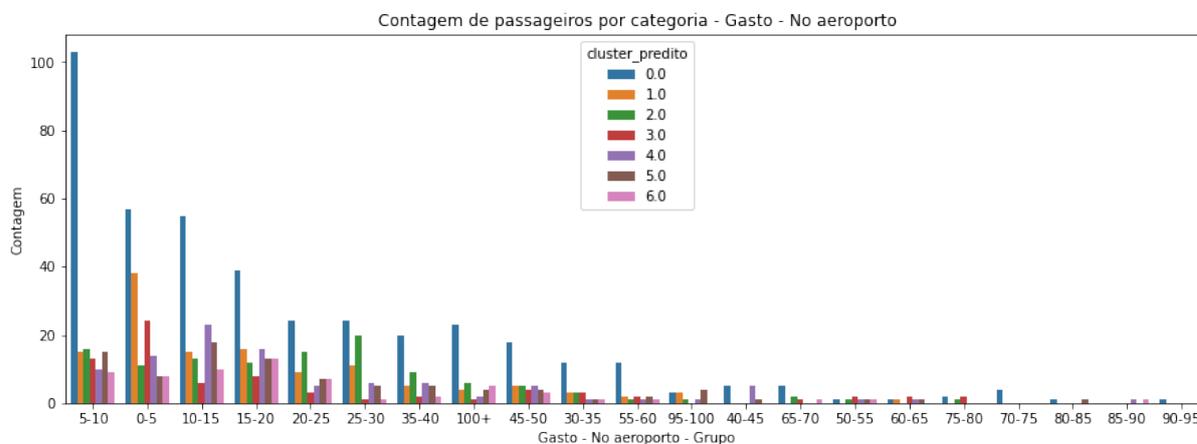


FIGURA 4.21 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto

Para o gasto realizado no aeroporto, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros que gastaram entre R\$5 e R\$10.
- 2^o Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 3^o Cluster: reúne passageiros que gastaram entre R\$25 e R\$30.
- 4^o Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 5^o Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 6^o Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 7^o Cluster: reúne passageiros que gastaram entre R\$15 e R\$20.

4.3.7 Antecedência de compra da passagem em dias

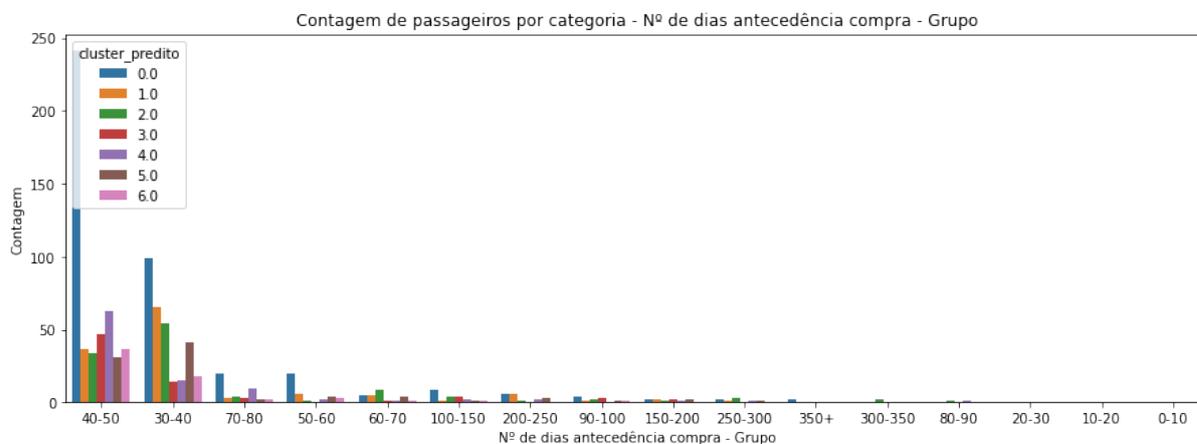


FIGURA 4.22 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias

Para a antecedência de compra da passagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 2º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 3º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 4º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 5º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 6º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 7º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.

4.3.8 Fileira e coluna dos assentos

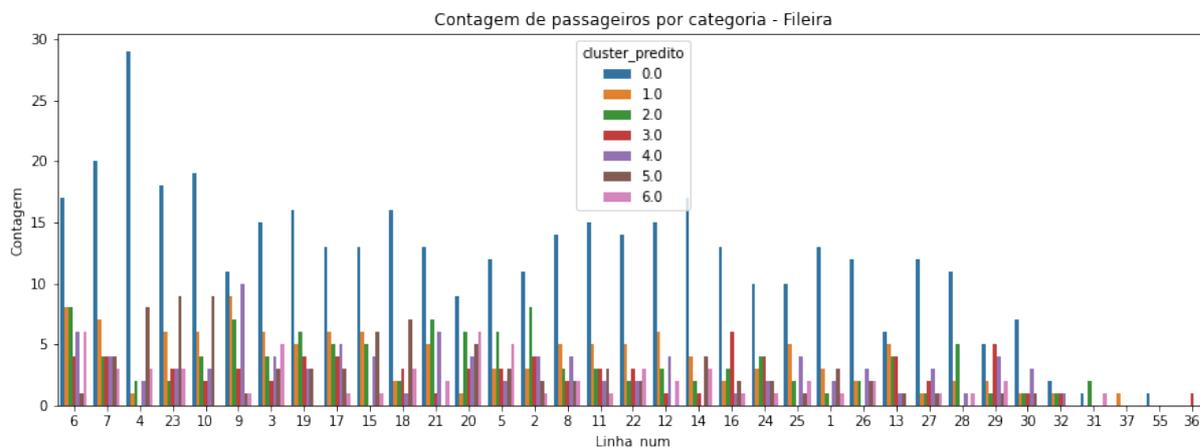


FIGURA 4.23 – Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos

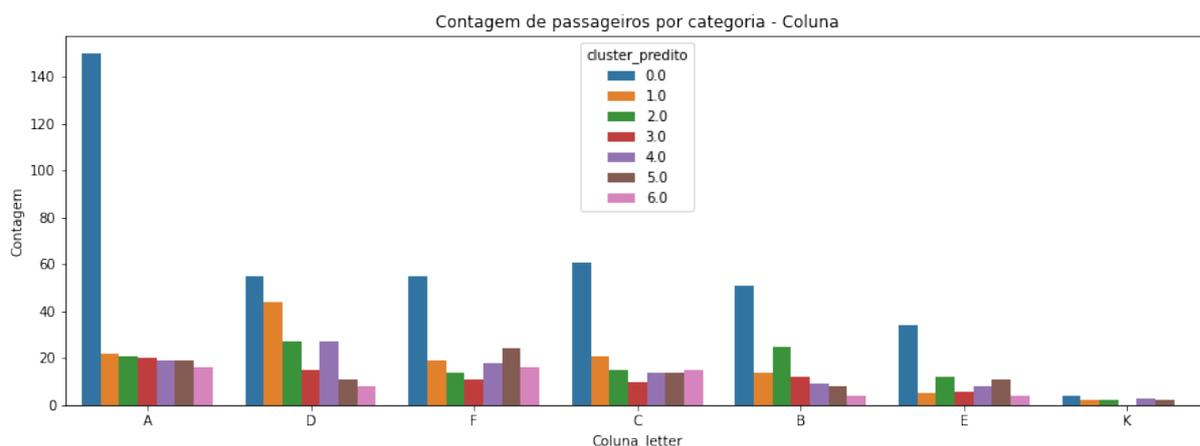


FIGURA 4.24 – Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos

Para a antecedência de compra da passagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileiras 4, 7 e 10.
- 2º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileiras 6, 9 e 7.
- 3º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e B e fileiras 2 e 6.
- 4º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileiras 16 e 29.

- 5^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileira 9.
- 6^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileiras 10, 23 e 4.
- 7^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A, C e F e fileiras 9 e 6.

4.3.9 7 Clusters - K-Modes - Resumo

TABELA 4.13 – Perfil identificado - 1^o Cluster - K-Modes

Variável	1 ^o Cluster - 410 pessoas
Método utilizado para fazer check-in	63,9% - Balcão
Motivo da viagem	35,12% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	90% - Comprada
Responsável pelo pagamento da passagem	64,39% - Eu mesmo
Profissão	43,66% - Funcionário de Empresa Privada
Renda familiar mensal bruta	29,51% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	7,07% - 4
Coluna do assento	36,59% - A
Idade-Grupo	41,22% - 30-40
Gasto - No aeroporto - Grupo	25,12% - 44474
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,27% - 0-10
N ^o de dias antecedência compra - Grupo	58,78% - 40-50

TABELA 4.14 – Perfil identificado - 2° Cluster - K-Modes

Variável	2° Cluster - 127 pessoas
Método utilizado para fazer check-in	40,94% - Internet
Motivo da viagem	52,76% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	85,83% - Comprada
Responsável pelo pagamento da passagem	46,46% - Empresa ou órgão que trabalha
Profissão	45,67% - Funcionário de Empresa Privada
Renda familiar mensal bruta	23,62% - 3) Entre R3.621eR 7.240 (Entre 5 e 10 sm)
Linha do assento	7,09% - 9
Coluna do assento	34,65% - D
Idade-Grupo	32,28% - 20-30
Gasto - No aeroporto - Grupo	29,92% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	81,1% - 0-10
Nº de dias antecedência compra - Grupo	51,18% - 30-40

TABELA 4.15 – Perfil identificado - 3° Cluster - K-Modes

Variável	3° Cluster - 116 pessoas
Método utilizado para fazer check-in	69,83% - Balcão
Motivo da viagem	47,41% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,93% - Comprada
Responsável pelo pagamento da passagem	56,03% - Eu mesmo
Profissão	30,17% - Funcionário de Empresa Privada
Renda familiar mensal bruta	20,69% - 2) Entre R1.449eR 3.620 (Entre 2 e 5 sm)
Linha do assento	6,9% - 2
Coluna do assento	23,28% - D
Idade-Grupo	31,9% - 40-50
Gasto - No aeroporto - Grupo	17,24% - 25-30
Número de vezes que realizou o mesmo trecho no último ano - Grupo	90,52% - 0-10
Nº de dias antecedência compra - Grupo	46,55% - 30-40

TABELA 4.16 – Perfil identificado - 4° Cluster - K-Modes

Variável	4° Cluster - 74 pessoas
Método utilizado para fazer check-in	40,54% - Balcão
Motivo da viagem	45,95% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	95,95% - Comprada
Responsável pelo pagamento da passagem	78,38% - Eu mesmo
Profissão	20,27% - Funcionário de Empresa Privada
Renda familiar mensal bruta	35,14% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	8,11% - 16
Coluna do assento	27,03% - A
Idade-Grupo	28,38% - 20-30
Gasto - No aeroporto - Grupo	32,43% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	93,24% - 0-10
Nº de dias antecedência compra - Grupo	63,51% - 40-50

TABELA 4.17 – Perfil identificado - 5° Cluster - K-Modes

Variável	5° Cluster - 98 pessoas
Método utilizado para fazer check-in	55,1% - Balcão
Motivo da viagem	44,9% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	89,8% - Comprada
Responsável pelo pagamento da passagem	57,14% - Eu mesmo
Profissão	27,55% - Estudante, estagiário ou bolsista
Renda familiar mensal bruta	20,41% - 4) Entre R7.241eR 10.860 (Entre 10 e 15 sm)
Linha do assento	10,2% - 9
Coluna do assento	27,55% - D
Idade-Grupo	40,82% - 20-30
Gasto - No aeroporto - Grupo	23,47% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	95,92% - 0-10
Nº de dias antecedência compra - Grupo	64,29% - 40-50

TABELA 4.18 – Perfil identificado - 6° Cluster - K-Modes

Variável	6° Cluster - 90 pessoas
Método utilizado para fazer check-in	60% - Balcão
Motivo da viagem	40% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	81,11% - Comprada
Responsável pelo pagamento da passagem	67,78% - Eu mesmo
Profissão	42,22% - Funcionário de Empresa Privada
Renda familiar mensal bruta	28,89% - 2) Entre R1.449eR 3.620 (Entre 2 e 5 sm)
Linha do assento	10,11% - 23
Coluna do assento	26,97% - F
Idade-Grupo	41,11% - 20-30
Gasto - No aeroporto - Grupo	20% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	84,44% - 0-10
Nº de dias antecedência compra - Grupo	45,56% - 30-40

TABELA 4.19 – Perfil identificado - 7° Cluster - K-Modes

Variável	7° Cluster - 63 pessoas
Método utilizado para fazer check-in	42,86% - Internet
Motivo da viagem	52,38% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,3% - Comprada
Responsável pelo pagamento da passagem	76,19% - Eu mesmo
Profissão	25,4% - Funcionário de Empresa Privada
Renda familiar mensal bruta	19,05% - 3) Entre R3.621eR 7.240 (Entre 5 e 10 sm)
Linha do assento	9,52% - 20
Coluna do assento	25,4% - A
Idade-Grupo	34,92% - 30-40
Gasto - No aeroporto - Grupo	20,63% - 15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	85,71% - 0-10
Nº de dias antecedência compra - Grupo	58,73% - 40-50

4.4 7 Clusters - K-Prototypes

Como resultado inicial, apresentam-se os centróides obtidos:

TABELA 4.20 – Perfil identificado - 1° Cluster - K-Prototypes

Variável	1° Cluster - 266 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	6
Coluna do assento	A
Idade-Grupo	32
Gasto - No aeroporto - Grupo	34.55
Número de vezes que realizou o mesmo trecho no último ano - Grupo	4
Nº de dias antecedência compra - Grupo	53

TABELA 4.21 – Perfil identificado - 2° Cluster - K-Prototypes

Variável	2° Cluster - 84 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	12
Coluna do assento	D
Idade-Grupo	36
Gasto - No aeroporto - Grupo	21.75
Número de vezes que realizou o mesmo trecho no último ano - Grupo	17
Nº de dias antecedência compra - Grupo	61

TABELA 4.22 – Perfil identificado - 3° Cluster - K-Prototypes

Variável	3° Cluster - 144 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	23
Coluna do assento	F
Idade-Grupo	37
Gasto - No aeroporto - Grupo	26.41
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	63

TABELA 4.23 – Perfil identificado - 4° Cluster - K-Prototypes

Variável	4° Cluster - 103 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	7
Coluna do assento	C
Idade-Grupo	43
Gasto - No aeroporto - Grupo	29.3
Número de vezes que realizou o mesmo trecho no último ano - Grupo	7
Nº de dias antecedência compra - Grupo	46

TABELA 4.24 – Perfil identificado - 5° Cluster - K-Prototypes

Variável	5° Cluster - 177 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	17
Coluna do assento	D
Idade-Grupo	39
Gasto - No aeroporto - Grupo	32.55
Número de vezes que realizou o mesmo trecho no último ano - Grupo	6
Nº de dias antecedência compra - Grupo	58

TABELA 4.25 – Perfil identificado - 6° Cluster - K-Prototypes

Variável	6° Cluster - 111 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Aposentado ou pensionista
Renda familiar mensal bruta	Não quis informar
Linha do assento	10
Coluna do assento	A
Idade-Grupo	59
Gasto - No aeroporto - Grupo	32.1
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	62

TABELA 4.26 – Perfil identificado - 7° Cluster - K-Prototypes

Variável	7° Cluster - 93 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	22
Coluna do assento	B
Idade-Grupo	38
Gasto - No aeroporto - Grupo	62.05
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	50

4.5 8 Clusters - K-Modes

TABELA 4.27 – Número de passageiros por cluster - N=8 - K-modes

Cluster	Contagem de passageiros
1º	350
2º	134
3º	125
4º	75
5º	96
6º	75
7º	64
8º	59

4.5.1 Método de Check-in

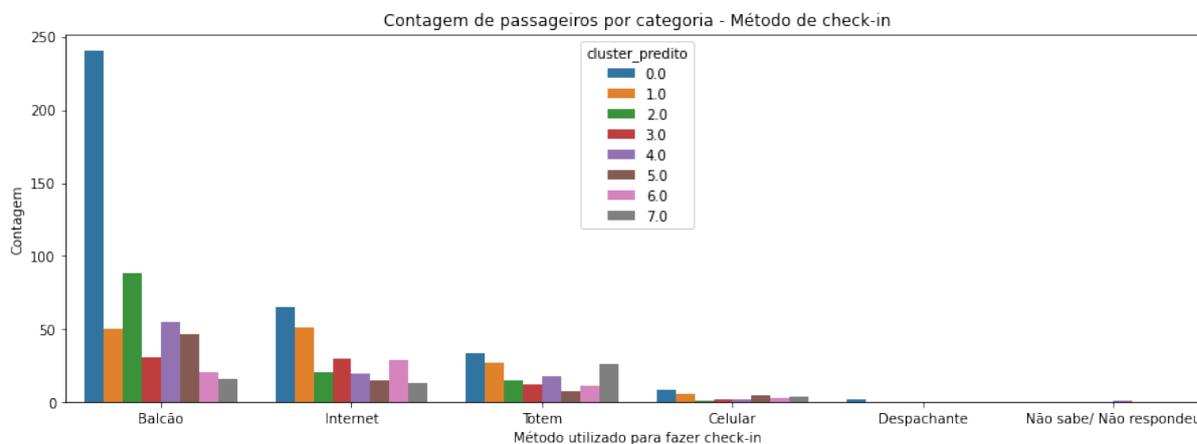


FIGURA 4.25 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in

Para o método de check-in utilizado, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 2º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 3º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 4º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 5º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 6º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 7º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet para realizar o Check-in.
- 8º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Totem para realizar o Check-in.

4.5.2 Motivo da viagem

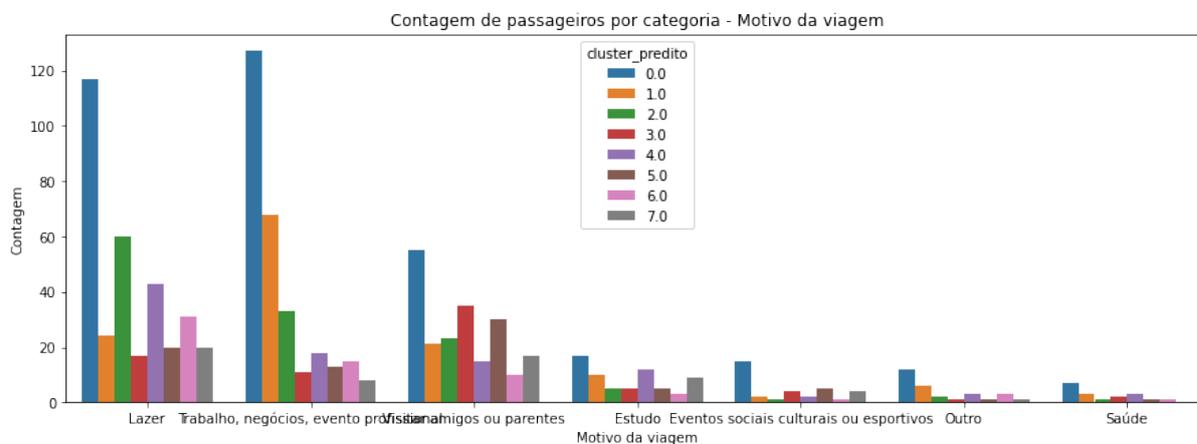


FIGURA 4.26 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem

Para a motivação da viagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais/Lazer.
- 2º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais.
- 3º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 4º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes.
- 5º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 6º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes.
- 7º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 8º Cluster: reúne passageiros cuja motivação para viajar é Lazer/visitar amigos ou parentes.

4.5.3 Compra de passagem e Responsável

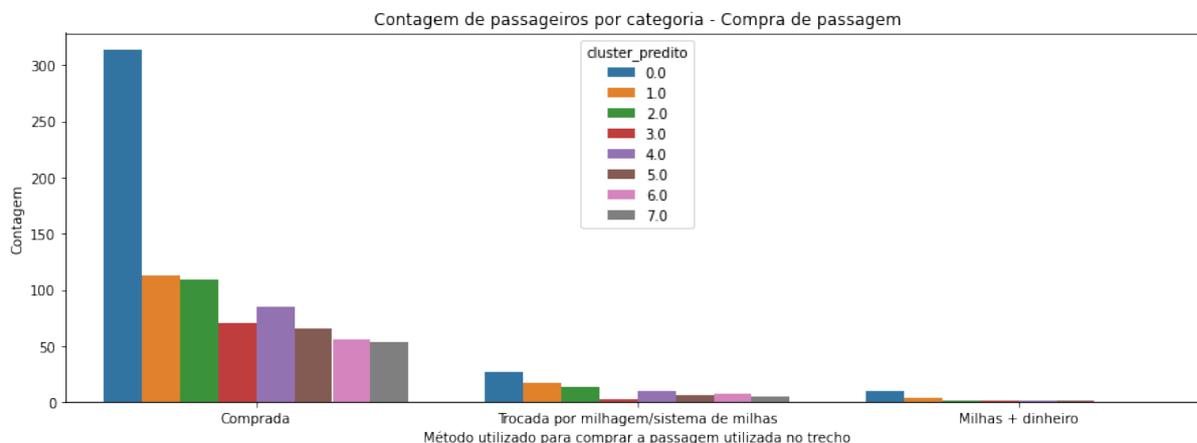


FIGURA 4.27 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método

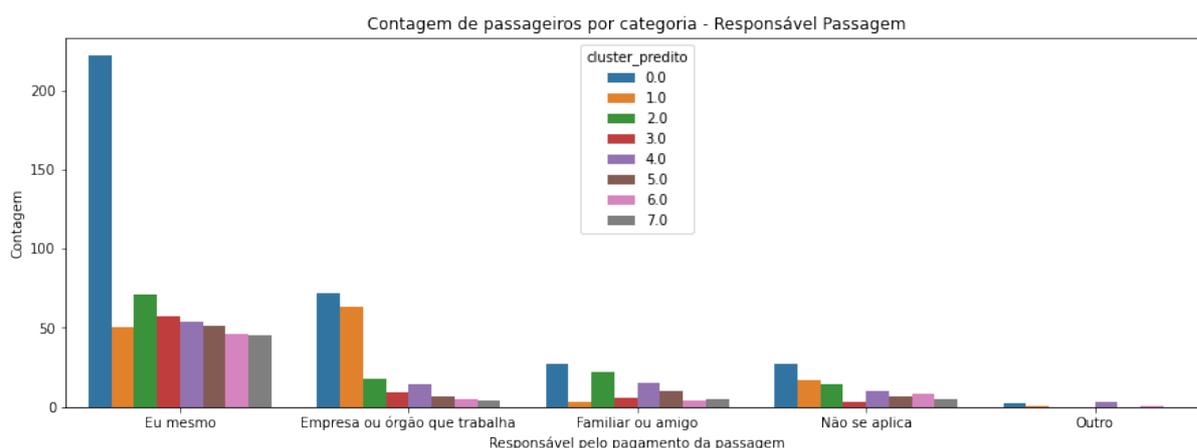


FIGURA 4.28 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável

Para a compra de passagem e seu responsável, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 2º Cluster: reúne passageiros que tiveram a passagem comprada pelo órgão ou empresa na qual trabalham sem auxílio de milhas.
- 3º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 4º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

- 5º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 6º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 7º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 8º Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.

4.5.4 Profissão, Renda familiar mensal bruta e Idade

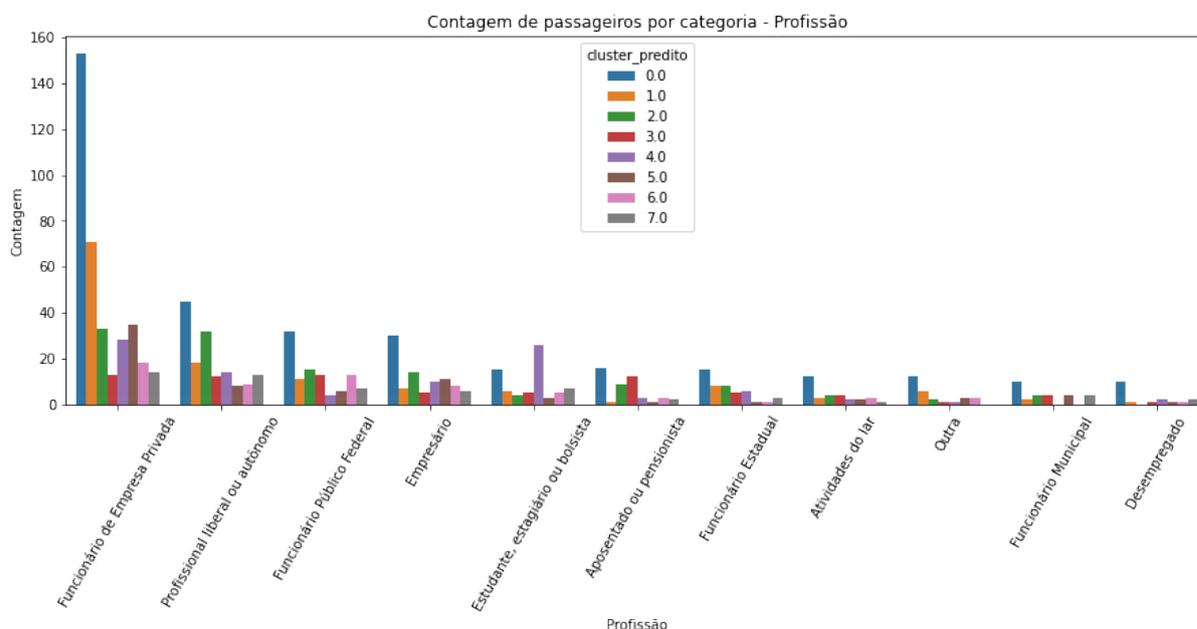


FIGURA 4.29 – Número de passageiros por cluster dentro de cada segmentação - Profissão

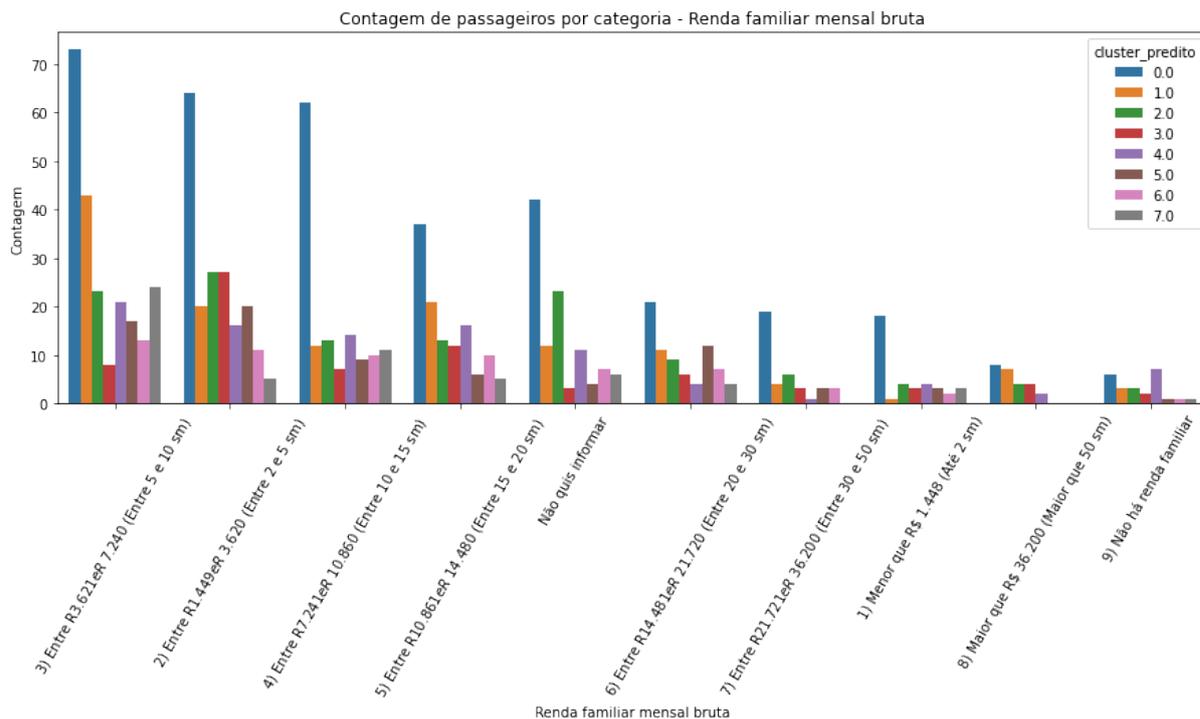


FIGURA 4.30 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta

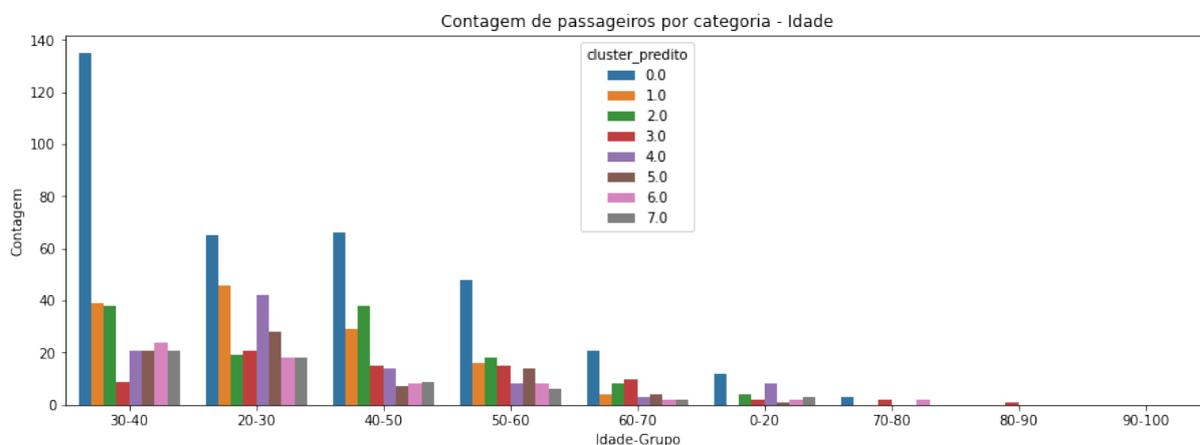


FIGURA 4.31 – Número de passageiros por cluster dentro de cada segmentação - Idade

Para os aspectos de renda, profissão e idade, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 2º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 20 a 30 anos.
- 3º Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 30 a 40 anos.

- 4^o Cluster: reúne passageiros que atuam como funcionários públicos federais, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 5^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 20 a 30 anos.
- 6^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 7^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 8^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.

4.5.5 Uso do transporte aéreo

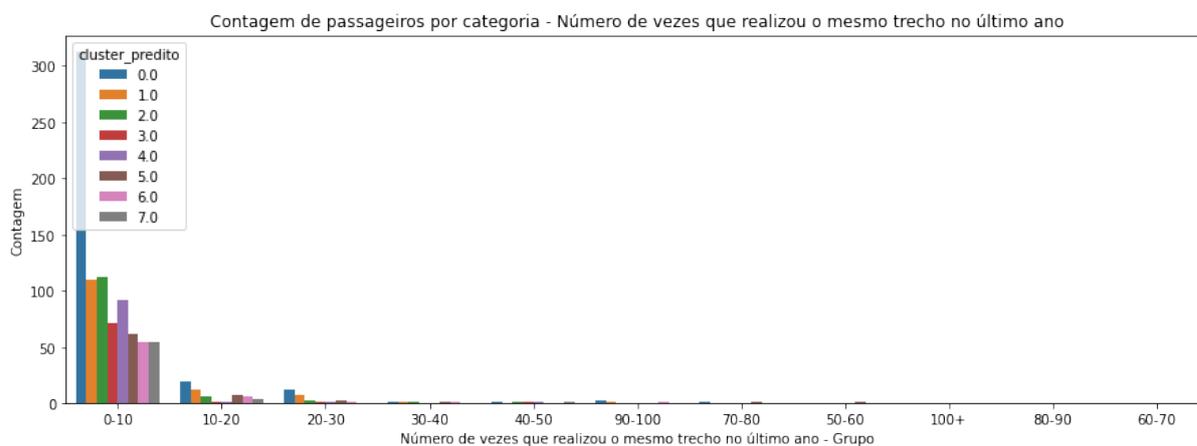


FIGURA 4.32 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano

Para os aspectos de uso do transporte aéreo, tem-se a seguinte observação:

- Todos os clusters: reúnem passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.

4.5.6 Gasto no aeroporto

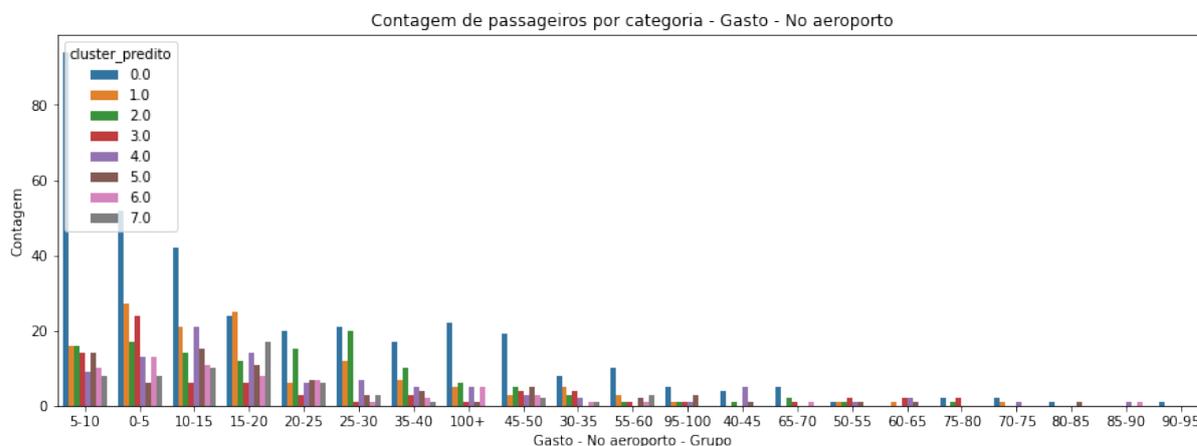


FIGURA 4.33 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto

Para o gasto realizado no aeroporto, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que gastaram entre R\$5 e R\$10.
- 2º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 3º Cluster: reúne passageiros que gastaram entre R\$25 e R\$30.
- 4º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 5º Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 6º Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 7º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 8º Cluster: reúne passageiros que gastaram entre R\$15 e R\$20.

4.5.7 Antecedência de compra da passagem em dias

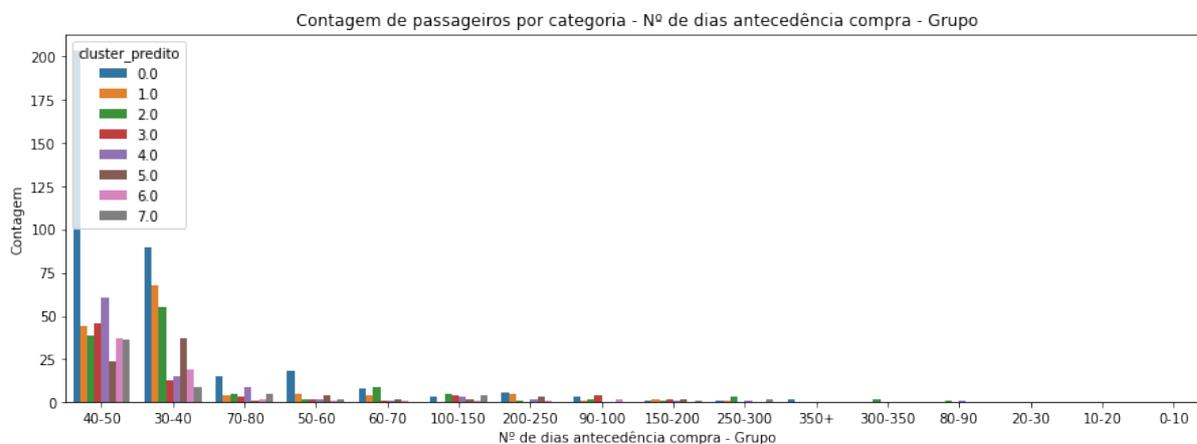


FIGURA 4.34 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias

Para a antecedência de compra da passagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 2º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 3º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 4º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 5º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 6º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 7º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 8º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.

4.5.8 Fileira e coluna dos assentos

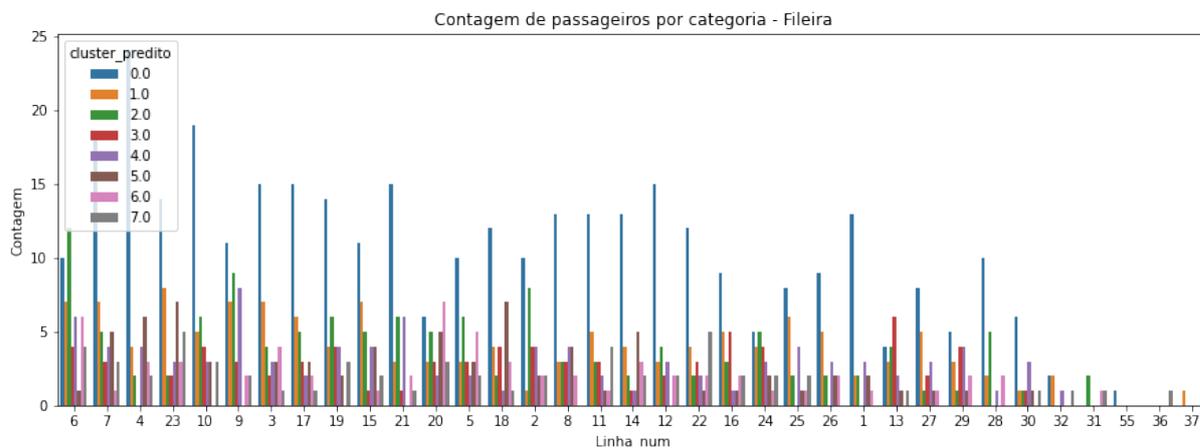


FIGURA 4.35 – Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos

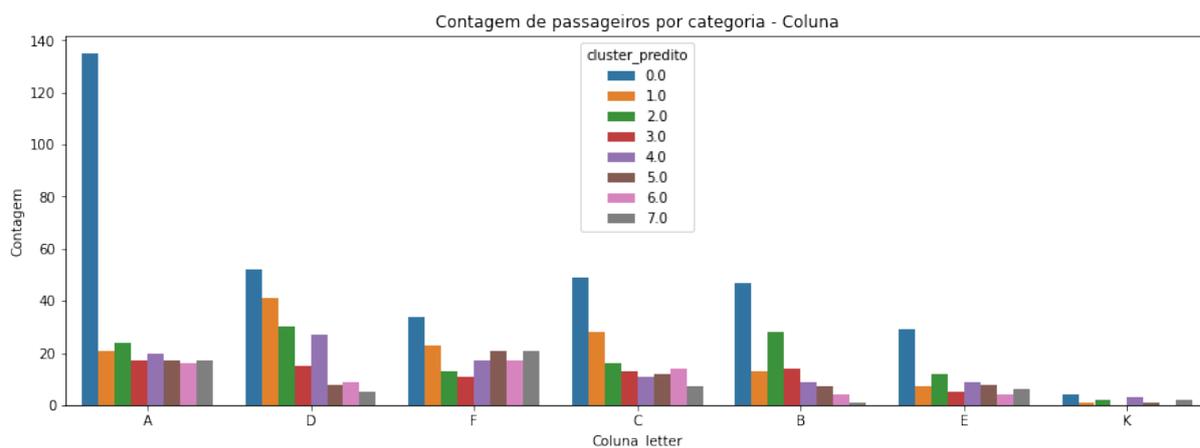


FIGURA 4.36 – Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos

Para a fileira e coluna dos assentos, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileiras 4, 7 e 10.
- 2º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileira 23.
- 3º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e B e fileiras 9 e 6.
- 4º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileira 13.

- 5º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileira 9.
- 6º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileira 18.
- 7º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e F e fileiras 20 e 5.
- 8º Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileira 22.

4.5.9 8 Clusters - K-Modes - Resumo

TABELA 4.28 – Perfil identificado - 1º Cluster - K-Modes

Variável	1º Cluster - 350 pessoas
Método utilizado para fazer check-in	68,57% - Balcão
Motivo da viagem	36,29% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	89,43% - Comprada
Responsável pelo pagamento da passagem	63,43% - Eu mesmo
Profissão	43,71% - Funcionário de Empresa Privada
Renda familiar mensal bruta	20,86% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	6,86% - 4
Coluna do assento	38,57% - A
Idade-Grupo	38,57% - 30-40
Gasto - No aeroporto - Grupo	26,86% - 44474
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,14% - 0-10
Nº de dias antecedência compra - Grupo	58% - 40-50

TABELA 4.29 – Perfil identificado - 2° Cluster - K-Modes

Variável	2° Cluster - 134 pessoas
Método utilizado para fazer check-in	38,06% - Internet
Motivo da viagem	50,75% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	84,33% - Comprada
Responsável pelo pagamento da passagem	47,01% - Empresa ou órgão que trabalha
Profissão	52,99% - Funcionário de Empresa Privada
Renda familiar mensal bruta	32,09% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	5,97% - 23
Coluna do assento	30,6% - D
Idade-Grupo	34,33% - 20-30
Gasto - No aeroporto - Grupo	20,15% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	82,09% - 0-10
Nº de dias antecedência compra - Grupo	50,75% - 30-40

TABELA 4.30 – Perfil identificado - 3° Cluster - K-Modes

Variável	3° Cluster - 125 pessoas
Método utilizado para fazer check-in	70,4% - Balcão
Motivo da viagem	48% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,2% - Comprada
Responsável pelo pagamento da passagem	56,8% - Eu mesmo
Profissão	26,4% - Funcionário de Empresa Privada
Renda familiar mensal bruta	21,6% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	9,6% - 6
Coluna do assento	24% - D
Idade-Grupo	30,4% - 30-40
Gasto - No aeroporto - Grupo	16% - 25-30
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,6% - 0-10
Nº de dias antecedência compra - Grupo	44% - 30-40

TABELA 4.31 – Perfil identificado - 4° Cluster - K-Modes

Variável	4° Cluster - 75 pessoas
Método utilizado para fazer check-in	41,33% - Balcão
Motivo da viagem	46,67% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	94,67% - Comprada
Responsável pelo pagamento da passagem	76% - Eu mesmo
Profissão	17,33% - Funcionário Público Federal
Renda familiar mensal bruta	36% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	8% - 13
Coluna do assento	22,67% - A
Idade-Grupo	28% - 20-30
Gasto - No aeroporto - Grupo	32% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	94,67% - 0-10
Nº de dias antecedência compra - Grupo	61,33% - 40-50

TABELA 4.32 – Perfil identificado - 5° Cluster - K-Modes

Variável	5° Cluster - 96 pessoas
Método utilizado para fazer check-in	57,29% - Balcão
Motivo da viagem	44,79% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	88,54% - Comprada
Responsável pelo pagamento da passagem	56,25% - Eu mesmo
Profissão	29,17% - Funcionário de Empresa Privada
Renda familiar mensal bruta	21,88% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	8,33% - 9
Coluna do assento	28,13% - D
Idade-Grupo	43,75% - 20-30
Gasto - No aeroporto - Grupo	21,88% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	95,83% - 0-10
Nº de dias antecedência compra - Grupo	63,54% - 40-50

TABELA 4.33 – Perfil identificado - 6° Cluster - K-Modes

Variável	6° Cluster - 75 pessoas
Método utilizado para fazer check-in	62,67% - Balcão
Motivo da viagem	40% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	88% - Comprada
Responsável pelo pagamento da passagem	68% - Eu mesmo
Profissão	46,67% - Funcionário de Empresa Privada
Renda familiar mensal bruta	26,67% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	9,46% - 18
Coluna do assento	28,38% - F
Idade-Grupo	37,33% - 20-30
Gasto - No aeroporto - Grupo	20% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	82,67% - 0-10
Nº de dias antecedência compra - Grupo	49,33% - 30-40

TABELA 4.34 – Perfil identificado - 7° Cluster - K-Modes

Variável	7° Cluster - 64 pessoas
Método utilizado para fazer check-in	45,31% - Internet
Motivo da viagem	48,44% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,5% - Comprada
Responsável pelo pagamento da passagem	71,88% - Eu mesmo
Profissão	28,13% - Funcionário de Empresa Privada
Renda familiar mensal bruta	20,31% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	10,94% - 20
Coluna do assento	26,56% - F
Idade-Grupo	37,5% - 30-40
Gasto - No aeroporto - Grupo	20,31% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	84,38% - 0-10
Nº de dias antecedência compra - Grupo	57,81% - 40-50

TABELA 4.35 – Perfil identificado - 8° Cluster - K-Modes

Variável	8° Cluster - 59 pessoas
Método utilizado para fazer check-in	44,07% - Totem
Motivo da viagem	33,9% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	91,53% - Comprada
Responsável pelo pagamento da passagem	76,27% - Eu mesmo
Profissão	23,73% - Funcionário de Empresa Privada
Renda familiar mensal bruta	40,68% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	8,47% - 22
Coluna do assento	35,59% - F
Idade-Grupo	35,59% - 30-40
Gasto - No aeroporto - Grupo	28,81% - 15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	91,53% - 0-10
Nº de dias antecedência compra - Grupo	61,02% - 40-50

4.6 8 Clusters - K-Prototypes

TABELA 4.36 – Perfil identificado - 1° Cluster - K-Prototypes

Variável	1° Cluster - 264 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	6
Coluna do assento	A
Idade-Grupo	35
Gasto - No aeroporto - Grupo	34.38
Número de vezes que realizou o mesmo trecho no último ano - Grupo	4
Nº de dias antecedência compra - Grupo	57

TABELA 4.37 – Perfil identificado - 2° Cluster - K-Prototypes

Variável	2° Cluster - 83 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	12
Coluna do assento	D
Idade-Grupo	40
Gasto - No aeroporto - Grupo	25.65
Número de vezes que realizou o mesmo trecho no último ano - Grupo	6
Nº de dias antecedência compra - Grupo	55

TABELA 4.38 – Perfil identificado - 3º Cluster - K-Prototypes

Variável	3º Cluster - 132 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	23
Coluna do assento	A
Idade-Grupo	55
Gasto - No aeroporto - Grupo	27.82
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	58

TABELA 4.39 – Perfil identificado - 4º Cluster - K-Prototypes

Variável	4º Cluster - 96 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	7
Coluna do assento	C
Idade-Grupo	43
Gasto - No aeroporto - Grupo	27.6
Número de vezes que realizou o mesmo trecho no último ano - Grupo	6
Nº de dias antecedência compra - Grupo	47

TABELA 4.40 – Perfil identificado - 5° Cluster - K-Prototypes

Variável	5° Cluster - 166 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	19
Coluna do assento	D
Idade-Grupo	37
Gasto - No aeroporto - Grupo	33.3
Número de vezes que realizou o mesmo trecho no último ano - Grupo	5
Nº de dias antecedência compra - Grupo	54

TABELA 4.41 – Perfil identificado - 6° Cluster - K-Prototypes

Variável	6° Cluster - 29 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	7
Coluna do assento	D
Idade-Grupo	43
Gasto - No aeroporto - Grupo	30.1
Número de vezes que realizou o mesmo trecho no último ano - Grupo	45
Nº de dias antecedência compra - Grupo	74

TABELA 4.42 – Perfil identificado - 7° Cluster - K-Prototypes

Variável	7° Cluster - 65 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	22
Coluna do assento	B
Idade-Grupo	44
Gasto - No aeroporto - Grupo	72
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	52

TABELA 4.43 – Perfil identificado - 8° Cluster - K-Prototypes

Variável	8° Cluster - 143 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	23
Coluna do assento	F
Idade-Grupo	29
Gasto - No aeroporto - Grupo	30.25
Número de vezes que realizou o mesmo trecho no último ano - Grupo	5
Nº de dias antecedência compra - Grupo	60

4.7 10 Clusters - K-Modes

TABELA 4.44 – Número de passageiros por cluster - N=10 - K-modes

Cluster	Contagem de passageiros
1 ^o	266
2 ^o	114
3 ^o	111
4 ^o	65
5 ^o	76
6 ^o	90
7 ^o	64
8 ^o	55
9 ^o	91
10 ^o	46

4.7.1 Método de Check-in

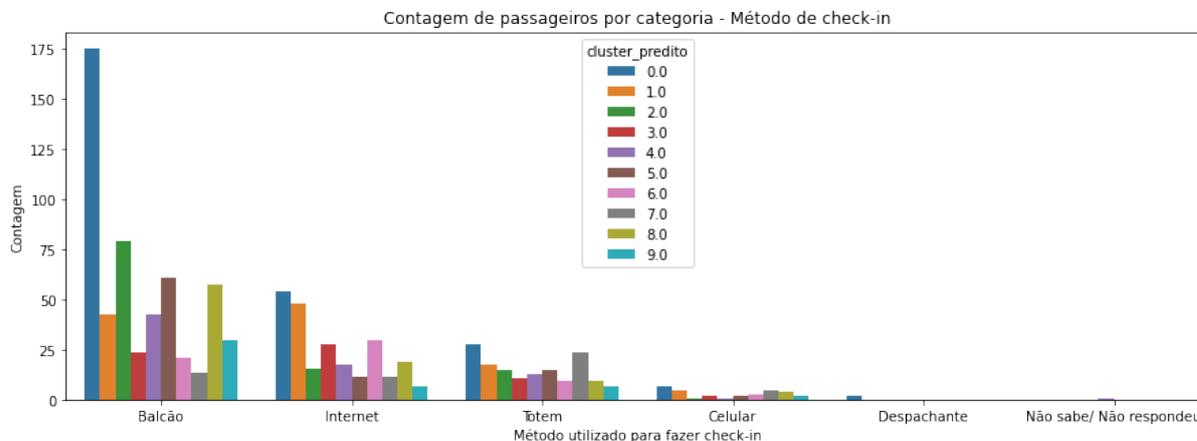


FIGURA 4.37 – Número de passageiros por cluster dentro de cada segmentação - Método de Check-in

Para o método de check-in utilizado, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 2^o Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.

- 3º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 4º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet/Balcão para realizar o Check-in.
- 5º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 6º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 7º Cluster: reúne passageiros que se utilizaram, majoritariamente, da Internet para realizar o Check-in.
- 8º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Totem para realizar o Check-in.
- 9º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.
- 10º Cluster: reúne passageiros que se utilizaram, majoritariamente, do Balcão para realizar o Check-in.

4.7.2 Motivo da viagem

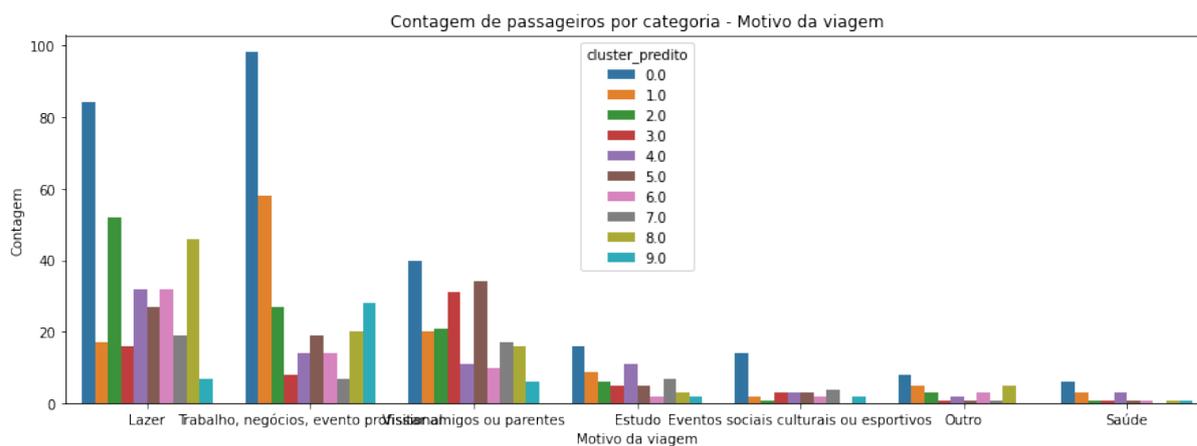


FIGURA 4.38 – Número de passageiros por cluster dentro de cada segmentação - Motivo da viagem

Para a motivação da viagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais/Lazer.
- 2º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais.
- 3º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 4º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes.
- 5º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 6º Cluster: reúne passageiros cuja motivação para viajar é visitar amigos ou parentes/Lazer.
- 7º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 8º Cluster: reúne passageiros cuja motivação para viajar é Lazer/visitar amigos ou parentes.
- 9º Cluster: reúne passageiros cuja motivação para viajar é Lazer.
- 10º Cluster: reúne passageiros cuja motivação para viajar é Trabalho, negócios ou eventos profissionais.

4.7.3 Compra de passagem e Responsável

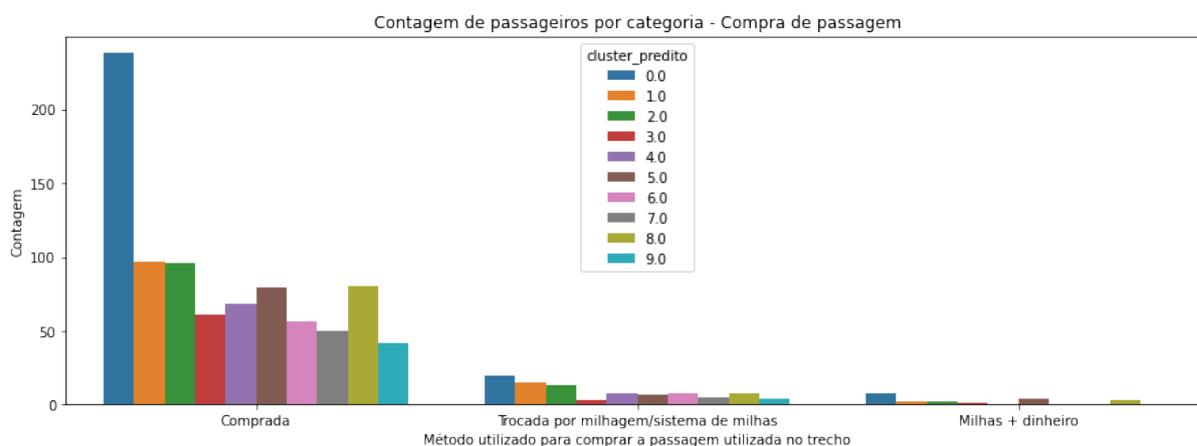


FIGURA 4.39 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - método

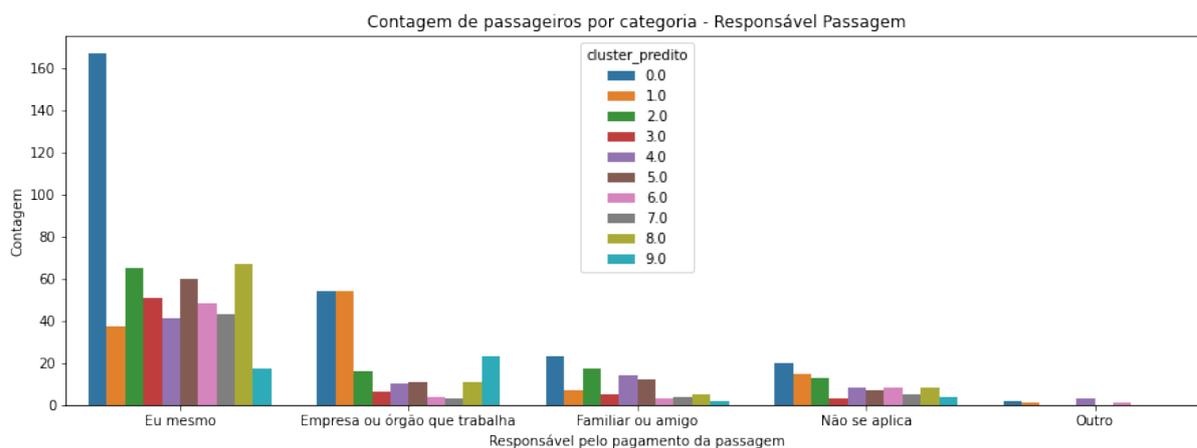


FIGURA 4.40 – Número de passageiros por cluster dentro de cada segmentação - Compra de passagem - responsável

Para a compra de passagem e seu responsável, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 2^o Cluster: reúne passageiros que tiveram a passagem comprada pelo órgão ou empresa na qual trabalham sem auxílio de milhas.
- 3^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 4^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 5^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 6^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 7^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 8^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 9^o Cluster: reúne passageiros que compraram a passagem sem auxílio de milhas por conta própria.
- 10^o Cluster: reúne passageiros que tiveram a passagem comprada pelo órgão ou empresa na qual trabalham sem auxílio de milhas.

4.7.4 Profissão, Renda familiar mensal bruta e Idade

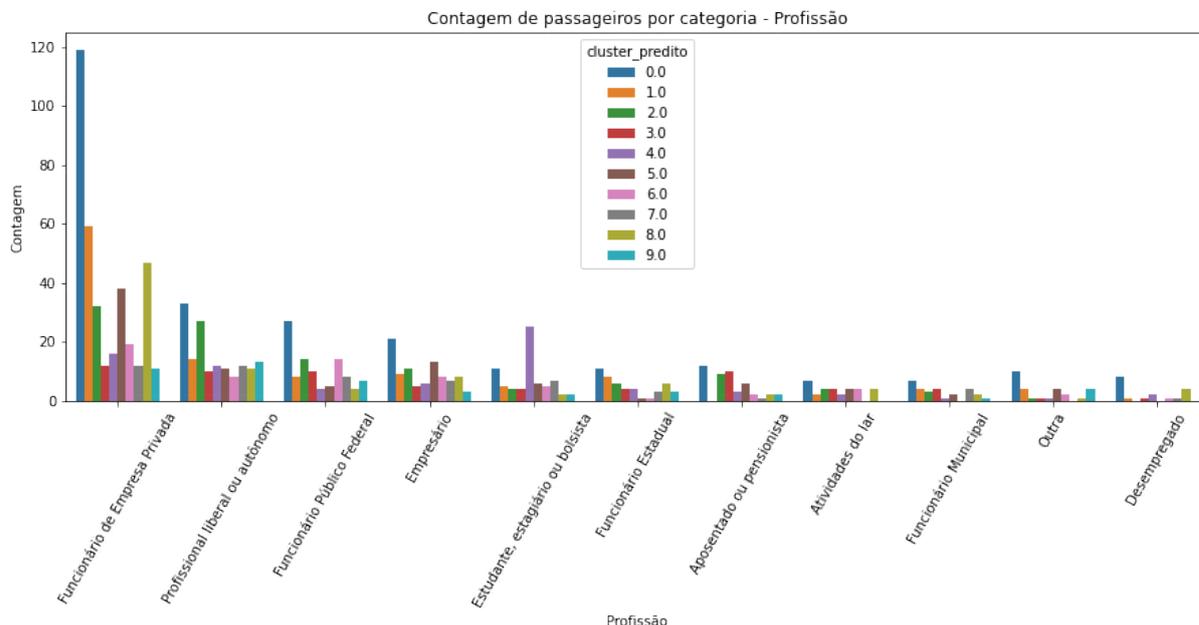


FIGURA 4.41 – Número de passageiros por cluster dentro de cada segmentação - Profissão

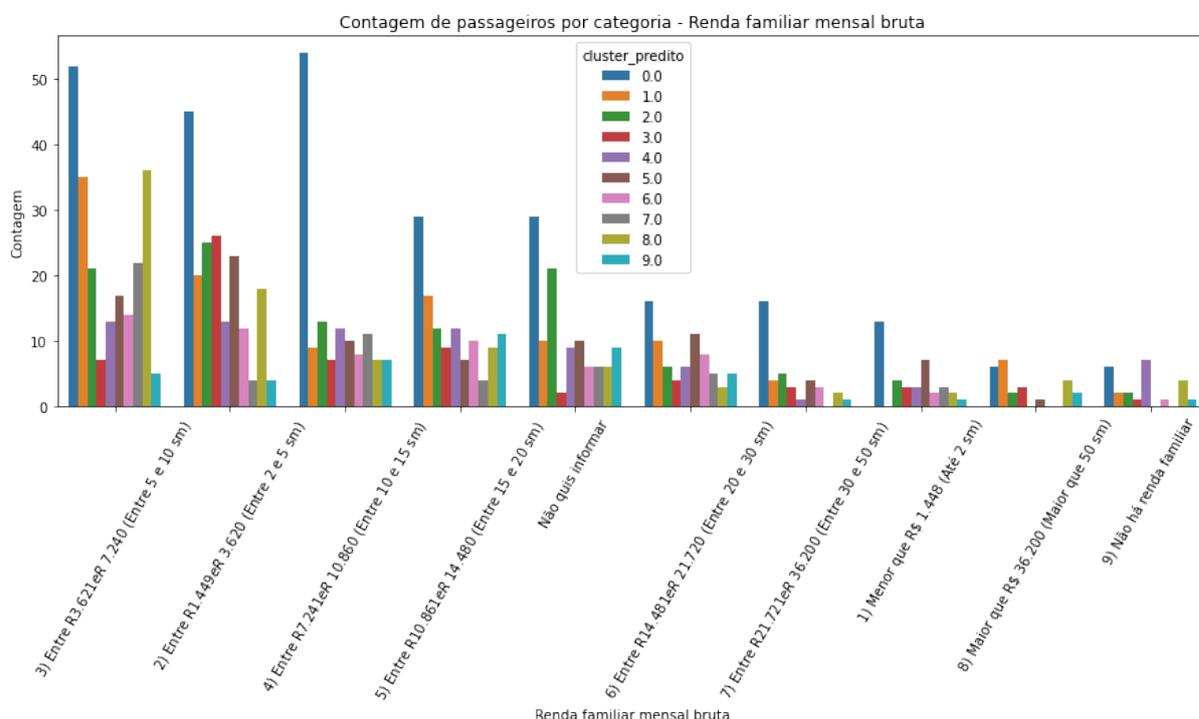


FIGURA 4.42 – Número de passageiros por cluster dentro de cada segmentação - Renda familiar mensal bruta

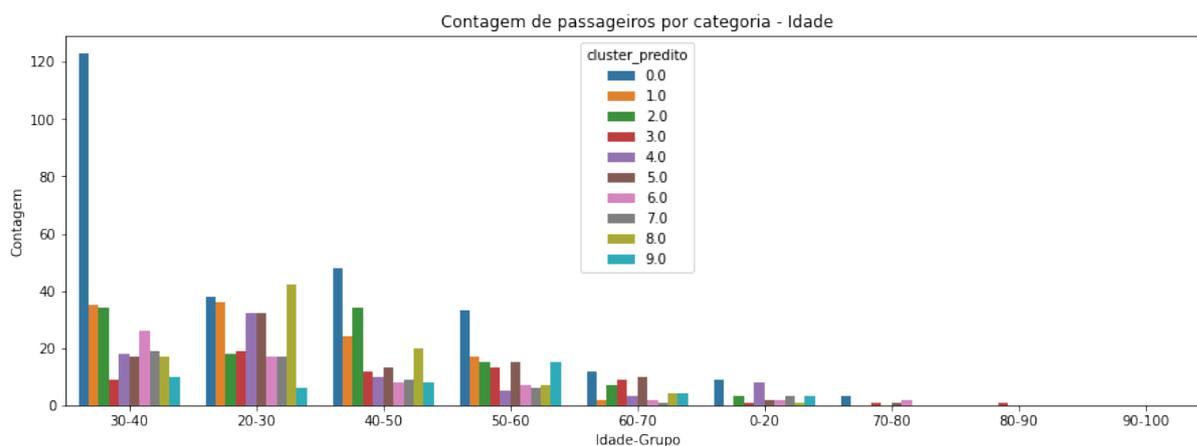


FIGURA 4.43 – Número de passageiros por cluster dentro de cada segmentação - Idade

Para os aspectos de renda, profissão e idade, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 10 e 15 salários mínimos e na faixa de idade de 30 a 40 anos.
- 2^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 20 a 30 anos.
- 3^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 40 a 50 anos.
- 4^o Cluster: reúne passageiros que atuam como funcionários públicos federais, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 5^o Cluster: reúne passageiros que são estudantes, estagiários ou bolsistas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 6^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 2 e 5 salários mínimos e na faixa de idade de 20 a 30 anos.
- 7^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 8^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 30 a 40 anos.
- 9^o Cluster: reúne passageiros contratados por empresas privadas, cuja renda familiar mensal bruta está entre 5 e 10 salários mínimos e na faixa de idade de 20 a 30 anos.

- 10^o Cluster: reúne passageiros que atuam como profissional liberal ou autônomo, cuja renda familiar mensal bruta está entre 15 e 20 salários mínimos e na faixa de idade de 50 a 60 anos.

4.7.5 Uso do transporte aéreo



FIGURA 4.44 – Número de passageiros por cluster dentro de cada segmentação - Utilização do mesmo trecho no último ano

Para os aspectos de uso do transporte aéreo, tem-se a seguinte observação:

- Todos os clusters: reúnem passageiros que realizaram o mesmo trecho entre 0 e 10 vezes no ano da pesquisa.

4.7.6 Gasto no aeroporto

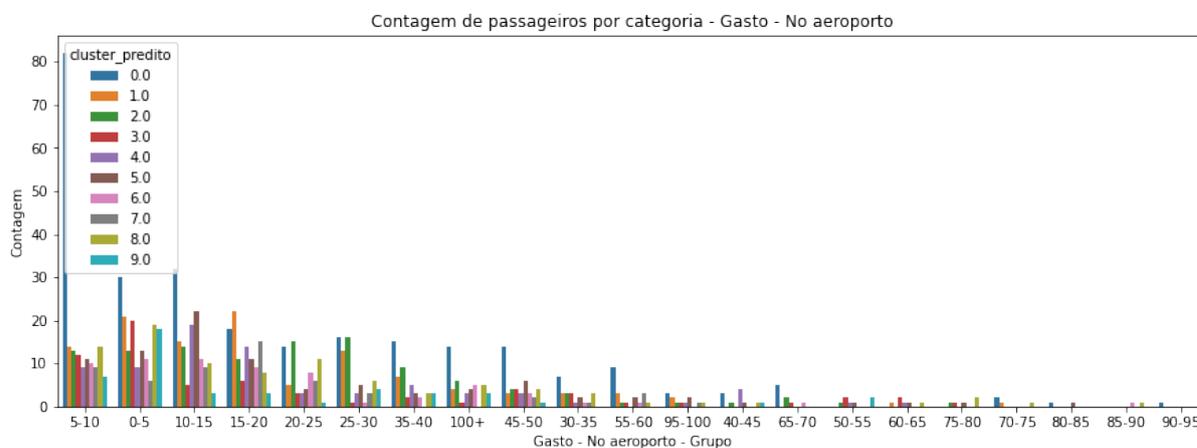


FIGURA 4.45 – Número de passageiros por cluster dentro de cada segmentação - Gasto no aeroporto

Para o gasto realizado no aeroporto, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que gastaram entre R\$20 e R\$25.
- 2º Cluster: reúne passageiros que gastaram entre R\$15 e R\$20.
- 3º Cluster: reúne passageiros que gastaram entre R\$25 e R\$30.
- 4º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 5º Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 6º Cluster: reúne passageiros que gastaram entre R\$10 e R\$15.
- 7º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 8º Cluster: reúne passageiros que gastaram entre R\$15 e R\$20.
- 9º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.
- 10º Cluster: reúne passageiros que gastaram entre R\$0 e R\$5.

4.7.7 Antecedência de compra da passagem em dias

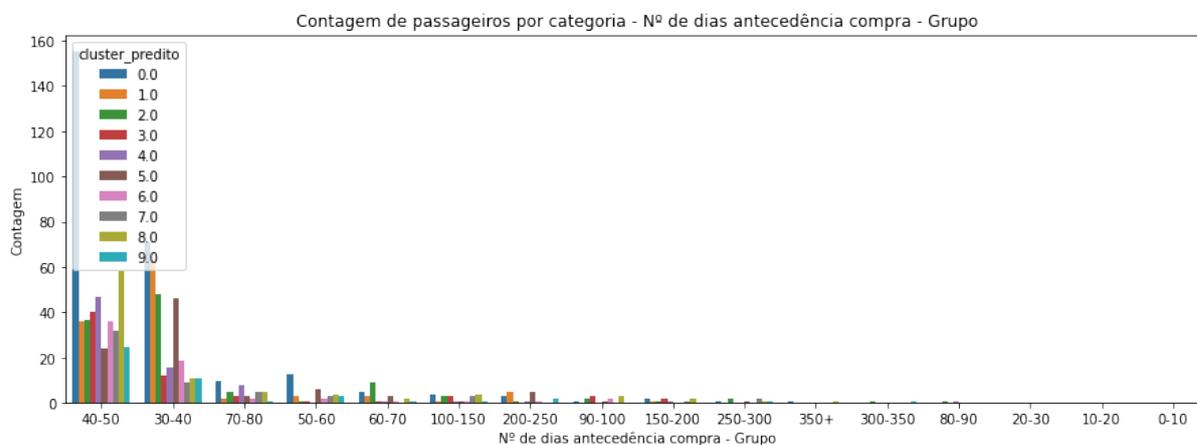


FIGURA 4.46 – Número de passageiros por cluster dentro de cada segmentação - Antecedência de compra da passagem em dias

Para a antecedência de compra da passagem, tem-se a seguinte observação:

- 1º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do voo.
- 2º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do voo.

- 3º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 4º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 5º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 6º Cluster: reúne passageiros que compraram a passagem entre 30 e 40 dias antes do vôo.
- 7º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 8º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 9º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.
- 10º Cluster: reúne passageiros que compraram a passagem entre 40 e 50 dias antes do vôo.

4.7.8 Fileira e coluna dos assentos

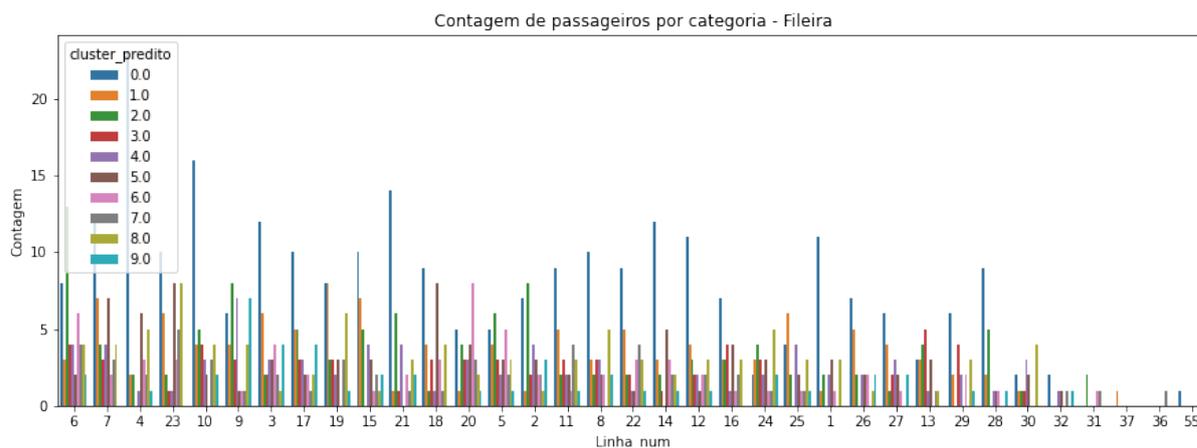


FIGURA 4.47 – Número de passageiros por cluster dentro de cada segmentação - Fileira dos assentos

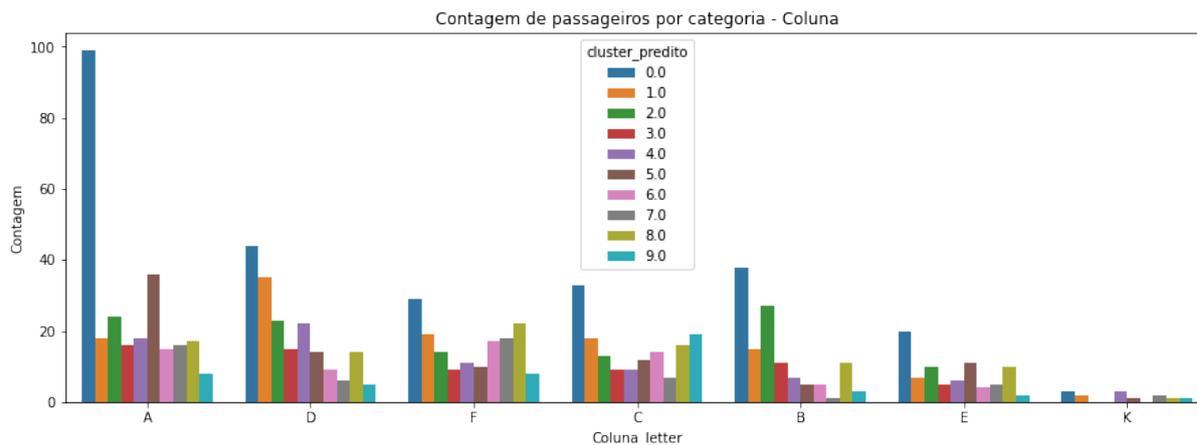


FIGURA 4.48 – Número de passageiros por cluster dentro de cada segmentação - Coluna dos assentos

Para a fileira e coluna dos assentos, tem-se a seguinte observação:

- 1^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileira 4.
- 2^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileira 19.
- 3^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna B e fileira 6.
- 4^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileira 13.
- 5^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna D e fileira 9.
- 6^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna A e fileira 23.
- 7^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileiras 20.
- 8^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileira 23.
- 9^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna F e fileira 23.
- 10^o Cluster: reúne passageiros cujo assento era majoritariamente distribuído na coluna C e fileira 9.

4.7.9 10 Clusters - K-Modes - Resumo

TABELA 4.45 – Perfil identificado - 1° Cluster - K-Modes

Variável	1° Cluster - 266 pessoas
Método utilizado para fazer check-in	65,79% - Balcão
Motivo da viagem	36,84% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	89,47% - Comprada
Responsável pelo pagamento da passagem	62,78% - Eu mesmo
Profissão	44,74% - Funcionário de Empresa Privada
Renda familiar mensal bruta	20,3% - 4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	8,65% - 4
Coluna do assento	37,22% - A
Idade-Grupo	46,24% - 30-40
Gasto - No aeroporto - Grupo	30,83% - 20-25
Número de vezes que realizou o mesmo trecho no último ano - Grupo	88,35% - 0-10
Nº de dias antecedência compra - Grupo	58,27% - 40-50

TABELA 4.46 – Perfil identificado - 2° Cluster - K-Modes

Variável	2° Cluster - 114 pessoas
Método utilizado para fazer check-in	42,11% - Internet
Motivo da viagem	50,88% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	85,09% - Comprada
Responsável pelo pagamento da passagem	47,37% - Empresa ou órgão que trabalha
Profissão	51,75% - Funcionário de Empresa Privada
Renda familiar mensal bruta	30,7% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	7,02% - 19
Coluna do assento	30,7% - D
Idade-Grupo	31,58% - 20-30
Gasto - No aeroporto - Grupo	19,3% - 15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	78,07% - 0-10
Nº de dias antecedência compra - Grupo	55,26% - 30-40

TABELA 4.47 – Perfil identificado - 3° Cluster - K-Modes

Variável	3° Cluster - 111 pessoas
Método utilizado para fazer check-in	71,17% - Balcão
Motivo da viagem	46,85% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	86,49% - Comprada
Responsável pelo pagamento da passagem	58,56% - Eu mesmo
Profissão	28,83% - Funcionário de Empresa Privada
Renda familiar mensal bruta	22,52% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	11,71% - 6
Coluna do assento	24,32% - B
Idade-Grupo	30,63% - 40-50
Gasto - No aeroporto - Grupo	14,41% - 25-30
Número de vezes que realizou o mesmo trecho no último ano - Grupo	90,09% - 0-10
Nº de dias antecedência compra - Grupo	43,24% - 30-40

TABELA 4.48 – Perfil identificado - 4° Cluster - K-Modes

Variável	4° Cluster - 65 pessoas
Método utilizado para fazer check-in	43,08% - Internet
Motivo da viagem	47,69% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	93,85% - Comprada
Responsável pelo pagamento da passagem	78,46% - Eu mesmo
Profissão	18,46% - Funcionário de Empresa Privada
Renda familiar mensal bruta	40% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	7,69% - 13
Coluna do assento	24,62% - A
Idade-Grupo	29,23% - 20-30
Gasto - No aeroporto - Grupo	30,77% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	95,38% - 0-10
Nº de dias antecedência compra - Grupo	61,54% - 40-50

TABELA 4.49 – Perfil identificado - 5° Cluster - K-Modes

Variável	5° Cluster - 76 pessoas
Método utilizado para fazer check-in	56,58% - Balcão
Motivo da viagem	42,11% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	89,47% - Comprada
Responsável pelo pagamento da passagem	53,95% - Eu mesmo
Profissão	32,89% - Estudante, estagiário ou bolsista
Renda familiar mensal bruta	17,11% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	9,21% - 9
Coluna do assento	28,95% - D
Idade-Grupo	42,11% - 20-30
Gasto - No aeroporto - Grupo	25% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	96,05% - 0-10
Nº de dias antecedência compra - Grupo	61,84% - 40-50

TABELA 4.50 – Perfil identificado - 6° Cluster - K-Modes

Variável	6° Cluster - 90 pessoas
Método utilizado para fazer check-in	67,78% - Balcão
Motivo da viagem	37,78% - Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	87,78% - Comprada
Responsável pelo pagamento da passagem	66,67% - Eu mesmo
Profissão	42,22% - Funcionário de Empresa Privada
Renda familiar mensal bruta	25,56% - 2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	8,99% - 23
Coluna do assento	40,45% - A
Idade-Grupo	35,56% - 20-30
Gasto - No aeroporto - Grupo	24,44% - 10-15
Número de vezes que realizou o mesmo trecho no último ano - Grupo	88,89% - 0-10
Nº de dias antecedência compra - Grupo	51,11% - 30-40

TABELA 4.51 – Perfil identificado - 7° Cluster - K-Modes

Variável	7° Cluster - 64 pessoas
Método utilizado para fazer check-in	46,88% - Internet
Motivo da viagem	50% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,5% - Comprada
Responsável pelo pagamento da passagem	75% - Eu mesmo
Profissão	29,69% - Funcionário de Empresa Privada
Renda familiar mensal bruta	21,88% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	12,5% - 20
Coluna do assento	26,56% - F
Idade-Grupo	40,63% - 30-40
Gasto - No aeroporto - Grupo	17,19% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	82,81% - 0-10
Nº de dias antecedência compra - Grupo	56,25% - 40-50

TABELA 4.52 – Perfil identificado - 8° Cluster - K-Modes

Variável	8° Cluster - 55 pessoas
Método utilizado para fazer check-in	43,64% - Totem
Motivo da viagem	34,55% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	90,91% - Comprada
Responsável pelo pagamento da passagem	78,18% - Eu mesmo
Profissão	21,82% - Funcionário de Empresa Privada
Renda familiar mensal bruta	40% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	9,09% - 23
Coluna do assento	32,73% - F
Idade-Grupo	34,55% - 30-40
Gasto - No aeroporto - Grupo	27,27% - 15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,09% - 0-10
Nº de dias antecedência compra - Grupo	58,18% - 40-50

TABELA 4.53 – Perfil identificado - 9° Cluster - K-Modes

Variável	9° Cluster - 91 pessoas
Método utilizado para fazer check-in	63,74% - Balcão
Motivo da viagem	50,55% - Lazer
Método utilizado para comprar a passagem utilizada no trecho	87,91% - Comprada
Responsável pelo pagamento da passagem	73,63% - Eu mesmo
Profissão	51,65% - Funcionário de Empresa Privada
Renda familiar mensal bruta	39,56% - 3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	8,79% - 23
Coluna do assento	24,18% - F
Idade-Grupo	46,15% - 20-30
Gasto - No aeroporto - Grupo	20,88% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	93,41% - 0-10
Nº de dias antecedência compra - Grupo	63,74% - 40-50

TABELA 4.54 – Perfil identificado - 10° Cluster - K-Modes

Variável	10° Cluster - 46 pessoas
Método utilizado para fazer check-in	65,22% - Balcão
Motivo da viagem	60,87% - Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	91,3% - Comprada
Responsável pelo pagamento da passagem	50% - Empresa ou órgão que trabalha
Profissão	28,26% - Profissional liberal ou autônomo
Renda familiar mensal bruta	23,91% - 5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	15,22% - 9
Coluna do assento	41,3% - C
Idade-Grupo	32,61% - 50-60
Gasto - No aeroporto - Grupo	39,13% - 0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	89,13% - 0-10
Nº de dias antecedência compra - Grupo	54,35% - 40-50

4.8 10 Clusters - K-Prototypes

Como resultado inicial, apresentam-se os centróides obtidos:

TABELA 4.55 – Perfil identificado - 1° Cluster - K-Prototypes

Variável	1° Cluster - 186 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	6
Coluna do assento	A
Idade-Grupo	36
Gasto - No aeroporto - Grupo	36.71
Número de vezes que realizou o mesmo trecho no último ano - Grupo	4
Nº de dias antecedência compra - Grupo	54

TABELA 4.56 – Perfil identificado - 2° Cluster - K-Prototypes

Variável	2° Cluster - 27 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	12
Coluna do assento	D
Idade-Grupo	42
Gasto - No aeroporto - Grupo	27.70
Número de vezes que realizou o mesmo trecho no último ano - Grupo	45
Nº de dias antecedência compra - Grupo	70

TABELA 4.57 – Perfil identificado - 3º Cluster - K-Prototypes

Variável	3º Cluster - 121 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	23
Coluna do assento	F
Idade-Grupo	43
Gasto - No aeroporto - Grupo	25.10
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	57

TABELA 4.58 – Perfil identificado - 4º Cluster - K-Prototypes

Variável	4º Cluster - 82 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)
Linha do assento	7
Coluna do assento	C
Idade-Grupo	43
Gasto - No aeroporto - Grupo	28.84
Número de vezes que realizou o mesmo trecho no último ano - Grupo	6
Nº de dias antecedência compra - Grupo	45

TABELA 4.59 – Perfil identificado - 5° Cluster - K-Prototypes

Variável	5° Cluster - 143 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)
Linha do assento	17
Coluna do assento	D
Idade-Grupo	39
Gasto - No aeroporto - Grupo	25.75
Número de vezes que realizou o mesmo trecho no último ano - Grupo	6
Nº de dias antecedência compra - Grupo	50

TABELA 4.60 – Perfil identificado - 6° Cluster - K-Prototypes

Variável	6° Cluster - 88 pessoas
Método utilizado para fazer check-in	Totem
Motivo da viagem	Trabalho, negócios, evento profissional
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)
Linha do assento	12
Coluna do assento	F
Idade-Grupo	32
Gasto - No aeroporto - Grupo	48.20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	5
Nº de dias antecedência compra - Grupo	61

TABELA 4.61 – Perfil identificado - 7º Cluster - K-Prototypes

Variável	7º Cluster - 73 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)
Linha do assento	22
Coluna do assento	B
Idade-Grupo	38
Gasto - No aeroporto - Grupo	60.68
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	63

TABELA 4.62 – Perfil identificado - 8º Cluster - K-Prototypes

Variável	8º Cluster - 92 pessoas
Método utilizado para fazer check-in	Internet
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Funcionário de Empresa Privada
Renda familiar mensal bruta	6) Entre R\$ 14.481 e R\$ 21.720 (Entre 20 e 30 sm)
Linha do assento	7
Coluna do assento	F
Idade-Grupo	32
Gasto - No aeroporto - Grupo	22.08
Número de vezes que realizou o mesmo trecho no último ano - Grupo	7
Nº de dias antecedência compra - Grupo	54

TABELA 4.63 – Perfil identificado - 9° Cluster - K-Prototypes

Variável	9° Cluster - 78 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Familiar ou amigo
Profissão	Estudante, estagiário ou bolsista
Renda familiar mensal bruta	Não quis informar
Linha do assento	6
Coluna do assento	A
Idade-Grupo	25
Gasto - No aeroporto - Grupo	31.73
Número de vezes que realizou o mesmo trecho no último ano - Grupo	2
Nº de dias antecedência compra - Grupo	54

TABELA 4.64 – Perfil identificado - 10° Cluster - K-Prototypes

Variável	10° Cluster - 88 pessoas
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Aposentado ou pensionista
Renda familiar mensal bruta	Não quis informar
Linha do assento	3
Coluna do assento	A
Idade-Grupo	60
Gasto - No aeroporto - Grupo	31.90
Número de vezes que realizou o mesmo trecho no último ano - Grupo	3
Nº de dias antecedência compra - Grupo	64

5 Análise espacial e de robustez dos resultados

5.1 Distribuição espacial dos resultados

Pra sintetizar os resultados obtidos na seção anterior, consolidou-se uma tabela representando o layout de uma aeronave padrão com 32 fileiras de assentos e 6 colunas, de tal maneira que os clusters encontrados estejam marcados pela sigla "xkpy" ou "xkmy". A primeira sigla refere-se ao x-ésimo cluster encontrado para o processo de clusterização com y clusters através do método k-Prototypes. A segunda sigla tem comportamento análogo, mas reflete os resultados das clusterizações realizadas pelo método k-Modes.

	A	B	C	D	E	F
Fileira 1	-	-	-	-	-	-
Fileira 2	3km5	3km5, 3km7	-	3km5, 3km7	-	-
Fileira 3	10kp10	-	4kp5	-	-	-
Fileira 4	1km5, 1km7, 1km8, 1km10	-	-	-	-	-
Fileira 5	7km8	-	-	-	-	7km8
Fileira 6	2km5, 4km5, 5km5, 1kp5, 7km7, 1kp7, 1kp8, 1kp10, 9kp10	3km8, 3km10	7km7	2km5, 5kp5, 2km7, 3km8	-	7km7
Fileira 7	1km5, 1km7, 1km8	-	4kp7, 4kp8, 4kp10	2km7, 6kp8	-	8kp10
Fileira 8	2km5	-	-	2km5	-	-
Fileira 9	2km5, 3km5, 5km5, 7km7	3km5, 3km8	7km7, 10km10	2km5, 3km5, 2km7, 5km7, 3km8, 5km8, 5km10	-	7km7
Fileira 10	1km5, 1km7, 6kp7, 1km8	-	-	-	-	-
Fileira 11	-	-	-	-	-	-
Fileira 12	-	-	-	2kp5, 2kp7, 2kp8, 2kp10	-	6kp10
Fileira 13	4km8, 4km10	-	-	-	-	-
Fileira 14	-	-	-	-	-	-
Fileira 15	-	-	-	-	-	-
Fileira 16	4km7	-	-	-	-	-
Fileira 17	-	-	-	5kp7, 5kp10	-	-
Fileira 18	-	-	-	-	-	6km8
Fileira 19	-	-	-	5kp8, 2km10	-	-
Fileira 20	7km8	-	-	-	-	7km8, 7km10
Fileira 21	-	-	-	-	-	-
Fileira 22	-	7kp7, 7kp8, 7kp10	-	-	-	8km8
Fileira 23	4km5, 3kp8, 6km10	-	-	2km8	-	3kp5, 3kp7, 8kp8, 8km10, 9km10
Fileira 24	-	-	-	-	-	-
Fileira 25	-	-	-	-	-	-
Fileira 26	-	-	-	-	-	-
Fileira 27	-	-	-	-	-	-
Fileira 28	-	-	-	-	-	-
Fileira 29	4km7	-	-	-	-	-
Fileira 30	-	-	-	-	-	-
Fileira 31	-	-	-	-	-	-
Fileira 32	-	-	-	-	-	-

FIGURA 5.1 – Distribuição espacial dos clusters encontrados

5.2 Análise de equivalências entre resultados

De modo complementar, dado que o intuito do estudo é estratificar o máximo de perfis possíveis de passageiros presentes na base de dados com alguma relevância, fez-se a agregação de grupos únicos encontrados por estratégia de clusterização.

Para o primeiro algoritmo, K-Modes, temos:

- Total de clusters gerados: 30
- Clusters únicos obtidos: 18
- Equivalência 1 - eq1km: $1km5 = 1km7 = 1km8$

- Equivalência 2 - eq2km: $2km5 = 2km10$
- Equivalência 3 - eq3km: $3km5 = 3km7$
- Equivalência 4 - eq4km: $4km5 = 4km7$
- Equivalência 5 - eq5km: $5km5 = 5km7$
- Equivalência 6 - eq6km: $2km7 = 2km8$
- Equivalência 7 - eq7km: $6km7 = 6km8 = 6km10$
- Equivalência 8 - eq8km: $5km8 = 9km10$
- Equivalência 9 - eq9km: $7km8 = 7km10$
- Equivalência 10 - eq10km: $8km8 = 8km10$

Para o segundo algoritmo, K-Prototypes, temos:

- Total de clusters gerados: 30
- Clusters únicos obtidos: 14
- Equivalência 1 - eq1kp: $1kp5 = 1kp7 = 1kp8 = 1kp10$
- Equivalência 2 - eq2kp: $2kp5 = 2kp7 = 2kp8 = 2kp10$
- Equivalência 3 - eq3kp: $3kp5 = 3kp7 = 3kp8 = 3kp10$
- Equivalência 4 - eq4kp: $5kp5 = 5kp8$
- Equivalência 5 - eq5kp: $4kp7 = 4kp8 = 4kp10$
- Equivalência 6 - eq6kp: $5kp7 = 5kp10$
- Equivalência 7 - eq7kp: $6kp7 = 10kp10$
- Equivalência 8 - eq8kp: $7kp7 = 7kp8 = 7kp10$

Faz-se, então, o cruzamento de resultados entre os dois métodos, novamente identificando os grupos já previstos e mapeando equivalências. Para os dois algoritmos:

- Total de clusters gerados: 60
- Clusters únicos obtidos: $14+18-2 = 30$
- Equivalência 1: $eq8km = eq1kp$

- Equivalência 2: $4km10 = 8kp8$

Assim, os resultados agora podem ser tratados como clusters "únicos" obtidos de todas as iterações de clusterizações, as legendas se modificam para que todos os clusters que podem ser considerados um só tenham a mesma nomenclatura.

5.3 Distribuição espacial dos resultados iterada

Dessa maneira, revisa-se a tabela apresentada na Figura 3.51 e apresenta-se nova tabela com agrupamento de classificações dado o processo efetuado:

	A	B	C	D	E	F
Fileira 1	-	-	-	-	-	-
Fileira 2	eq3km	eq3km	-	eq3km	-	-
Fileira 3	eq7kp	-	4kp5	-	-	-
Fileira 4	eq1km, 1km10	-	-	-	-	-
Fileira 5	eq9km	-	-	-	-	eq9km
Fileira 6	eq2km, eq4km, eq5km, eq1kp, 7km7, 9kp10	3km8, 3km10	7km7	eq2km, eq4kp, eq6km, 3km8	-	7km7
Fileira 7	eq1km	-	eq5kp	eq6km, 6kp8	-	8kp10
Fileira 8	eq2km	-	-	eq2km	-	-
Fileira 9	eq2km, eq3km, eq5km, 7km7	eq3km	7km7, eq7kp	eq2km, eq5km, eq6km, eq8km, 3km8, 5km10	-	7km7
Fileira 10	eq1km, eq7kp	-	-	-	-	-
Fileira 11	-	-	-	-	-	-
Fileira 12	-	-	-	eq2kp	-	6kp10
Fileira 13	4km8, 4km10	-	-	-	-	-
Fileira 14	-	-	-	-	-	-
Fileira 15	-	-	-	-	-	-
Fileira 16	eq4km	-	-	-	-	-
Fileira 17	-	-	-	eq6kp	-	-
Fileira 18	-	-	-	-	-	eq7km
Fileira 19	-	-	-	eq2km, 5kp8	-	-
Fileira 20	eq9km	-	-	-	-	eq9km
Fileira 21	-	-	-	-	-	-
Fileira 22	-	eq8kp	-	-	-	eq10km
Fileira 23	eq4km, eq7km, eq3kp	-	-	eq6km	-	eq3kp, 8kp8, eq10km, eq8km
Fileira 24	-	-	-	-	-	-
Fileira 25	-	-	-	-	-	-
Fileira 26	-	-	-	-	-	-
Fileira 27	-	-	-	-	-	-
Fileira 28	-	-	-	-	-	-
Fileira 29	4km7	-	-	-	-	-
Fileira 30	-	-	-	-	-	-
Fileira 31	-	-	-	-	-	-
Fileira 32	-	-	-	-	-	-

FIGURA 5.2 – Distribuição espacial dos clusters encontrados após revisão e agrupamento de resultados

Como é perceptível na Figura 4.2 e 4.1, a cobertura do mapeamento de perfil foi restrita nas colunas B, C e E, de maneira que poucas regiões entre as fileiras 1 e 32 tenham sido identificadas. Já nas colunas A, D e F, houve maior cobertura de identificação de perfis, possibilitando inferir mais características em mais regiões.

Para fim de sintetização e agrupamento, escolheu-se dividir o corpo da aeronave em 4 regiões do sentido longitudinal, cada região contando com 8 fileiras em sequência.

Por fim, faz-se uma consolidação dos resultados de perfis obtidos por coluna de assentos e regiões de fileiras.

5.3.1 Coluna A

5.3.1.1 Fileiras 1-8

- K-Modes: eq1km, eq2km, eq3km, eq4km, eq5km, eq9km, 7km7, 1km10
- K-Prototypes: eq1kp, eq7kp, 9kp10

TABELA 5.1 – Perfis identificados - eq1km, eq2km, eq3km

Variável	eq1km	eq2km	eq3km
Método utilizado para fazer check-in	Balcão	Internet	Balcão
Motivo da viagem	Trabalho...	Trabalho...	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 5 e 10 sm	Entre 2 e 5 sm
Idade-Grupo	30-40	20-30	40-50
Gasto - No aeroporto - Grupo	5-10	15-20	25-30
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	30-40	30-40

TABELA 5.2 – Perfis identificados - eq4km, eq5km, eq9km

Variável	eq4km	eq5km	eq9km
Método utilizado para fazer check-in	Balcão	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	Estudante, estag. ou bolsista	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 10 e 15 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30	30-40
Gasto - No aeroporto - Grupo	0-5	10-15	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50	40-50

TABELA 5.3 – Perfis identificados - 7km7, 1km10

Variável	7km7	1km10
Método utilizado para fazer check-in	Internet	Balcão
Motivo da viagem	Lazer	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 10 e 15 sm
Idade-Grupo	30-40	30-40
Gasto - No aeroporto - Grupo	15-20	20-25
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50

TABELA 5.4 – Perfis identificados - eq1kp, eq7kp, 9kp10

Variável	eq1kp	eq7kp	9kp10
Método utilizado para fazer check-in	Balcão	Balcão	Balcão
Motivo da viagem	Lazer	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Familiar/Amigo
Profissão	F. Empresa Privada	Aposentado/Pensionista	Estudante, estag. ou bolsista
Renda familiar mensal bruta	Entre 5 e 10 sm	-	-
Idade	32	59	25
Gasto - No aeroporto - Grupo	R\$36,20	R\$32,10	R\$31,73
Núm. vezes - mesmo trecho no último ano - Grupo	4	3	2
Nº de dias antecedência compra - Grupo	54	62	54

5.3.1.2 Fileiras 9-16

- K-Modes: eq1km, eq2km, eq3km, eq4km, eq5km, 7km7, 4km8, 4km10
- K-Prototypes: eq7kp

TABELA 5.5 – Perfis identificados - eq1km, eq2km, eq3km

Variável	eq1km	eq2km	eq3km
Método utilizado para fazer check-in	Balcão	Internet	Balcão
Motivo da viagem	Trabalho...	Trabalho...	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 5 e 10 sm	Entre 2 e 5 sm
Idade-Grupo	30-40	20-30	40-50
Gasto - No aeroporto - Grupo	5-10	15-20	25-30
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	30-40	30-40

TABELA 5.6 – Perfis identificados - eq4km, eq5km, 7km7

Variável	eq4km	eq5km	7km7
Método utilizado para fazer check-in	Balcão	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	Estudante, estag. ou bolsista	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 10 e 15 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30	30-40
Gasto - No aeroporto - Grupo	0-5	10-15	15-20
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50	40-50

TABELA 5.7 – Perfis identificados - 4km8, 4km10

Variável	4km8	4km10
Método utilizado para fazer check-in	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Visitar amigos/parentes
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo
Profissão	F. Público Federal	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 2 e 5 sm
Idade-Grupo	20-30	20-30
Gasto - No aeroporto - Grupo	0-5	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50

TABELA 5.8 – Perfil identificado - eq7kp

Variável	eq7kp
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Trabalho...
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Empresa ou órgão que trabalha
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	Entre 15 e 20 sm
Idade-Grupo	50-60
Gasto - No aeroporto - Grupo	0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

5.3.1.3 Fileiras 17-24

- K-Modes: eq4km, eq7km, eq9km, 7km7, 4km8, 4km10
- K-Prototypes: eq3kp

TABELA 5.9 – Perfis identificados - eq4km, eq7km, eq9km

Variável	eq4km	eq7km	eq9km
Método utilizado para fazer check-in	Balcão	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Visitar amigos/parentes	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 2 e 5 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30	30-40
Gasto - No aeroporto - Grupo	0-5	10-15	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	30-40	40-50

TABELA 5.10 – Perfis identificados - 7km7, 4km8, 4km10

Variável	7km7	4km8	4km10
Método utilizado para fazer check-in	Internet	Balcão	Internet
Motivo da viagem	Lazer	Visitar amigos/parentes	Visitar amigos/parentes
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Público Federal	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 2 e 5 sm	Entre 2 e 5 sm
Idade-Grupo	30-40	20-30	20-30
Gasto - No aeroporto - Grupo	15-20	0-5	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50	40-50

TABELA 5.11 – Perfil identificado - eq3kp

Variável	eq3kp
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	Entre 2 e 5 sm
Idade	38
Gasto - No aeroporto	R\$28.12
Núm. vezes - mesmo trecho no último ano	4
Nº de dias antecedência compra	59

5.3.1.4 Fileiras 25-32

- K-Modes: 4km7

TABELA 5.12 – Perfil identificado - 4km7

Variável	4km7
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Visitar amigos ou parentes
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm
Idade-Grupo	20-30
Gasto - No aeroporto - Grupo	0-5
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

5.3.2 Coluna B

5.3.2.1 Fileiras 1-8

- K-Modes: eq3km, 3km8, 3km10

TABELA 5.13 – Perfis identificados - eq3km, 3km8, 3km10

Variável	eq3km	3km8	3km10
Método utilizado para fazer check-in	Balcão	Balcão	Balcão
Motivo da viagem	Lazer	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 2 e 5 sm	Entre 2 e 5 sm
Idade-Grupo	40-50	30-40	40-50
Gasto - No aeroporto - Grupo	25-30	25-30	25-30
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	30-40	30-40	30-40

5.3.2.2 Fileiras 9-16

- K-Modes: eq3km

TABELA 5.14 – Perfil identificado - eq3km

Variável	eq3km
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método compra da passagem utilizada no trecho	Comprada
Responsável pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm
Idade-Grupo	40-50
Gasto - No aeroporto - Grupo	25-30
Núm. vezes - mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	30-40

5.3.2.3 Fileiras 17-24

- K-Prototypes: eq8kp

TABELA 5.15 – Perfil identificado - eq8kp

Variável	eq8kp
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Empresário
Renda familiar mensal bruta	Entre 2 e 5 sm
Idade-Grupo	38
Gasto - No aeroporto	R\$60.68
Núm. de vezes mesmo trecho no último ano	3
Nº de dias antecedência compra	63

5.3.3 Coluna C

5.3.3.1 Fileiras 1-8

- K-Modes: 7km7
- K-Prototypes: eq5kp, 4kp5

TABELA 5.16 – Perfil identificado - 7km7

Variável	7km7
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm
Idade-Grupo	30-40
Gasto - No aeroporto - Grupo	15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

TABELA 5.17 – Perfis identificados - eq5kp, 4kp5

Variável	4km8	4km10
Método utilizado para fazer check-in	Balcão	Balcão
Motivo da viagem	Trabalho...	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo
Profissão	Empresário	Empresário
Renda familiar mensal bruta	Entre 15 e 20 sm	Entre 15 e 20 sm
Idade-Grupo	43	55
Gasto - No aeroporto - Grupo	R\$29,30	R\$32,05
Núm. vezes - mesmo trecho no último ano - Grupo	7	4
Nº de dias antecedência compra - Grupo	46	54

5.3.3.2 Fileiras 9-16

- K-Modes: 7km7
- K-Prototypes: eq7kp

TABELA 5.18 – Perfil identificado - 7km7

Variável	7km7
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm
Idade-Grupo	30-40
Gasto - No aeroporto - Grupo	15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

TABELA 5.19 – Perfil identificado - eq7kp

Variável	eq7kp
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Aposentado/Pensionista
Renda familiar mensal bruta	-
Idade-Grupo	60
Gasto - No aeroporto	R\$31,90
Núm. de vezes mesmo trecho no último ano	3
Nº de dias antecedência compra	64

5.3.4 Coluna D

5.3.4.1 Fileiras 1-8

- K-Modes: eq2km, eq3km, eq6km, 3km8
- K-Prototypes: eq4kp, 6kp8

TABELA 5.20 – Perfis identificados - eq2km, eq3km, eq6km

Variável	eq2km	eq3km	eq6km
Método utilizado para fazer check-in	Internet	Balcão	Internet
Motivo da viagem	Trabalho...	Lazer	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Empresa ou órgão que trabalha
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 2 e 5 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	40-50	20-30
Gasto - No aeroporto - Grupo	15-20	25-30	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	30-40	30-40	30-40

TABELA 5.21 – Perfil identificado - 3km8

Variável	3km8
Método utilizado para fazer check-in	Balcão
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm
Idade-Grupo	30-40
Gasto - No aeroporto - Grupo	25-30
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	30-40

TABELA 5.22 – Perfis identificados - eq4kp, 6kp8

Variável	eq4kp	6kp8
Método utilizado para fazer check-in	Balcão	Internet
Motivo da viagem	Trabalho...	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Empresa ou órgão que trabalha	Empresa ou órgão que trabalha
Profissão	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 15 e 20 sm
Idade-Grupo	40	43
Gasto - No aeroporto - Grupo	R\$34,30	R\$30,10
Núm. vezes - mesmo trecho no último ano - Grupo	7	45
Nº de dias antecedência compra - Grupo	56	74

5.3.4.2 Fileiras 9-16

- K-Modes: eq2km, eq5km, eq6km, eq8km, 3km8, 5km10
- K-Prototypes: eq2kp

TABELA 5.23 – Perfis identificados - eq2km, eq5km, eq6km

Variável	eq2km	eq5km	eq6km
Método utilizado para fazer check-in	Internet	Balcão	Internet
Motivo da viagem	Trabalho...	Lazer	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Empresa ou órgão que trabalha
Profissão	F. Empresa Privada	Estudante, estag. ou bolsista	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 10 e 15 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30	20-30
Gasto - No aeroporto - Grupo	15-20	10-15	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	30-40	40-50	30-40

TABELA 5.24 – Perfis identificados - eq8km, 3km8, 5km10

Variável	eq8km	3km8	5km10
Método utilizado para fazer check-in	Balcão	Balcão	Balcão
Motivo da viagem	Lazer	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	Estudante, estag. ou bolsista
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 2 e 5 sm	Entre 2 e 5 sm
Idade-Grupo	20-30	30-40	20-30
Gasto - No aeroporto - Grupo	10-15	25-30	10-15
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	30-40	40-50

TABELA 5.25 – Perfil identificado - eq2kp

Variável	eq2kp
Método utilizado para fazer check-in	Internet
Motivo da viagem	Trabalho...
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	Profissional liberal ou autônomo
Renda familiar mensal bruta	Entre 10 e 15 sm
Idade-Grupo	34
Gasto - No aeroporto	R\$35,74
Núm. de vezes mesmo trecho no último ano	15
Nº de dias antecedência compra	57

5.3.4.3 Fileiras 17-24

- K-Modes: eq2km, eq6km
- K-Prototypes: eq6kp, 5kp8

TABELA 5.26 – Perfis identificados - eq2km, eq6km

Variável	eq2km	eq6km
Método utilizado para fazer check-in	Internet	Internet
Motivo da viagem	Trabalho...	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Empresa ou órgão que trabalha
Profissão	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30
Gasto - No aeroporto - Grupo	15-20	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10
Nº de dias antecedência compra - Grupo	30-40	30-40

TABELA 5.27 – Perfis identificados - eq6kp, 5kp8

Variável	eq6kp	5kp8
Método utilizado para fazer check-in	Balcão	Balcão
Motivo da viagem	Trabalho...	Trabalho...
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Empresa ou órgão que trabalha	Empresa ou órgão que trabalha
Profissão	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 2 e 5 sm
Idade-Grupo	39	37
Gasto - No aeroporto - Grupo	R\$32,55	R\$33,30
Núm. vezes - mesmo trecho no último ano - Grupo	6	5
Nº de dias antecedência compra - Grupo	58	54

5.3.5 Coluna F

5.3.5.1 Fileiras 1-8

- K-Modes: eq9km, 7km7
- K-Prototypes: 8kp10

TABELA 5.28 – Perfis identificados - eq9km, 7km7

Variável	eq9km	7km7
Método utilizado para fazer check-in	Totem	Internet
Motivo da viagem	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm	Entre 5 e 10 sm
Idade-Grupo	30-40	30-40
Gasto - No aeroporto - Grupo	15-20	15-20
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10
Nº de dias antecedência compra - Grupo	40-50	40-50

TABELA 5.29 – Perfil identificado - 8kp10

Variável	8kp10
Método utilizado para fazer check-in	Internet
Motivo da viagem	Visitar amigos/parentes
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 20 e 30 sm
Idade-Grupo	32
Gasto - No aeroporto	R\$22,08
Núm. de vezes mesmo trecho no último ano	7
Nº de dias antecedência compra	54

5.3.5.2 Fileiras 9-16

- K-Modes: 7km7
- K-Prototypes: 6kp10

TABELA 5.30 – Perfil identificado - 7km7

Variável	7km7
Método utilizado para fazer check-in	Internet
Motivo da viagem	Lazer
Método compra da passagem utilizada no trecho	Comprada
Responsável pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm
Idade-Grupo	30-40
Gasto - No aeroporto - Grupo	15-20
Núm. vezes - mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

TABELA 5.31 – Perfil identificado - 6kp10

Variável	6kp10
Método utilizado para fazer check-in	Totem
Motivo da viagem	Trabalho...
Método compra da passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 10 e 15 sm
Idade-Grupo	32
Gasto - No aeroporto	R\$48,20
Núm. de vezes mesmo trecho no último ano	5
Nº de dias antecedência compra	61

5.3.5.3 Fileiras 17-24

- K-Modes: eq7km, eq8km, eq9km, eq10km
- K-Prototypes: eq3kp, 8kp8

TABELA 5.32 – Perfis identificados - eq7km, eq8km, eq9km

Variável	eq7km	eq8km	eq9km
Método utilizado para fazer check-in	Balcão	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Lazer	Lazer
Método compra da passagem utilizada no trecho	Comprada	Comprada	Comprada
Responsável pagamento da passagem	Eu mesmo	Eu mesmo	Eu mesmo
Profissão	F. Empresa Privada	F. Empresa Privada	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 5 e 10 sm	Entre 5 e 10 sm
Idade-Grupo	20-30	20-30	30-40
Gasto - No aeroporto - Grupo	10-15	10-15	0-5
Núm. vezes - mesmo trecho no último ano - Grupo	0-10	0-10	0-10
Nº de dias antecedência compra - Grupo	30-40	40-50	40-50

TABELA 5.33 – Perfil identificado - eq10km

Variável	eq10km
Método utilizado para fazer check-in	Totem
Motivo da viagem	Lazer
Método utilizado para comprar a passagem utilizada no trecho	Comprada
Responsável pelo pagamento da passagem	Eu mesmo
Profissão	F. Empresa Privada
Renda familiar mensal bruta	Entre 5 e 10 sm
Idade-Grupo	30-40
Gasto - No aeroporto - Grupo	15-20
Número de vezes que realizou o mesmo trecho no último ano - Grupo	0-10
Nº de dias antecedência compra - Grupo	40-50

TABELA 5.34 – Perfis identificados - eq3kp, 8kp8

Variável	eq3kp	8kp8
Método utilizado para fazer check-in	Balcão	Internet
Motivo da viagem	Visitar amigos/parentes	Visitar amigos/parentes
Método compra da passagem utilizada no trecho	Comprada	Comprada
Responsável pagamento da passagem	Empresa ou órgão que trabalha	Eu mesmo
Profissão	Profissional liberal ou autônomo	F. Empresa Privada
Renda familiar mensal bruta	Entre 2 e 5 sm	Entre 2 e 5 sm
Idade-Grupo	38	29
Gasto - No aeroporto - Grupo	R\$28,12	R\$30,25
Núm. vezes - mesmo trecho no último ano - Grupo	4	5
Nº de dias antecedência compra - Grupo	59	60

5.4 Distribuição espacial dos resultados - variável renda

Para tentar correlacionar a distribuição espacial e a característica do grupo quanto a renda familiar mensal bruta, utiliza-se o seguinte esquema de cores:

1) Menor que R\$ 1.448 (Até 2 sm)	
2) Entre R\$ 1.449 e R\$ 3.620 (Entre 2 e 5 sm)	
3) Entre R\$ 3.621 e R\$ 7.240 (Entre 5 e 10 sm)	
4) Entre R\$ 7.241 e R\$ 10.860 (Entre 10 e 15 sm)	
5) Entre R\$ 10.861 e R\$ 14.480 (Entre 15 e 20 sm)	
6) Entre R\$ 14.481 e R\$ 21.720 (Entre 20 e 30 sm)	
7) Entre R\$ 21.721 e R\$ 36.200 (Entre 30 e 50 sm)	
8) Maior que R\$ 36.200 (Maior que 50 sm)	
9) Não há renda familiar	
Não quis informar	

FIGURA 5.3 – Legenda - Distribuição espacial dos clusters encontrados - variável renda

Então, utiliza-se o mesmo mapa iterado obtido na Figura 5.2 para que se represente como as categorias de renda se distribuem ao longo do corpo da aeronave:

	A	B	C	D	E	F
Fileira 1	-	-	-	-	-	-
Fileira 2	eq3km	eq3km	-	eq3km	-	-
Fileira 3	eq7kp	-	4kp5	-	-	-
Fileira 4	eq1km, 1km10	-	-	-	-	-
Fileira 5	eq9km	-	-	-	-	eq9km
Fileira 6	eq2km, eq4km, eq5km, eq1kp, 7km7, 9kp10	3km8, 3km10	7km7	eq2km, eq4kp, eq6km, 3km8	-	7km7
Fileira 7	eq1km	-	eq5kp	eq6km, 6kp8	-	8kp10
Fileira 8	eq2km	-	-	eq2km	-	-
Fileira 9	eq2km, eq3km, eq5km, 7km7	eq3km	7km7, eq7kp	eq2km, eq5km, eq6km, eq8km, 3km8, 5km10	-	7km7
Fileira 10	eq1km, eq7kp	-	-	-	-	-
Fileira 11	-	-	-	-	-	-
Fileira 12	-	-	-	eq2kp	-	6kp10
Fileira 13	4km8, 4km10	-	-	-	-	-
Fileira 14	-	-	-	-	-	-
Fileira 15	-	-	-	-	-	-
Fileira 16	eq4km	-	-	-	-	-
Fileira 17	-	-	-	eq6kp	-	-
Fileira 18	-	-	-	-	-	eq7km
Fileira 19	-	-	-	eq2km, 5kp8	-	-
Fileira 20	eq9km	-	-	-	-	eq9km
Fileira 21	-	-	-	-	-	-
Fileira 22	-	eq8kp	-	-	-	eq10km
Fileira 23	eq4km, eq7km, eq3kp	-	-	eq6km	-	eq3kp, 8kp8, eq10km, eq8km
Fileira 24	-	-	-	-	-	-
Fileira 25	-	-	-	-	-	-
Fileira 26	-	-	-	-	-	-
Fileira 27	-	-	-	-	-	-
Fileira 28	-	-	-	-	-	-
Fileira 29	4km7	-	-	-	-	-
Fileira 30	-	-	-	-	-	-
Fileira 31	-	-	-	-	-	-
Fileira 32	-	-	-	-	-	-

FIGURA 5.4 – Distribuição espacial dos clusters encontrados - variável renda

Como se verifica na Figura 5.4, encontram-se as seguintes inferências:

- 1) Coluna A: majoritariamente ocupada pelos passageiros com renda familiar entre 5 e 10 salários mínimos concentrados nas fileiras de 4-10 e fileira 20. Ocupada também

por passageiros com renda familiar entre 2 e 5 salários mínimos dispersos ao longo das fileiras remanescentes.

- 2) Coluna B: majoritariamente ocupada pelos passageiros com renda familiar entre 2 e 5 salários mínimos dispersos ao longo das fileiras, com maior incidência nas iniciais.
- 3) Coluna C: ocupada igualmente pelos passageiros com renda familiar entre 5 e 10 salários mínimos e 15 e 20 salários mínimos dispersos ao longo das fileiras iniciais.
- 4) Coluna D: coluna com maior variabilidade de perfis encontrados. Nela, há passageiros na faixa de 2 a 20 salários mínimos distribuídos de maneira disforme, de acordo com a figura.
- 5) Coluna F: coluna, da mesma forma que a anterior, com maior variabilidade de perfis encontrados. Nela, há passageiros na faixa de 2 a 30 salários mínimos distribuídos de maneira disforme, de acordo com a figura.

5.5 Distribuição espacial dos resultados - variável profissão

Para tentar correlacionar a distribuição espacial e a característica do grupo quanto a profissão, utiliza-se o seguinte esquema de cores:

Aposentado ou pensionista
Atividades do lar
Desempregado
Empresário
Estudante, estagiário ou bolsista
Funcionário de Empresa Privada
Funcionário Estadual
Funcionário Municipal
Funcionário Público Federal
Outra
Profissional liberal ou autônomo

FIGURA 5.5 – Legenda - Distribuição espacial dos clusters encontrados - variável profissão

Então, utiliza-se o mesmo mapa iterado obtido na Figura 5.2 para que se represente como as categorias de renda se distribuem ao longo do corpo da aeronave:

	A	B	C	D	E	F
Fileira 1	-	-	-	-	-	-
Fileira 2	eq3km	eq3km	-	eq3km	-	-
Fileira 3	eq7kp	-	4kp5	-	-	-
Fileira 4	eq1km, 1km10	-	-	-	-	-
Fileira 5	eq9km	-	-	-	-	eq9km
Fileira 6	eq2km, eq4km, eq5km, eq1kp, 7km7, 9kp10	3km8, 3km10	7km7	eq2km, eq4kp, eq6km, 3km8	-	7km7
Fileira 7	eq1km	-	eq5kp	eq6km, 6kp8	-	8kp10
Fileira 8	eq2km	-	-	eq2km	-	-
Fileira 9	eq2km, eq3km, eq5km, 7km7	eq3km	7km7, eq7kp	eq2km, eq5km, eq6km, eq8km, 3km8, 5km10	-	7km7
Fileira 10	eq1km, eq7kp	-	-	-	-	-
Fileira 11	-	-	-	-	-	-
Fileira 12	-	-	-	eq2kp	-	6kp10
Fileira 13	4km8, 4km10	-	-	-	-	-
Fileira 14	-	-	-	-	-	-
Fileira 15	-	-	-	-	-	-
Fileira 16	eq4km	-	-	-	-	-
Fileira 17	-	-	-	eq6kp	-	-
Fileira 18	-	-	-	-	-	eq7km
Fileira 19	-	-	-	eq2km, 5kp8	-	-
Fileira 20	eq9km	-	-	-	-	eq9km
Fileira 21	-	-	-	-	-	-
Fileira 22	-	eq8kp	-	-	-	eq10km
Fileira 23	eq4km, eq7km, eq3kp	-	-	eq6km	-	eq3kp, 8kp8, eq10km, eq8km
Fileira 24	-	-	-	-	-	-
Fileira 25	-	-	-	-	-	-
Fileira 26	-	-	-	-	-	-
Fileira 27	-	-	-	-	-	-
Fileira 28	-	-	-	-	-	-
Fileira 29	4km7	-	-	-	-	-
Fileira 30	-	-	-	-	-	-
Fileira 31	-	-	-	-	-	-
Fileira 32	-	-	-	-	-	-

FIGURA 5.6 – Distribuição espacial dos clusters encontrados - variável profissão

Como se verifica na Figura 5.6, encontram-se as seguintes inferências:

- 1) Coluna A: majoritariamente ocupada pelos passageiros atuantes como funcionários de empresa privada. Ocupada também por passageiros aposentados / pensionistas, estudantes / estagiários / bolsistas, profissionais liberais / autônomos e funcionários públicos federais em poucas posições.
- 2) Coluna B: majoritariamente ocupada pelos passageiros atuantes como funcionários de empresa privada, com maior incidência nas fileiras iniciais. Também ocupada por empresários nas posições em torno da fileira 22.

- 3) Coluna C: ocupada pelos passageiros atuantes como funcionários de empresa privada, empresários e aposentados / pensionistas.
- 4) Coluna D: majoritariamente ocupada pelos passageiros atuantes como funcionários de empresa privada, com maior incidência. Também ocupada por estudantes / estagiários / bolsistas e profissionais liberais / autônomos.
- 5) Coluna F: majoritariamente ocupada pelos passageiros atuantes como funcionários de empresa privada, com maior incidência. Também ocupada por profissionais liberais / autônomos em baixa concentração.

6 Conclusões

6.1 Proposta de integração entre trabalhos

Como citado ao longo do trabalho de graduação, essa tese é abraçada pela proposta de pesquisa submetida ao CNPq "Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial". Desse modo, espera-se que os resultados aqui obtidos sejam de utilidade para serem acumulados ao modelo econométrico proposto pela pesquisa.

Existem algumas maneiras de aproveitar os resultados obtidos por este trabalho de graduação, dentre elas, citam-se:

1. Para cada clusterização, o resultado cru obtido é a classificação de cada observação na base de dados utilizada de acordo com os clusters calculados, portanto, cria-se uma relação um pra um em todo passageiro. Basta que se escolha um dos métodos utilizados (k-modes ou k-prototypes) e gere-se a coluna de classificação de clusters obtidos.
2. Como resultado fim do trabalho, obtém-se uma dispersão de classificações de passageiros de acordo com a posição ocupada a bordo de aeronaves. Uma possível abordagem é que se considere em cada região da aeronave o perfil mapeado na região de acordo com os resultados deste trabalho. Um possível impedimento do uso desta estratégia é que existem regiões em que nenhum perfil foi predominante, portanto, ficariam sem correlação direta com os perfis obtidos.
3. Ainda tratando dos resultados obtidos neste trabalho, poderia ser construído um "perfil médio" da região que pudesse ser utilizado como variável do modelo, tal região poderia ser qualquer, desde uma simples divisão por colunas até um mix entre colunas e fileiras.
4. De maneira complementar, também poderia se obter através dos resultados deste trabalho uma classificação entre os clusters "únicos" obtidos, tentando fitar o máximo

de passageiros na base de dados de acordo com cada cluster único e características gerais dos passageiros.

Em resumo, há várias maneiras de conectar ambos modelos, e fica a cargo do operador do modelo econométrico decidir que tipo de agregação e análise de perfil faz mais sentido de ser utilizada. No entanto, sugere-se a adoção de uma das práticas citadas acima, já que são desdobramentos mais diretos a partir do fluxo que guia este trabalho de graduação.

6.2 Limitações do trabalho de graduação

Ao longo deste trabalho de graduação, encontram-se alguns pontos que poderiam ser melhorados / iterados. Com o intuito de possibilitar a resolução futura desses detalhes e o próprio encaminhamento da solução, discutem-se alguns deles:

1. Filtro de assentos utilizado: neste trabalho, foram escolhidos os assentos que se encaixavam num padrão especificado, porém, neste filtro foram excluídos, por exemplo, os assentos da modalidade "Conforto" dado que não trariam nenhuma referência espacial do posicionamento. Seria interessante em um estudo futuro que se traçasse um panorama entre os resultados de perfis obtidos utilizando esse conjunto de assentos escolhidos e outros tipos, como por exemplo, os assentos do tipo "Conforto".
2. Escolha de variáveis socioeconômicas: dentre as 107 perguntas, utilizaram-se 12 para a realização de agrupamento entre as pessoas, as que em tese trariam maiores insights sobre o perfil dos passageiros, porém, pode ser que hajam outras variáveis que sirvam de maneira otimizada à realização de agrupamentos ou tenham maior valor significativo neste processo.
3. Tamanho da base de dados: dadas as variáveis escolhidas, a redução do número de observações a serem clusterizadas foi expressiva, talvez com menos variáveis ou com outras variáveis, o resultado em número de observações seria melhorado/aumentado.
4. Métodos de clusterização: foram implementados dois tipos de algoritmos particionais, porém existem muitas outras possibilidades. Um possível avanço seria aumentar o número de métodos utilizados e avaliar os resultados obtidos de maneira distinta.
5. Escolha do número de clusters: foi utilizado como estratégia o "Elbow Method". Pela sua natureza empírica, é possível que não tenha sido ajustado o número ótimo de clusters. É um dos pontos passíveis de validação no modelo implementado.

6. Inclusão ou não das informações de assentos completas: neste trabalho, a abordagem quanto ao posicionamento foi decompor as dimensões em que se posicionam os assentos em linha e coluna, de maneira que se pudessem analisar a frequência dos grupos em linhas e colunas. Ainda assim, podem existir abordagens melhores quanto ao posicionamento e cruzamento dessa informação com o grupo de informações socioeconômicas.

Referências

AHMAD, A.; KHAN, S. Survey of state-of-the-art mixed data clustering algorithms. In: IEEE ACCESS, 7., 2019. United Arab Emirates: IEEE, 2019. p. 1–20.

ANJANI. The effect of aircraft seat pitch on comfort. In: . [S.l.]: Applied Ergonomics, 88, 103132, 2020.

AVIOINTERIORS. Skyrider 2.0 concept product. In: . [S.l.: s.n.], 2021.
<https://aviointeriors.it/portfolio/skyrider-2-0/>.

BUTTERFLY. Checkerboard aircraft seating. In: . [S.l.: s.n.], 2021.
<https://www.butterflyseating.com/checkerboard>.

CONCLI, M. Aircraft seats - clustering passengers according to socioeconomic information - 8 clusters. In: . [S.l.: s.n.], 2021.
https://colab.research.google.com/drive/1r5Vw3_ZfVn3FHy3n9b0T23bbDhePGGU1.

GARSTEN, E. Lighter, more comfortable airline seats built with formula one materials. In: . [S.l.: s.n.], 2019. <https://www.forbes.com/sites/edgarsten/2019/04/01/lighter-more-comfortable-airline-seats-built-with-formula-one-materials/?sh=34cc316d3e9c>.

HUANG, Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. In: DATA MINING AND KNOWLEDGE DISCOVERY, 2., 1998. Canberra, 1998. p. 283–304.

KREMSER. Aircraft seating comfort: the influence of seat pitch on passengers' well-being. In: . [S.l.: s.n.], 2012. p. 4936–4942.

KYPARISIS; KOULAMAS. Optimal pricing and seat allocation for a two-cabin airline revenue management problem. In: . [S.l.]: International Journal of Production Economics, 201, 2018. p. 18–25.

LOGÍSTICA, E. Empresa de Planejamento e Pesquisa de origem/destino do transporte aéreo de passageiros - relatório técnico. In: . Brasília, DF: [s.n.], 2014.
<https://www.ep1.gov.br/pesquisa-de-origem-e-destino-nos-aeroportos>.

MUMBOWER; GARROW; NEWMAN. Investigating airline customers' premium coach seat purchases and implications for optimal pricing strategies. In: . [S.l.]: Transportation Research Part A: Policy and Practice, 73, 2015. p. 53–69.

- OLIVEIRA, A. V. M. de. Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial. In: NÚCLEO DE ECONOMIA DO TRANSPORTE AÉREO (NECTAR-ITA). São José dos Campos, 2021. p. 1–48.
- ROSSI. Contribuições do entretenimento a bordo no conforto e desconforto em voos comerciais. dissertação (mestrado em engenharia de produção). universidade federal de são carlos. In: . [S.l.: s.n.], 2011.
- ROUNCIPELL; TIMMIS; ISON. Willingness to pay for preferred seat selection on uk domestic flights. In: . [S.l.]: Journal of Air Transport Management, 70, 2018. p. 57–61.
- SHAO; KAUERMANN; SMITH. Whether, when and which: Modelling advanced seat reservations by airline passengers. In: . [S.l.]: Transportation Research Part A: Policy and Practice, 132, 2020. p. 490–514.
- SILVA; MONTEIRO. Levantamento do perfil antropométrico da população brasileira usuária do transporte aéreo nacional: projeto conhecer. relatório de atividades de pesquisa. agência nacional de aviação civil, superintendência de segurança operacional, gerência de fatores humanos em aviação e medicina de aviação. In: . [S.l.: s.n.], 2009.
- SOUZA. Parâmetros para o projeto de poltronas aeronáuticas: revisão da literatura e as práticas da indústria no setor de transportes. dissertação (mestrado em engenharia de produção). universidade federal de são carlos. In: . [S.l.: s.n.], 2010.

FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO <p style="text-align: center;">TC</p>	2. DATA 22 de novembro de 2021	3. REGISTRO N° DCTA/ITA/TC-102/2021	4. N° DE PÁGINAS 150
5. TÍTULO E SUBTÍTULO: Características socioeconômicas e comportamentais de passageiros: influências e correlações com o assento ocupado a bordo de aeronaves.			
6. AUTOR(ES): Marcelo Gonçalves Concli			
7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): Instituto Tecnológico de Aeronáutica – ITA			
8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: Clusterização; Layout e disposição de assentos; Perfil socioeconômico; Passageiros.			
9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: Transporte de passageiros; Assentos; Layouts; Análise de aglomerados; Projetos; Indicadores socioeconômico; Transporte aéreo; Transportes.			
10. APRESENTAÇÃO: <input checked="" type="checkbox"/> Nacional <input type="checkbox"/> Internacional ITA, São José dos Campos. Curso de Graduação em Engenharia Civil-Aeronáutica. Orientador: Prof. Dr. Alessandro Vinícius Marques de Oliveira. Publicado em 2021			
11. RESUMO: Desde seu princípio, a indústria da aviação experimenta de maneira constante evolução e modernização em seus diversos pilares. Um dos assuntos que entra nessa jornada evolutiva é o layout ou disposição dos assentos, bem como a modernização dos materiais utilizados na construção deles. No aspecto da utilização de materiais, a busca sempre foi direcionada à redução de peso da aeronave. No aspecto de posicionamento, a otimização proposta e atualmente percebida na prática é o aumento da capacidade de transporte da aeronave, de maneira que coloquem maior quantidade de assentos em menor área, para que se haja melhor diluição de custos da viagem pelo número de passageiros. Os recentes avanços dessa estrutura de pensamento estão sendo levados por uma geração que idealiza novos tipos de configuração. Como destaques e objeto de estudo, estão as configurações Skyrider e Checkerboard. A primeira traz uma proposta que tem por base alterar o assento para uma estrutura similar a um banco de bicicleta, na qual o encosto e assentos se alinham numa posição quase vertical, possibilitando o aumento de capacidade das aeronaves. A segunda proposta de configuração pode ser comparada a um tabuleiro de damas, na qual se obtém liberdade e flexibilidade para modelar a configuração de acordo com a demanda do voo, de maneira que seja possível transitar entre assentos mais espaçados (considerados premium) e assentos mais juntos (considerados econômicos), de maneira que se ganhem possibilidades de combinação entre áreas premium e áreas econômicas. Nessa linha de pensamento, a pesquisa "Estudo da densificação e segmentação de cabine pela introdução de conceitos inovadores de assentos em aeronaves da aviação comercial" em seu escopo global tem o objetivo de analisar a viabilidade econômica das configurações citadas de layout e posicionamento de assentos dentro de aeronaves. Como parte fundamental dessa análise, reside o estudo direcionado ao longo deste trabalho de graduação sobre a segmentação de perfis dos passageiros que ocupam os assentos, de maneira que se identifiquem dentro de um conjunto de características socioeconômicas e comportamentais possíveis agrupamentos e correlações com o posicionamento daquele passageiro. A base de dados que orienta o trabalho é resultado de um questionário aplicado entre os meses de Fevereiro e Setembro no ano de 2014 em 66 aeroportos, com 122.039 respondentes em um total de 3.793 voos. O questionário é composto por 107 perguntas. Dentre as respostas coletadas, obtêm-se informações que variam desde informações sobre trecho, gastos envolvidos durante os trajetos e estadias, perfil do consumidor (faixa salarial, profissão e outros) até um mapeamento sobre a opinião do respondente acerca do segmento de infraestrutura aeroportuária. Através de análise na base de dados proveniente do questionário, será possível a identificação de "clusters" ou grupos de passageiros e relações claras de influência / correlação destes nas posições ocupadas por cada um dos passageiros. O método de clusterização a ser utilizado deve se aplicar tanto à variáveis categóricas (como profissão, motivo da viagem, grupo salarial, meios de transporte utilizados nos trajetos e assento ocupado) quanto à variáveis não-categóricas (como gastos no interior dos aeroportos, nos trajetos, idades e outros). Como proposta inicial, serão exploradas as metodologias "K-Modes" e "K-Prototypes", derivadas da "K-Means" (aplicável para variáveis não-categóricas) com as devidas modificações para incluir variáveis categóricas.			
12. GRAU DE SIGILO: <p style="text-align: center;"><input checked="" type="checkbox"/> OSTENSIVO <input type="checkbox"/> RESERVADO <input type="checkbox"/> SECRETO</p>			