



UNIFEOB

CENTRO UNIVERSITÁRIO DA FUNDAÇÃO DE ENSINO
OCTÁVIO BASTOS

PROJETO INTEGRADO

InvestAER - Investir na Concessão de Aeroportos
**Suporte à decisão de priorização entre aeroportos
públicos em leilão de concessão**

SÃO JOÃO DA BOA VISTA, SP

SETEMBRO, 2022

UNIFEOB

CENTRO UNIVERSITÁRIO DA FUNDAÇÃO DE ENSINO
OCTÁVIO BASTOS

PROJETO INTEGRADO

InvestAER - Investir na Concessão de Aeroportos
**Suporte à decisão de priorização entre aeroportos
públicos em leilão de concessão**

MÓDULO: IHC e ESTATÍSTICA

INTERFACE HUMANO-COMPUTADOR- PROF. MSC. RODRIGO
MARUDI

PROBABILIDADE DE ESTATÍSTICA - PROF. ESP. CARLOS
COLLOZZO

ESTUDANTES:

Patrick Dutra - 1012022100534

José Carlos Martins - 1012022100497

Thais de Moraes Vieira - 1012022100664

Bruno Salim - 1012022100006

Rainer Ferraz Passos - 1012022100326

SÃO JOÃO DA BOA VISTA, SP

SETEMBRO, 2022

SUMÁRIO

| | |
|--|-----------|
| 1. INTRODUÇÃO | 3 |
| 2. DESCRIÇÃO DO TEMA | 5 |
| 3. PROJETO INTEGRADO | 6 |
| 3.1 INTERFACE HUMANO-COMPUTADOR | 7 |
| 3.2 PROBABILIDADE DE ESTATÍSTICA | 8 |
| 4. CERTIFICAÇÃO DO PI E COMPETÊNCIAS | 10 |
| 5. CONCLUSÃO | 12 |
| REFERÊNCIAS | 13 |
| APÊNDICE A - Pairplot voos internacionais | 14 |
| APÊNDICE B - Pairplot voos domésticos | 17 |
| APÊNDICE C - Heatmap voos internacionais | 21 |
| APÊNDICE D - Heatmap voos domésticos | 26 |
| APÊNDICE E - Links das fontes de dados e artefatos utilizados | 33 |
| APÊNDICE F - Código utilizado no Colab para os dados de voos internacionais | 34 |
| APÊNDICE G - Código utilizado no Colab para os dados de voos domésticos | 37 |
| APÊNDICE H - Relatórios criados e apresentados no site do projeto | 40 |
| ANEXO A - Artigo de origem da ideação do tema | 43 |

1. INTRODUÇÃO

O Projeto Integrado do terceiro trimestre de 2022, conforme esclarecido em seu Plano Pedagógico, tem como objetivo desenvolver um projeto de análise e apresentação de dados reais, utilizando fontes de dados fornecidos pelos docentes, utilizando conceitos de estatística e interface humano-computador.

Nesse sentido, o referido plano pedagógico contextualiza o trabalho da seguinte forma:

TEMA GERADOR/ EIXO TEMÁTICO/ PROJETO INTEGRADO

Desafio do módulo

O Projeto Integrado é um componente curricular multidisciplinar, interdisciplinar e transdisciplinar de cunho teórico-prático que engloba temas abordados por todas as unidades cursadas pelo estudante, alinhado ao eixo temático de cada módulo.

O título do PI é “ANÁLISE ESTATÍSTICA DE DADOS”, cuja descrição é: “Você é um profissional que foi convocado para fazer parte de uma equipe que deve desenvolver um sistema de análise de dados estatísticos voltados a área de projetos institucionais. Seu trabalho será desenvolver um projeto de análise e apresentação de dados reais, utilizando fontes de dados fornecidos pelos docentes, utilizando conceitos de estatística e interface humano-computador.”

Diante desse desafio, foram iniciadas as interações entre os membros do Grupo 10, optando por tratar de um caso hipotético, de uma equipe alocada ao suporte para a análise de dados de satisfação dos usuários de aeroportos públicos brasileiros, com a finalidade de complementar os dados financeiros e estruturais que os decisores de uma empresa disponham, tudo a fim de facilitar uma tomada de decisão embasada nesses diversos pontos de vista (financeiro, estrutural e satisfação de usuários), para priorizar quais aeroportos buscar nos leilões de concessão.

Esses processos de concessão dos aeroportos públicos, naturalmente, envolvem investimentos de grande monta, por isso a decisão sobre participar ou não e, ainda mais, as decisões sobre até quanto se pode investir na concessão e sua expectativa de retorno, são imprescindíveis para o sucesso dos processos.

Não é interessante nem para as empresas, nem para os usuários e para a gestão pública, que uma empresa embarque nesse processo e depois desista. Nessa hipótese, a empresa provavelmente ficaria com um grande prejuízo, os usuários poderiam ficar com um ambiente de prestação de serviços de baixa qualidade e a gestão pública teria que

arcar com processos de concessão repetidos, além de ter que assumir a manutenção das instalações interinamente.

Assim, o projeto ilustra como os Processos de Gerenciamento de Recursos de Dados podem gerar valor para as organizações, neste caso específico, em uma futura tomada de decisão ao entrar em leilão de aeroporto para concessão. Cenário em que é importante, senão imprescindível, que o investidor obtenha várias informações e faça comparação entre os indicadores, extraídos, por exemplo, até com Inteligência Artificial (IA), em busca de garantir a competitividade da empresa no mercado. Essa é mais uma situação que ilustra o benefício do alinhamento da TI às áreas estratégicas de negócios.

2. DESCRIÇÃO DO TEMA

A proposta deste trabalho foi explorar e apresentar dados reais de uma avaliação realizada anualmente, em 20 aeroportos brasileiros. A pesquisa contava com 37 critérios específicos, desde “limpeza dos sanitários” e “custo-benefício do estacionamento”, até “tempo na fila do check-in” e “integração com transporte público”, e mais a avaliação geral do aeroporto. A caracterização dos respondentes também é bem completa, incluindo idade, média salarial e motivo e frequência de viagem, num total de sete quesitos descritivos dos entrevistados.

Neste Projeto Integrado (PI) 03 de 2022, o Grupo 10 simulou assumir a função de construir uma interface para a visualização de dados e sua integração em informações, no cenário hipotético de apoio à tomada de decisão, de empresários interessados em obter o direito de exploração de aeroportos, nos leilões de concessão.

A ideia chegou a esse formato em função da notícia replicada no Anexo A (LEILÃO, 2022).

Para atingir o objetivo, foram utilizados dados abertos de pesquisa de satisfação já citada, obtidos a partir de página da administração federal (PESQUISA, 2022), expostos e explorados nos vários dashboards e gráficos inseridos na página web de apresentação do exercício:

<https://sites.google.com/sou.unifeob.edu.br/pi202203grupo10/in%C3%ADcio>

Assim, considerou-se que, para esse tema, seria natural que os empresários investidores tivessem interesse em avaliar a relação de “custo X benefício” do investimento no direito de exploração de cada aeroporto.

Em suas reuniões e análises para decidir o valor limite, até o qual cada instalação se manteria vantajosa, os usuários precisariam, além dos dados financeiros e de projeção de movimento aéreo, de indicações um pouco mais qualitativas da percepção dos usuários. Ao final, essa informação, condensada dos dados de pesquisas com passageiros, deveria compor a visão mais detalhada de cada instalação.

Deve-se ressaltar a restrição, deste exercício educativo, de simular que os principais aeroportos listados estariam disponíveis para leilão, o que não é a realidade, pois vários leilões já ocorreram, em diferentes lotes.

3. PROJETO INTEGRADO

A proposta deste trabalho, como descrito anteriormente, foi apresentar uma pesquisa que abordou vários critérios, com a possibilidade de relacioná-los com as características dos entrevistados que emitiram suas percepções.

No intuito de utilizar metodologia estatística para capturar dados e construir informações, por meio de métodos que permitissem comparar os aeroportos entre si, conforme a visão dos passageiros, buscando alguma correlação estatística, e facilitar que os usuários do sistema tivessem maior facilidade para encontrar algum padrão e interpretar variáveis quantitativas e qualitativas e obter decisões mais adequadas.

Nesse processo, constam do Apêndice E os links de todos os artefatos, utilizados e construídos por nós, para a execução da atividade e sua metodologia. Na ordem desses links pode-se observar:

- (1) a origem dos dados;
- (2) depois a planilha Google onde limpamos os dados e construímos médias das avaliações por áreas;
- (3) em seguida o ambiente do Data Studio, onde construímos os relatórios / dashboards, ilustrados aqui no Apêndice H;
- (4) esses relatórios foram inseridos no site construído no ambiente Google Sites;
- (5) cujo endereço de publicação também foi disponibilizado por link;
- (6) depois, os dados foram separados entre os voos domésticos e internacionais, em planilhas excel, também disponíveis para verificação por seus links; e
- (7) finalmente, os trechos de código utilizados no ambiente Colab do Google para obter visualizações gerais e gráficas da avaliação de correlação entre as notas dadas por áreas e as notas gerais, atribuídas pelos passageiros.

Esse cuidado na construção de visualizações mais fáceis de obter a visão do todo e intuir hipóteses de tendências, acabaram por expressar a prática de empatia com o usuário do sistema, conforme descrito em mais detalhes a seguir.

3.1 INTERFACE HUMANO-COMPUTADOR

Este trabalho apresenta uma integração de uma base de dados em um sistema de informação, com a finalidade de facilitar a visualização desses dados em diferentes formas de organização. Essa variação de formas de organização deve permitir identificar aquelas que melhor se adequem às necessidades ou hábitos dos decisores, os quais seriam os consumidores da ferramenta de apoio à decisão.

Essa possibilidade de variação representa a empatia dos formuladores da ferramenta com os seus “consumidores”, na medida em que evitam “cristalizar” o sistema na forma mais habitual apenas para seu desenvolvedor.

Todo esse processo representa a estrutura que deve permitir a conexão da gestão com os dados e a informação que se pode construir com eles. Dessa forma contribui para a tomada de decisão assertiva pelos gestores. Mais especificamente, nas palavras de autores que tratam da relação entre os sistemas de informação e a business intelligence (BI):

[...] por meio do uso das tecnologias e dos processos de BI, as organizações podem compreender elementos e fatores fundamentais — internos e externos — que afetam seus negócios e sua produtividade no mercado. O BI possui parâmetros e análises sofisticados para “enxergar dentro dos dados” e encontrar relacionamentos e oportunidades que possam ser transformados em benefícios (O'BRIEN; MARAKAS, 2013, p. 9).

O processo de construção dos relatórios, apresentados para consumo pelos decisores, se iniciou com uma tentativa de deixar em uma única tela (dashboard) todos os dados que imaginamos que seriam úteis para que um usuário pudesse construir suas perguntas sobre os dados.

Rapidamente, ficou claro que, apesar de esse formato parecer otimizado, não facilitava que o usuário encontrasse o que buscava, nem enxergasse bem o retorno oferecido em cima dos dados do dashboard.

Por isso, realizamos a divisão já descrita no site, em que as visualizações foram divididas em 04 "questões básicas":

(1) Localização e avaliação - permitindo a pessoa sair do geral para o específico de forma referenciada à geografia do mundo real;

(2) Estratificação dos usuários - permitindo “mergulhar” nos números e dados, inicialmente com dados gerais das pessoas e depois com dados voltados às possibilidades de gasto das pessoas nos aeroportos;

(3) Áreas com espaço a investir - destacando as médias das avaliação dos aeroportos, ou seja, não mais olhando para as pessoas, mas para as oportunidades de melhoria das instalações, equipes e serviços; e

(4) Correlação entre essas áreas e a satisfação geral - onde a estatística é explorada, mas com visualização gráfica (cores e proximidade/formas de curvas), para permitir intuir diferenças nas percepções dos entrevistados e diferenças ou similaridades dos aeroportos entre si.

Esse cuidado de dividir o que se poderia observar dos dados em “questões básicas” e organizá-las de forma a saírem do geral para o específico, sempre que possível de forma geográfica ou gráfica, representa o nosso esforço concreto de empatia com o usuário de nosso sistema.

3.2 PROBABILIDADE DE ESTATÍSTICA

Os dados originais coletados vieram no formato open office e contavam com mais de 600.000 avaliações, isso fazia tanto o excel, quanto o Google planilhas travar. Tivemos que fazer uma filtragem dos dados com a retirada de algumas informações como os anos de 2013 a 2017 e todas as linhas que não tinham a respostas correspondentes às informações do entrevistado, restando 107.686 respostas analisadas.

Ainda assim, identificar possíveis relações, usando 37 notas distintas, por entrevista, seria difícil de observar, por isso, essas 37 respostas foram tratadas por meio de 05 médias por área, de cada respondente.

A pesquisa demonstrou que o critério que interferiu mais diretamente na nota de satisfação do aeroporto foi a infraestrutura, assim os outros critérios utilizados na pesquisa não tiveram relevância para alterar diretamente a média de satisfação geral dos respondentes com cada aeroporto.

Essa ideia de influência de variáveis, como as notas das áreas, levou à ideia de que poderiam ser variáveis independentes, influenciando uma variável dependente, a nota de satisfação geral.

Sendo esse o caso, deveria ser possível fazer a regressão, identificando coeficientes para as notas que levariam à nota geral. No entanto, a interpretação do grupo foi de que, apesar de serem números de 1 a 5, as notas na realidade estariam em uma escala ordinal, “(...) onde a ordem ou classificação dos dados for [é] significativa”, mas sem “(...) intervalo entre os valores expresso em uma unidade fixa” (ANDERSON *et al.*, 2021, p. 5), pois seriam equivalentes a uma escala do tipo “discordo completamente, discordo parcialmente, indiferente, concordo parcialmente e concordo plenamente”, e, por isso, a tentativa de obter uma regressão linear numérica não seria precisa.

4. CERTIFICAÇÃO DO PI E COMPETÊNCIAS

O processo descrito na seção anterior, com as considerações específicas sobre as áreas de estudo envolvidas, levaram a algumas percepções enquanto aprendizado sobre o caminho que se percorre, na busca de tratar dados e facilitar sua visualização e análise:

(1) Grandes quantidades de dados necessitam de ferramentas (softwares) adequados para processar, interpretar e construir visualizações de informação. Quando os dados atingem quantidades que realmente fazem diferença para a extração de informação, as ferramentas gratuitas ou mesmo o MS Excel, têm dificuldades para realizar os cálculos (mesmo simples como médias) e gerar gráficos. Organizar e limpar os dados consumiu mais da metade do tempo de todo o projeto.

(2) Mesmo quando se consegue grande quantidade de dados relevantes, abertos e interessantes, não se pode usar todos, porque muitas vezes a coleta deixa campos vazios. Por isso a limpeza seguiu um critério que foi o de priorizar a capacidade do dado de nota ser relacionável com a descrição dos passageiros, porque o objetivo, neste caso específico, era que os empresários pudessem fazer essa estimativa de que tipo de pessoas utilizam e buscam melhorias nos aeroportos.

(3) De forma semelhante ao processo de limpar os dados, foi necessário um de “enriquecer” os dados com novas métricas, sem as quais ficava difícil entender a relação entre as expectativas dos passageiros e suas notas de avaliação geral dos aeroportos. Neste caso específico, bastou a média, por áreas, para permitir ver as notas específicas de forma agrupada em possíveis áreas de interesse dos empresários.

(4) O uso do Data Studio deixou evidente a utilidade de visualizações por gráficos dinâmicos, que se recalculam conforme o estrato selecionado dos respondentes (Apêndice H). Percebemos que não só a informação adequada, mas também a visualização mais clara, acessados pela organização, facilitam o posicionamento e a decisão, contribuindo na estratégia e no crescimento do mercado.

(5) Ainda que esses relatórios fiquem adequados, sua disponibilização ao usuário, via site publicado por meio do Google Sites, teve que seguir o cuidado de inserir instruções breves, antes de cada relatório, orientando a seleção de opções e “onde olhar” em cada dashboard.

(6) Além do que se podia obter com o Data Studio, o teste em busca de eventuais correlações, entre as notas das áreas e as notas gerais, levou ao uso de outra ferramenta, o Colab Google, onde foram obtidos esses índices de correlação. Mais uma vez percebeu-se que era mais fácil observar essa métrica estatística por meio de representações gráficas e com cores (os mapas de calor e *pairplots*, replicados aqui nos Apêndices A, B, C e D).

(7) Finalmente, percebeu-se que, para gerenciar, é necessário o conhecimento de ferramentas para processar e organizar informações em tempo oportuno, de modo efetivo e tornando preciso o sistema tratado pela organização.

5. CONCLUSÃO

O projeto Interface humano computador, abordado, apresenta a importância da tecnologia contemporânea. Esse processo, no ambiente real, contempla necessidade do executivo e gestores contarem com uma ferramenta de apoio para compreender as informações disponíveis na IA, nos dados armazenados em central (ou nuvem), nas Data Warehouses, ou seja, nos meios de Big Data em geral.

Já no ambiente de ensino, nosso projeto percorreu os passos desde a busca da base de dados e sua “limpeza”, parte que demandou maior esforço não esperado, até a exposição de relatórios dinâmicos, passando pelo aprendizado do código Python mínimo necessário para explorar o ambiente Colab e obter a métrica estatística de correlação e sua representação em formas empáticas para visualização pelo usuário.

Essa atividade trouxe, além da experiência de entender que a preparação dos dados é a etapa basilar para todo o processo, também a percepção de que as ferramentas de planilhas, representação e análise devem ser muito bem escolhidas, pois o aumento da quantidade de dados rapidamente torna inviável o uso das ferramentas mais populares.

Finalmente, percebeu-se que empatia torna a visualização mais clara e eficaz. É preciso entender que grande quantidade de dados leva a grande variedade de formas de visualização possíveis, no entanto, para se manter útil, é necessário balancear o que se pergunta e quais respostas se deseja, conforme a demanda específica de cada caso e dos usuários esperados.

REFERÊNCIAS

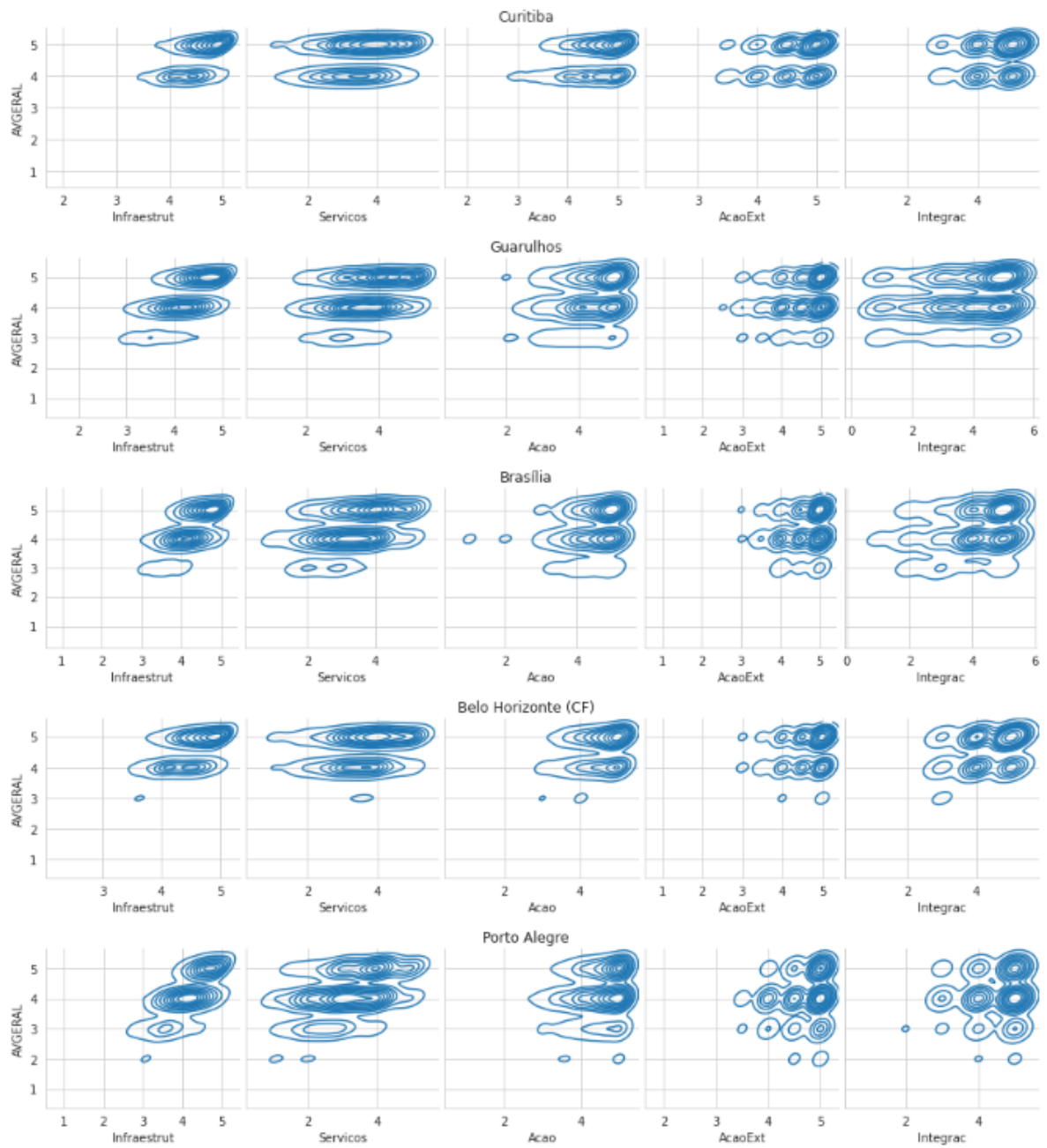
ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A; CAMM, J. D.; COCHRAN, J. J. **Estatística aplicada à administração e economia**. 8. ed. São Paulo: Cengage Learning, 2021.

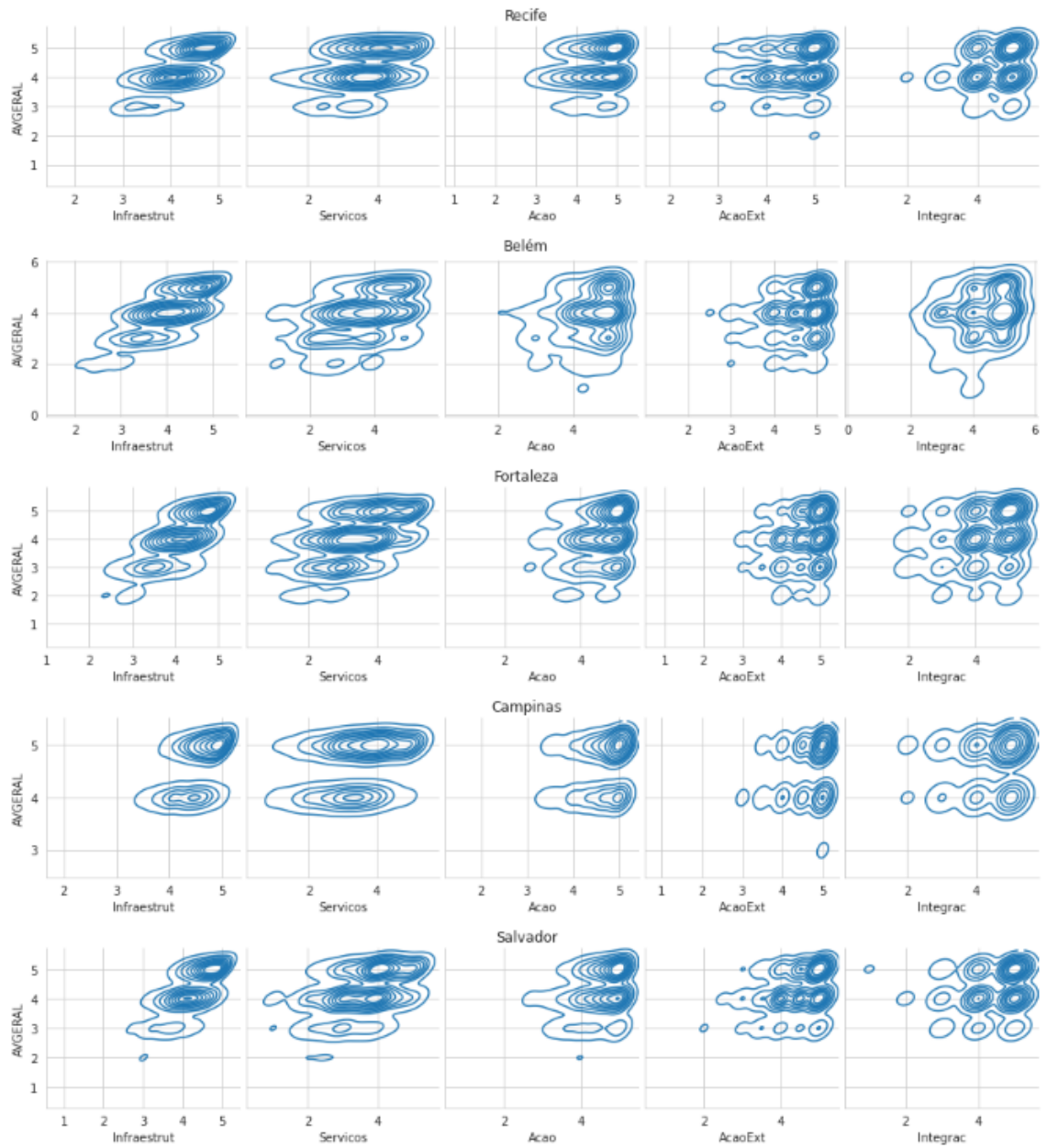
LEILÃO de aeroportos liderado por Congonhas gera R\$2,72 bilhões em contribuição inicial. **gov.br Ministério da Infraestrutura, Agência Nacional Aviação Civil (ANAC)** 2022. Disponível em: <https://www.gov.br/anac/pt-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>. Acesso em: 24 ago. 2022.

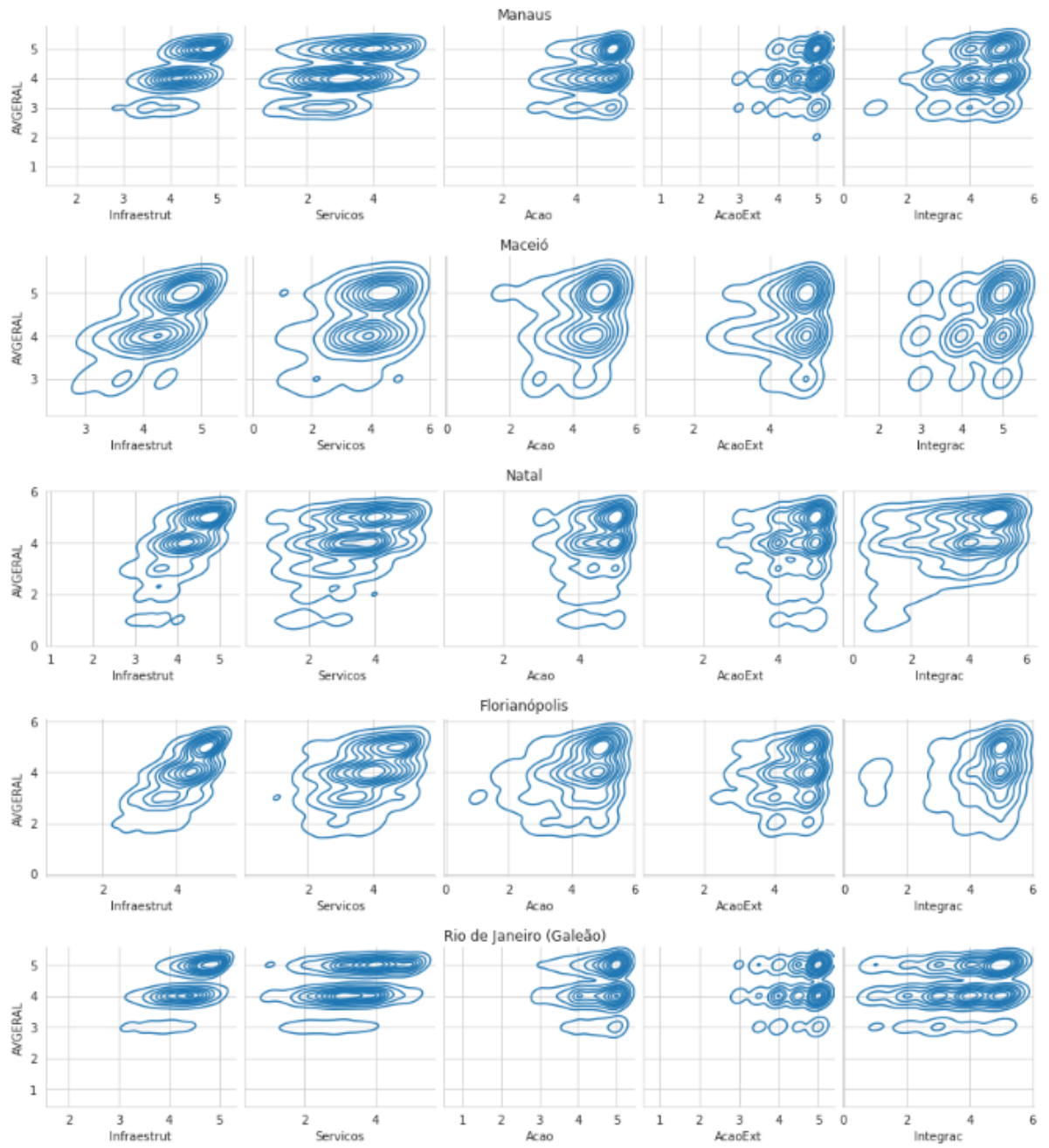
O'BRIEN, J. A.; MARAKAS, G. M. **Administração de Sistemas de Informação**. Porto Alegre: McGraw-Hill, 2012.

PESQUISA de satisfação do passageiro em aeroportos. **Portal Brasileiro de Dados Abertos**. 2022. Disponível em: <https://dados.gov.br/dataset/pesquisa-de-satisfacao-do-passageiro-em-aeroportos>. Acesso em: 05 set. 2022.

APÊNDICE A - Pairplot voos internacionais

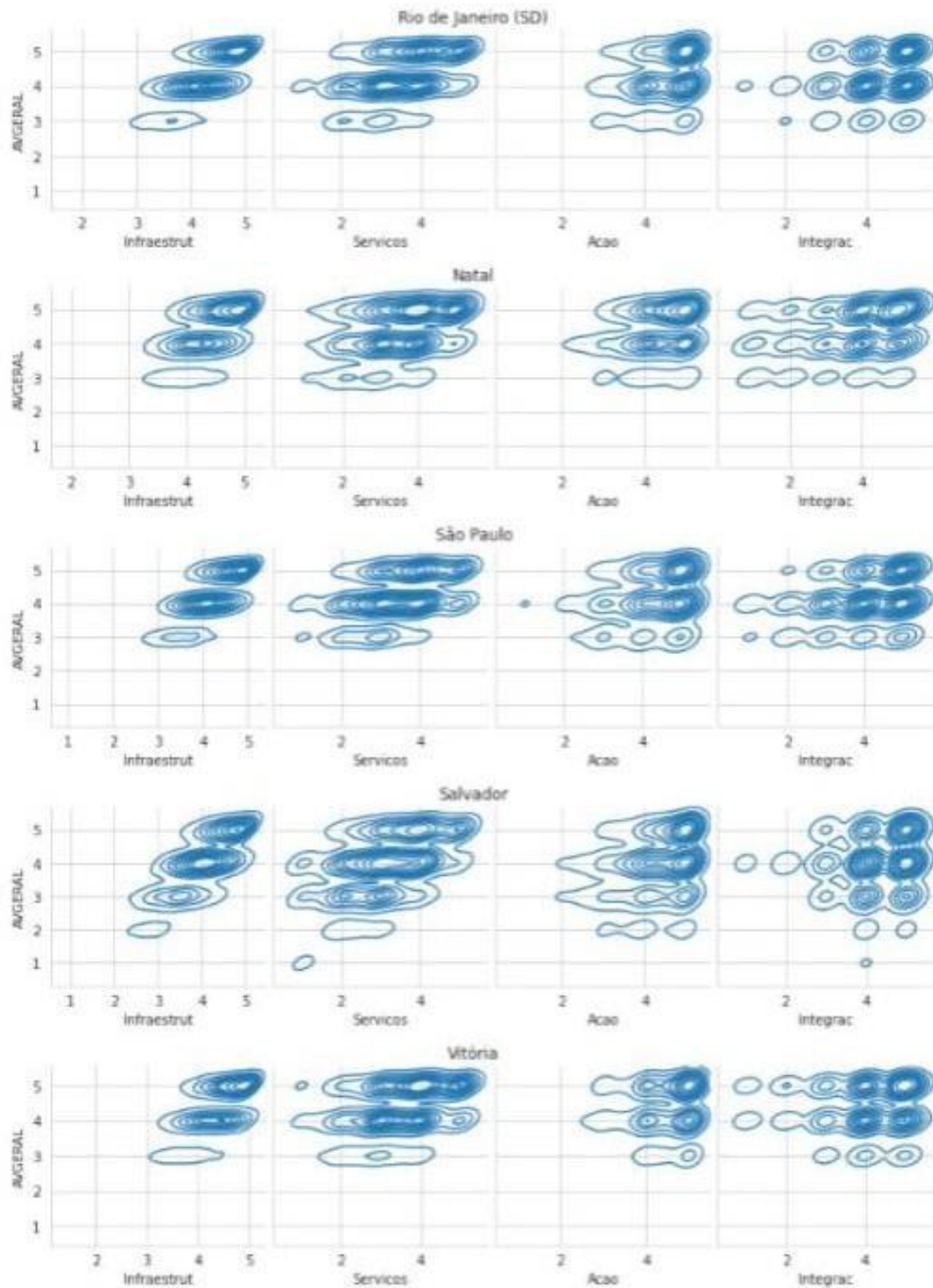


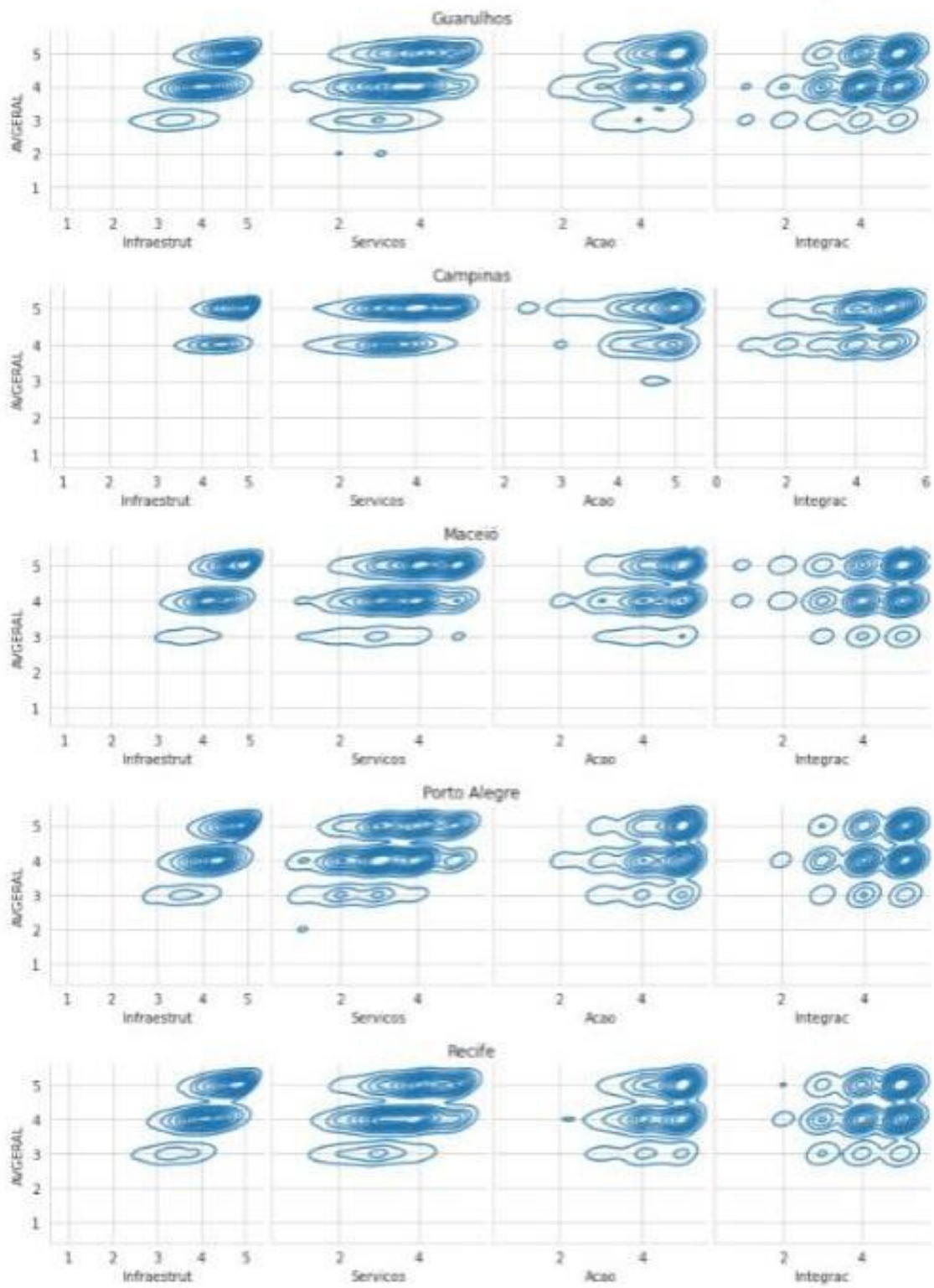


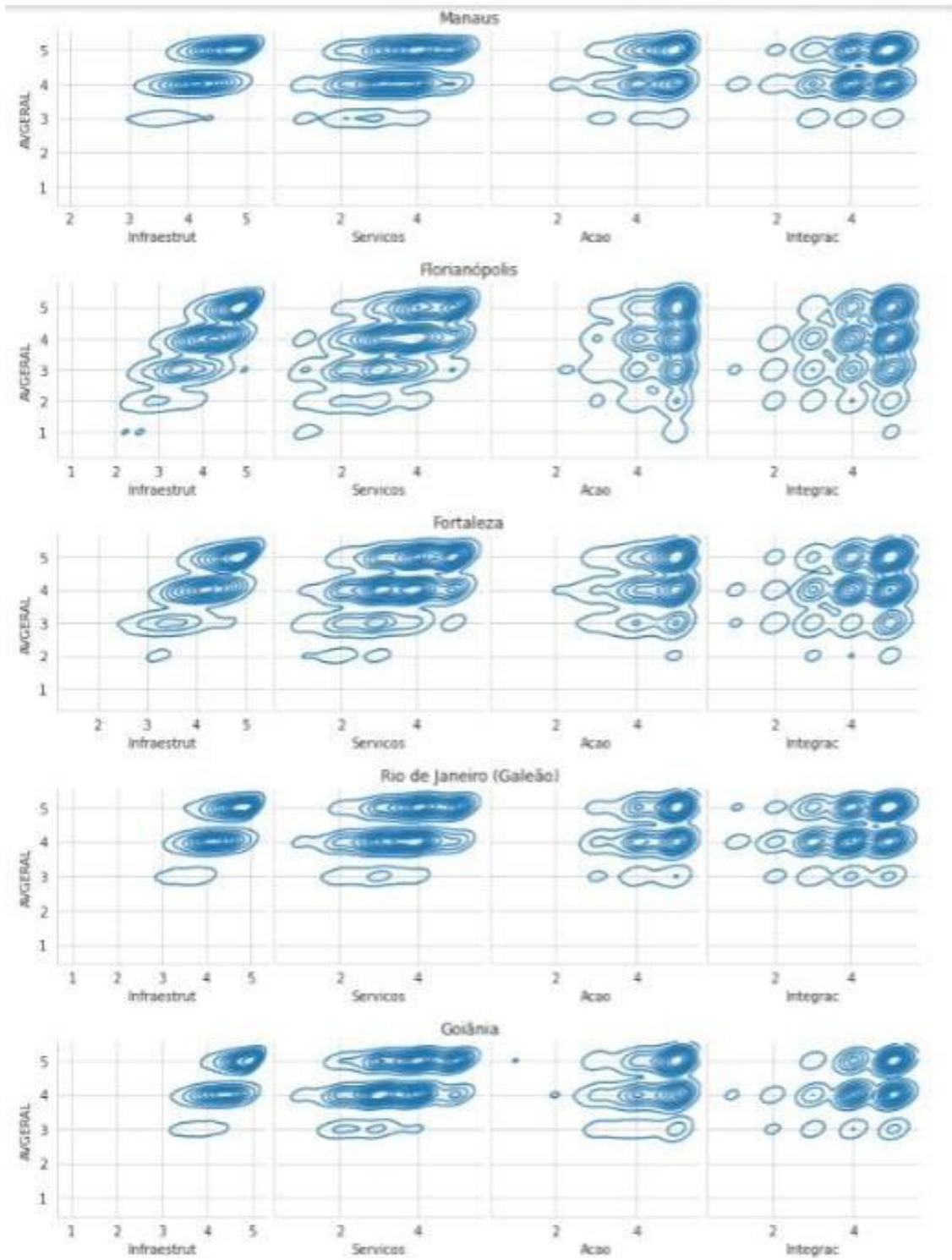


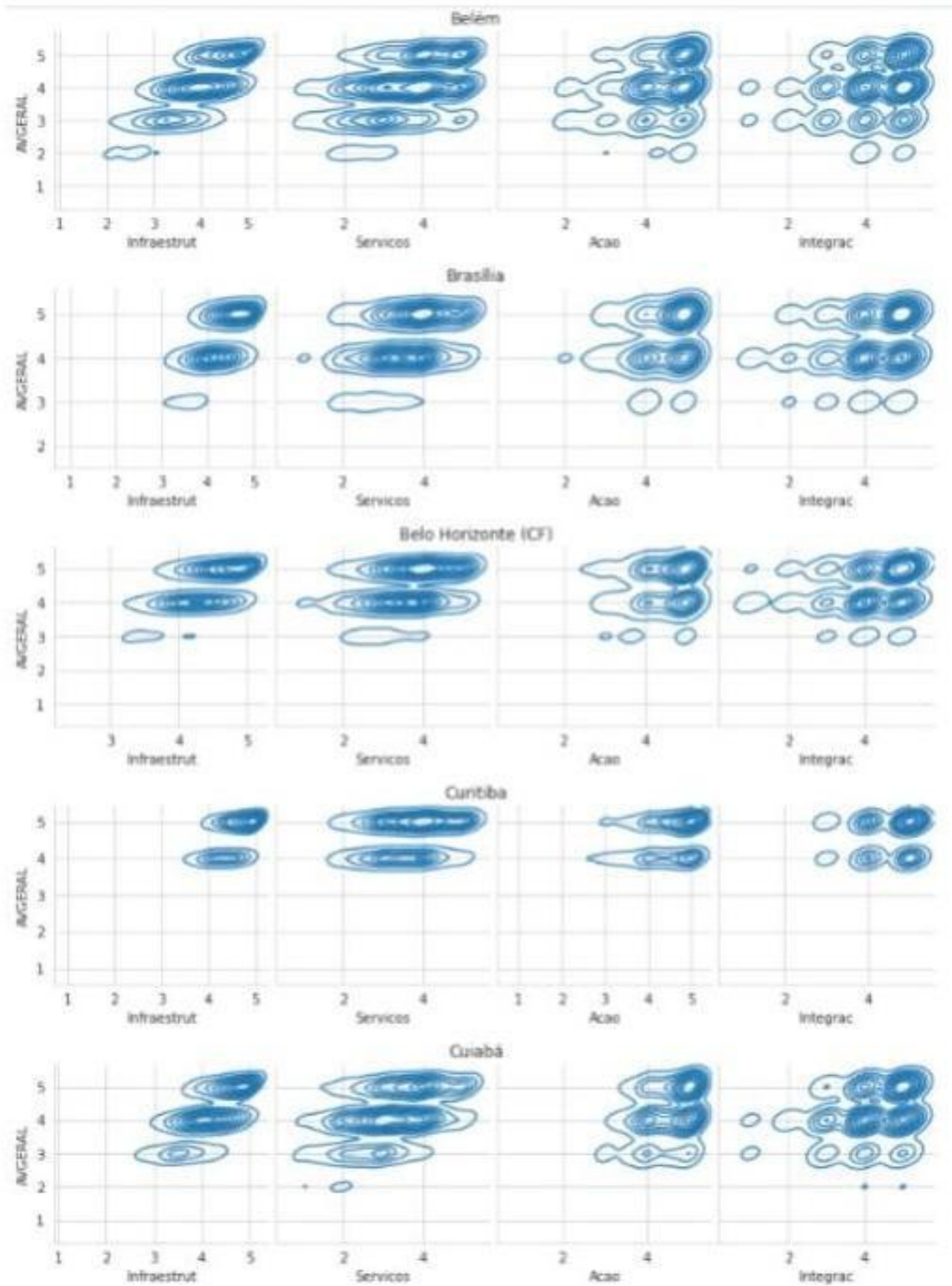
✓ 2m8s conclusão: 19:13

APÊNDICE B - Pairplot voos domésticos



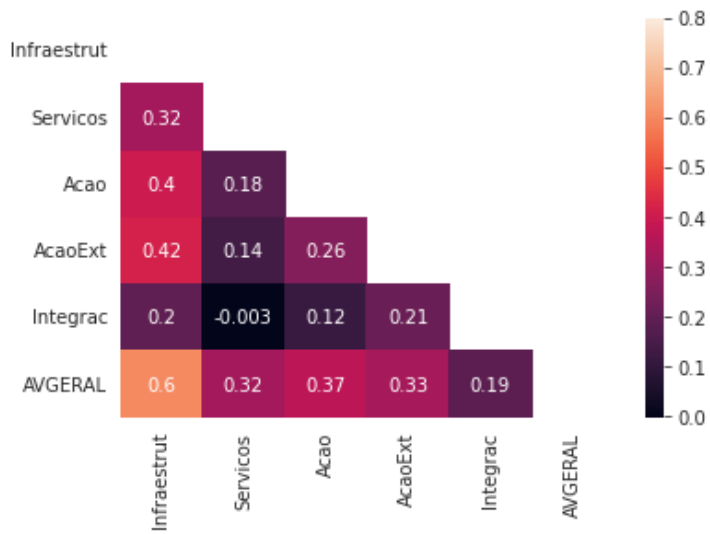




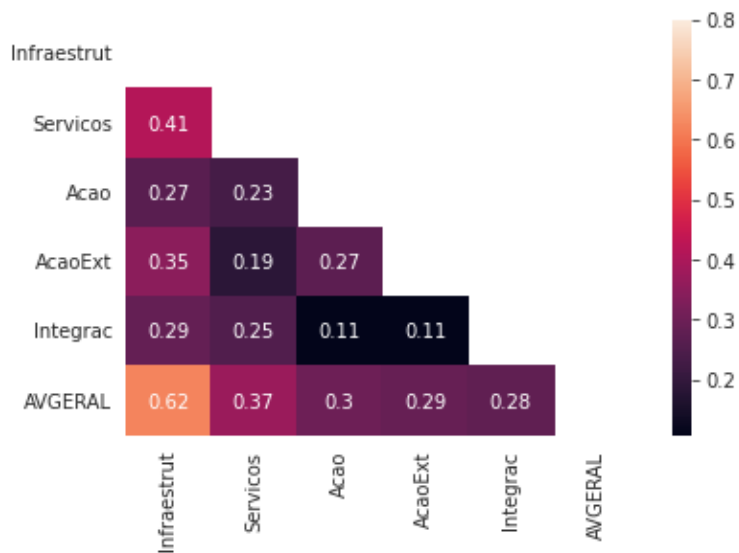


APÊNDICE C - Heatmap voos internacionais

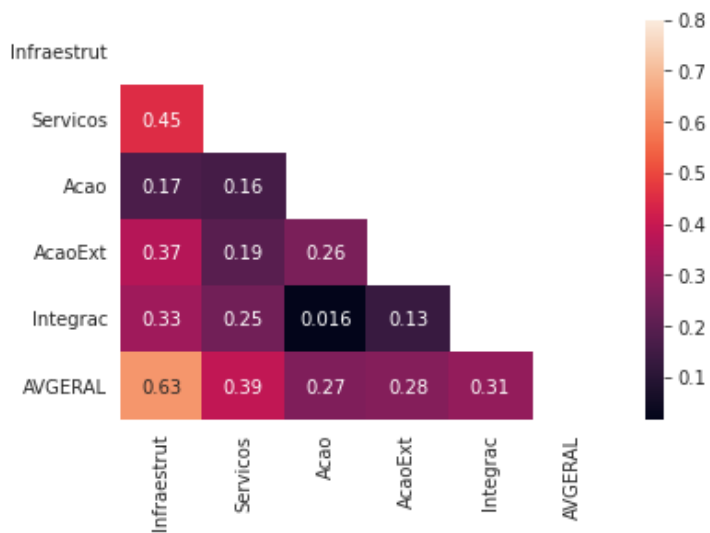
Curitiba



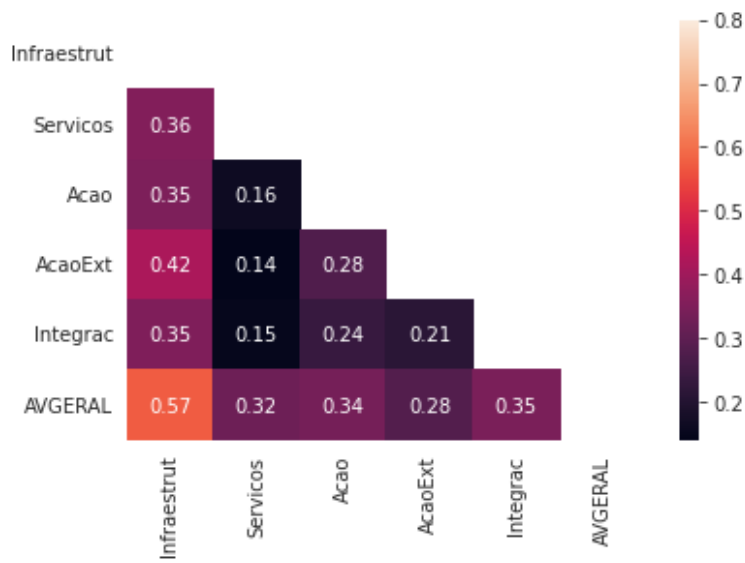
Guarulhos



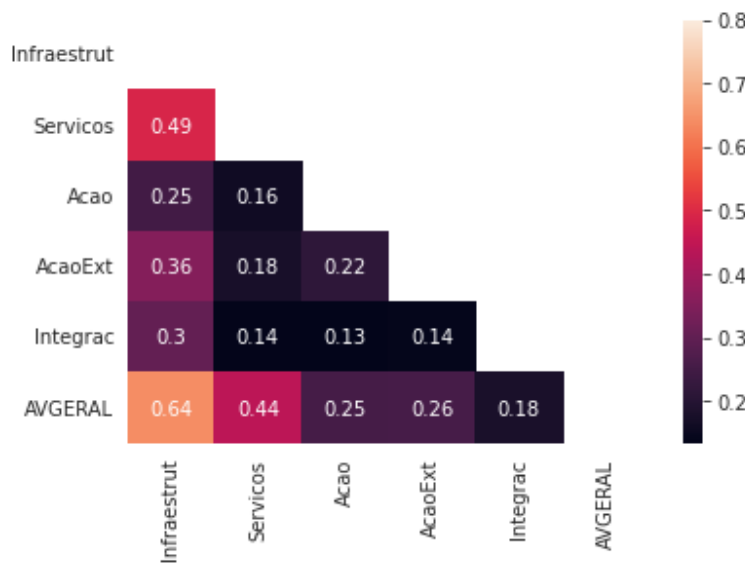
Brasília



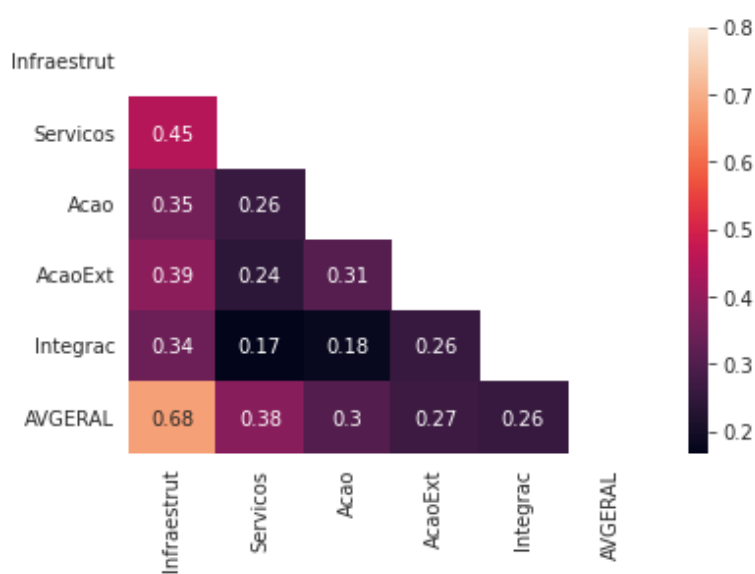
Belo Horizonte (CF)



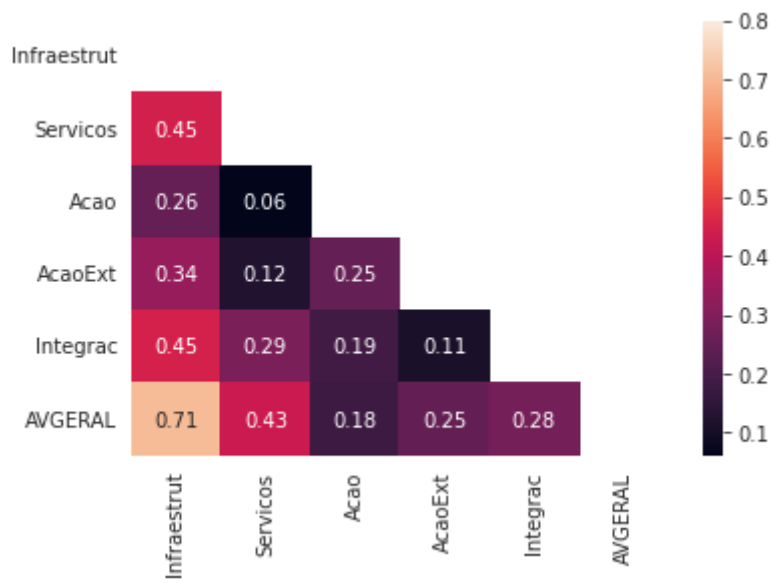
Porto Alegre



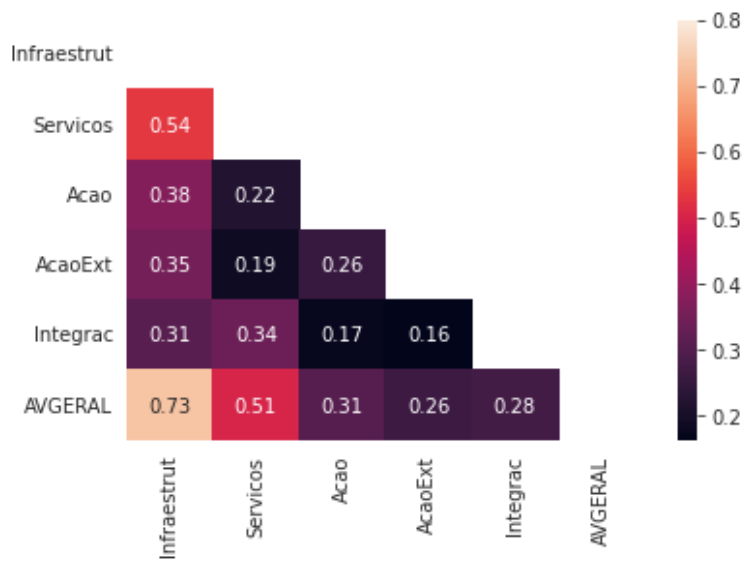
Recife



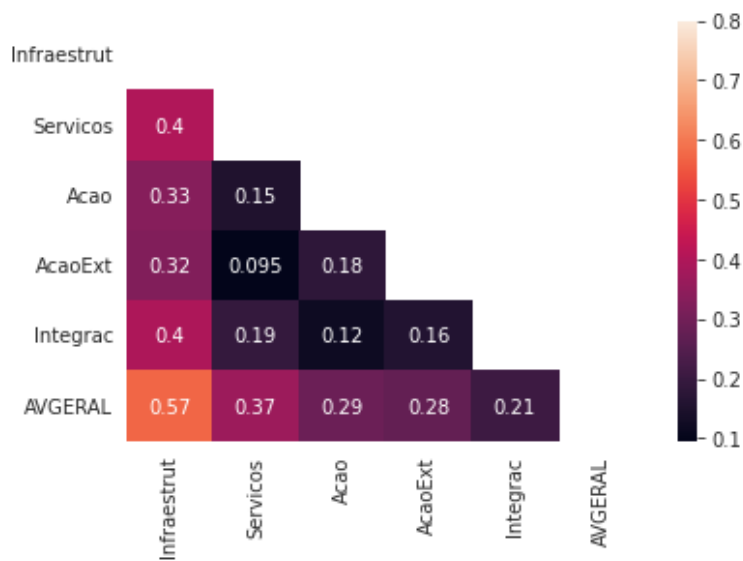
Belém



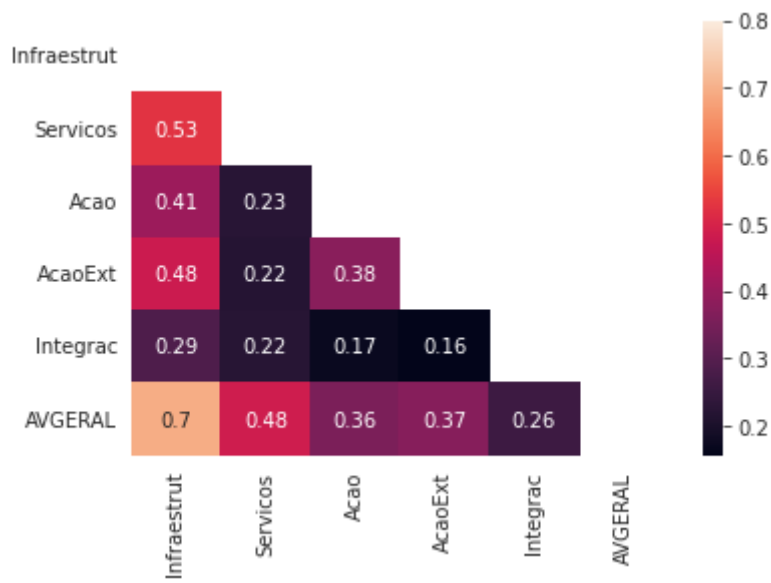
Fortaleza



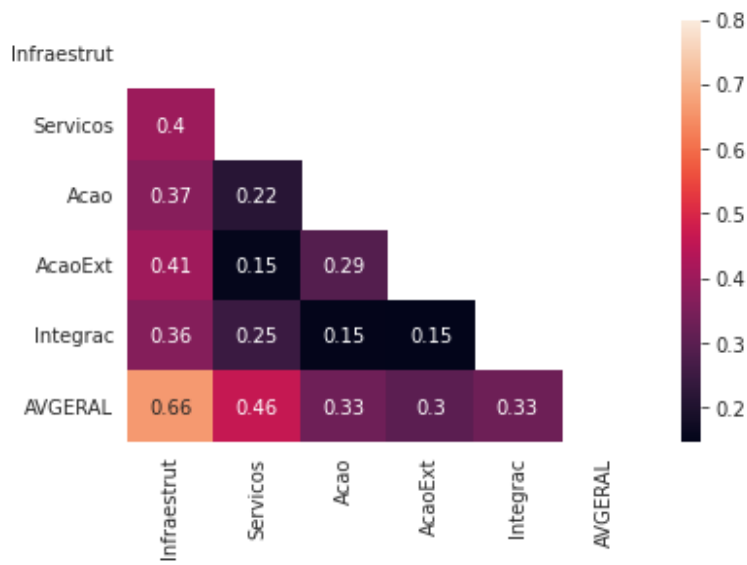
Campinas



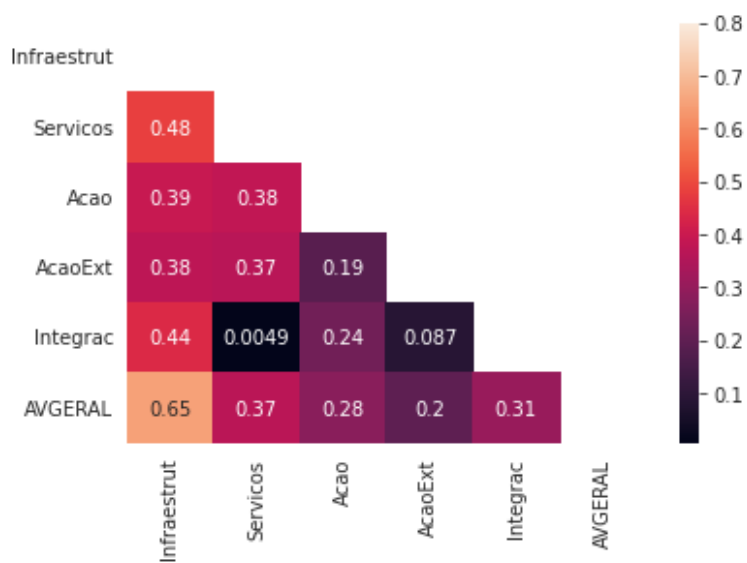
Salvador



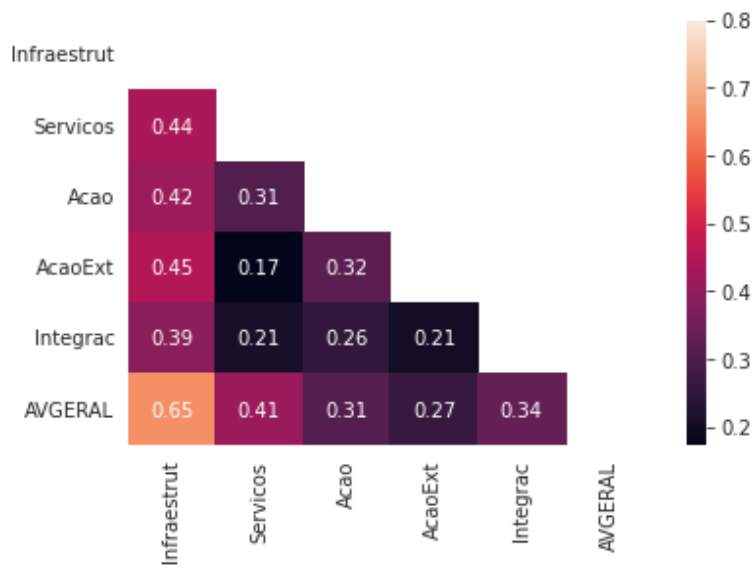
Manaus



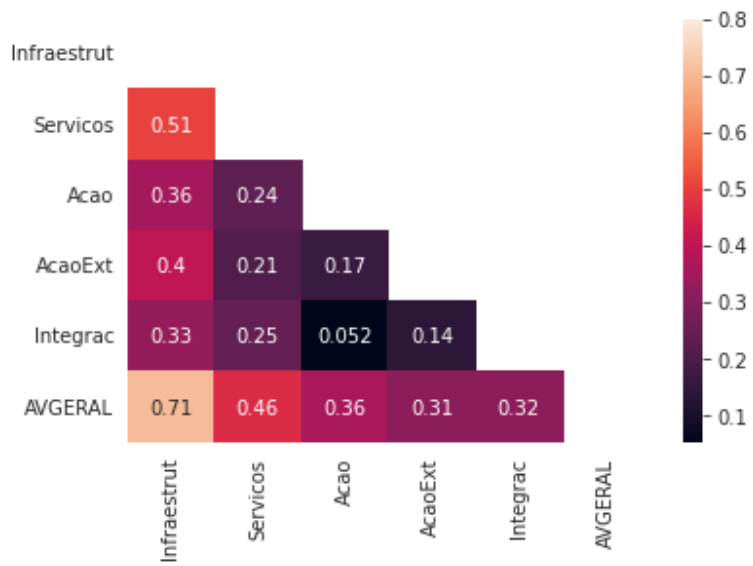
Maceió



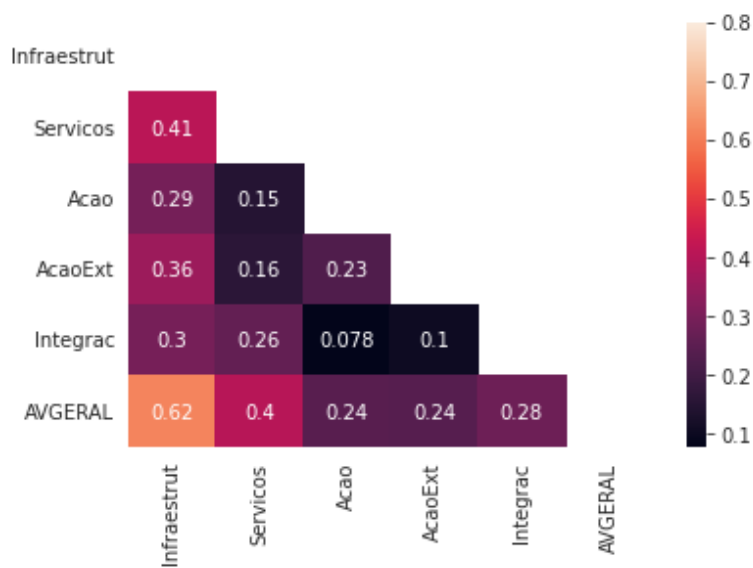
Natal



Florianópolis

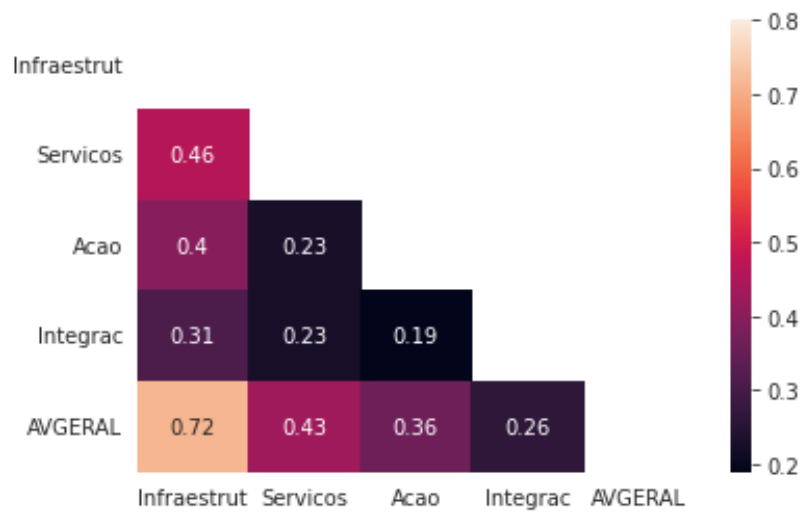


Rio de Janeiro (Galeão)

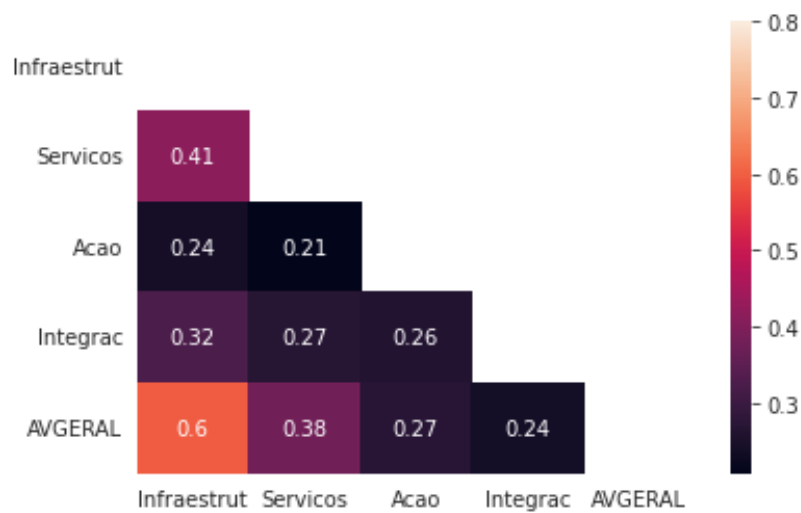


APÊNDICE D - Heatmap voos domésticos

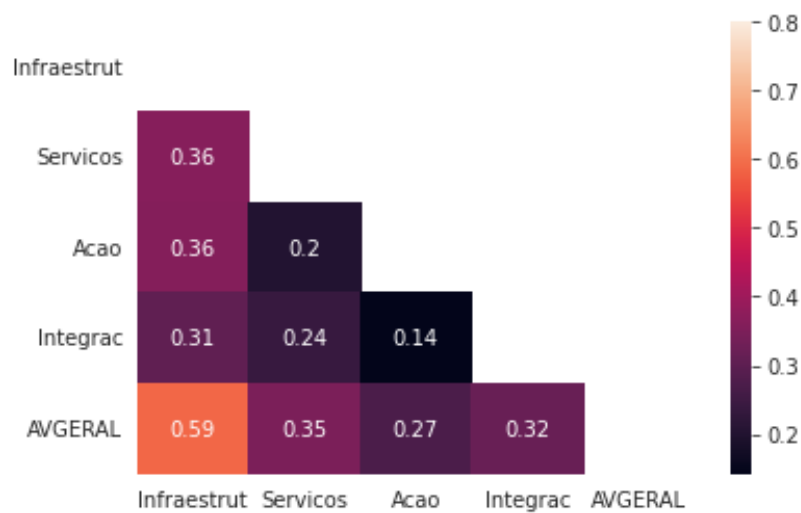
Belém



Brasília



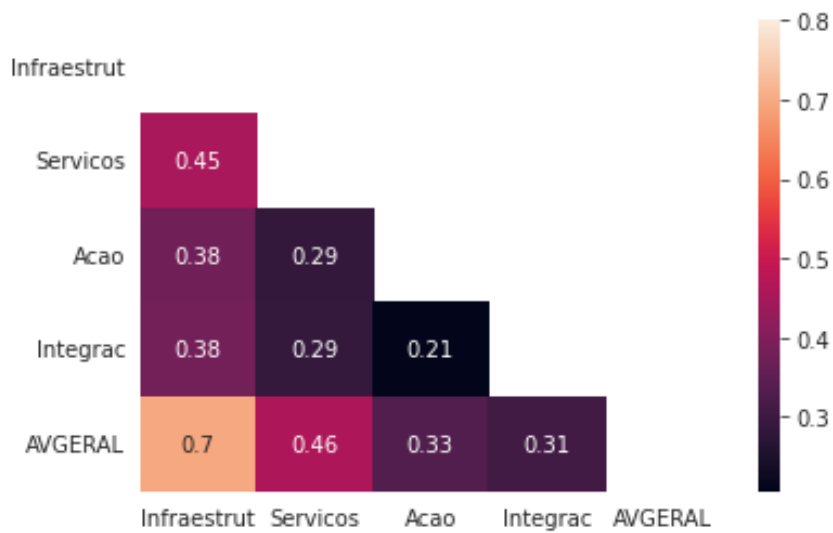
Belo Horizonte (CF)



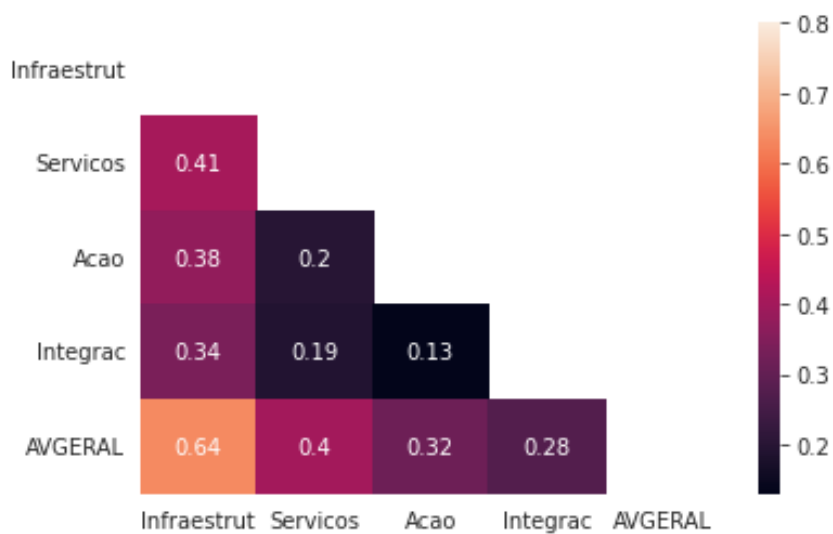
Curitiba



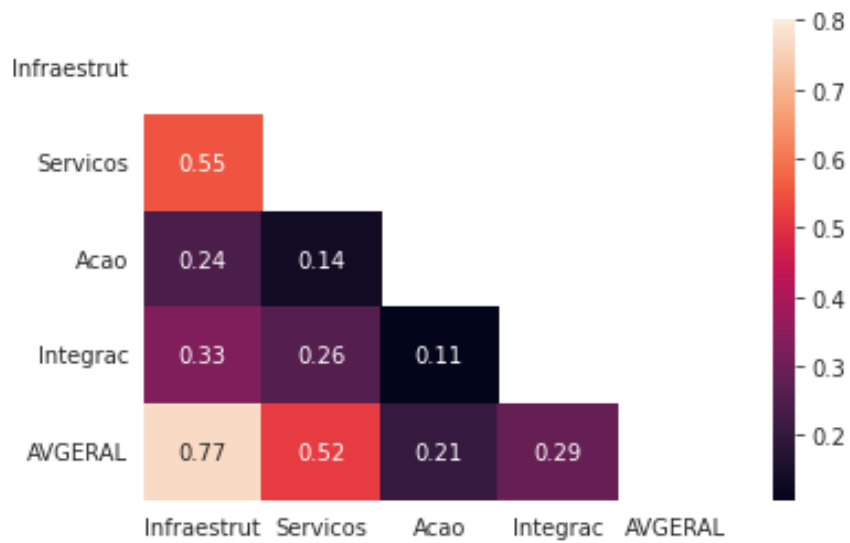
Cuiabá



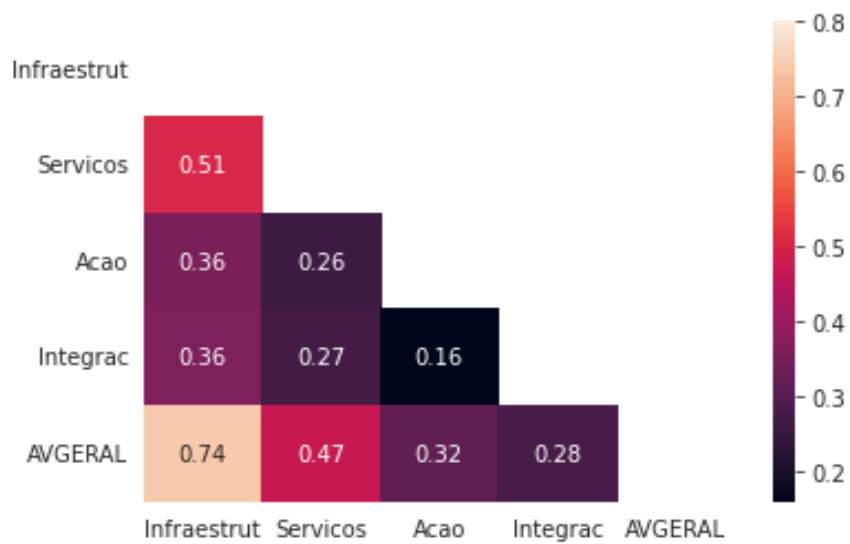
Manaus



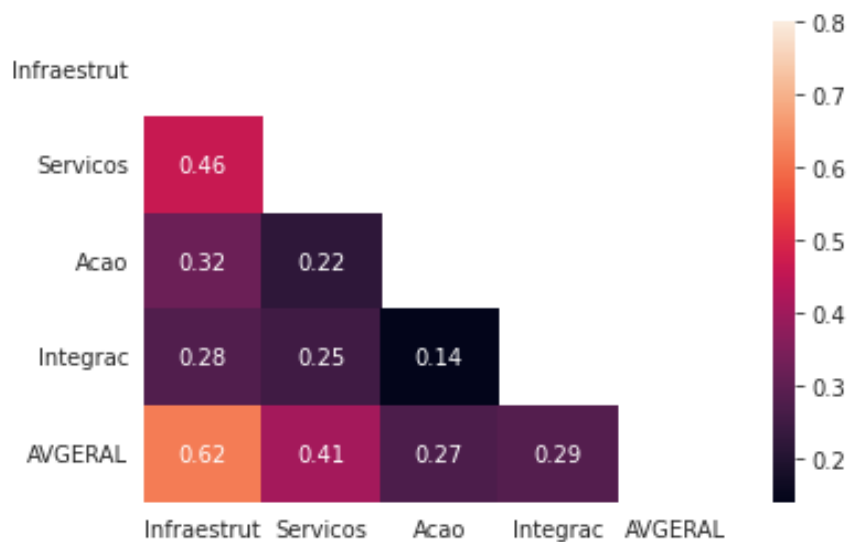
Florianópolis



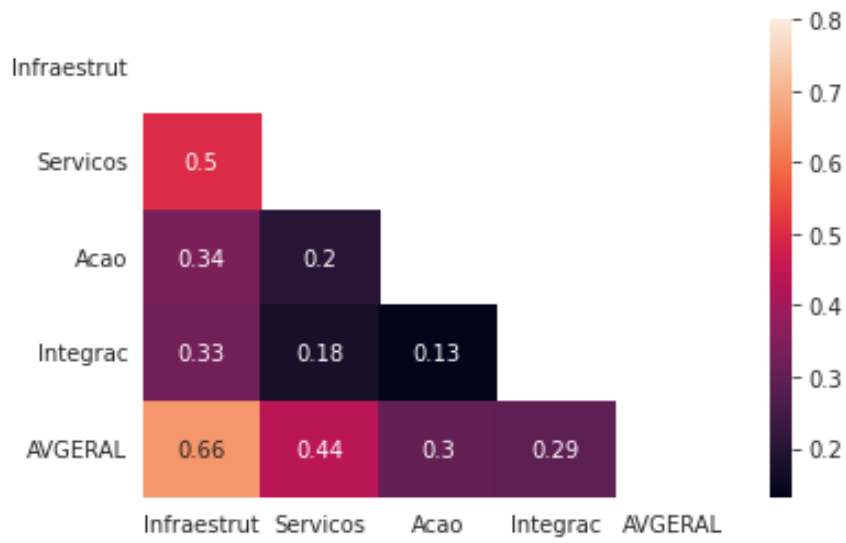
Fortaleza



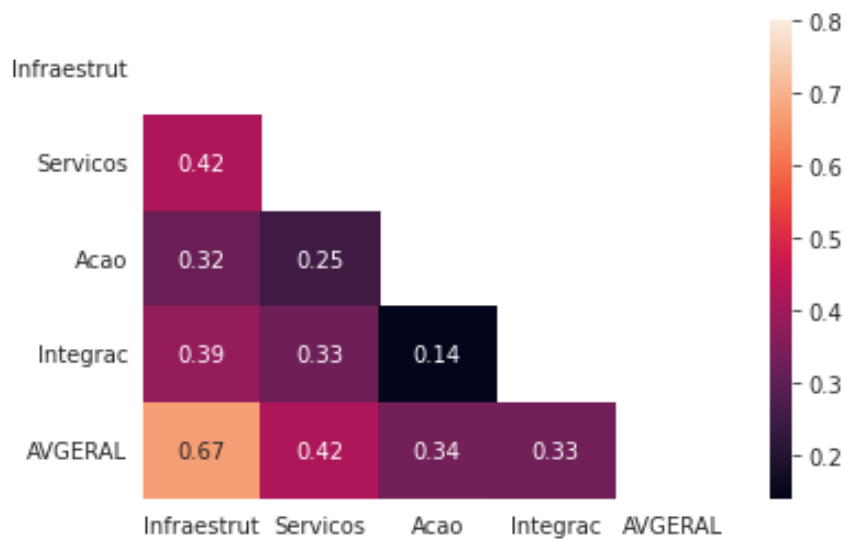
Rio de Janeiro (Galeão)



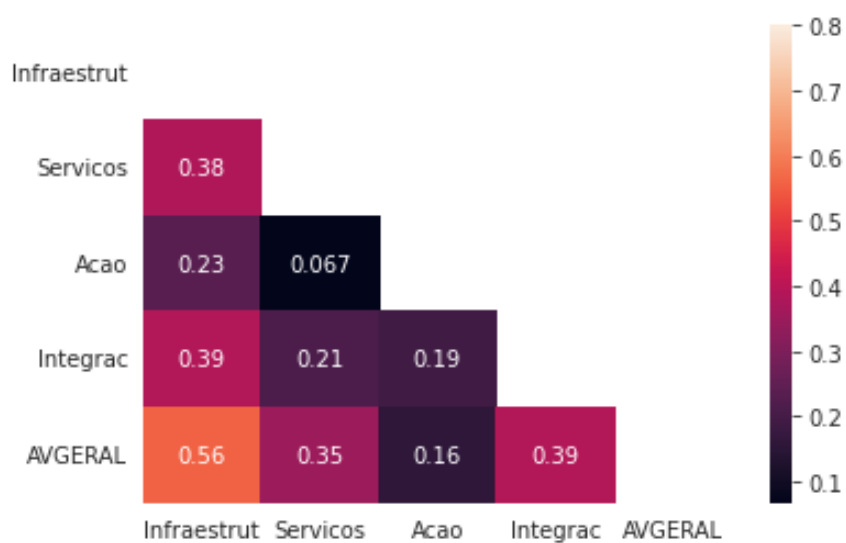
Goiânia



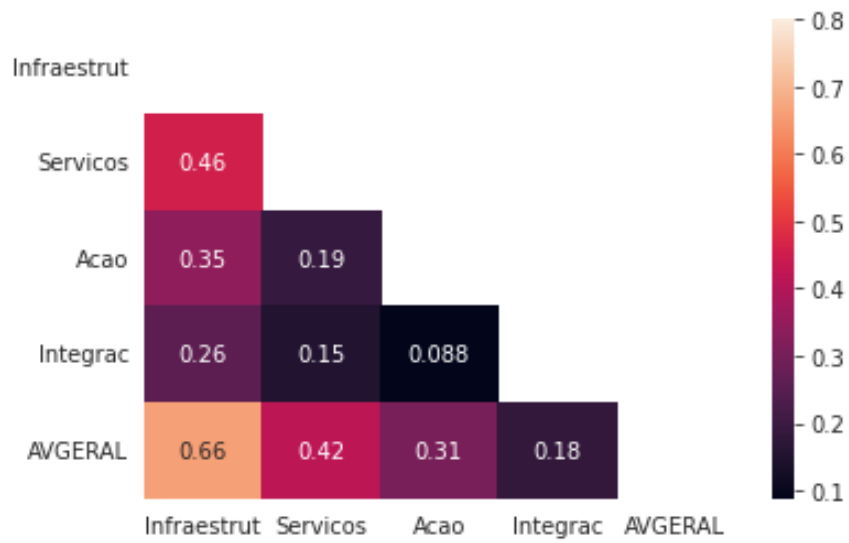
Guarulhos



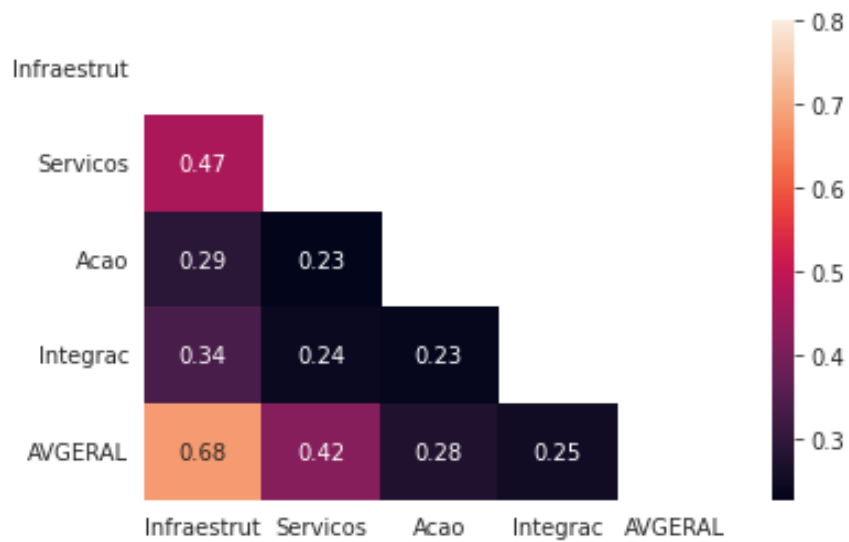
Campinas



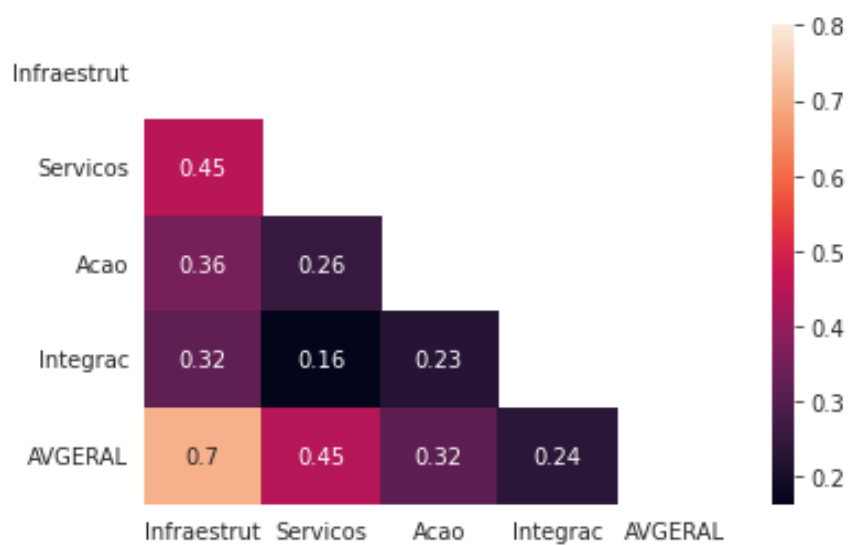
Maceió



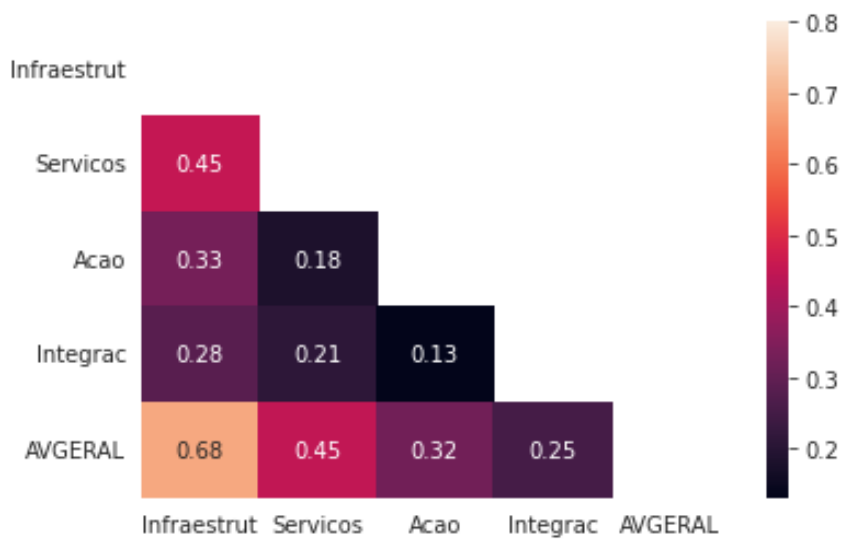
Porto Alegre



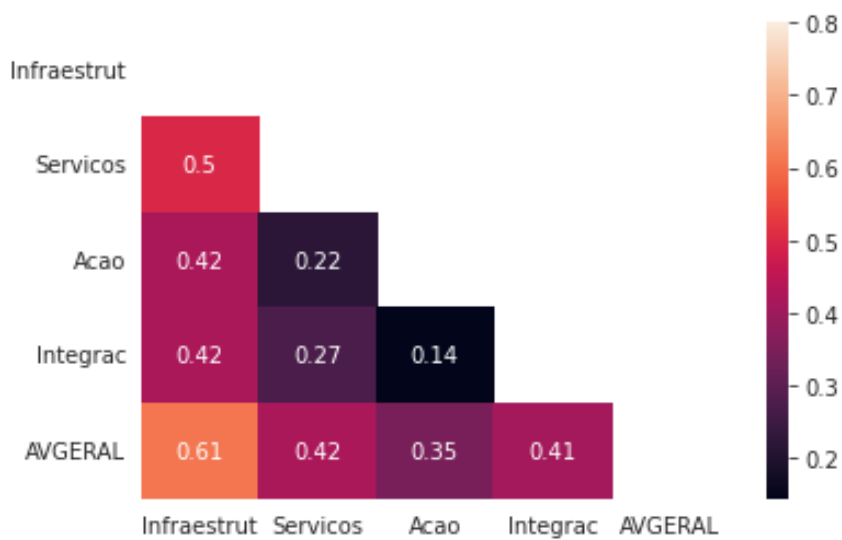
Recife



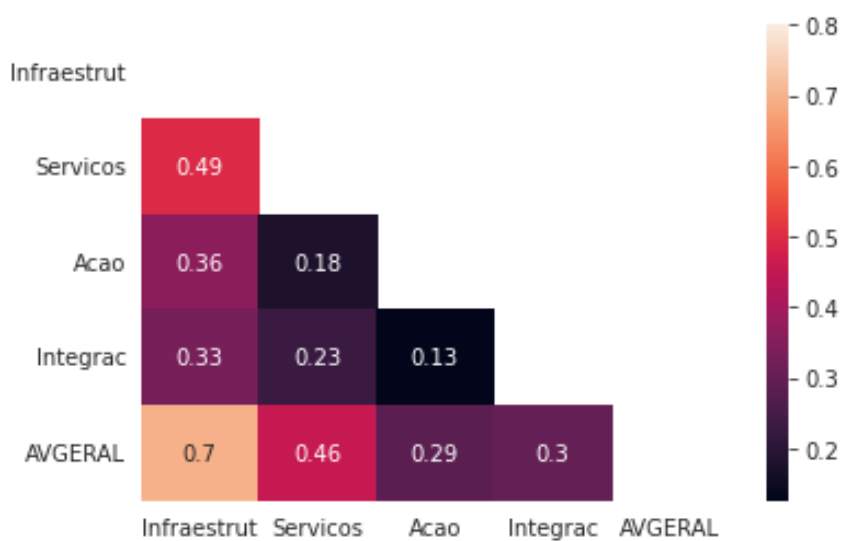
Rio de Janeiro (SD)



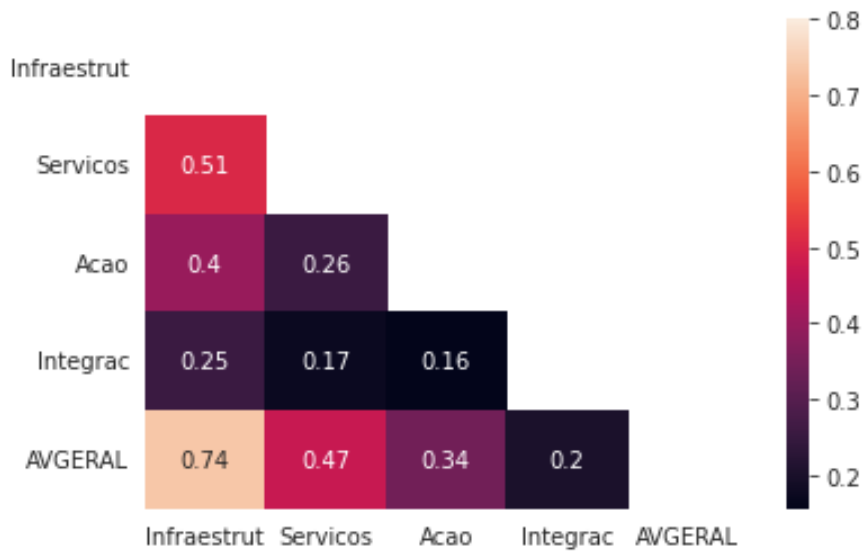
Natal



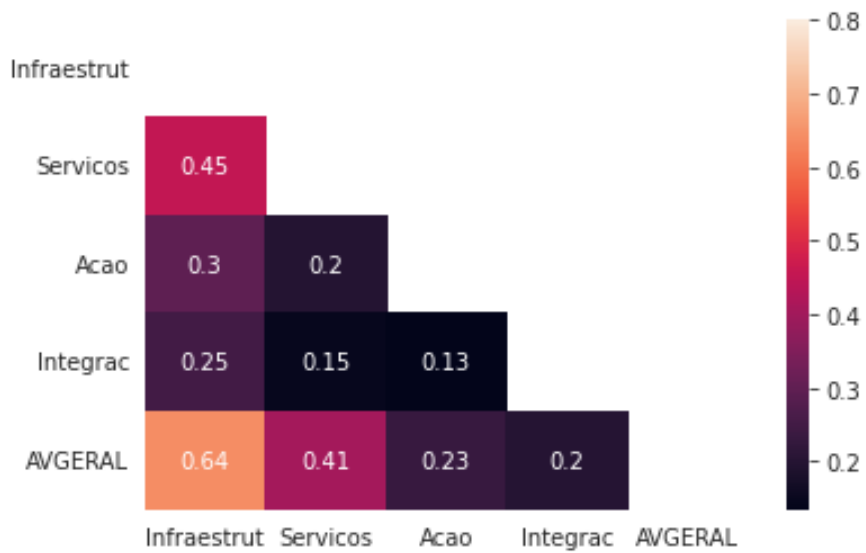
São Paulo



Salvador



Vitória



APÊNDICE E - Links das fontes de dados e artefatos utilizados

Link do vídeo do grupo apresentando o site com os relatórios (dashboards)

qualidade alta:

<https://drive.google.com/file/d/1zDqhEfVdkWx29ykHReaPqsrdaqIAlpQ9L/view?usp=sharing>

qualidade média:

<https://drive.google.com/file/d/1SBqBJdPmOqsJI70Tim-yE65y7MWbtgU-/view?usp=sharing>

Endereço da origem dos dados:

<https://dados.gov.br/dataset/pesquisa-de-satisfacao-do-passageiro-em-aeroportos>

Planilha de dados geral em uso e base dos dashboards:

<https://docs.google.com/spreadsheets/d/1GrcEpcbwlAuOjKC3Xx3nYd65Api3QnH0VpJS1WptxcM/edit?usp=sharing>

Área do Data Studio em uso para elaborar os dashboards:

https://datastudio.google.com/s/tj_UV6tPXMk

Área de edição do site:

https://sites.google.com/d/1bUI9CQvW7feaDexogiwJpGLcVRu2srRj/p/1Bu6NN_5K-xQIIVd3jEdqf-CGnOHzxOtT/edit

Site publicado:

<https://sites.google.com/sou.unifeob.edu.br/pi202203grupo10/in%C3%ADcio>

Planilha dados voos domésticos, para análise estatística no colab:

<https://docs.google.com/spreadsheets/d/1ysoAvh4hPkKLVqfla2eEASsT05hYOyWo/edit?usp=sharing&oid=104410613881675180839&rtpof=true&sd=true>

Colab para estatísticas dos voos domésticos:

<https://colab.research.google.com/drive/1Bxq1f73rNRbHaSsaKsDFkpin6NGrIrPk?usp=sharing>

Planilha dados voos internacionais, para análise estatística no colab:

<https://docs.google.com/spreadsheets/d/1qCtXcRI1Sf7w5hz3xsFVtGyeaZwwnVe2/edit?usp=sharing&oid=104410613881675180839&rtpof=true&sd=true>

Colab para estatísticas dos voos internacionais:

<https://colab.research.google.com/drive/1MViPKwNtUZrshrz0vEcgxaZUjAQvM7G?usp=sharing>

APÊNDICE F - Código utilizado no Colab para os dados de voos internacionais

```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

#from google.colab import drive
#drive.mount('/content/drive/')

#satisfacao = pd.read_csv('/content/20220909internacional.csv', delimiter=';',
encoding='latin-1')
satisfacao = pd.read_excel('/content/20220913internacional.xlsx')
satisfacao.head()

satisfacao.shape

satisfacao.describe()

satisfacao['CIDADE'].value_counts

tot_cidades = satisfacao['CIDADE'].unique()

tot_cidades

tot_cidades.size

colunas_de_notas = ['Infraestrut', 'Serviços', 'Acao', 'AcaoExt', 'Integrac', 'AVGERAL']
notas_satisfacao = satisfacao[colunas_de_notas]

notas_satisfacao.head()

notas_satisfacao.columns = ['Infraestrut', 'Servicos', 'Acao', 'AcaoExt', 'Integrac',
'AVGERAL']
correlacao = notas_satisfacao.corr()
correlacao

sns.heatmap(correlacao)

colunas_de_pareamento = ['CIDADE', 'Infraestrut', 'Serviços', 'Acao', 'AcaoExt',
'Integrac', 'AVGERAL']
pareamento = satisfacao[colunas_de_pareamento]
```

```
pareamento.columns = ['CIDADE', 'Infraestrut', 'Servicos', 'Acao', 'AcaoExt', 'Integrac',
'AVGERAL']
pareamento.shape
```

```
pareamento.head()
```

```
corr = pareamento.corr()
```

```
corr
```

```
#colunas_df1 = ['CIDADE', 'Infraestrut', 'Serviços', 'Acao', 'AcaoExt', 'Integrac']
#colunas_df2 = ['AVGERAL']
#df1 = satisfacao[colunas_df1]
#df2 = satisfacao[colunas_df2]
```

```
#corr2 = df1.corrwith(df2, axis=1)
#corr2
```

```
sns.pairplot(pareamento, hue='CIDADE')
```

```
dependente = ['AVGERAL']
independente = ['Infraestrut', 'Servicos', 'Acao', 'AcaoExt', 'Integrac']
```

```
sns.pairplot(pareamento, x_vars=independente, y_vars=dependente, kind="reg")
```

```
sns.pairplot(pareamento, x_vars=independente, y_vars=dependente, kind="kde")
```

Tentar código para iterar pelas cidades e fazer o pairplot acima, só que por cidades (talvez também o boxplot).

```
def get_casos_cidade(cidade):
    paircidade = pareamento.groupby(['CIDADE']).get_group(cidade)
    #pp = sns.pairplot(paircidade, x_vars=independente, y_vars=dependente, kind="kde")
    return paircidade
```

```
for c in range(tot_cidades.size):
    cidade = tot_cidades[c]
    #print (cidade)
    paircidade = get_casos_cidade(cidade)
    sns.set_style("whitegrid")
    pp = sns.pairplot(paircidade, x_vars=independente, y_vars=dependente, kind="kde")
    pp.fig.suptitle(cidade, y=1.05)
```

Antes dos boxplot, experimentar se é possível/basta utilizar os heatmap por cidade.

```
colunas_numericas = ['Infraestrut', 'Servicos', 'Acao', 'AcaoExt', 'Integrac', 'AVGERAL']

for c in range(tot_cidades.size):
    cidade = tot_cidades[c]
    # chamandao a funcao para pegar os dados de uma cidade de cada vez
    paircidade = get_casos_cidade(cidade)
    notas_paircidade = paircidade[colunas_numericas]
    correlapair = notas_paircidade.corr()

    # imprimindo os resultados
    print ('\n',cidade)
    # formato de máscara copiado do stackoverflow para
    # mostrar só metade de baixo da matriz no heatmap
    mask = np.zeros_like(correlapair)
    mask[np.triu_indices_from(mask)] = True
    # heatmap de cada cidade, só com a metade de baixo das correlações
    hm = sns.heatmap(correlapair, mask = mask, vmax=.8, annot=True, square=False)
    plt.show()
    #print ('\n')
```

Tentar código para iterar pelas cidades e fazer o boxplot, com a finalidade de avaliar se vale a pena testar diferenciação entre as cidades.

APÊNDICE G - Código utilizado no Colab para os dados de voos domésticos

```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

#from google.colab import drive
#drive.mount('/content/drive/')

pd.set_option('display.precision', 3)

satisfacao = pd.read_excel('/content/20220916domestico.xlsx')
satisfacao.head()

satisfacao.shape

satisfacao.describe()

satisfacao['CIDADE'].value_counts

tot_cidades = satisfacao['CIDADE'].unique()

tot_cidades

tot_cidades.size

colunas_de_notas = ['Infraestrut', 'Servicos', 'Acao', 'Integrac', 'AVGERAL']
notas_satisfacao = satisfacao[colunas_de_notas]

notas_satisfacao.head()

correlacao = notas_satisfacao.corr()
correlacao

sns.heatmap(correlacao)

colunas_de_pareamento = ['CIDADE', 'Infraestrut', 'Servicos', 'Acao', 'Integrac',
'AVGERAL']
pareamento = satisfacao[colunas_de_pareamento]
pareamento.shape
```

```
pareamento.head()
```

```
corr = pareamento.corr()
```

```
corr
```

```
dependente = ['AVGERAL']
```

```
independente = ['Infraestrut', 'Servicos', 'Acao', 'Integrac']
```

```
sns.pairplot(pareamento, x_vars=independente, y_vars=dependente, kind="reg")
```

```
sns.pairplot(pareamento, x_vars=independente, y_vars=dependente, kind="kde")
```

Tentar código para iterar pelas cidades e fazer o pairplot acima, só que por cidades (talvez também o boxplot).

```
def get_casos_cidade(cidade):
```

```
    paircidade = pareamento.groupby(['CIDADE']).get_group(cidade)
```

```
    return paircidade
```

```
for c in range(tot_cidades.size):
```

```
    cidade = tot_cidades[c]
```

```
    paircidade = get_casos_cidade(cidade)
```

```
    sns.set_style("whitegrid")
```

```
    pp = sns.pairplot(paircidade, x_vars=independente, y_vars=dependente, kind="kde")
```

```
    pp.fig.suptitle(cidade, y=1.05)
```

Antes dos boxplot, experimentar se é possível/basta utilizar os heatmap por cidade.

```
colunas_numericas = ['Infraestrut', 'Servicos', 'Acao', 'Integrac', 'AVGERAL']
```

```
for c in range(tot_cidades.size):
```

```
    cidade = tot_cidades[c]
```

```
    # chamando a funcao para pegar os dados de uma cidade de cada vez
```

```
    paircidade = get_casos_cidade(cidade)
```

```
    notas_paircidade = paircidade[colunas_numericas]
```

```
    correlapair = notas_paircidade.corr()
```

```
    # imprimindo os resultados
```

```
    print ("\n",cidade)
```

```
    # formato de máscara copiado do stackoverflow para
```

```
    # mostrar só metade de baixo da matriz no heatmap
```

```
    mask = np.zeros_like(correlapair)
```

```
    mask[np.triu_indices_from(mask)] = True
```

```
    # heatmap de cada cidade, só com a metade de baixo das correlações
```

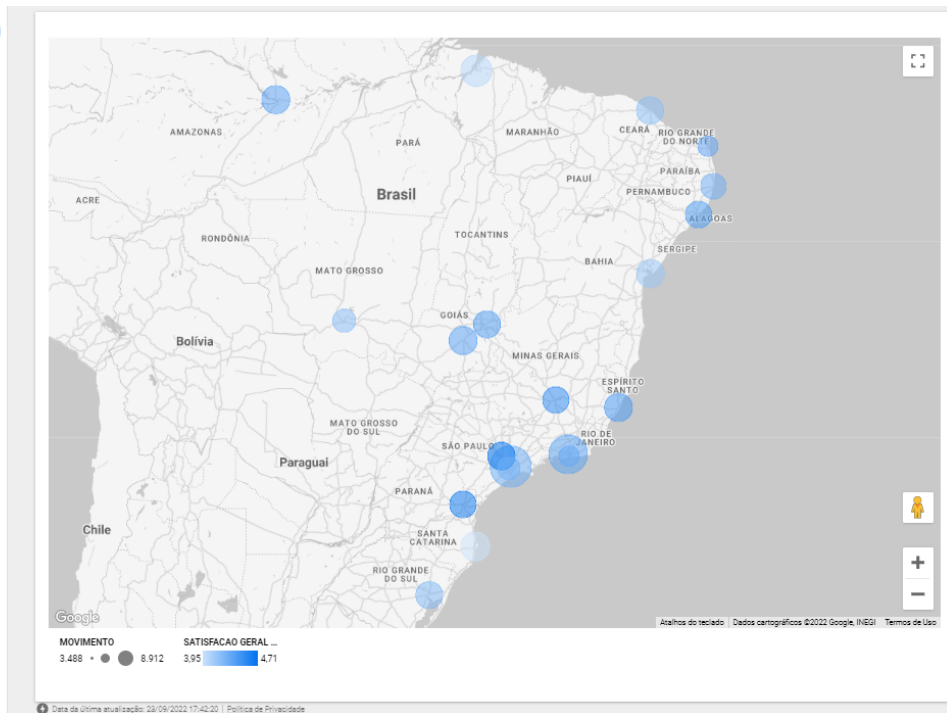
```
hm = sns.heatmap(correlapair, mask = mask, vmax=.8, annot=True, square=False)
plt.show()
#print ('\n')
```

Tentar código para iterar pelas cidades e fazer o boxplot, com a finalidade de avaliar se vale a pena testar diferenciação entre as cidades.

APÊNDICE H - Relatórios criados e apresentados no site do projeto

MOVIMENTO E SATISFAÇÃO ...

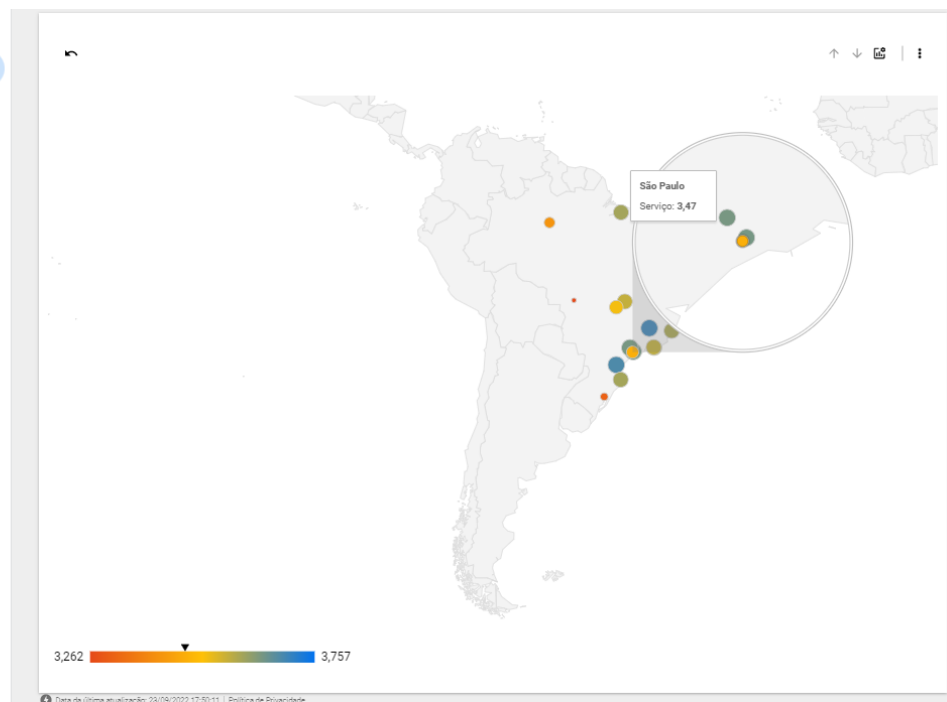
- AEROPORTOS POR CRITÉRIOS
- ESTRATIFICAÇÃO SOCIAL
- ESTRATIFICAÇÃO FINANCEIRA
- DETALHA ÁREAS INVESTIR
- COMPARA ÁREAS INVESTIR
- Página sem título

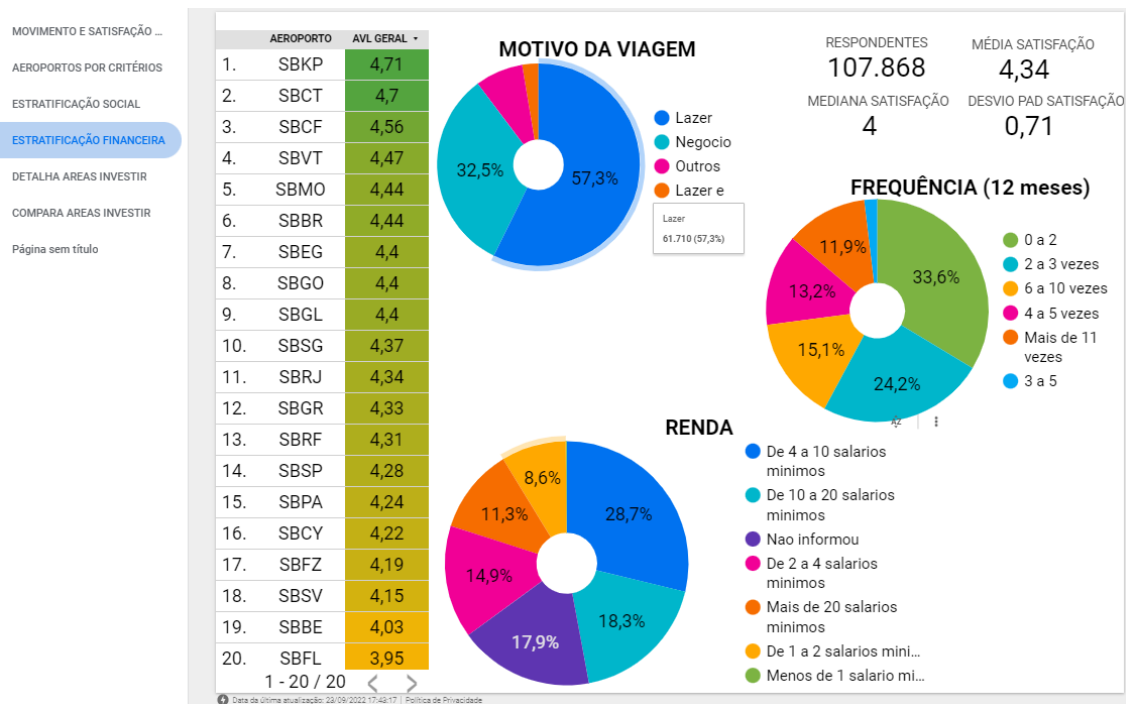
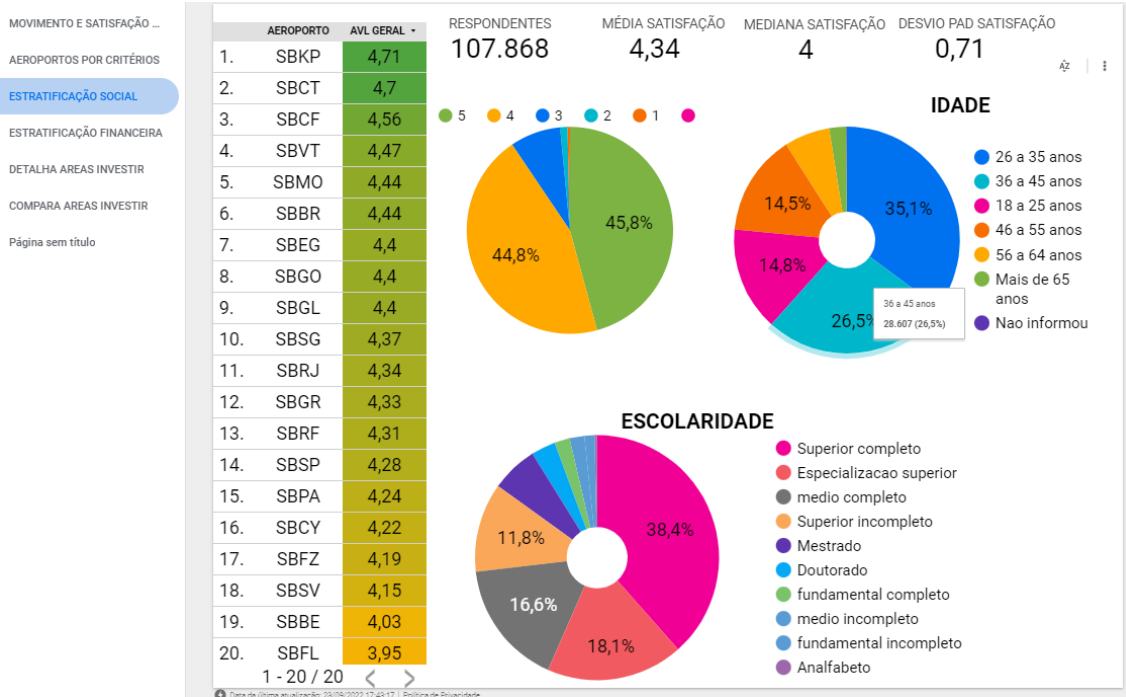


MOVIMENTO E SATISFAÇÃO ...

AEROPORTOS POR CRITÉRIOS

- ESTRATIFICAÇÃO SOCIAL
- ESTRATIFICAÇÃO FINANCEIRA
- DETALHA ÁREAS INVESTIR
- COMPARA ÁREAS INVESTIR
- Página sem título





MOVIMENTO E SATISFAÇÃO ...

AEROPORTOS POR CRITÉRIOS

ESTRATIFICAÇÃO SOCIAL

ESTRATIFICAÇÃO FINANCEIRA

DETALHA AREAS INVESTIR

COMPARA AREAS INVESTIR

Página sem título



MOVIMENTO E SATISFAÇÃO ...

AEROPORTOS POR CRITÉRIOS

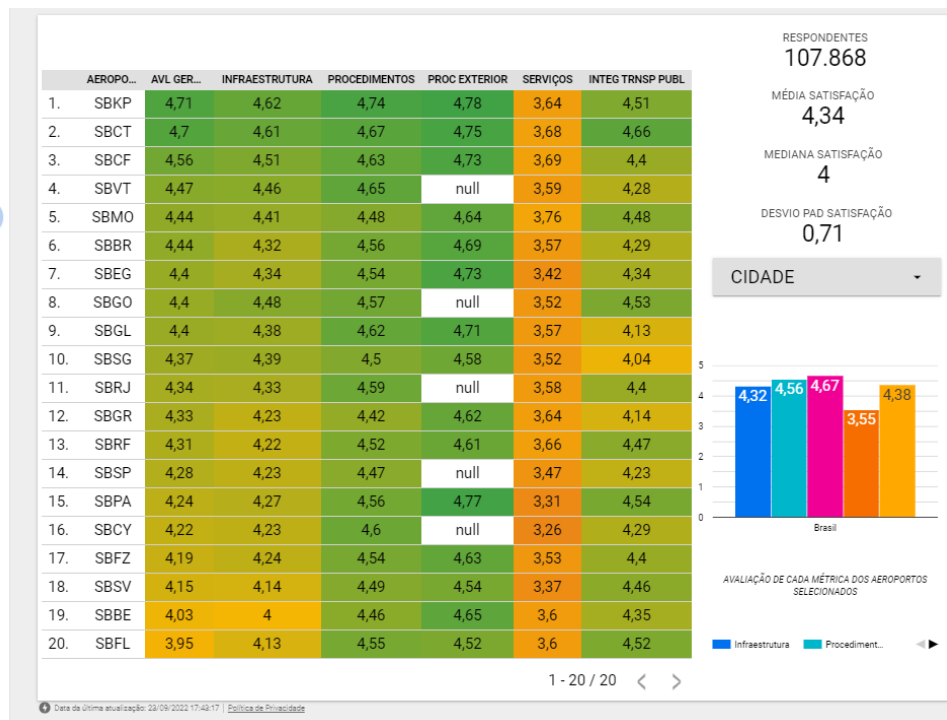
ESTRATIFICAÇÃO SOCIAL

ESTRATIFICAÇÃO FINANCEIRA

DETALHA AREAS INVESTIR

COMPARA AREAS INVESTIR

Página sem título



ANEXO A - Artigo de origem da ideação do tema

24/08/2022 16:33

Leilão de aeroportos liderado por Congonhas gera R\$ 2,72 bilhões em contribuição inicial — Português (Brasil)

O Bloco SP-MS-PA-MG, liderado por Congonhas (SP), e composto ainda pelos aeroportos Campo Grande, Corumbá e Ponta Porã, no Mato Grosso do Sul (MS); Santarém, Marabá, Parauapebas e Altamira, no Pará (PA); Uberlândia, Uberaba e Montes Claros, em Minas Gerais (MG), foi arrematado pela Aena Desarrollo Internacional SME SA por R\$ 2,45 bilhões, com ágio de 231,02% em relação ao lance mínimo inicial de R\$ 740,1 milhões.

Integrado pelos aeroportos de Campo de Marte, em São Paulo (SP) e Jacarepaguá, no Rio de Janeiro (RJ), o Bloco Aviação Geral teve como vencedor a XP Infra IV FIP EM INFRAESTRUTURA, com ágio de 0,01% em relação ao lance mínimo inicial de R\$ 141,3 milhões. O bloco foi arrematado por R\$ 141,4 milhões.

Já o Bloco Norte II, formado pelos aeroportos de Belém (PA) e Macapá (AP), foi arrematado pelas empresas Dix e Socicam, integrantes do Consórcio Novo Norte. O grupo pagou R\$ 125 milhões pelos dois aeroportos do bloco, com ágio de 119,78% em relação ao lance mínimo inicial de R\$ 56,9 milhões.

Tabela-resumo da licitação da 7ª rodada de aeroportos

| | Bloco SP-MS-PA-MG | Bloco Aviação Geral | Bloco Norte II |
|---|--|--|------------------------------------|
| | Consórcio Aena Desarrollo Internacional | Consórcio XP Infra IV FIP em Infraestrutura | Consórcio Novo Norte Aeroportos |
| Nº de propostas | 1 | 1 | 2 |
| Lance mínimo | R\$ 740.132.333,48 | R\$ 141.388.607,98 | R\$ 56.875.878,09 |
| Lance vencedor | R\$ 2.450.000.000,00 | R\$ 141.400.000,00 | R\$ 125.000.000,00 |
| Valor do ágio | R\$ 1.709.867.666,52 | R\$ 11.392,02 | R\$ 68.124.121,91 |
| Ágio sobre lance mínimo | 231,02% | 0,01% | 119,78% |
| Valor total a ser pago na assinatura dos contratos | 2.716.400.000,00 bilhões | | |
| Ágio médio do leilão | 116,94% | | |

O leilão da 7ª rodada de concessões de aeroportos foi realizado na B3, em São Paulo, e contou com a concorrência de 4 proponentes habilitados. O certame teve início às 14h e foi concluído por volta de 15h, após 30 minutos de disputa de lances em viva voz. Confira, a seguir, a lista com todos os proponentes que participaram do leilão da 7ª rodada de aeroportos.

CONTEÚDO 1 PÁGINA INICIAL 2 NAVEGAÇÃO 3 BUSCA 4 MAPA DO SITE 5

<https://www.gov.br/anac/pt-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>

2/6

24/08/2022 16:33

Leilão de aeroportos liderado por Congonhas gera R\$ 2,72 bilhões em contribuição inicial — Português (Brasil)

Proponentes vencedores da 7ª rodada de concessão de aeroportos

| Bloco de aeroportos | Proponente | Tipo | Integrantes do consórcio | Assistente Técnico |
|---------------------|--------------------------------------|-----------------------|--------------------------|--------------------|
| Bloco SP-MS-PA-MG | AENA Desarrollo Internacional SME SA | Proponente individual | | |
| Bloco Aviação Geral | XP Infra IV FIP EM INFRAESTRUTURA | Proponente individual | | Egis |
| Bloco Norte II | Consórcio Novo Norte Aeroportos | Consórcio | Dix e Socicam | Socicam |

Demais proponentes participantes

| Bloco de aeroportos | Proponente | Tipo | Integrantes do consórcio | Assistente Técnico |
|---------------------|----------------|-----------------------|--------------------------|--------------------|
| Bloco Norte II | Vinci Airports | Proponente individual | | |

Regras da 7ª rodada

A 7ª rodada de concessão de aeroportos propõe regulação flexível, compatível e proporcional ao porte de cada aeroporto em relação a tarifas, investimentos e qualidade dos serviços, a exemplo do que já ocorreu na 5ª e na 6ª rodadas. A exigência quanto ao nível de serviço será proporcional ao porte do aeroporto, sempre visando o melhor atendimento ao usuário.

Habilitação e homologação

A etapa seguinte do leilão, no dia 25 de agosto, será o recebimento dos documentos de habilitação dos proponentes vencedores de cada bloco. A assinatura dos contratos de concessão deverá ocorrer após a homologação do resultado pela Diretoria da ANAC, em data ainda a ser definida. Confira os próximos eventos do cronograma da 7ª rodada:

Passos da 7ª rodada de concessão

| Data | Evento |
|----------------|------------------|
| CONTEÚDO 1 | PÁGINA INICIAL 2 |
| NAVEGAÇÃO 3 | BUSCA 4 |
| MAPA DO SITE 5 | |

<https://www.gov.br/anac/pt-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>

3/8

| | |
|--------------------------|--|
| 24/08/2022 16:33 | Leilão de aeroportos liderado por Congonhas gera R\$ 2,72 bilhões em contribuição inicial — Português (Brasil) |
| 25/08/2022: | Recebimento dos documentos de habilitação dos proponentes vencedores |
| 23/09/2022: | Publicação da ata de julgamento relativa à análise dos documentos de habilitação da proponente classificada em 1º lugar |
| 26/09/2022 a 30/09/2022: | Prazo para interposição dos recursos de que trata o item 5.29 do Edital |
| Em aberto: | Homologação do resultado e adjudicação do objeto pela Diretoria da ANAC |
| Em aberto: | Prazo final para comprovação de atendimento, pela proponente vencedora, das obrigações previstas na Seção I do Capítulo VI do Edital |
| Em aberto: | Convocação para celebração do contrato de concessão do respectivo bloco de aeroportos |

Contribuição Variável

Além da contribuição inicial a ser paga na assinatura dos contratos, as novas concessionárias deverão pagar também outorga variável sobre a receita bruta, estabelecida em percentuais crescentes calculados do 5º ao 9º ano do contrato, tornando-se constantes a partir de então até o final da concessão (confira abaixo informações de cada bloco). Esse mecanismo busca adequar os contratos às oscilações de demanda e receita ao longo da concessão. A outorga variável estimada para o Bloco SP-MS-PA-MG, liderado pelo Aeroporto de Congonhas, começará em R\$ 33,6 milhões a partir do 5º ano da concessão e chegará a R\$ R\$ 234 milhões no último ano do contrato.

Os valores projetados para os contratos contemplam uma receita estimada para toda a concessão de R\$ 15,2 bilhões (para os 15 aeroportos), sendo R\$ 11,6 bilhões para o Bloco SP-MS-PA-MG; R\$ 1,7 bilhão para o Bloco Aviação Geral; e R\$ 1,9 bilhão para o Bloco Norte II.

BLOCO SP-MS-PA-MG

Contribuição Inicial Mínima (paga no leilão): R\$ 740.132.333,48 + ágio

Contribuição Variável (parcelas anuais conforme percentuais da receita)

| | |
|--------|-------|
| 5º ano | 3,23% |
| 6º ano | 6,46% |

CONTEÚDO 1 PÁGINA INICIAL 2 NAVEGAÇÃO 3 BUSCA 4 MAPA DO SITE 5

<https://www.gov.br/anac/p-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>

4/6

24/08/2022 16:33

Leilão de aeroportos liderado por Congonhas gera R\$ 2,72 bilhões em contribuição inicial — Português (Brasil)

| | |
|--------|--------|
| 8º ano | 12,92% |
|--------|--------|

| | |
|--------|--------|
| 9º ano | 16,15% |
|--------|--------|

Valor do contrato (receita estimada ao longo da concessão): R\$ 11.608.831.026,23

Investimento previsto em EVTEA: R\$ 5.808.778.318,08

BLOCO AVIAÇÃO GERAL

Contribuição Inicial Mínima (paga no leilão): R\$ 141.388.607,98 + ágio

Contribuição Variável (parcelas anuais conforme percentuais da receita)

| | |
|--------|-------|
| 5º ano | 3,05% |
|--------|-------|

| | |
|--------|-------|
| 6º ano | 6,10% |
|--------|-------|

| | |
|--------|-------|
| 7º ano | 9,15% |
|--------|-------|

| | |
|--------|--------|
| 8º ano | 12,20% |
|--------|--------|

| | |
|--------|--------|
| 9º ano | 15,25% |
|--------|--------|

Valor do contrato (receita estimada ao longo da concessão): R\$ 1.710.204.752,13

Investimento previsto em EVTEA: R\$ 552.013.358,66

BLOCO NORTE II

Contribuição Inicial Mínima (paga no leilão): R\$ 56.875.878,09 + ágio

[CONTEÚDO](#) **1** [PÁGINA INICIAL](#) **2** [NAVEGAÇÃO](#) **3** [BUSCA](#) **4** [MAPA DO SITE](#) **5**

<https://www.gov.br/anac/pt-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>

5/6

24/08/2022 16:33

Leilão de aeroportos liderado por Congonhas gera R\$ 2,72 bilhões em contribuição inicial — Português (Brasil)

| | |
|--|-------|
| 5º ano | 1,42% |
| 6º ano | 2,84% |
| 7º ano | 4,25% |
| 8º ano | 5,67% |
| 9º ano | 7,09% |
| Valor do contrato (receita estimada ao longo da concessão): R\$ 1.931.983.096,09 | |
| Investimento previsto em EVTEA: R\$ 874.656.126,52 | |

Investimentos e melhorias

Os novos concessionários dos 15 aeroportos leiloados na quinta-feira deverão fazer investimentos da ordem de R\$ 7,2 bilhões durante os 30 anos da concessão. De acordo com os Estudos de Viabilidade Técnica, Econômica e Ambiental (EVTEAs), os investimentos estimados por bloco de aeroportos serão de R\$ 5,8 bilhões para o Bloco SP-MS-PA-MG; R\$ 552 milhões para o Bloco Aviação Geral; e R\$ 875 milhões para o Bloco Norte II. Só no Aeroporto de Congonhas, a previsão é que sejam investidos um total de R\$ 3,35 bilhões nos 30 anos da concessão, sendo 75,4% desse montante (R\$ 2,53 bilhões) aplicados na ampliação da infraestrutura nos primeiros cinco anos da vigência do contrato.

Nos 36 meses contados a partir da data de eficácia do contrato (Fase I-B) para os aeroportos do Bloco Aviação Geral e Bloco Norte II e nos 60 meses para os aeroportos do Bloco SP-MS-PA-MG, os 15 aeroportos concedidos da 7ª rodada deverão realizar os investimentos necessários na infraestrutura atual para a prestação do serviço adequado aos usuários.

Além de investimentos específicos definidos conforme as características de cada aeroporto ([confira mais detalhes na página especial com informações dos blocos de aeroportos da 7ª rodada](#), clicando no link para acessar), as novas concessões terão que adequar sua capacidade de processamento de passageiros, bagagens e estacionamento de veículos; observar especificações mínimas da infraestrutura aeroportuária e indicadores de qualidade de serviço.

Mais informações sobre o leilão da 7ª rodada podem ser [obtidas na página de acompanhamento do leilão da sétima rodada](#) (clique no link para acessar).

Assessoria de Comunicação Social da ANAC

Compartilhe: [f](#) [t](#) [s](#)

<https://www.gov.br/anac/p-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial>

6/6

[Leilão de aeroportos liderado por Congonhas gera R\\$ 2,72 bilhões em contribuição inicial — Português \(Brasil\) \(www.gov.br\)](https://www.gov.br/anac/p-br/noticias/2022/leilao-de-aeroportos-liderado-por-congonhas-gera-rs-2-72-bilhoes-em-contribuicao-inicial)