



Universidade de Brasília

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas  
Públicas - FACE

Departamento de Administração

PEDRO SOARES BONISOLO

**Análise exploratória de dados de histórico de voos da ANAC: Uma  
aplicação com a malha aérea da empresa GOL**

Brasília – DF  
2023

PEDRO SOARES BONISOLO

**Análise exploratória de dados de histórico de voos da ANAC:  
Uma aplicação com a malha aérea da empresa GOL**

Monografia apresentada ao  
Departamento de Administração  
como requisito parcial à obtenção  
do título de Bacharel em  
Administração.

Professor Orientador: Professor  
Dr., Victor Rafael Rezende  
Celestino

Brasília– DF  
2023

PEDRO SOARES BONISOLO

**Análise exploratória de dados de histórico de voos da ANAC:  
Uma aplicação com a malha aérea da empresa GOL**

A Comissão Examinadora, abaixo identificada, aprova o Trabalho de Conclusão do  
Curso de Administração da Universidade de Brasília do (a) aluno (a)

**Pedro Soares Bonisolo**

Professor Dr., Victor Rafael Rezende Celestino  
Professor-Orientador

Profa. Dra., Silvia Araújo dos Reis  
Professora-Examinadora

Profa. Dra., Vanessa Cabral Gomes  
Professora-Examinadora

Brasília, 14 de fevereiro de 2023

Aos companheiros da aviação civil que estimulam o seu desenvolvimento com segurança.

À minha família, que é tudo em minha vida.

Agradeço, primeiramente, à Deus que me deu determinação e sabedoria para concluir este trabalho.

Agradeço ao meu orientador, Professor Dr. Victor Rafael Rezende Celestino que me mostrou mais uma belíssima face da aviação e não poupou esforços para compartilhar comigo todo o conhecimento por ele adquirido ao longo de sua invejável carreira.

Agradeço aos meus pais que me incentivaram em todos os momentos da minha vida.

Agradeço aos meus companheiros de turma tanto na universidade quanto no aeroclube que participaram e me apoiaram nessa pesquisa.

Agradeço aos meus amigos que me apoiaram em todos os momentos da minha jornada.

Agradeço à minha companheira, Lara, que me apoiou incondicionalmente, e que mesmo de longe, contribuiu indiretamente para que esse trabalho se realizasse.

Enfim, agradeço a todos que fizeram parte dessa etapa decisiva em minha vida.

“O homem há de voar.”

(Alberto Santos Dumont)

## RESUMO

O objetivo desta pesquisa foi realizar uma análise exploratória dos dados abertos relativos ao histórico de voos disponibilizados pela Agência Nacional de Aviação Civil, a ANAC; e, posteriormente, apresentar os resultados examinados sob forma de uma releitura da visualização de dados utilizando a linguagem de programação Python. Atualmente, dados relativos a histórico de voos são fornecidos pela agência em formato de dados abertos, porém sem grandes representações gráficas e visuais, tornando a sua leitura pouco estimulante. Além disso, considera-se também que a linguagem de programação Python permite um aperfeiçoamento da visualização numérica e gráfica, proporcionando diferentes configurações e uma narrativa com dados mais voltada aos interesses do usuário, facilitando sua apresentação, por conseguinte sua análise. Logo, o escopo repousa numa reapresentação dos dados de uma forma mais assertiva. Por fim, a metodologia utilizada foi a de Design Science, e esta pesquisa pretende obedecer às 12 etapas desta aplicação metodológica. Os resultados incluem a apresentação de dados em um formato de mapa interativo.

**Palavras-chave:** Modelagem, malha aérea, aviação civil, análise de dados, python.

## ABSTRACT

The objective of this research is to carry out an exploratory analysis of the open data related to the history of flights made available by the Brazilian National Civil Aviation Agency, also known as ANAC; and, later, to present the examined results in the form of a reinterpretation of the data visualization using the Python programming language. Currently, data related to flight history are provided by the aviation agency in open data format, but without large graphic and visual representations, making reading less stimulating. In addition, it is also considered that the Python programming language allows an improvement of the numeric and the graphic visualization, providing different configurations and a narrative with data more focused on the interests of the user, facilitating its presentation, therefore, its analysis. Therefore, the scope rests on a more assertive representation of the data. Finally, the methodology used is Design Science, and this research intends to follow the 12 steps of this methodological application. Results include the presentation of data in an interactive map format.

**Keywords:** Modeling, air transport network, civil aviation, data analysis, python.



## LISTA DE FIGURAS

Figura 1 - Malha aérea doméstica da Varig Rio-Sul.....	4
Figura 2 - Dados e Estatísticas da Agência Nacional de Aviação Civil.....	5
Figura 3 - Index dos dados abertos da ANAC.....	6
Figura 4 - Rota do voo RG-254 da Varig.....	10
Figura 5 - Exemplo de rota circular.....	11
Figura 6 - Hubs do grupo AirFrance-KLM.....	12
Figura 7 - Pirâmide das etapas de transformação do dado.....	12
Figura 8 - Etapas da estatística descritiva.....	14
Figura 9 - Tipos de Artefatos.....	18
Figura 10 - Representação gráfica do número de decolagens por região.....	28
Figura 11 - Representação gráfica do percentual de decolagens por região.....	29
Figura 12 - Representação gráfica do número de decolagens por HUBs.....	30
Figura 13 - Malha aérea do HUB de Congonhas (SBSP).....	30
Figura 14 - Malha aérea do HUB de Guarulhos (SBGR).....	31
Figura 15 - Malha aérea do HUB de Brasília (SBBR).....	31
Figura 16 - Malha aérea do HUB de Galeão.....	32
Figura 17 - Malha aérea do HUB de Salvador (SBSV).....	32
Figura 18 - Malha aérea do HUB de Fortaleza (SBCZ).....	33
Figura 19 - Malha aérea do HUB de Belo Horizonte (SBCF).....	33

## LISTA DE TABELAS

Tabela 1 - Abordagens estatísticas e estratégias utilizadas na AED.....	13
Tabela 2 - Definição dos dados .....	24
Tabela 3 - Dados de voos doméstico da GOL para Outubro.....	25
Tabela 4 - Dados de voos doméstico da GOL para Novembro .....	25
Tabela 5 - Dados de voos doméstico da GOL para Dezembro .....	26
Tabela 6 - Destinos e HUBs da GOL LINHAS AÉREAS .....	27
Tabela 7 - Número de decolagens por região .....	28
Tabela 8 - Número de decolagens por HUBs.....	29

## LISTA DE ABREVIATURAS E SIGLAS

ANAC – Agência Nacional de Aviação Civil  
RAB – Registro Aeronáutico Brasileiro  
OACI - Organização Internacional da Aviação Civil  
ICAO - Organização Internacional da Aviação Civil  
IATA - Associação Internacional de Transportes Aéreos  
VARIG - Companhia Aérea Viação Aérea Rio-Grandense  
GOL – Companhia aérea GOL LINHAS AÉREAS  
GLO - Código ICAO da GOL LINHAS AÉREAS  
KLM - Koninklijke Luchtvaart Maatschappij, companhia aérea real holandesa  
AF - Código IATA da companhia aérea Air France  
VRG - Código ICAO da Viação Aérea Rio-Grandense  
RG - Código IATA da Viação Aérea Rio-Grandense  
HUB - aeroporto que serve como centro de distribuição de passageiros através de uma rede extensa de voos  
SBGL - Código ICAO do Aeroporto Internacional Antonio Carlos Jobim (Galeão) no Rio de Janeiro  
AED - Análise exploratória de dados  
DS - Design Science  
DSR - Design Science Research  
SAF - Superintendência de Administração e Finanças  
SPI - Superintendência de Planejamento Institucional  
STI - Superintendência de Tecnologia da Informação  
SFI - Superintendência de Ação Fiscal  
SPL - Superintendência da Superintendência de Pessoal da Aviação Civil  
SRA - Superintendência de Regulação Econômica de Aeroportos  
SAR - Superintendência de Aeronavegabilidade  
SIA - Superintendência de Infraestrutura Aeroportuária  
SGP - Superintendência de Gestão de Pessoas  
SAS - Superintendência de Acompanhamento de Serviços Aéreos  
SPO - Superintendência de Padrões Operacionais

## SUMÁRIO

1	INTRODUÇÃO .....	1
1.1	Contextualização .....	3
1.2	Formulação do problema.....	5
1.3	Objetivo Geral .....	7
1.4	Objetivos Específicos .....	7
1.5	Justificativa.....	7
2	REVISÃO TEÓRICA .....	9
2.1	Malha aérea doméstica .....	9
2.1.1	Rota em linha (point-to-point).....	10
2.1.2	Rotas circulares .....	11
2.1.3	Rota com centro de distribuição (Hub & Spoke) .....	11
2.2	Visualização de dados .....	12
3	METODOLOGIA.....	15
3.1	Tipologia e descrição geral dos métodos de pesquisa .....	15
3.2	Caracterização da organização, setor ou área, indivíduos objeto do estudo .....	15
3.3	População e amostra ou Participantes da pesquisa.....	17
3.4	Caracterização e descrição dos instrumentos de pesquisa.....	18
3.5	Procedimentos de coleta e de análise de dados .....	22
3.6	Análise de dados.....	22
4	RESULTADOS E DISCUSSÕES .....	23
4.2	Execução .....	23
4.3	Análise geográfica.....	26
5	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS .....	35
	REFERÊNCIA .....	36
	APÊNDICES.....	39
	Apêndice A – Notebook utilizado para o manuseio dos dados.....	39
	ANEXOS.....	41
	ANEXO I – Lei de Criação da ANAC nº 11.182, de 27 de setembro de 2005.....	41
	ANEXO II – Decreto de Instalação da ANAC nº 5.731, de 20 de março de 2006 .....	41
	ANEXO III – Regimento Interno: Resolução nº 381, de 14 de junho de 2016.....	41
	ANEXO IV – Alteração do RI: Resolução nº 581, de 21 de agosto de 2020 .....	41
	ANEXO V – Instrução Normativa nº 127, de 5 de outubro de 2018 .....	41

## 1 INTRODUÇÃO

A aviação civil revolucionou o jeito de se deslocar de um ponto a outro. Seja transportando passageiros ou cargas diversas, hoje é extremamente difícil se imaginar um mundo sem aviões, aeroportos e rotas aéreas. Segundo Lopes (2005), “O setor de transporte aéreo possui segmentos característicos, que são as empresas aéreas, os órgãos reguladores e os aeroportos”. Estudar a aviação como um todo é extremamente importante para a sociedade, pois permite observar, compreender e projetar o que é visto como um das linhas dorsais do transporte mundial.

Compreender os voos, suas frequências, suas particularidades e tendências é uma excelente forma de se adentrar neste tão vasto e importante estudo. Quando se fala em malha aérea, é possível logo imaginar em um emaranhado de linhas sobrepostas em um mapa que liga dois pontos cuja linha representa o trajeto e os pontos, as cidades de origem e destino. Como forma de visualizar uma malha aérea com mais facilidade, pode-se imaginar um esquema de grafos que é obtido a partir da associação de cada vértice a um ponto em uma determinada área (um mapa) na qual os vértices podem ser ligados por meio de retas representando assim uma associação de vértices (Boaventura Netto, 1996).

As rotas que se conhecem nos dias de hoje, porém, são frutos de uma série de variáveis que, juntas, moldaram a malha aérea doméstica. Guerras, crises de petróleo, planos econômicos são algumas das variáveis mais tangíveis em que se pode imaginar e que tiveram um impacto profundo na aviação civil mundial e também, na aviação nacional.

Entender a evolução da malha aérea doméstica do país, é uma tarefa difícil, porém, extremamente importante para compreender o passado e o presente, para assim, facilitar a imaginação ou a idealização da aviação doméstica do futuro. Para tal empreitada, é importante se munir de dados confiáveis e de fácil acesso, afinal, em um mercado tão competitivo como o da aviação, qualquer informação pode ser a origem de uma vantagem competitiva, logo as empresas as guardam sobre sete chaves.

No Brasil, a aviação civil possui uma longa tradição. É evidente que tal tradição é marcada por erros e acertos. E talvez o maior acerto de todos seja a confiável regulação da aviação pelo Estado, tornando-se assim o Brasil como um dos países

mais seguros de se voar e também com transparência em informações conhecidas por serem tão inacessíveis em outros países.

## 1.1 Contextualização

A Agência Nacional de Aviação Civil (ANAC), é uma das agências reguladoras federais do País, foi criada para regular, normatizar, supervisionar e fiscalizar as atividades da aviação civil e a infraestrutura aeronáutica e aeroportuária no Brasil. Instituída pela lei federal nº 11.182 de 27 de setembro de 2005 (BRASIL, 2005), começou a atuar após o decreto federal nº 5.731 de 20 de março de 2006 (BRASIL, 2006) instituindo sua instalação, estrutura organizacional e substituição do Departamento de Aviação Civil (DAC). É uma autarquia federal de regime especial e está vinculada ao Ministério da Infraestrutura. Suas ações se enquadram nos macroprocessos de certificação, fiscalização, normatização e representação institucional.

Dentro dessas atribuições, cabe à ANAC editar, revisar e atualizar regulamentos técnicos e também os regulamentos relacionados a aspectos econômicos. Suas normas estão em acordo com a Convenção de Chicago, e à OACI, a principal organização internacional de aviação, cujo Brasil é membro. A instrução normativa número 127, de 5 de outubro de 2018 (BRASIL, 2018), determina regras e diretrizes para edição do regimento interno, organização das unidades organizacionais e para os processos de modificação da estrutura organizacional da agência, de acordo com o site da agência.

Entender sobre a regulação aérea é fundamental para entender a evolução da malha aérea no país já que a aviação civil no Brasil possui um amplo histórico de interferências estatais. Segundo (MENDES, 2011), diversas intervenções por parte do estado foram feitas na aviação civil nacional, pois esse modal, por muito tempo, foi e ainda continua sendo visto como estratégico para o país. Foi a partir dessa ideia que foi estabelecido um amplo controle de rotas, horários de voos, frequências e tarifas sob pena de multa de descumprimento para as companhias aéreas por meio do Decreto-Lei Nº 32 de 1966 (BRASIL, 1966).

Foi somente na década de 90 que o Estado adotou políticas de flexibilização das regulamentações de diversos setores da economia, incluindo a aviação. Segundo (CASTRO & LAMY, 1993), essa foi a porta de entrada para novas empresas no mercado e o início do fim dos monopólios existentes. E é a partir deste momento em que a malha aérea tal qual é conhecida nos dias de hoje toma forma. Como pôde ser

visto anteriormente, a malha aérea é uma representação, sobre um mapa, dos voos realizados em uma ou mais regiões. Por consequência, quando se diz malha aérea doméstica, se abrange somente os voos comerciais realizados no território nacional.

Como observa-se na Figura 2, onde tem-se a malha aérea da extinta companhia aérea Varig RioSul. Cada cidade em que a companhia atua é representada por um círculo, e os voos que partem das cidades são representados por linhas. Quando se chega a uma grande quantidade de voos, a representação no mapa remete à imagem de uma malha aérea, semelhante à Figura 1, abaixo:

*Figura 1 - Malha aérea doméstica da Varig Rio-Sul*



Fonte: [varig-airlines.com](http://varig-airlines.com)



## 1.2 Formulação do problema

Sabe-se que é possível encontrar as mais variadas informações no campo dos dados abertos disponíveis no portal da ANAC, conforme presente na Figura 2, abaixo:

*Figura 2 - Dados e Estatísticas da Agência Nacional de Aviação Civil*



*Fonte: Portal ANAC*

Algumas dessas informações, porém, não se apresentam de forma intuitiva, e nem de fácil entendimento, sendo necessário, na maioria das vezes, que o próprio interessado possua conhecimentos avançados em análise de dados para poder extrair e analisar as informações de forma detalhada. E esse é o caso dos dados abertos referentes ao histórico de voos, que se apresentam de uma forma pouco estimulante, visualmente falando, para o interessado, conforme a Figura 3, abaixo:

Figura 3 - Index dos dados abertos da ANAC

## Index of /dadosabertos/Voos e operações aéreas/Dados

<a href="#">../</a>		
<a href="#">Dados Estatisticos.csv</a>	24-Feb-2023 01:12	312668333
<a href="#">Dados Estatisticos 2000 a 2010.json</a>	24-Feb-2023 01:12	482435892
<a href="#">Dados Estatisticos 2011 a 2020.json</a>	24-Feb-2023 01:12	431644761
<a href="#">Dados Estatisticos 2021 a 2030.json</a>	24-Feb-2023 01:13	70276441

Fonte: Portal ANAC

Diante dessa situação, surge uma real necessidade de suprir uma deficiência que compreende a apresentação superficial dos dados de algo que merece um olhar tão minucioso. A fim de aprimorar a visualização dos tão recônditos dados disponíveis pela ANAC, a necessidade de modelar e desenvolver uma ferramenta que possa extrair e traduzir essas informações para expô-las ao público de forma mais dinâmica e seguindo um padrão cronológico, torna-se vital para encurtar a longa distância que separa os dados dos leitores e interessados. Logo, de que maneira é possível coletar e analisar os dados abertos disponíveis na Agência Nacional de Aviação Civil sobre os voos domésticos e exibi-los de forma precisa e eficiente para a comunidade?

### **1.3 Objetivo Geral**

O objetivo deste trabalho é desenvolver uma ferramenta de extração, processamento e análise de informações a partir de uma base de dados disponível da empresa GOL LINHAS AÉREAS no portal da ANAC, e dispô-las de forma clara e intuitiva por meio de um mapeamento de conexões aeroportuárias no Brasil.

### **1.4 Objetivos Específicos**

Os três objetivos específicos deste trabalho são:

- 1) Extrair os dados referentes ao histórico de voos no portal da ANAC;
- 2) Compor e desenvolver um modelo de notebook em Python que analise os dados extraídos e gere um modelo gráfico de fácil compreensão a partir dos dados coletados; e
- 3) Realizar uma apresentação gráfica e espacial dos dados extraídos e analisados.

### **1.5 Justificativa**

A modelagem de dados da malha aérea doméstica da GOL LINHAS AÉREAS no quarto trimestre de 2019 serve como um modelo básico para entender de forma inicial a disposição dos voos no Brasil para aquele período que compreende os meses de outubro, novembro e dezembro. Ter um modelo desses em mãos é vital para sua posterior replicação para períodos mais extensos e, conseqüentemente, compreender a evolução desse modal ao longo dos anos. De fato, os dados, aqui recolhidos e analisados, atrelados a outros materiais poderão compreender informações ainda mais profundas tais como o desenvolvimento das regiões do país, avanço tecnológico da aviação e das infraestruturas aeroportuárias, bem como diversos outros dados, evidenciando-se assim a sua grande relevância acadêmica. Vale ressaltar, também, que essa pesquisa pode contribuir para que ferramentas e plataformas completas

sejam criadas, de forma que seja possível exceder as barreiras da ANAC e chegar até mesmo em outras organizações.

Apesar de tal estudo possuir uma relevância acadêmica, nota-se também que a relevância técnico-econômica também se faz fortemente presente. Se por um lado, as informações aqui geradas podem refletir uma ampla observação dos aspectos socioeconômicos, por outro lado, é evidente a importância dos resultados aqui apresentados para todo o mercado da aviação civil no país.

Por fim, e não menos importante, no que tange a sua relevância social, pode-se imaginar o uso deste breve trabalho por parte da agência e de tantas outras organizações. De fato, estima-se que a ferramenta aqui utilizada poderá ser facilmente replicada em outros projetos a fim de fortalecer os diferentes tipos de governanças, para que as mesmas consigam tirar proveito da melhor forma possível nas análises dos dados e em diversas outras demandas em que se adeque às suas aplicações, mesmo que mínimo.

## 2 REVISÃO TEÓRICA

Nesta pesquisa os conceitos teóricos de malha de transporte aéreo doméstico e visualização de dados serão revisados. Abaixo está a conceituação de cada um deles de acordo com a literatura acadêmica.

### 2.1 Malha aérea doméstica

A malha aérea aeronáutica é uma representação em forma de grafos cuja trajetória é representada por meio de retas ligando dois ou mais pontos, pontos esses que representam os aeroportos de origem, destino e escala caso se encontre mais de duas retas. Bergiante (2003) discutiu sobre a importância dada à pesquisa da evolução das rotas aéreas no Brasil. Costa, Lohmann e Oliveira (2010) analisaram o impacto das rotas aéreas nos aeroportos brasileiros, identificando os aeroportos por número e concentração de voos.

O pilar principal na compreensão da malha aérea é a tipologia de rotas existentes na aviação civil. Segundo Corrêa (2013), tem-se três principais classificações de rotas quanto ao tipo de execução:

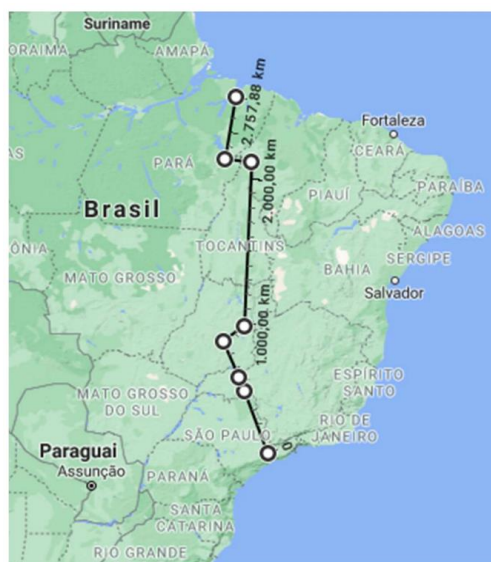
- Rota em linha (point-to-point);
- Rotas circulares; e
- Rota com centro de distribuição (Hub & Spoke).

É possível observar que cada tipo de rota teve o seu auge em épocas diferentes, evidenciando assim uma evolução da malha aérea. Os fatores que influenciaram nesta evolução são diversos: evolução tecnológica das aeronaves, crises econômicas, desenvolvimento regional e outros com graus diferentes de tangibilidade. Além disso, segundo Bazargan (2010), toda companhia aérea possui ao menos um hub e diversos spokes, podendo elas optarem por ter também diversos hubs.

### 2.1.1 Rota em linha (point-to-point)

A rota em linha atende um grupo de cidades e é realizada em uma sequência de escalas. Pouco utilizada na atualidade, já foi amplamente empregada para atender cidades com baixa demanda de passageiros e localizadas entre cidades de grande demanda. Um bom exemplo é o voo RG-254 da extinta Varig, aqui representada na Figura 3. Cada ponto branco representa uma parada da aeronave. O voo partia de Guarulhos, em São Paulo, com destino a Belém, no Pará, com paradas em Uberaba, Uberlândia, Goiânia, Brasília, Imperatriz e Marabá.

*Figura 4 - Rota do voo RG-254 da Varig*



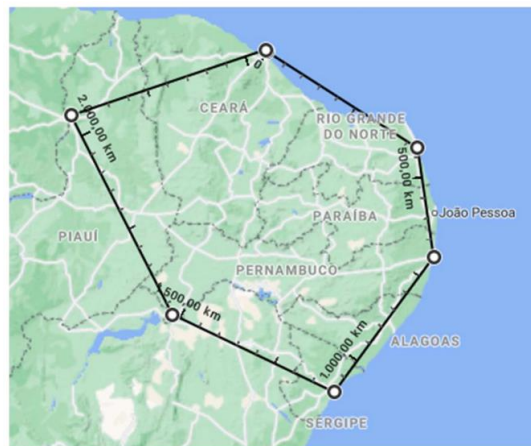
*Fonte: Google MyMaps.*

A rota em linha faz parte da estratégia de rota point-to-point, que traduzido ao português, representa os voos diretos entre dois pontos. Essa estratégia dispensa um hub (que é um aeroporto principal de escala) pois conecta dois pontos diretamente. Segundo Cook e Goodwin (2008), o modelo point-to-point é antagônico ao modelo Hub & Spoke, porém, não é incomum encontrar empresas aéreas que adotam as duas estratégias simultâneas de forma complementar.

### 2.1.2 Rotas circulares

As rotas circulares são semelhantes às rotas em linha e são dificilmente encontradas nos dias de hoje. Esse tipo de percurso também atende cidades cuja demanda não justifica voos diretos entre cada uma delas. A diferença é que a disposição das cidades, ao invés de se dar próximo a uma linha reta, se dá em formato de um círculo. As rotas circulares possuem as mesmas vantagens e desvantagens da rota em linha: atendem cidades com baixa demanda de passageiros e localizadas entre cidades de grande demanda. Na Figura 4, tem-se um exemplo hipotético de rota circular ligando as cidades de Fortaleza, Natal, Recife, Aracaju e Teresina.

*Figura 5 - Exemplo de rota circular*



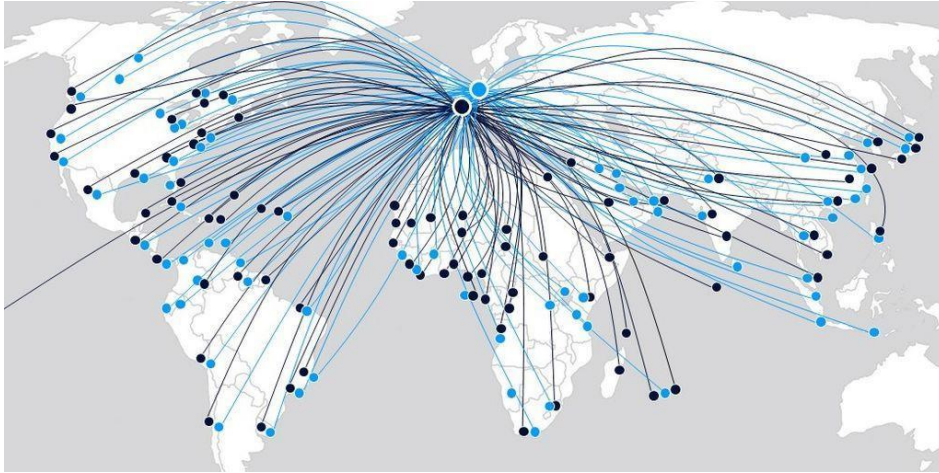
Fonte: Google MyMaps

### 2.1.3 Rota com centro de distribuição (Hub & Spoke)

O modelo tal como é conhecido hoje por *Hub & Spoke* surgiu na década de 70 nos Estados Unidos e é o modelo mais comumente encontrado nos dias de hoje. Esse sistema consiste na escolha de um determinado aeroporto como centro de operações (também conhecido como hub). Segundo Cook e Goodwin (2008), esse modelo ganha no quesito de conectividade entre os destinos, porém perde no quesito de dependência de destinos em comparação com o modelo point-to-point. Nesse modelo, a malha é elaborada de forma a ter um grande número de voos diretos ligando os hubs aos demais destinos atendidos por uma companhia. O *Hub & Spoke* apresenta a vantagem de aumentar a taxa de ocupação dos voos, tornando, portanto,

o uso da aeronave mais eficiente. Na Figura 5 tem-se um exemplo dos Hubs do grupo AirFrance-KLM.

*Figura 6 - Hubs do grupo AirFrance-KLM*

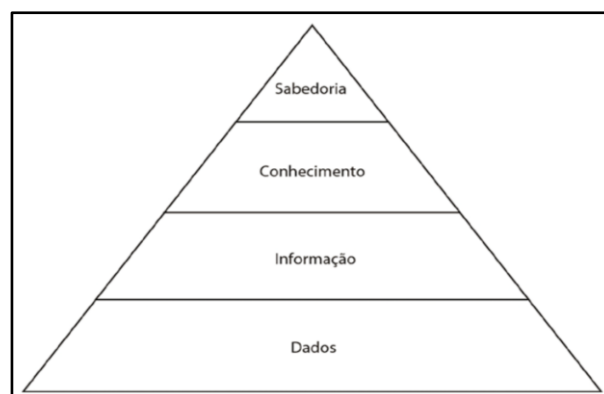


*Fonte: International Transport Journal, 2018.*

## 2.2 Visualização de dados

A análise exploratória de dados (AED) possui um papel fundamental nesse estudo. De fato, de nada adianta extrair dados sem apresentá-los de forma adequada. Segundo (FERREIRA; MIRANDA; PINTO, 2021), o uso correto dos dados pode transformar informação em conhecimento, e a utilização desses conhecimentos em benefício da tomada de decisão gera o que é conhecido como sabedoria. Na Figura 6, tem-se um exemplo da pirâmide de conhecimento a partir das diferentes etapas da transformação dos dados.

*Figura 7 - Pirâmide das etapas de transformação do dado*



*Fonte: Preparação e Análise Exploratória de Dados, pág.14, 2021.*



Ainda segundo Ferreira, Miranda e Pinto (2021), sabe-se que os dados não se apresentam em pacotes prontos e formatados para análises, logo, é preciso transformar todas as fontes de dados de forma que a análise possa ser feita.

A análise exploratória de dados (AED) tem por objetivo final verificar os dados para qualquer aplicação estatística. É somente a partir do entendimento dos dados coletados que se torna possível observar assertivamente sobre as relações existentes entre as variáveis analisadas. É importante ressaltar que, diferentemente dos outros modelos de estatísticas, a AED tenta olhar para os dados de forma mais profunda sem resumir muito a quantidade de informações. Na Tabela 1, tem-se um exemplo das abordagens estatísticas e estratégias utilizadas na AED.

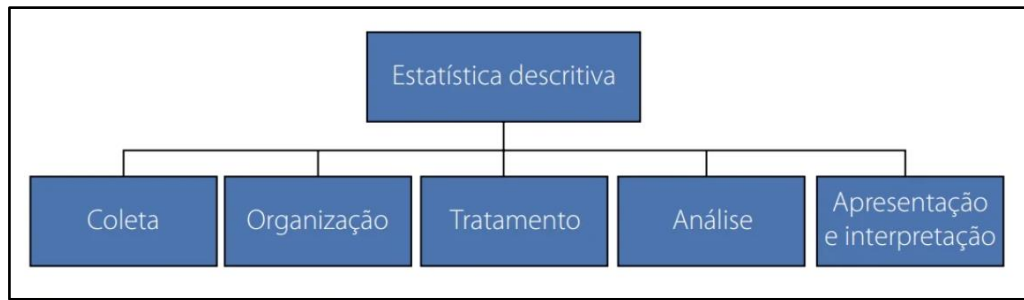
*Tabela 1 - Abordagens estatísticas e estratégias utilizadas na AED*

<b>Abordagem</b>	<b>Estratégia</b>
Estatística clássica	Problema → Dados → Modelo → Análise
Estatística bayesiana	Problema → Dados → Modelo Priori → Análise
AED	Problema → Dados → Análise → Modelo

*Fonte: Preparação e Análise Exploratória de Dados, pág.15, 2021*

No que tange à questão de valores numéricos e não numéricos, tem-se a distinção de dados quantitativos e qualitativos respectivamente. Vale ressaltar também a importância da representação dos dados de forma fácil para o leitor, o que segundo Ferreira, Miranda e Pinto (2021), é mais facilmente atingido por meio de representações gráficas e tabelas; se por um lado, os gráficos facilitam a visualização dos dados, as tabelas, por sua vez, demonstram as precisões dos números em si. Segundo Navidi (2012), uma amostragem possui um grande número de valores e, para destacar as características dessa amostra, usualmente se calculam algumas medidas estatísticas. A estatística descritiva, portanto, repousa na interpretação dos dados, na sua visualização e, por fim, relaciona-os com as variáveis. Ela está diretamente relacionada com os objetivos da pesquisa e é composta por etapas, conforme apresenta a Figura 8.

*Figura 8 - Etapas da estatística descritiva*



*Fonte: Preparação e Análise Exploratória de Dados, pág.15, 2021*

### **3 METODOLOGIA**

Na metodologia científica, são discutidos a tipologia e descrição dos métodos de pesquisa, a caracterização do objeto de estudo, a população e caracterização dos instrumentos de pesquisa e os procedimentos de coleta e análise dos dados.

#### **3.1 Tipologia e descrição geral dos métodos de pesquisa**

O tipo de pesquisa proposto é o da modelagem de dados utilizando a metodologia do Design Science Research, usando como ferramenta de linguagem de programação, a linguagem Python. Para criação de códigos para esta pesquisa foi utilizado a plataforma Google Colab, portal de hospedagem de códigos para controle de versão e colaboração. O portal permite que pessoas trabalhem colaborativamente no desenvolvimento de códigos para diferentes ferramentas, entre elas a de modelagem.

#### **3.2 Caracterização da organização, setor ou área, indivíduos objeto do estudo**

Como pôde ser visto de forma breve na contextualização, o trabalho da ANAC consiste em elaborar normas, certificar empresas, oficinas, escolas, profissionais da aviação civil, aeródromos e aeroportos e fiscalizar as operações de aeronaves, de empresas aéreas, de aeroportos e de profissionais do setor, certificar e validar novas aeronaves, normatizar e supervisionar a infraestrutura aeroportuária, fiscalizar as relações econômicas de consumo no âmbito da aviação civil. A principal atribuição da Agência consiste na regulação das atividades de aviação civil, atuando com foco na segurança e na qualidade do transporte aéreo, para estimular a concorrência e a melhoria da prestação dos serviços no setor.

Sua missão é garantir a segurança e a excelência da aviação civil. A agência tem como visão ser referência na promoção da segurança e no desenvolvimento da aviação civil. Ela possui os seguintes valores:

- Segurança é o nosso propósito;
- Atuamos com foco no resultado e no interesse público;
- Trabalhamos com autonomia e competência técnica;
- Agimos com integridade, comprometimento e transparência;
- Valorizamos as pessoas e suas competências;
- Incentivamos a inovação e a cooperação no setor de aviação civil; e
- Temos orgulho de trabalhar na ANAC.

Além disso, a dentre as principais competências destacam-se também, atribuídas à agência:

- Representar o Brasil junto a organismos internacionais de aviação e negociar acordos e tratados sobre transporte aéreo internacional;
- Emitir regras sobre segurança em área aeroportuária e a bordo de aeronaves civis;
- Conceder, permitir ou autorizar a exploração de serviços aéreos e de infraestrutura aeroportuária;
- Estabelecer o regime tarifário da exploração da infraestrutura aeroportuária;
- Administrar o Registro Aeronáutico Brasileiro (RAB);
- Homologar, registrar e cadastrar os aeródromos;
- Emitir certificados de aeronavegabilidade atestando aeronaves, produtos e processos aeronáuticos e oficinas de manutenção;
- Fiscalizar serviços aéreos e aeronaves civis;
- Certificar licenças e habilitações dos profissionais de aviação civil;
- Autorizar, regular e fiscalizar atividades de aeroclubes e escolas e cursos de aviação civil; e
- Reprimir infrações às normas do setor, inclusive quanto aos direitos dos usuários, aplicando as sanções cabíveis.

O objeto de estudo desta pesquisa é o mercado de transporte aéreo no Brasil. Especificamente, levantou-se os dados coletados e produzidos pela Agência Nacional de Aviação Civil do Brasil, a ANAC.

No Anexo I a este trabalho consta a Lei no 11.182, de 27 de dezembro de 2005, que cria a Agência Nacional de Aviação Civil. Também o Decreto no 5.731, de 20 de março de 2006, que dispõe sobre a instalação, a estrutura organizacional da ANAC e aprova o seu regulamento. O Regimento Interno da ANAC, através da Resolução no 381, de 14 de junho de 2016. A Resolução no 581, de 21 de agosto de 2020, que altera o regimento interno da agência. A Instrução Normativa no 127, de 05 de outubro de 2018, que estabelece regras e diretrizes para a edição do regimento interno, para a organização interna das unidades organizacionais e para os processos de modificação da estrutura organizacional da Agência.

### **3.3 População e amostra ou Participantes da pesquisa**

A população analisada foi os voos da companhia aérea GOL LINHAS AÉREAS, que é regulada pela ANAC. A ANAC tem a prerrogativa de regular o setor de transporte aéreo no Brasil, editar normas e decretos sobre o processo de regulação das atividades do setor, diretrizes e procedimentos para garantir a segurança do transporte aéreo no país, bem como fiscalizar a infraestrutura aeronáutica e aeroportuária do setor.

O período escolhido para a pesquisa é o quarto trimestre de 2019, período esse que compreende os meses de outubro, novembro e dezembro (também chamado no 4T). A escolha desse período se deu por causa de alguns fatores importantes como:

- Período antes da pandemia, sem interferências externas significativas à malha aérea doméstica nacional e internacional causadas pelo vírus COVID-19;
- Período extremamente intenso e com altos índices de voos no Brasil e na grande maioria dos países;
- Dados registrados tanto por parte da reguladora (ANAC) e da regulada (GOL LINHAS AÉREAS).

Dos dados coletados pela ANAC, uma amostra será selecionada para análise. Entre os dados que podem ser analisados usando a ferramenta proposta estão disponíveis em Dados Abertos, Voos e Operações Aéreas, Registro de Serviços Aéreos.

### 3.4 Caracterização e descrição dos instrumentos de pesquisa

As metodologias utilizadas para desenvolver o trabalho foram a Design Science (DS) e a Design Science Research (DSR). De acordo com Dresch, Lacerda e Cauchick-Miguel (2019, v. 1, p. 109) “A DS/DSR pode ser aplicada para a formalização, prescrição e o projeto de soluções satisfatórias para as organizações”. O desenvolvimento da Design Science se deu a partir da necessidade de estruturar e resolver problemas complexos, ou seja, aqueles em que a solução não poderiam ser consideradas comuns (DRESCH; LACERDA; CHAUCHICK-MIGUEL, 2019).

Essas metodologias partem do pressuposto de que para se chegar a um resultado satisfatório, deve-se utilizar de instrumentos criados pelos homens, os chamados artefatos. É possível observar as definições e exemplos dos diferentes tipos de artefatos na Figura 9.

*Figura 9 - Tipos de Artefatos*

Tipo de artefato	Definição	Exemplo de artefato desenvolvido	Referências dos exemplos
Constructo*	Conceitos utilizados para descrever problemas ou especificar as respectivas soluções	Constructos para descrever os fluxos de informação de um processo de negócio	<a href="#">Rosenkranz e Holten (2011)</a>
Modelo	Conjunto de elementos e relações que representam a estrutura geral da realidade	Modelo de controle integrado da produção e da qualidade para construção civil	<a href="#">Leão, Isatto e Formoso (2016)</a>
Método	Conjunto de passos lógicos necessários para a efetivação de determinada atividade	Método para resolver desafios concernentes à arquitetura organizacional	<a href="#">Nakakawa, Van Bommel e Proper (2011)</a>
Instanciação	Execução do(s) artefato (s) em seu ambiente real, evidenciando a viabilidade e eficácia dos artefatos	Instanciação de um método para desenvolver habilidades de resiliência em eletricitistas	<a href="#">Saurin et al. (2014)</a>
<i>Design Proposition</i>	Regras tecnológicas ou regras de projeto, consideradas contribuições teóricas da <i>Design Science</i> .	<i>Design proposition</i> para uma abordagem de inovação aberta para micro e pequenas empresas	<a href="#">Krause e Schutte (2016)</a>

*Fonte: Metodologia Científica para Engenharia (Dresch et al., 2019)*

Nesta pesquisa os tipos de artefatos utilizados foram método, modelo e instanciação. O modelo utilizado, isto é, a representação da estrutura geral da

realidade, foram os dados obtidos através da coleta em campo e dos Dados Abertos presentes no portal da agência. O método utilizado, ou seja, o conjunto de passos lógicos necessários para a realização de uma atividade, foram a coleta de dados no portal da ANAC em Dados Abertos, o seu manuseio e análise por meio do notebook em Python disponível no Apêndice A e exibição dos resultados de forma precisa e eficiente. A instanciação foi executada através do modelo e método sendo adaptados para o Python.

O problema desenvolvido precisou ser separado em dois conceitos: validade científica e pragmática. O primeiro é muito importante para a Design Science, pois diz respeito à maneira com que as estruturações do problema e sua resolução são feitas e formalizadas considerando os modelos e métodos científicos envolvidos. O segundo, porém, é essencial já que se configura sendo a aplicação do que foi cientificamente desenvolvido. Isso significa que para a DS (Design Science) a prática é muito relevante. É importante frisar que dentro da metodologia, a solução não precisa ser necessariamente ótima, mas sim satisfatória, visto que em diversos casos a solução ótima nem sempre se torna viável. Este princípio será considerado na estruturação da resolução do problema de pesquisa (DRESCH; LACERDA; CHAUCHICK- MIGUEL, 2019).

Segundo Dresch et al. (2015), a Design Science Research possui 12 etapas para sua execução. Abaixo estão evidenciadas todas essas etapas, assim como o resumo e/ou objetivo de cada uma delas:

i) Identificação do problema;

Entender exatamente a problemática sobre a qual o estudo foi desenvolvido. No caso deste trabalho, como dito na seção introdutória, foi a criação de uma ferramenta de extração, processamento e análise de informações a partir de uma base de dados disponível no portal da ANAC, e dispô-las de forma clara e intuitiva por meio de um mapeamento de dos voos da GOL LINHAS AÉREAS, no período de outubro à dezembro de 2019 (4T19).

ii) Conscientização do problema e Revisão sistemática da literatura;

Estudo do conteúdo já produzido sobre ou relacionado ao problema de pesquisa. Neste trabalho esse tópico foi aprofundado durante a revisão teórica da literatura.

iii) Identificação de artefatos e configuração das classes de problemas;

Identificação dos artefatos criados pelos pesquisadores durante a revisão da literatura. Artefatos são artificialidades ou artifícios criados, ou seja, não naturais, para o atingimento de determinado objetivo do projeto. Um exemplo claro de artefato são os números, que não existem naturalmente, mas foram criados pelo homem para “quantificar” o mundo.

A partir da identificação dos artefatos que podem ser utilizados, há também a configuração das classes de problema para organizar e facilitar a estruturação e execução do restante do projeto. Neste trabalho, essa etapa consistiu na identificação dos dados que foram escolhidos para análise no site da ANAC.

iv) Proposição de artefatos para resolver o problema específico;

Ainda com base na revisão da literatura, propor mecanismos não naturais para a resolução do problema de pesquisa. No caso do presente projeto, foram utilizados os conceitos de modelo, método e instanciação para a criação do notebook em Python.

v) Projeto do artefato selecionado;

Serão levados em consideração tanto a literatura e conceitos conhecidos sobre o artefato escolhido, como também o contexto em que o problema de pesquisa está proposto. Neste trabalho, essa etapa consistiu em entender a estratégia de rota utilizada pela GOL LINHAS AÉREAS, para ajustar a criação do notebook em Python. O notebook pode ser encontrado no Apêndice A.

vi) Desenvolvimento do artefato;

Os artefatos foram desenvolvidos com o auxílio da linguagem de programação Python na plataforma Google Colab. O notebook pode ser encontrado no Apêndice A.

vii) Avaliação do artefato;



É a etapa em que se avalia se o artefato desenvolvido de fato atende a demanda e aos objetivos do projeto. Nesta etapa, verificou-se a efetividade e funcionamento do notebook em Python. O notebook pode ser encontrado no Apêndice A.

viii) Explicitação das aprendizagens;

Ao longo do desenvolvimento do projeto diversas iterações devem acontecer, de forma com que aprendizados sejam registrados e melhorias sejam aplicadas. A documentação de tudo isso é uma das etapas do processo da Design Science Research, pois contribui para a evolução do projeto, e conseqüentemente, a resolução do problema. No presente trabalho, a explicitação das aprendizagens refletiu na “lapidação” do código, tornando-o bem amoldado com o objetivo do trabalho.

ix) Conclusões;

Resumo da pesquisa e dos resultados obtidos com ela, principalmente a partir da avaliação do próprio artefato. Nesta etapa, foram divulgados os resultados encontrados ao longo do trabalho.

x) Generalização para uma classe de problemas;

Sintetização de tudo o que foi desenvolvido na pesquisa, evidenciando o seu passo a passo, suas etapas e o detalhamento de todas as decisões tomadas, sendo estas acompanhadas de suas respectivas justificativas. Dessa forma, é possível realizar a generalização do artefato para uma classe de problemas específica.

xi) Comunicação dos resultados;

Comunicar os resultados obtidos para a comunidade acadêmica e para profissionais que possuam relação com o problema que foi desenvolvido. É muito importante dentro da Design Science e da Design Science Research que se busque dar acessibilidade aos resultados da pesquisa para essas pessoas.

### **3.5 Procedimentos de coleta e de análise de dados**

Nesta pesquisa os dados coletados para análise foram colhidos de forma manual na aba de dados abertos no site da ANAC. Uma amostra foi selecionada para análise. Entre os dados que podem ser analisados usando a ferramenta proposta estão disponíveis em Dados Abertos, Voos e Operações Aéreas, Dados Estatísticos do transporte aéreo.

Os procedimentos para coleta dos dados se deu por meio do acesso ao site da ANAC, na área de “Dados e Estatísticas”. Portanto a amostra desta pesquisa será por conveniência. Após a coleta os dados foram dispostos em planilhas excel.

As variáveis coletadas foram de valor quantitativo, o que permitiu uma boa adaptação ao notebook em Python e uma maior facilidade análise por se tratar do código por se tratar de valores numéricos.

### **3.6 Análise de dados**

A análise destes dados será orientada a partir da utilização de notebooks desenvolvidos na etapa de programação dos instrumentos e ocorrerão somente após a coleta de todos os dados.

Em um primeiro momento, os dados serão extraídos da base de dados abertos da ANAC de forma manual (por se tratar de um período de tempo relativamente curto, a extração manual pôde ser possível, o que é inviável para períodos muito extensos, que exige programas de raspagem de dados) Os códigos usados neste trabalhos poderão ser encontrados no apêndice A deste trabalho. Em um segundo momento, se fará necessário a utilização da biblioteca “pandas” para manipular os dados tabulares extraídos pela ANAC. Por fim, após a manipulação, diversas planilhas serão geradas e analisadas. Os códigos usados neste trabalho poderão ser encontrados no apêndice A deste trabalho.

Os dados serão expostos em formas de gráficos bem como dispostos também em gráficos inteligentes contendo os principais aeroportos e as incidências das rotas.

## 4 RESULTADOS E DISCUSSÕES

Seguindo um padrão semelhante aos modelos de visualização de rotas de voos desenvolvidos pelas empresas aéreas, esse modelo utilizou os dados abertos relativos aos voos domésticos da GOL LINHAS AÉREAS disponíveis no portal da ANAC para construir um modelo próprio. Um dos muitos desafios desse trabalho consistiu na junção de arquivos referentes às informações sobre os voos no país. Com um grande número de arquivos disponíveis para análise, o trabalho manual se fez inviável, visto que há uma limitação de tempo em analisar os mais de noventa arquivos diferentes contendo informações de mais de sessenta mil voos.

Além dos dados recolhidos da ANAC, também foram adquiridos dados divulgados pela companhia aérea GOL LINHAS AÉREAS, para o quarto trimestre de 2019 (4T19), para ter uma noção inicial do ponto de vista da companhia aérea e também para possíveis comparações futuras.

Como pode-se observar anteriormente, as informações necessárias foram obtidas junto à ANAC e à companhia aérea GOL LINHAS AÉREAS, através da extração de dados feitos manualmente. Em um segundo momento, fez-se a necessidade de um código em Python, para refinar as informações e reduzi-las ao mínimo necessário para facilitar e otimizar as análises.

A análise foi realizada de acordo com todos os 12 critérios necessários de análise da Design Science, voltada para identificar padrões, tendências e anomalias e em um último momento, utilizar tabelas e ilustrações para ajudar na interpretação dos dados.

### 4.2 Execução

No apêndice A desta tese se encontra o código utilizado na plataforma Google Colab para posteriores consultas.

Após o aprimoramento dos dados colhidos junto à ANAC na seção dos dados estatísticos do transporte aéreo, grande parte das variáveis foram cortadas visto que elas não acrescentaram fatos relevantes para a análise da malha aérea doméstica selecionada para o período. Os principais dados tidos como relevantes para a

pesquisa foram:

- Companhia aérea;
- Data;
- Status do voo (Grupo de voo);
- Aeroporto de origem;
- Estado do aeroporto de origem;
- Aeroporto de destino;
- Estado do aeroporto de destino; e
- Número de decolagens de cada trecho.

Logo abaixo na Tabela 03 é possível identificar uma ilustração exemplificando as informações aprimoradas a partir da extração da seção dos dados estatísticos do transporte aéreo.

*Tabela 2 - Definição dos dados*

empresa_sigla	ano	mes	aeroporto_de_origem	uf_origem	aeroporto_de_destino	uf_destino	grupo_de_voo	decolagens
GLO	2019	12	SBCF	MG	SBBR	DF	REGULAR	102
GLO	2019	12	SBCF	MG	SBCJ	PA	REGULAR	19
GLO	2019	12	SBCF	MG	SBEG	AM	REGULAR	1
GLO	2019	12	SBCF	MG	SBGL	RJ	REGULAR	74
GLO	2019	12	SBCF	MG	SBGR	SP	REGULAR	137

Segundo os dados estatísticos do transporte aéreo colhidos, para o quarto trimestre de 2019, o período escolhido no trabalho, a companhia aérea “GOL LINHAS AÉREAS” cuja sigla é reconhecida por “GLO” tanto na aviação como um todo mas também para a ANAC, realizou um total de 62.087 voos (descrito como “decolagens” na planilha) entre 56 destinos. É importante frisar que dos 62.087 voos, 61.608 são voos regulares e 479 voos improdutivos (voos sem passageiros geralmente causados por imprevistos e operações não programadas). É importante destacar neste momento os voos improdutivos, pois para os lançamentos de resultados oficiais da companhia aérea, tais voos são contabilizados pois possuem um impacto nos custos da empresa. Abaixo nas Tabela 3, 4 e 5, pode-se observar os resultados divulgados pela empresa para o quarto trimestre de 2019.

Tabela 3 - Dados de voos doméstico da GOL para Outubro

GOL Doméstico									
Decolagens	20.487	19.122	7,1%	198.620	195.293	1,7%	239.189	236.826	1,0%
Assentos (milhares)	3.669	3.345	9,7%	34.845	33.362	4,4%	41.929	40.293	4,1%
ASK (milhões)	3.695	3.437	7,5%	35.949	34.963	2,8%	43.413	42.388	2,4%
RPK (milhões)	3.060	2.749	11,3%	29.838	27.988	6,6%	36.126	34.064	6,1%
Taxa de Ocupação	82,8%	80,0%	2,8 p.p	83,0%	80,0%	3,0 p.p	83,2%	80,4%	2,8 p.p
Pax transportados (milhares)	2.950	2.615	12,8%	28.110	25.740	9,2%	33.935	31.243	8,6%

Fonte: GOL LINHAS AÉREAS, 2019

Na Tabela 3, logo acima, foi possível observar os dados divulgados pela GOL LINHAS AÉREAS para o mês de outubro. Para o presente trabalho, somente a informação das decolagens é relevante, pois permite a comparação com os resultados encontrados a partir da análise dos dados abertos da ANAC. No mês de outubro, segundo a companhia aérea, 20.487 decolagens foram realizadas, 1.365 decolagens a mais que o mesmo período em 2018 (um ano antes) de 19.122 decolagens.

Tabela 4 - Dados de voos doméstico da GOL para Novembro

GOL Doméstico									
Decolagens	20.524	18.923	8,5%	219.144	214.216	2,3%	240.790	236.353	1,9%
Assentos (milhares)	3.679	3.295	11,6%	38.524	36.658	5,1%	42.312	40.354	4,9%
ASK (milhões)	3.716	3.413	8,9%	39.682	38.377	3,4%	43.733	42.423	3,1%
RPK (milhões)	3.036	2.858	6,3%	32.886	30.845	6,6%	36.317	34.146	6,4%
Taxa de Ocupação	81,7%	83,7%	-2,0 p.p	82,9%	80,4%	2,5 p.p	83,0%	80,5%	2,5 p.p
Pax transportados (milhares)	2.939	2.696	9,0%	31.021	28.437	9,1%	34.149	31.358	8,9%

Fonte: GOL LINHAS AÉREAS, 2019

Na tabela 4, logo acima foi possível observar os dados divulgados pela GOL LINHAS AÉREAS para o mês de novembro. Com um total de 20.524 decolagens, houve mais decolagens no mês de novembro em comparação com o mês de outubro.

O mês de dezembro, porém, representado na tabela 6, abaixo, apresentou um total de 23.259 voos, com um aumento significativo em comparação com os meses de outubro e novembro. Tal aumento é justificado pela sazonalidade, de fato, por dezembro ser o mês das férias e festividades de fim de ano, o fluxo de passageiros aumenta. Todos resultados das decolagens para os três meses (outubro, novembro e dezembro) são semelhantes aos encontrados nos dados abertos da ANAC, comprovando assim a efetividade dos resultados divulgados.

Tabela 5 - Dados de voos doméstico da GOL para Dezembro

GOL Doméstico									
Decolagens	23.259	21.646	7,5%	64.265	59.691	7,7%	242.398	235.862	2,8%
Assentos (milhares)	4.108	3.788	8,5%	11.455	10.429	9,8%	42.631	40.446	5,4%
ASK (milhões)	4.260	4.051	5,2%	11.677	10.901	7,1%	43.939	42.428	3,6%
RPK (milhões)	3.535	3.431	3,1%	9.638	9.037	6,6%	36.421	34.276	6,3%
Taxa de Ocupação	83,0%	84,7%	-1,7 p.p	82,5%	82,9%	-0,4 p.p	82,9%	80,8%	2,1 p.p
Pax transportados (milhares)	3.341	3.128	6,8%	9.189	8.440	8,9%	34.336	31.565	8,8%

Fonte: GOL LINHAS AÉREAS, 2019

Para a construção do panorama da malha aérea neste trabalho, os voos improdutivos serão descartados pois não são relevantes para a construção da malha aérea visto que esses não transportam passageiros. Vale ressaltar que como o período de análise se resume ao quarto trimestre de 2019, todos os 61.608 voos serão observados nos meses de outubro, novembro e dezembro de 2019. Além disso, a companhia aérea GOL LINHAS AÉREAS possui sete HUBs espalhados em três regiões do país. A lista completa de todos os HUBs se encontra abaixo, acompanhada dos códigos ICAO de cada aeroporto.

- Na região sudeste: Belo Horizonte - Confins (SBCF), Rio de Janeiro - Galeão (SBGL), e dois aeroportos em São Paulo, Congonhas (SBSP) e Guarulhos (SBGR).
- Na região centro-oeste: Brasília (SBBR); e
- Na região nordeste: Fortaleza (SBFZ) e Salvador (SBSV).

Com essas informações selecionadas, foi possível construir toda uma análise geográfica dos voos, o que é a base da construção de um panorama de uma malha aérea.

### 4.3 Análise geográfica

Com todos os dados em mãos, foi optado por uma análise da distribuição geográfica dos voos bem como a concentração dos voos nos HUBs e nas determinadas regiões do país. Na Tabela 06 é possível observar todos os destinos separados por suas respectivas regiões.

Tabela 6 - Destinos e HUBs da GOL LINHAS AÉREAS

Destinos e HUBs da GOL LINHAS AÉREAS no quarto trimestre de 2019 (4T19)				
Sudeste	Sul	Centro-oeste	Nordeste	Norte
<ul style="list-style-type: none"> <li>• Belo Horizonte (SBCF)</li> <li>• Pres. Prudente (SBDN)</li> <li>• Rio de Janeiro - Galeão (SBGL)</li> <li>• São Paulo - Guarulhos (SBGR)</li> <li>• Campinas (SBKP)</li> <li>• Montes Claros (SBMK)</li> <li>• Vitória (SBVT)</li> <li>• Zona da Mata (SBZM)</li> <li>• Uberlândia (SBUL)</li> <li>• Rio de Janeiro - Santos Dumont (SBRJ)</li> <li>• São Paulo - Congonhas (SBSP)</li> </ul>	<ul style="list-style-type: none"> <li>• Cascavel (SBCA)</li> <li>• Chapecó (SBCH)</li> <li>• Curitiba (SBCT)</li> <li>• Caxias do Sul (SBCX)</li> <li>• Foz do Iguaçu (SBFI)</li> <li>• Florianópolis (SBFL)</li> <li>• Joinville (SBJV)</li> <li>• Londrina (SBLO)</li> <li>• Maringá (SBMG)</li> <li>• Navegantes (SBNF)</li> <li>• Porto Alegre (SBPA)</li> <li>• Passo Fundo (SBPF)</li> </ul>	<ul style="list-style-type: none"> <li>• Brasília (SBBR)</li> <li>• Campo Grande (SBCG)</li> <li>• Caldas Novas (SBCN)</li> <li>• Cuiabá (SBCY)</li> <li>• Goiânia (SBGO)</li> </ul>	<ul style="list-style-type: none"> <li>• Aracaju (SBAR)</li> <li>• Fernando de Noronha (SBFN)</li> <li>• Fortaleza (SBFZ)</li> <li>• Ilhéus (SBIL)</li> <li>• Jericoacoara (SBJE)</li> <li>• João Pessoa (SBJP)</li> <li>• Juazeiro do Norte (SBJU)</li> <li>• Campina Grande (SBKG)</li> <li>• Maceió (SBMO)</li> <li>• São Luiz (SBSL)</li> <li>• Teresina (SBTE)</li> <li>• Recife (SBRF)</li> <li>• Petrolina (SBPL)</li> <li>• Natal (SBSG)</li> <li>• Salvador (SBSV)</li> <li>• Porto Seguro (SBPS)</li> <li>• Vitória da Conquista (SBVC)</li> </ul>	<ul style="list-style-type: none"> <li>• Belém (SBBE)</li> <li>• Boa Vista (SBBV)</li> <li>• Carajás (SBCJ)</li> <li>• Cruzeiro do Sul (SBCZ)</li> <li>• Manaus (SBEG)</li> <li>• Marabá (SBMA)</li> <li>• Macapá (SBMQ)</li> <li>• Palmas (SBPJ)</li> <li>• Rio Branco (SBRB)</li> <li>• Porto Velho (SBPV)</li> <li>• Santarém (SBSN)</li> </ul>
<b>HUBs:</b> <ul style="list-style-type: none"> <li>• Belo Horizonte (SBCF)</li> <li>• Rio de Janeiro - Galeão (SBGL)</li> <li>• São Paulo - Guarulhos (SBGR)</li> <li>• São Paulo - Congonhas (SBSP)</li> </ul>	<b>HUBs:</b>  <p style="text-align: center;"><b>Nenhum</b></p>	<b>HUBs:</b> <ul style="list-style-type: none"> <li>• Brasília (SBBR)</li> </ul>	<b>HUBs:</b> <ul style="list-style-type: none"> <li>• Fortaleza (SBFZ)</li> <li>• Salvador (SBSV)</li> </ul>	<b>HUBs:</b>  <p style="text-align: center;"><b>Nenhum</b></p>

No período analisado, a GOL LINHAS AÉREAS atendia 56 destinos no país, com 5 destinos na região centro-oeste, 11 destinos nas regiões norte e sudeste, 12 destinos da região sul e 17 destinos na região nordeste.

Após a identificação de cada aeroporto, foi feita uma seleção para descobrir o número total de voos para cada região nos meses de outubro, novembro e dezembro. A identificação dos números totais de decolagens por cada região por ser vista na Tabela 7, logo abaixo.

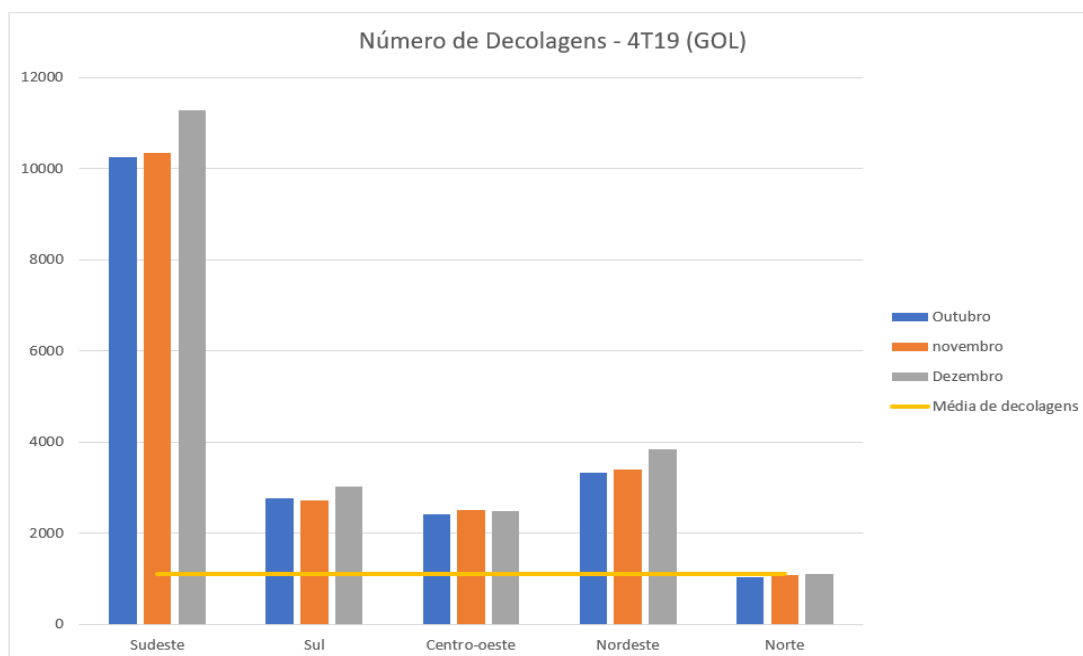
Tabela 7 - Número de decolagens por região

	Sudeste	Sul	Centro-oeste	Nordeste	Norte	Total
Outubro	<b>10.246</b>	<b>2.766</b>	<b>2.416</b>	<b>3.331</b>	<b>1.044</b>	<b>19.803</b>
Novembro	<b>10.349</b>	<b>2.717</b>	<b>2.516</b>	<b>3.395</b>	<b>1.087</b>	<b>20.064</b>
Dezembro	<b>11.287</b>	<b>3.018</b>	<b>2.493</b>	<b>3.840</b>	<b>1.103</b>	<b>21.741</b>
Total	<b>31.882</b>	<b>8.501</b>	<b>7.425</b>	<b>10.566</b>	<b>3.234</b>	<b>61.608</b>

Pode-se observar na Tabela 7 uma grande diferença do número de voos realizados em cada região, como é o exemplo da diferença entre a região sudeste e norte, cuja diferença beira os 10%.

Na Figura 10, logo abaixo, pode-se ver uma representação do número de voos em cada região. Note que a média de voos por destino é de 1.100 (número obtido pela razão entre o número de voos totais pelo número de destinos).

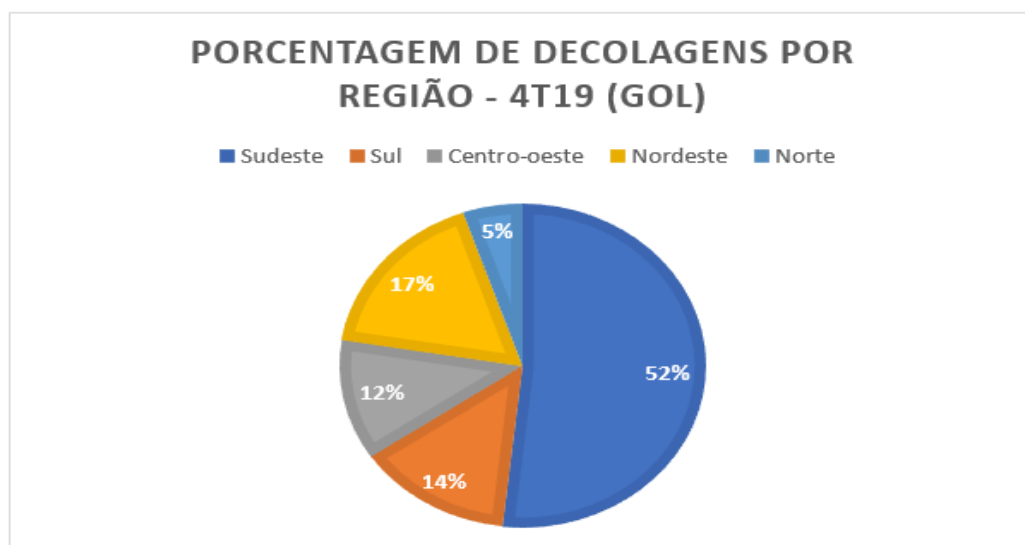
Figura 10 - Representação gráfica do número de decolagens por região



Além disso, com essas informações, foi possível também calcular o peso que cada região possui na malha aérea doméstica da GOL LINHAS AÉREAS. A representação gráfica dos voos por região é exposta na Figura 11.



Figura 11 - Representação gráfica do percentual de decolagens por região



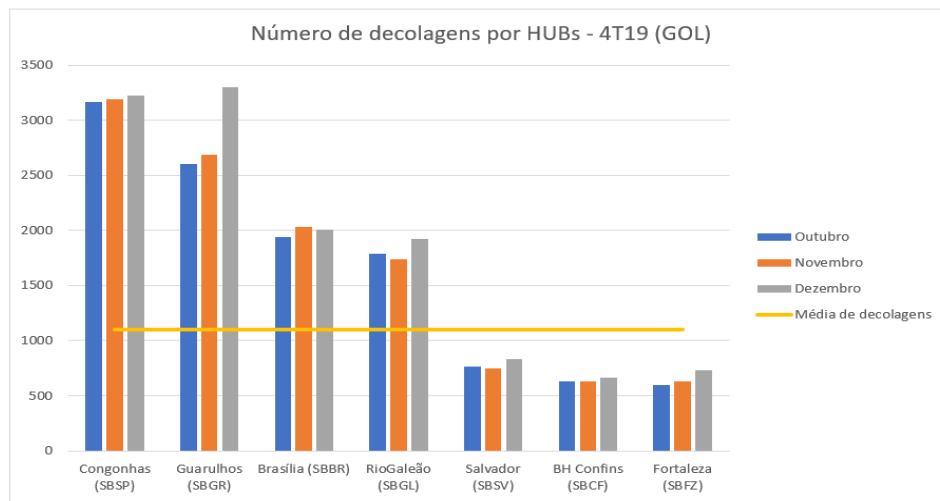
Além dos voos de cada região, foi interessante também identificar o número de voos totais de cada HUB da companhia aérea, pois segundo Corrêa (2013) facilitaria na identificação das principais rotas da companhia aérea para a visualização da malha aérea. Na Tabela 8, é possível ver a identificação de cada hub por número de decolagens.

Tabela 8 - Número de decolagens por HUBs

Hub	Outubro	Novembro	Dezembro
Congonhas (SBSP)	<b>3.168</b>	<b>3.193</b>	<b>3.228</b>
Guarulhos (SBGR)	<b>2.607</b>	<b>2.691</b>	<b>3.302</b>
Brasília (SBBR)	<b>1.940</b>	<b>2.036</b>	<b>2.006</b>
Rio - Galeão (SBGL)	<b>1.789</b>	<b>1.740</b>	<b>1.921</b>
Salvador (SBSV)	<b>764</b>	<b>751</b>	<b>832</b>
BH - Confins (SBCF)	<b>629</b>	<b>634</b>	<b>662</b>
Fortaleza (SBFZ)	<b>597</b>	<b>634</b>	<b>730</b>

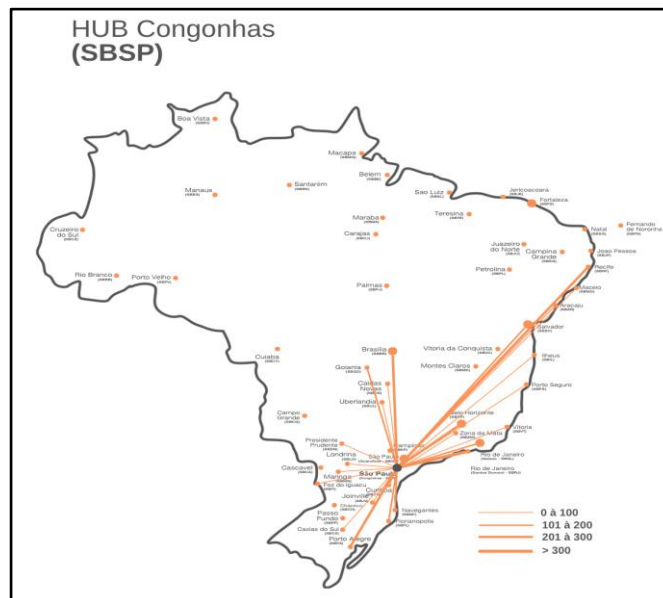
Seguindo nesse prisma, foi elaborado um gráfico com o número de decolagens por HUBs para facilitar a visualização. Na Figura 12 há também a média de 1100 decolagens por destinos.

Figura 12 - Representação gráfica do número de decolagens por HUBs



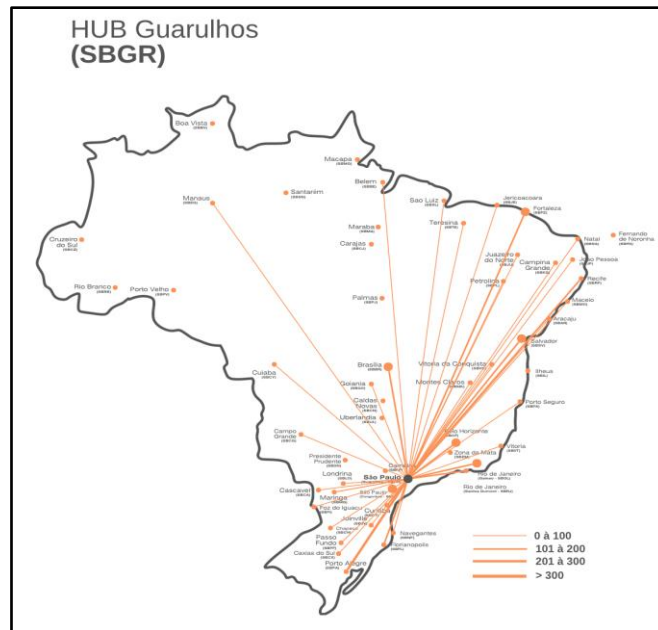
Por fim, com todos os resultados em mãos (dados dos voos, destinos e frequência de voos) foi possível elaborar a construção definitiva da representação em mapa da malha aérea para cada HUB da empresa. Diante dos sete Hubs que a GOL LINHAS AÉREAS opera, foi elaborado um mapa para cada aeroporto, cuja ordem vai do HUB mais movimentado ao HUB menos movimentado.

Figura 13 - Malha aérea do HUB de Congonhas (SBSP)



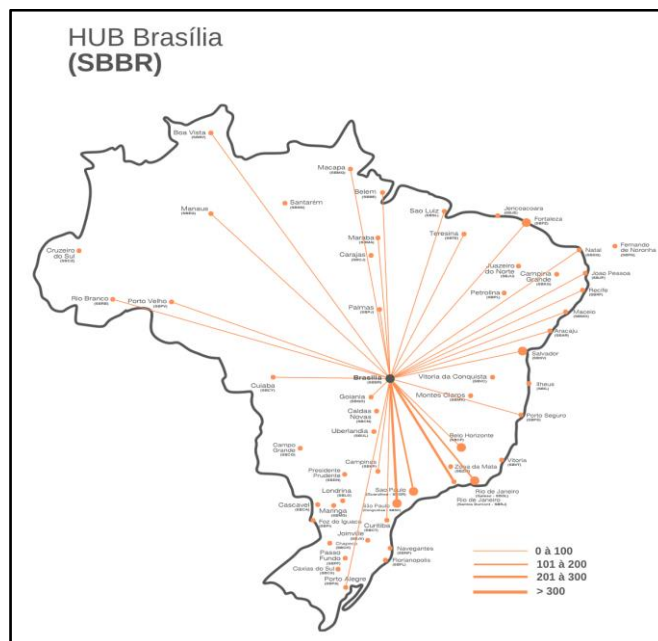
Na Figura 13 é representado o HUB de Congonhas (código ICAO: SBSP), o mais movimentado da empresa, com um total de 9.589 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

Figura 14 - Malha aérea do HUB de Guarulhos (SBGR)



Na Figura 14 é representado o HUB de Guarulhos (código ICAO: SBGR), o segundo mais movimentado da empresa, com um total de 8.600 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

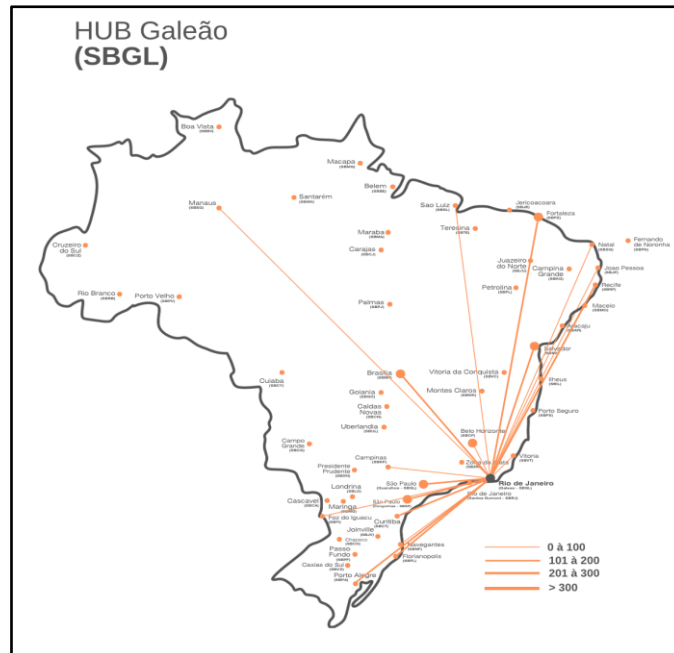
Figura 15 - Malha aérea do HUB de Brasília (SBBR)



Na Figura 15 é representado o HUB de Brasília (código ICAO: SBBR), o terceiro

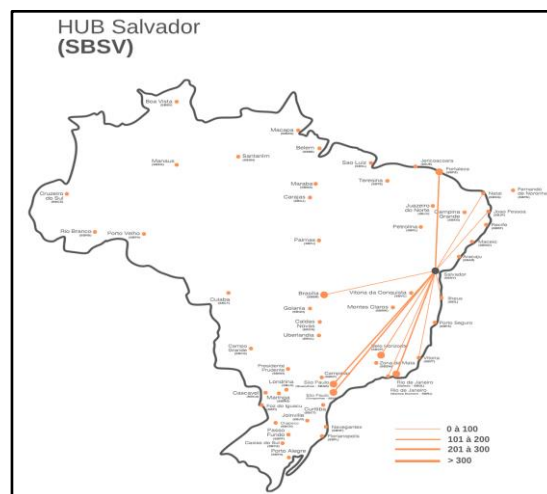
mais movimentado da empresa, com um total de 5.982 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

*Figura 16 - Malha aérea do HUB de Galeão*



Na Figura 16 é representado o HUB do aeroporto do Rio de Janeiro - Galeão (código ICAO: SBGL), o quarto mais movimentado da empresa, com um total de 5.450 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

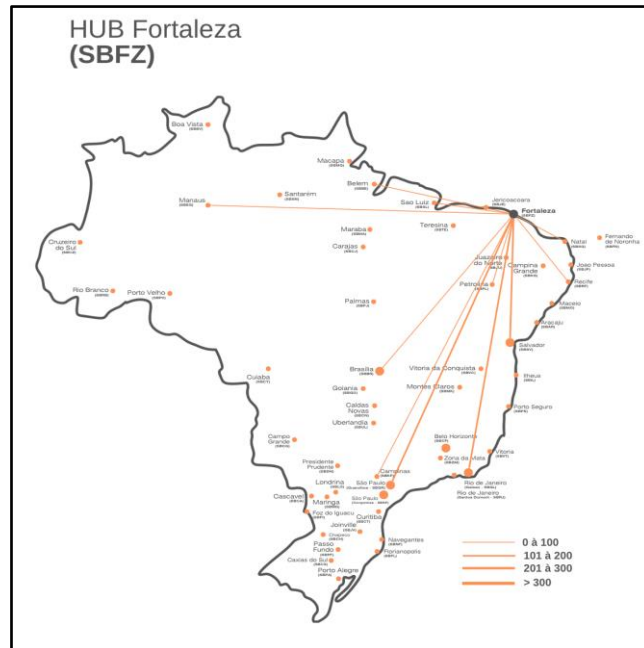
*Figura 17 - Malha aérea do HUB de Salvador (SBSV)*



Na Figura 17 é representado o HUB de Salvador (código ICAO: SBSV), o quinto mais movimentado da empresa, com um total de 2.347 decolagens para o período de

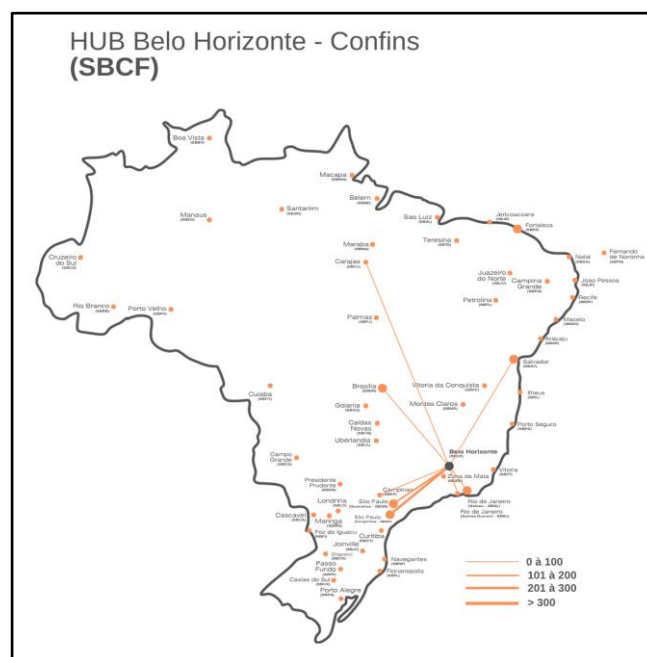
outubro, novembro e dezembro de 2019 (4T19).

*Figura 18 - Malha aérea do HUB de Fortaleza (SBCZ)*



Na Figura 18 é representado o HUB de Fortaleza (código ICAO: SBCF), o sexto mais movimentado da empresa, com um total de 1.961 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

*Figura 19 - Malha aérea do HUB de Belo Horizonte (SBCF)*



Na Figura 19 é representado o HUB de Belo Horizonte - Confins (código ICAO: SBCF), o sétimo e o menos movimentado da empresa, com um total de 1.925 decolagens para o período de outubro, novembro e dezembro de 2019 (4T19).

## 5 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O trabalho executou o seu objetivo com sucesso. De fato, todas as informações colhidas junto à ANAC foram devidamente analisadas e apresentadas de forma clara e intuitiva, tanto em gráficos e números quanto em ilustrações de rotas voadas por HUBs e diferenciando-as por frequências, cumprindo assim o objetivo geral em compor e desenvolver uma ferramenta de extração de dados e, em seguida, apresentá-los de forma simplificada por meio de tabelas, gráficos e figuras. É importante ressaltar que o atingimento do objetivo geral só foi possível por meio do cumprimento dos objetivos específicos de extrair os dados da agência reguladora, de tratá-los por meio de um notebook em python e por fim, analisá-los e ilustrá-los de forma clara e intuitiva.

Em linhas gerais, pode-se afirmar que os 61.608 voos realizados no quarto trimestre de 2019, período esse que compreende os meses de outubro, novembro e dezembro são indicativos claros da importância e da relevância que a GOL LINHAS AÉREAS possui na malha aérea doméstica do país. Pode-se observar também outro importante êxito no trabalho: a aproximação dos resultados obtidos via dados abertos da ANAC em relação aos dados expostos pela companhia aérea, que são extremamente próximos, comprovando uma assertividade nas contabilizações de informações de voos tanto da companhia quanto da agência reguladora do país, a ANAC.

Em adição, pode-se afirmar também que a linguagem Python teve um papel fundamental neste trabalho, pois por meio desta linguagem e de seus programas, foi possível juntar todos os dados extraídos e compilá-los a um nível ideal de análise, nível ideal esse que possibilitou a criação de gráficos e análises consideravelmente intuitivas.

No que tange às limitações do trabalho, pode-se dizer que cada empresa aérea possui uma filosofia de modelo de voos, a filosofia hub and spoke e a filosofia point-to-point, o que pode dificultar as comparações entre empresas e períodos distintos.

Por fim e não menos importante, sugere-se para trabalhos futuros o estudo da malha aérea doméstica em outros períodos de tempo utilizando a mesma dinâmica de análise para maiores comparações. Outra sugestão se refere à correlação da malha aérea com outros dados econômicos e socioeconômicos.

## REFERÊNCIA

AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL (ANAC). Acesso à informação: Institucional. Brasília, 09/2022. Disponível em: <<https://www.gov.br/anac/pt-br/aceso-a-informacao/institucional>> Acesso em 27 de Agosto de 2022.

BAZARGAN, M. **Airline operations and scheduling**. [S.l.]: Ashgate, 2010.

BERGIANTE, N.C.R., Soares de Mello, J.C.C.B., Nunes, M.V.R., Paschoalino F.F. (2011) **Aplicação de uma proposta de medida de centralidade para avaliação de malha aérea de uma empresa do setor de transporte aéreo brasileiro**. Journal of Transport Literature, 5(4), 119-135.

BOAVENTURA NETTO, P.O. (1996) **Grafos: Teoria, Modelos, Algoritmos**. Editora Edgard Blücher Ltda, São Paulo, Brasil. 1996.

BRASIL. **Lei 11.182, de 27 de setembro de 2005**. Brasília: 2005. Disponível em:<[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2005/lei/l11182.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/lei/l11182.htm)>. Acesso em 29 Jan. 2023.

BRASIL. **Decreto nº 5.731, de 20 de março de 2006**. Brasília: 2006. Disponível em:<[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2006/decreto/d5731.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/decreto/d5731.htm)>. Acesso em 29 Jan. 2023.

BRASIL. **Decreto-Lei Nº 32 de 1966**. Brasília: 1966. Disponível em:<[BRASIL. \*\*Instrução normativa número 127, de 5 de outubro de 2018\*\*. Brasília: 2018. Disponível em: <<https://www.anac.gov.br/assuntos/legislacao/legislacao->](https://www2.camara.leg.br/legin/fed/declei/1960-1969/decreto-lei-32-18-novembro-1966-375943-publicacaooriginal-1-pe.html#:~:text=Institui%20o%20C%C3%B3digo%20Brasileiro%20do%20Ar.&text=Art.,ratificado%20e%20pelo%20presente%20C%C3%B3digo.></a>>. Acesso em 29 Jan. 2023.</p></div><div data-bbox=)



1/instrucoes-normativas/2018/instrucao-normativa-no-127-05-10-2018#:~:text=Instru%C3%A7%C3%A3o%20Normativa%20n%C2%BA%20127%2C%20DE,da%20estrutura%20organizacional%20da%20Ag%C3%AAncia.>. Acesso em 29 Jan. 2023.

CASTRO, N.; LAMY, P. **Desregulamentação do Setor de Transporte e SubSetor Transporte Aéreo de Passageiros**. Brasília, p. 36-50, 1993.

CHEN, Daniel Y. **Análise de dados com Python e Pandas**. Novatec Editora, 2018.

COOK, Gerald N.; GOODWIN, Jeremy. Airline networks: **A comparison of hub-and-spoke and point-to-point systems**. Journal of Aviation/Aerospace Education & Research, v. 17, n. 2, p. 1, 2008.

CORRÊA, F. **Planejamento do transporte aéreo**. Palhoça, p.112-122, 2013.

COSTA, T. F. G., LOHMANN, G. e OLIVEIRA, A. V. M. **A model to identify airport hubs and their importance to tourism in Brazil**. Research in Transportation Economics, v. 26, p. 3-11, 2010.

DRESCH, Aline. et al. **Metodologia Científica para Engenharia**. 1a ed. Rio de Janeiro: Elsevier, 2019.

FERREIRA, Rafael G C.; MIRANDA, Leandro B. A D.; PINTO, Rafael A.; et al. **Preparação e Análise Exploratória de Dados**. [Digite o Local da Editora]: Grupo A, 2021. E-book. ISBN 9786556902890. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556902890/>. Acesso em: 19 set. 2022.

GOL LINHAS AÉREAS. **Central de Resultados**. Disponível em: <[https://ri.voegol.com.br/conteudo\\_pt.asp?idioma=0&conta=28&tipo=53858&ano=2019](https://ri.voegol.com.br/conteudo_pt.asp?idioma=0&conta=28&tipo=53858&ano=2019)>. Acesso em: 22 Dez. 2022.

INTERNATIONAL TRANSPORT JOURNAL. **Six new destinations for Air France KLM Martinair Cargo.** Disponível em: <<https://www.transportjournal.com/en/home/news/artikeldetail/six-new-destinations-for-air-france-klm-martinair-cargo.html>>. Acesso em: 19 Dez. 2022

LOPES, F. S. (2005) **Estudo da evolução da estrutura de rotas das empresas aéreas no Brasil.** 74f. Trabalho de Conclusão de Curso de Graduação - Instituto Tecnológico de Aeronáutica, São José dos Campos.

MENDES, M. **Por que é importante investir em infraestrutura?** Brasília, p. 1-3, 2011. VARIG-AIRLINES. **Rotas.** Disponível em: <<https://www.varig-airlines.com/pt/rotas.htm>> . Acesso em: 15 Dez. 2022

## APÊNDICES

### Apêndice A – Notebook utilizado para o manuseio dos dados

#### MONTAGEM DO GOOGLE DRIVE

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

%cd /content/drive/My Drive/Colab Notebooks/TCC/Pedro

!pwd
```

#### INSTALAÇÃO DE PACOTES DO PYTHON

```
import pandas as pd

import time
import os
import sys
from zipfile import ZipFile

# Somente se precisar extrair dados de pdf
#!pip install -q --upgrade camelot-py[cv] tabula-py tabulate

# Somente se precisar extrair dados de pdf
#import tabula as tb
#import tabulate
#import camelot
```

#### IDENTIFICANDO OS ARQUIVOS DE ETA BAIXADO

```
dadosETA = pd.read_csv("./Dados Estatisticos/Dados_Estatisticos.csv",
delimiter=';', decimal = ',', skiprows=1)
```

ANÁLISE EXPLORATÓRIA DOS DADOS DE ETA

```
dadosETA.shape
dadosETA.info()
dadosETA.EMPRESA_SIGLA.unique()
dadosETA.NATUREZA.unique()

dadosETA_Gol = dadosETA[(dadosETA.EMPRESA_SIGLA == 'GLO') &
(dadosETA.NATUREZA == 'DOMÉSTICA')]

dadosETA_Gol_2019 = dadosETA_Gol[(dadosETA_Gol.ANO == 2019)]
```

```
dadosETA_Gol_2019 =
dadosETA_Gol_2019.drop(columns=['EMPRESA_NACIONALIDADE',
'AEROPORTO_DE_ORIGEM_NOME', 'AEROPORTO_DE_ORIGEM_REGIAO',
'AEROPORTO_DE_ORIGEM_PAIS', 'AEROPORTO_DE_ORIGEM_CONTINENTE',
'AEROPORTO_DE_DESTINO_NOME', 'AEROPORTO_DE_DESTINO_REGIAO',
'AEROPORTO_DE_DESTINO_PAIS', 'AEROPORTO_DE_DESTINO_CONTINENTE',
'NATUREZA'])

dadosETA_Gol_2019_4T = dadosETA_Gol_2019[(dadosETA_Gol_2019.MES == 10)
| (dadosETA_Gol_2019.MES == 11) | (dadosETA_Gol_2019.MES == 12)]

dadosETA_Gol_2019_4T.shape
dadosETA_Gol_2019_4T.head()
dadosETA_Gol_2019_4T.MES.value_counts()
dadosETA_Gol_2019_4T.describe()
dadosETA_Gol_2019_4T.groupby(['MES']).sum()
dadosETA_Gol_2019_4T.groupby(['GRUPO_DE_VOO']).sum()
dadosETA_Gol_2019_4T[dadosETA_Gol_2019_4T.GRUPO_DE_VOO ==
'REGULAR'].shape
```

## ANEXOS

ANEXO I – [Lei de Criação da ANAC nº 11.182, de 27 de setembro de 2005](#)

ANEXO II – [Decreto de Instalação da ANAC nº 5.731, de 20 de março de 2006](#)

ANEXO III – [Regimento Interno: Resolução nº 381, de 14 de junho de 2016](#)

ANEXO IV – [Alteração do RI: Resolução nº 581, de 21 de agosto de 2020](#)

ANEXO V – [Instrução Normativa nº 127, de 5 de outubro de 2018](#)