



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLOGIA DA INFORMAÇÃO
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

VICTOR BEZERRA DE MENEZES DOS SANTOS

**MINERAÇÃO E ANÁLISE DOS DADOS DAS OPERAÇÕES NOS AEROPORTOS
BRASILEIROS DURANTE A PANDEMIA DE COVID-19**

ANGICOS

2021

VICTOR BEZERRA DE MENEZES DOS SANTOS

**MINERAÇÃO E ANÁLISE DOS DADOS DAS OPERAÇÕES NOS AEROPORTOS
BRASILEIROS DURANTE A PANDEMIA DE COVID-19**

Trabalho de conclusão de curso apresentado a Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Sistemas de Informação, sob orientação do Prof. Dr. Francisco de Assis P. V. de Arruda.

ANGICOS

2021

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

S237m Santos, Victor Bezerra de Menezes dos .
MINERAÇÃO E ANÁLISE DOS DADOS DAS OPERAÇÕES NOS
AEROPORTOS BRASILEIROS DURANTE A PANDEMIA DE
COVID-19 / Victor Bezerra de Menezes dos Santos.
- 2021.
43 f. : il.

Orientador: Francisco de Assis Pereira
Vasconcelos Arruda.
Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Sistemas de
Informação, 2021.

1. Covid-19. 2. ANAC. 3. INFRAERO. 4. mineração
de dados. 5. visualização de dados. I. de Assis
Pereira Vasconcelos Arruda, Francisco, orient.
II. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo(a) autor(a).

Biblioteca Campus Angicos
Bibliotecário: Helder Romero Maia Duarte
CRB: 15/673

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

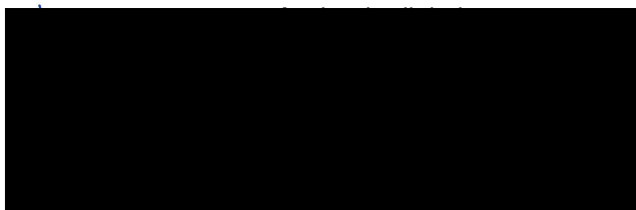
VICTOR BEZERRA DE MENEZES DOS SANTOS

**MINERAÇÃO E ANÁLISE DOS DADOS DAS OPERAÇÕES NOS AEROPORTOS
BRASILEIROS DURANTE A PANDEMIA DE COVID-19**

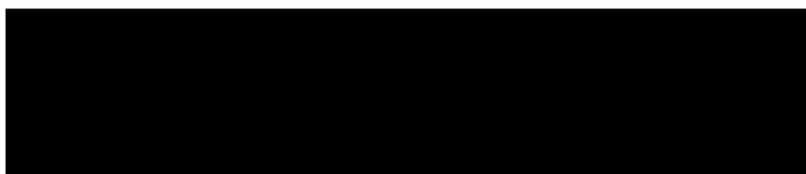
Trabalho de conclusão de curso apresentado a Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Sistemas de Informação, sob orientação do Prof. Dr. Francisco de Assis P. V. de Arruda.

Defendida em: 19/11/2021.

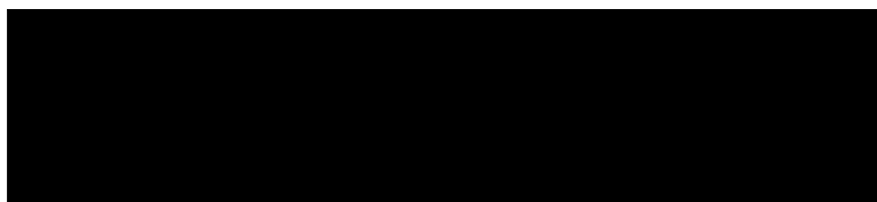
BANCA EXAMINADORA



Prof. Dr. Francisco de Assis Pereira Vasconcelos de Arruda (UFERSA)
Orientador-Presidente



Profa. Sara Guimarães Negreiros (UFERSA)
Membro Examinador



Prof. Me. Marcos Alvarenga Oliveira (UFERSA)
Membro Examinador

RESUMO

O ano de 2020 o mundo foi impactado pela pandemia da covid-19, a doença teve origem na China e espalhou-se pelo mundo tendo como a sua principal forma de transmissão pelas vias aéreas de pessoas infectadas. O avião, um dos principais meios de transporte do mundo e interliga diversas regiões distantes do Brasil e do mundo, é utilizado no mundo todo por meio da aviação comercial. A aviação comercial foi um dos diversos setores atingidos durante o período da pandemia, para combater a pandemia foram adotadas formas de prevenção e uma delas é o isolamento social. O isolamento social, tido como uma das principais formas de prevenção, impactou as companhias aéreas, diminuindo a quantidade de passageiros. Por meio das bases de dados da INFRAERO e ANAC, e também das tecnologias utilizadas como Python, Pandas e Google Colab, o trabalho estuda como a redução de passageiros afetou o número de voos nos aeroportos coordenados de GRU e REC e aeroportos administrados pela INFRAERO, por meio da mineração de dados e também da visualização de dados.

Palavras-chave: Covid-19; ANAC; INFRAERO; mineração de dados; visualização de dados.

ABSTRACT

The year 2020 the world was impacted by the covid-19 pandemic, the disease originated in China and spread around the world having as its main form of transmission through the airways of infected people. The plane, one of the main means of transport in the world and connecting several distant regions of Brazil and the world, is used all over the world by means of commercial aviation. Commercial aviation was one of the several sectors affected during the period of the pandemic, to fight the pandemic, ways of prevention were adopted and one of them is social isolation. Social isolation, seen as one of the main forms of prevention, impacted airlines, reducing the number of passengers. Through the INFRAERO and ANAC databases, and also the technologies used such as Python, Pandas and Google Colab, the work studies how the reduction of passengers affected the number of flights in the coordinated airports of GRU and REC and airports managed by INFRAERO, through data mining as well as data visualization.

Keywords: Covid-19; ANAC; INFRAERO; data mining; data visualization.

LISTA DE FIGURAS

Figura 1: Obtendo o link do arquivo	19
Figura 2: Código que realiza extração dos arquivos por meio do <i>link</i>	20
Figura 3: Página do aeroporto de Guarulhos	21
Figura 4: Dados disponibilizados pela Aena Brasil.....	22
Figura 5: Página da ANAC com os dados disponíveis.....	23
Figura 6: Arquivo .JSON da ANAC.....	23
Figura 7: Arquivo .CSV da ANAC	23
Figura 8: Peça do Arquivo da INFRAERO	25
Figura 9: Imagem do Google Colab	26
Figura 10: Passo a passo para extração dos dados.....	28
Figura 11: Pousos e decolagens nos aeroportos da INFRAERO.....	28
Figura 12: Operações realizadas nos aeroportos analisados no ano de 2020	29
Figura 13: Trecho do código responsável por fazer primeiro tratamento	30
Figura 14: Tabela da ANAC.....	30
Figura 15: Operações no aeroporto de Guarulhos (GRU) em 2020	31
Figura 16: Operações no aeroporto de Recife de 2020	31
Figura 17: Trecho do código que trata o arquivo da INFRAERO.....	32
Figura 18: Arquivo .xlsx da INFRAERO sem tratamentos.....	32
Figura 19: <i>Sheetname</i> do arquivo da INFRAERO.	33
Figura 20: Colunas da base de dados da INFRAERO	33
Figura 21: Trecho do código responsável por selecionar coluna da INFRAERO.....	33
Figura 22: Operações na INFRAERO no ano de 2019	34
Figura 23: Operações na INFRAERO no ano de 2020	34
Figura 24: Operações na INFRAERO no ano de 2021	35
Figura 25: Operações mês a mês na INFRAERO	35
Figura 26: Pousos e decolagens nos anos de 2020 e 2021	36
Figura 27: Trecho do código que seleciona os tipos de serviços.....	36
Figura 28: Casos de COVID-19 entre março de 2020 e julho de 2021	37

LISTA DE TABELAS

Tabela 1: Tabela na página de dados abertos da INFRAERO	24
Tabela 2: Número de operações	29
Tabela 3: Dataframe após a importação dos dados da INFRAERO no Google Colab	32
Tabela 4: Operações nos anos 2019 e 2020.....	37
Tabela 5: Distribuição dos voos em GRU e REC em 2020.....	38
Tabela 6: Melhor e pior mês de 2020 em GRU e REC	38

LISTA DE ABREVIATURAS E SIGLAS

ANAC	Agência Nacional de Aviação Civil
GRU	Aeroporto de Guarulhos
INFRAERO	Empresa Brasileira de Infraestrutura Aeroportuária
REC	Aeroporto de Recife

SUMÁRIO

1. INTRODUÇÃO	10
1.1. Justificativa	11
1.2. Objetivo Geral.....	12
1.3. Objetivos Específicos.....	12
1.4. Organização do Documento.....	12
2. FUNDAMENTAÇÃO TEÓRICA	14
2.2. Aeroporto Coordenado.....	14
2.3. Slots.....	14
2.4. ANAC	15
2.5. INFRAERO.....	15
2.6. Mineração e Visualização de Dados	15
3. TRABALHOS CORRELACIONADOS	16
4. MATERIAIS E MÉTODOS.....	19
4.1. Arquivos da ANAC.....	22
4.2. Arquivos da INFRAERO	24
4.3. Google Colaboratory, Pandas e Matplotlib.....	26
5. Resultados Obtidos.....	27
5.1. Extração e Análise de Dados	27
5.2. Processamento dos Dados da ANAC.....	29
5.3. Processamento dos Dados da INFRAERO	31
5.4. Análise dos Resultados Obtidos.....	36
5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS.....	40
REFERÊNCIAS.....	41

1. INTRODUÇÃO

A pandemia de COVID-19, transmitida pelo vírus SARS-CoV-2, impôs restrições de interações entre os humanos, como principal maneira de proteção, causando impactos em diversos setores da economia de diversos países, inclusive o Brasil. Um dos setores afetados foi a aviação comercial, responsável por integrar regiões do país e também fazer integração internacional.

O setor aéreo ajuda o turismo diminuindo as distâncias, facilita o comércio tanto internacional quanto nacional e gera empregos diretos e indiretos. De acordo com a ICAO (*International Civil Aviation Organization* - Organização Internacional da Aviação Civil), o setor aéreo representou 3,6% do PIB (Produto Interno Bruto) mundial e movimentou US\$ 2,7 trilhões no ano de 2019.

De acordo com a ANAC (Agência Nacional de Aviação Civil), no ano de 2018, a aviação comercial brasileira representou 1,9% do PIB, impactando a economia com R\$ 131 bilhões. A INFRAERO (Empresa Brasileira de Infraestrutura Aeroportuária) registrou 83,9 milhões de passageiros no ano de 2019. Adicionalmente, de acordo com a Associação Brasileira das Empresas Aéreas, aviação comercial brasileira contribuiu com 1,4% do PIB nacional em 2019.

O Ministério da Saúde declarou que a transmissão do vírus causador da COVID-19 já era sustentada em solo brasileiro no mês de março de 2020, e a OMS (Organização Mundial da Saúde) decretou pandemia mundial, impactando as vidas das pessoas. A ANAC divulgou nos meses de abril, maio e junho de 2020 quedas significativas nos números de voos planejados pelas empresas aéreas entre os estados. Comparado o ano de 2019, o mês de junho de 2020 teve uma queda de 85% em relação aos meses de abril e maio de 2020.

Com uma análise em relação ao período que o mundo está passando causada pela pandemia descrita anteriormente, o mundo precisou adequar-se a restrições de interações entre pessoas. Diante do contexto da pandemia e diante das restrições impostas ao setor aéreo, o trabalho apresentado neste documento busca analisar o movimento de aeronaves nos aeroportos a partir de dados disponibilizados pelos órgãos ANAC e INFRAERO.

O capítulo é dividido em quatro seções. A primeira seção é apresentada a justificativa do trabalho. O segundo capítulo é apresentado o objetivo geral e o terceiro os objetivos específicos para a realização do trabalho e a última e quarta seção descreve como o

documento é organizado e os capítulos que serão apresentados no decorrer do trabalho.

1.1. Justificativa

O Governo Federal do Brasil, por meio do portal de dados abertos¹, disponibiliza diversas bases de dados sobre várias áreas e vários órgãos públicos, desde a esfera federal até a municipal, contendo informações que vão desde agrotóxicos registrados, passando por amostras de porções de rocha ou controle de produção de sementes e mudas. Os dados são disponibilizados em vários formatos de arquivo, tais quais .pdf, .json, HTML, .csv ou outros.

O arquivo no formato .csv (comma-separated-values) apresenta-se como o mais comumente utilizado, sendo este um tipo de arquivo que pode ser usado para transferência entre aplicações de software. É um tipo de arquivo em que a sua estrutura de dados tem seus valores separados por vírgulas (dado1,dado2,dado3) ou ponto-e-vírgula (dado1;dado2;dado3).

Esse formato torna os dados de fácil manuseio e é intercambiável, por ser texto simples (plain-text) e sem formatação, e é especialmente adequado para a transferência de informações entre sistemas distintos. Em um arquivo pode haver inúmeras linhas de dados, o que pode ser de difícil extração manual ou de visualização de forma facilitada sem as ferramentas adequadas.

Adicionalmente, viu-se um movimento ascendente na área de mineração, análise e visualização de grandes massas de dados nos anos recentes, incluindo o desenvolvimento de técnicas, ferramentas e aprofundamento de estudos nessas áreas. De acordo com a Qiang (2018), o movimento crescente aconteceu na área de Big Data, para processamento e análise de massas de dados, bem como a disponibilização de bibliotecas de software por parte da Google ou iniciativas como Python+Pandas. Ainda de acordo com Qiang (2018):

A convergência de IA e big data começou no início dos anos 2000, quando o Google e Baidu...passaram a utilizar sistemas de recomendação para propagandas alimentados por IA, e descobriram que os resultados eram ainda melhores que o esperado.

A visualização de dados é uma importante área na análise de dados, pois auxilia na geração de informação para tomada de decisão. A visualização de dados visa resumir, organizar, apresentar e possibilitar que o usuário interprete uma grande quantidade de dados. Os dados utilizados podem ser extraídos, por exemplo, de arquivos .csv, sendo possível a transformação das muitas linhas de dados em uma imagem resumida e objetiva.

¹ Disponível em: <https://dados.gov.br/dataset>

Assim sendo, o trabalho apresentado justifica-se, pois, apesar dos órgãos ANAC e INFRAERO disponibilizarem dados sobre as operações aéreas em aeroportos coordenados por elas em seus sites, tais dados são de difícil extração, visualização e entendimento por uma grande parcela do público geral, visto que são fornecidos em arquivos no formato .csv, .json ou .xlsx, apresentados sem formatação, não agrupados em gráficos ou expostos em tabelas. Os dados apresentados pelos órgãos citados podem conter divergências a depender do ano de obtenção, seja por readequação a novos formatos ou inclusão/supressão de colunas na base de dados.

O trabalho também é possível analisar os casos de COVID-19 e os números de operações de pousos e decolagens nos aeroportos no ano de 2019, ano pré-pandemia, e durante, podendo ser analisado por meio das informações extraídas, se no início os voos podem ter sido um dos vetores da disseminação da doença e o quanto o setor aéreo é afetado pelo aumento ou diminuição do número de casos da doença e as restrições para combater a doença.

1.2. Objetivo Geral

Analisar dados fornecidos pela ANAC e INFRAERO por meio de técnicas de mineração de dados, a fim de verificar, com o agravamento da doença gerando uma pandemia, a movimentação nos aeroportos brasileiros no período pré-pandemia e no período da pandemia até o mês de agosto de 2021.

1.3. Objetivos Específicos

A seção irá tratar quais objetivos específicos serão trabalhados:

- Desenvolver uma pesquisa quantitativa sobre os aeroportos.
- Coletar dados da ANAC e INFRAERO.
- Importar dados usando linguagem Python e Google Colab.
- Pré-processar os dados.
- Realizar análise dos dados obtidos.
- Aplicar técnicas de visualização de dados.

1.4. Organização do Documento

O restante do trabalho é composto por 5 capítulos, organizados da seguinte maneira. No capítulo 2 é apresentada a fundamentação teórica, dividida em duas seções, a primeira a dos aeroportos e a segunda sobre o processo de mineração e visualização de dados. No

capítulo 3 são apresentados trabalhos correlacionados. No capítulo 4 são apresentadas a metodologia e as ferramentas utilizadas. No capítulo 5 são apresentados os resultados obtidos. Por fim, no capítulo 6 são apresentadas as considerações finais e perspectivas de trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo destina-se à apresentação da fundamentação teórica do trabalho proposto, e está dividido em duas seções. A primeira seção destina-se a exploração do conteúdo sobre as definições da área aeroportuária e a definição dos órgãos de controle de tráfego aéreo, mais especificamente a INFRAERO e a ANAC. A segunda seção destina-se a apresentar o conteúdo relacionado a análise de dados.

2.1. Aeroporto Coordenado

De acordo com a resolução² nº 338, de 22/07/2014, no capítulo III, Artº 5, quando houver um comprometimento de seus componentes aeroportuários (pista, pátio ou terminal), seja em determinadas horas, dia da semana ou período do ano, o aeroporto pode ser declarado coordenado.

Portanto, um aeroporto coordenado é uma área aeroportuária onde há necessidade de expansão para atender melhor a uma determinada demanda. Em curto prazo de tempo torna-se impossível a ampliação física de um aeroporto, o que, por consequência, acarreta a saturação das operações da área já construída e em operação. Com a alta demanda, a depender da sazonalidade, torna-se difícil a gerência das operações, tanto por parte das empresas aéreas que recebem *slots* para operarem, quanto dos órgãos de fiscalização e controle.

2.2. Slots

De acordo com a ANAC, *slots* (espaços) são um espaço de tempo estabelecido para uma aeronave realizar uma operação em um aeroporto e utilizar a sua infraestrutura. Então *slot* é uma permissão a uma empresa aérea para que possa utilizar todas as instalações aeroportuária com data e hora predeterminado.

A alocação de *slots* é realizada por meio de conferências realizadas pela *International Air Transportation Association* (IATA), que acontece duas vezes por ano reunindo representantes de 213 aeroportos e companhias aéreas internacionais. No Brasil a regulamentação e processos de alocação de slots é por meio da ANAC e por sua resolução nº 338.

² Resolução nº 338, de 22/07/2014 da ANAC: <https://bit.ly/3jMD1Iw>

2.3. ANAC

A ANAC é um órgão do poder executivo federal vinculada ao Ministério da Infraestrutura, criada em por meio da lei federal nº 11.182 de 27 de setembro de 2005, que regula e fiscaliza a aviação civil, regula a infraestrutura aeronáutica e aeroportuária no Brasil, possui competências de fiscalizar, normatizar, representação institucional e dar orientações para obter de certificação. Antes, essas atribuições eram distribuídas em outros departamentos e eram de responsabilidade do Comando da Aeronáutica vinculada ao Ministério da Defesa.

2.4. INFRAERO

A INFRAERO, empresa pública, criada em 1972 por meio da lei de nº 5.862, que tem objetivo implantar, administrar, operar e explorar industrial e comercialmente a infraestrutura aeroportuária e de apoio a navegação aérea. Slots na aviação, horário estabelecido para uma aeronave realizar uma operação de chegada ou partida em um aeroporto coordenado.

2.5. Mineração e Visualização de Dados

A mineração de dados é um conjunto de técnicas de reconhecimento de padrões, análise estatística, armazenamento e visualização de dados com a finalidade de extrair informações (CABENA, 1998). A mineração de dados tem como objetivo a busca de informações, analisando e explorando dados, fazendo uso de técnicas de computação e estatísticas. É aplicada em diversas áreas onde queira se extrair informações ou observar comportamentos durante períodos, como por exemplo na aviação.

A visualização de dados utiliza dos canais visuais com o objetivo de transformar conjunto de dados em representações visuais para que a compreensão e análise possam ser concluídas de forma eficiente (CHEN, 2015). A visualização de dados tem como objetivo principal, transformar grandes volumes de dados difíceis de serem compreendidos, em uma interface de fácil compreensão.

3. TRABALHOS CORRELACIONADOS

Nesta seção são apresentados trabalhos que se relacionem mais diretamente aos objetivos apresentados neste documento.

O primeiro trabalho, o autor Vinícius Modolo Teixeira (2020) refere-se aos impactos da Covid-19 no setor aéreo, com o título “O IMPACTO DA CRISE DA COVID-19 NO SETOR AERONÁUTICO”, no artigo é explorado o impacto e as perspectivas da crise causada pela COVID-19 no setor aeronáutico. O autor explora os impactos com a implantação das restrições de movimentação, visto que o transporte aéreo é o meio de transporte que interliga as regiões mais distante do mundo e também do Brasil. Em março de 2020, diversos países restringiram a entrada de passageiros vindo de outros países, paralisando voos.

Teixeira (2020) descreve que a paralização dos voos, as companhias aéreas serão afetadas em curto e médio prazo, como por exemplo: redução de frotas, enxugamento da malha aérea, número de passageiros e também a redução do quadro de funcionários.

O segundo trabalho, dos autores Fabio Teixeira, Felipe Massao Higa, Felipe Massao Higa e Rodrigo Moreira Ribeiro (2020), com o título “2020 - COVID 19 Impact on Aircraft Ground Time at Congonhas Airport (CGH)”, é um trabalho que estuda os impactos da pandemia utilizando o aeroporto de Congonhas como referência, que tem como objetivo simular a duração do tempo de solo antes e após a implementação de novos procedimentos num dos principais aeroportos do Brasil.

A pandemia reduziu drasticamente os números de voos em todos os aeroportos, incluindo o aeroporto de Congonhas, que é um aeroporto na cidade de São Paulo, cidade considerado centro comercial do país e também por sua localização devido ao seu acesso aos centros de negócios e também conhecido por ser a principal ponte aérea entre Rio de Janeiro e São Paulo.

De modo geral, qualquer aeronave tem um custo durante toda a operação de transporte, desde o tempo de voo e também a sua permanência em solo. A implementação de novos procedimentos causada pela pandemia afetou as companhias aéreas com o aumento do tempo durante o solo, impactando diretamente o tempo e o financeiro.

Os autores realizaram simulações para avaliar os impactos dos novos procedimentos por meio do novo tempo de que as aeronaves terão para realizarem os processos de solo, e chegaram a conclusão de que o tempo de solo das aeronaves subiria de 30 a 40 minutos para

50 a 60 minutos. Apesar dos acréscimos no tempo, o aeroporto CGH tem grande capacidade de slots e o impacto nas novas regras de operação pode ser mitigado. Vale ressaltar ainda que a linguagem de programação Python e o Jupyter Notebook aparecem como ferramentas chave utilizadas pelos autores, tanto para análise de dados estatísticos como para a visualização de dados por meio de gráficos.

No terceiro trabalho, De Oliveira e Castilho (2020) ressaltam a amplitude e alcance da rede de transporte aéreo, que é ampla e envolve vários setores da economia, abrangendo fluxos de passageiros e de carga. Os autores ressaltam ainda que, globalmente, entre os dias 30 de dezembro de 2019 e 30 de junho de 2020, houve um decréscimo de 86,37% no fluxo aeronáutico. Ainda, de acordo com os autores, a articulação das redes de transportes, atinge lugares diferentes de formas distintas. Tais redes acabaram sendo não apenas vetores do fluxo de pessoas e mercadorias, mas também da COVID-19. A rede de transporte aéreo foi um dos grandes vetores de transmissão da COVID-19, pelo mundo, especialmente nos primeiros meses de 2020, e foi, provavelmente, a rede de transporte mais atingida. A relevância do setor aéreo e o impacto sofrido e relatado por De Oliveira e Castilho (2020) serviram também de inspiração para as análises apresentadas neste trabalho de conclusão de curso.

Já no quarto trabalho, Rojas (2021) ressalta que, devido a pandemia, da Covid-19, houve aumento nos cancelamentos de voos e que, apesar da diminuição do número de passageiros, a autoridade de transporte aéreo definiu uma malha aérea mínima, com a preocupação de que as companhias aéreas sejam utilizadas para conectar estados da federação com pelo menos um voo. O autor ressalta ainda que as autoridades brasileiras de transporte aéreo também permitiram a mudança das configurações das aeronaves, da configuração de transporte de passageiros para transporte de carga. Essa permissão auxiliou as companhias aéreas a alavancar o uso das aeronaves e a melhorar as receitas do período.

Rojas (2021) afirma que a ANAC calculou um prejuízo de R\$ 6,2 bilhões de reais no segundo trimestre do ano de 2020, quando conjugadas as companhias aéreas Azul, Gol e Latam. Tal resultado figura uma margem líquida negativa de 399,6%, contraposto ao lucro líquido de R\$ 191,8 milhões obtido no mesmo período de 2019. O resultado líquido é dito como o de maior impacto negativo trimestral, sendo o pior resultado de toda a série histórica, iniciada em 2015. Evidencia-se aqui vários aspectos a serem levados em consideração ao analisar o setor aéreo, seja o financeiro, seja a quantidade de voos, seja o impacto diferente entre companhias aéreas, seja a tentativa dos órgãos de fiscalização e controle para tentar

ajudar na contenção dos danos durante a pandemia.

Rojas (2021) afirma ainda que, no mês de julho de 2020, a ANAC prorrogou a medida “waiver”. *Waiver* é o abono do cancelamento de slots (horários de chegada e partida em aeroportos coordenados) do cálculo do índice de regularidade para a obtenção de direitos históricos pelas companhias aéreas. Tal abono acena na tentativa de mitigar prejuízos para companhias aéreas. O autor ressalta ainda outros aspectos que podem não estar aparente nas planilhas de operações disponibilizadas pela INFRAERO ou ANAC, que é o aspecto das demissões ou licença dos empregados das companhias aéreas. O autor ressalta várias medidas adotadas pelas empresas, quer seja licença não remunerada voluntária, programa de demissão voluntária, aposentadoria incentivada ou até licenças remuneradas compulsórias.

No trabalho de Rojas (2021) estão elencadas várias recomendações ao operador aéreo, aos passageiros e aos administradores aeroportuários para tentar conter o avanço da pandemia, advindas de recomendações dos órgãos de controle. As medidas elencadas vão desde a desinfecção de voos, imposição do uso de máscaras e o distanciamento social. Outras medidas de intervenções não farmacêuticas são procedimentos de triagem e testagem, adesão ao distanciamento, medidas de higiene e uso de máscara em aeroportos, durante o voo e durante toda a viagem, juntamente com testes pragmáticos pós-voo e rastreamento. Por fim, o autor afirma que “as companhias aéreas experimentarão um retorno gradual e desigual às operações, que requer um esforço logístico sem precedentes. Mal feita, sua estratégia pode custar tanto quanto a própria crise.”

4. MATERIAIS E MÉTODOS

Neste capítulo são apresentados os materiais utilizados e os métodos definidos para a realização do trabalho. Para a realização do trabalho é necessário extrair dados encontrados nos sites da INFRAERO e ANAC. Para extrair os dados é preciso obter o *link* que realiza o *download* dos arquivos, o passo para a obtenção desse link é por meio do *menu* criado com o clique do botão direito do *mouse* e depois clicando sobre a opção “Copiar o endereço do link” como está na Figura 1.

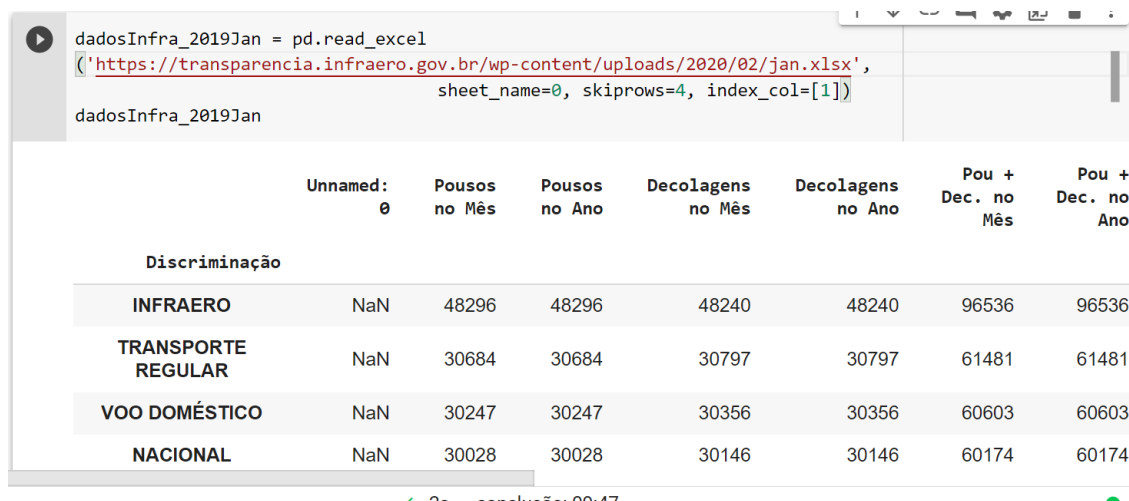
Figura 1: Obtendo o link do arquivo



Fonte: Autoria própria (2021)

Após obter o link do arquivo no site, o próximo para realizar a extração é utilizá-lo na ferramenta Google Colab por meio da linguagem de programação Python e a biblioteca Pandas para que finalmente possa ser extraído os dados dos arquivos, conforme a Figura 2.

Figura 2: Código que realiza extração dos arquivos por meio do link



```
dadosInfra_2019Jan = pd.read_excel(
    'https://transparencia.infraero.gov.br/wp-content/uploads/2020/02/jan.xlsx',
    sheet_name=0, skiprows=4, index_col=[1])
dadosInfra_2019Jan
```

	Unnamed: 0	Pousos no Mês	Pousos no Ano	Decolagens no Mês	Decolagens no Ano	Pou + Dec. no Mês	Pou + Dec. no Ano
Discriminação							
INFRAERO	NaN	48296	48296	48240	48240	96536	96536
TRANSPORTE REGULAR	NaN	30684	30684	30797	30797	61481	61481
VOO DOMÉSTICO	NaN	30247	30247	30356	30356	60603	60603
NACIONAL	NaN	30028	30028	30146	30146	60174	60174

Fonte: Autoria própria (2021)

Há outros tipos de dados que as instituições que administram os aeroportos disponibilizam. Foi constatado que as instituições ANAC e INFRAERO disponibilizam dados dos aeroportos, que estão sob concessão, que já passaram por algum tipo de limpeza ou tratamento. A página que está acessível para o público, do aeroporto de Guarulhos, que contém os links para os dados está apresentada na Figura 3.

Na Figura 3 percebe-se que, para comparar dados entre períodos é difícil, pois o site não permite que o usuário abra abas de dois períodos diferentes, o que dificulta também a visualização. Outro agravante é que o valor destacado em azul não tem escala, ou seja, o usuário não consegue entender se a quantidade de passageiros transportados é grande em comparação com outros meses. Mesmo que tais dados sejam apresentados de forma aparentemente tratada e consolidada, nessa página não há acesso para os dados brutos.

Figura 3: Página do aeroporto de Guarulhos

Resumo de movimentação aeroportuária - RMA:

Abril 2021			^
Tipo	Aeronaves (Unid.)	Passageiros (Unid.)	
Doméstico	7.715	933.968	
Internacionais	2.036	131.047	
Total	9.751	1.065.015	
Março 2021			✓
Fevereiro 2021			✓
Janeiro 2021			✓

Fonte: GRU Airport³, 2021

Outra forma do público ter acesso aos dados é por meio dos microdados de movimentação aeroportuária disponibilizados em arquivos *plain-text* ou minimamente formatados, tais quais *.txt*, *.xlsx*, *.csv*. Esses arquivos podem ser difícil acesso, leitura e compreensão para o usuário comum e até para usuário experientes, e pode não ser possível extrair informações destes dados facilmente. O aeroporto internacional do Recife, que é administrado pela empresa Aena Brasil, é um exemplo que a empresa disponibiliza dados sobre as operações aéreas em formato *.txt*.

Na Figura 4 a base sobre as operações está no formato *.txt*, os dados fornecidos são de difícil entendimento e necessitam de um processamento mínimo e do auxílio do dicionário de dados para entendê-los, já que os dados não contêm informação sobre o que significa cada linha.

³ Dados da instituição GRU Airport: <https://www.gru.com.br/pt/institucional/informacoes-operacionais>

slots alocados. Os dados são fornecidos tanto no formato *.csv* e *.json*. Na Figura 5 apresenta-se uma tela do site da instituição de onde os arquivos podem ser obtidos.

Figura 5: Página da ANAC com os dados disponíveis

Index of /dadosabertos/Voos e operações aéreas/Slots Alocados/2020/REC

[LIVE_REC_W20-SD.csv](#)
[LIVE_REC_W20-SD.json](#)

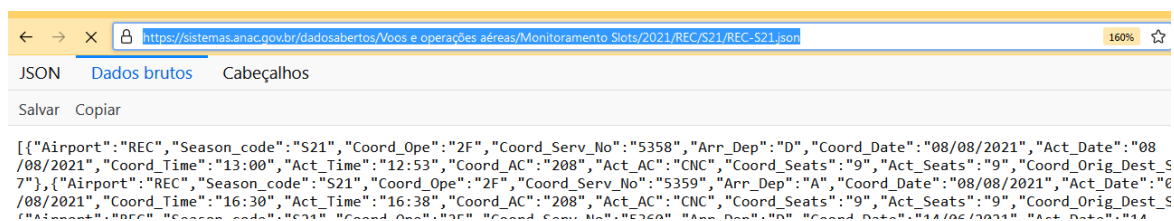
16-Mar-2021 16:32
16-Mar-2021 16:32

2232451
9181100

Fonte: ANAC (sistemas.anac.gov.br/dadosabertos), 2021

Percebe-se que os arquivos estão num formato *.csv* e *.json*. Entretanto, tais arquivos só estão disponíveis após exploração do site da ANAC, e os dados não aparecem de forma direta em *links* de mecanismos de busca como Google. Ademais, ao abrir quaisquer desses arquivos, a visualização apresenta-se excessivamente extensa e confusa para um usuário leigo, conforme pode ser visto na Figura 6 e Figura 7.

Figura 6: Arquivo .JSON da ANAC



Fonte: ANAC (sistemas.anac.gov.br/dadosabertos), 2021

No arquivo *.json*, visualmente a base está numa estrutura de um *array*, no *.csv* cada linha significa um dado, no *.json* os dados são separados por chaves {...}. Para uma pessoa sem conhecimento de programação ou para uma pessoa sem as ferramentas necessárias para o tratamento de dados, a extração de dados e informação pode demandar muito tempo.

Figura 7: Arquivo .CSV da ANAC

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Atualizado em:	2021-06-08													
2	TipodeOperacao	CodEmpresaAere	NumerodoVo	HorariodoVoo	DatadoVoo	DiadaSemana	Equipamento	Assento	Aeroporto	Aeroporto	Tipodeser	Terminal	PaísdeOri	Paísdaetapaanterior	
3	A	0L	101	13:00	06/08/2020	4	735	120	UIO	UIO	C	INT2-3	EC	EC	
4	A	0L	101	16:00	09/05/2020	6	735	120	UIO	UIO	C	INT2-3	EC	EC	
5	A	0L	101	17:40	07/07/2020	2	735	120	UIO	UIO	C	INT2-3	EC	EC	
6	A	2F	5500	02:30	04/05/2020	1	208	0	QDV	QDV	P	NULL	BR	BR	
7	A	2F	5502	23:50	13/04/2020	1	208	0	VIX	VIX	F	CARGO-DI	BR	BR	
8	A	2F	5502	23:50	15/04/2020	3	208	0	VIX	VIX	F	CARGO-DI	BR	BR	

Fonte: ANAC (sistemas.anac.gov.br/dadosabertos), 2021

A quantidade de dados dispostos no arquivo fornecido dificulta a extração de algum

tipo de informação. As informações que podem ser obtidas por meio de uma análise das colunas e com o auxílio do dicionário de dados⁵, e sobre um voo específico, sendo difícil extrair uma informação sobre o conjunto total dos dados. Neste trabalho foram analisados 6 arquivos da ANAC, os arquivos possuíam mais de 12000 linhas, onde cada linha representou uma operação realizada por uma aeronave.

Para a realização do trabalho foi utilizado a coluna “TipodeOperacao” destacada na Figura 7, que, de acordo com o dicionário de dados, a coluna destaca sobre o tipo de operação realizada por aquele voo, se é decolagem ou pouso. Os dados são referentes aos aeroportos REC (Recife) e GRU (Guarulhos), encontrados na base de dados da ANAC e não são encontrados na base de dados da INFRAERO, pois os aeroportos estão sob concessão. Nesse contexto, para cada código de empresa foi realizada uma busca no site da IATA. Essa busca mostrou-se muito exaustiva pelos resultados imprecisos no site da IATA, pois a descrição do site não estava atualizada, a referência para o hyperlink estava quebrada ou não estava de fácil acesso na busca.

4.2. Arquivos da INFRAERO

No site da INFRAERO os dados estão dispostos de acordo com a Tabela 1 onde estão separados por mês e ano. Os dados estão disponíveis em fontes de dados no formato .xlsx conforme a Figura 8, com o dicionário de dados localizado em separado⁶.

Tabela 1: Tabela na página de dados abertos da INFRAERO

21	Jan	Fev										
20	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
19	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
18	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
17	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
16	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
15	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
14	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez

Fonte: INFRAERO (transparencia.infraero.gov.br/estatisticas), 2021

Ao abrir o arquivo, os dados estão apresentados no arquivo tipo .xlsx e formatados

⁵ Dicionário de dados da ANAC: bit.ly/3xjC4LV

⁶ Dicionário de dados da INFRAERO: transparencia.infraero.gov.br/estatisticas/

como visto na Figura 8. Os arquivos contêm vários elementos, incluindo imagens, filtros, créditos e diversas informações sobre os dados, mas são apresentados isoladamente. Destaca-se, por exemplo, o número 1, que representa os tipos de voos que foram realizados nas unidades que são representadas nas linhas de cores azul como a INFRAERO e SBBE que estão na Figura 8, havendo outras unidades no arquivo não sendo mostrados. O número 2 é a coluna “Discriminação” que contém a unidade aeroportuária e os tipos de voos praticados nas unidades.

Figura 8: Pedaço do Arquivo da INFRAERO

Discriminação	Pousos no Mês	Pousos no Ano	Decolagens no Mês	Decolagens no Ano	Pou + Dec. no Mês	Pou + Dec. no Ano
INFRAERO	34.370	34.370	34.248	34.248	68.618	68.618
TRANSPORTE REGULAR	14.863	14.863	14.863	14.863	29.726	29.726
VOO DOMÉSTICO	14.775	14.775	14.771	14.771	29.546	29.546
NACIONAL	14.577	14.577	14.573	14.573	29.150	29.150
REGIONAL	198	198	198	198	396	396
VOO INTERNACIONAL	88	88	92	92	180	180
TRANSPORTE NÃO REGULAR	19.507	19.507	19.385	19.385	38.892	38.892
VOO DOMÉSTICO	19.366	19.366	19.251	19.251	38.617	38.617
VOO INTERNACIONAL	141	141	134	134	275	275
SBBE	1.637	1.637	1.629	1.629	3.266	3.266
TRANSPORTE REGULAR	1.046	1.046	1.042	1.042	2.088	2.088
VOO DOMÉSTICO	1.045	1.045	1.041	1.041	2.086	2.086
NACIONAL	1.045	1.045	1.041	1.041	2.086	2.086

Fonte: <https://transparencia.infraero.gov.br/estatisticas/>, 2021

Já o número 3, Pousos no mês, destaca a coluna que é o número de pousos de aeronaves que aconteceram no aeroporto no mês. No número 4, Pousos no Ano, é uma coluna que é a somatória de todos os pousos dos meses anteriores do respectivo ano, o valor da coluna “Pousos no Ano” é igual a coluna “Pousos no Mês” pois o arquivo é do mês de janeiro e consequentemente não havendo meses anteriores para o somatório.

O número 5, Decolagens no Mês, mostra a soma todas as decolagens de aeronaves realizadas no respectivo mês e o número 6, Decolagens no Ano, as decolagens durante todo o ano. Os números 7 e 8, Pou+Dec. no Mês e Pou+Dec no Ano, são as somatórias dos pousos e decolagens, o 7 a soma do mês (soma das colunas 3 & 5) e o 8 a soma do ano (soma das colunas 4 & 6).

Percebe-se que o dado apresentado pela INFRAERO não se apresenta correlato aos dados apresentados pela ANAC, isso dificultou bastante a extração e visualização de dados de bases de dados tão distintas. Percebe-se ainda que alguns dados da INFRAERO são derivados (uma coluna é somatória de outras duas), apresentam-se formatados com cores, não formam um banco de microdados e não estão em *plain-text*. Para a realização do estudo foi utilizado somente a coluna “Pou. + Dec. no Mês”, que é representada pelo número 7, e também da

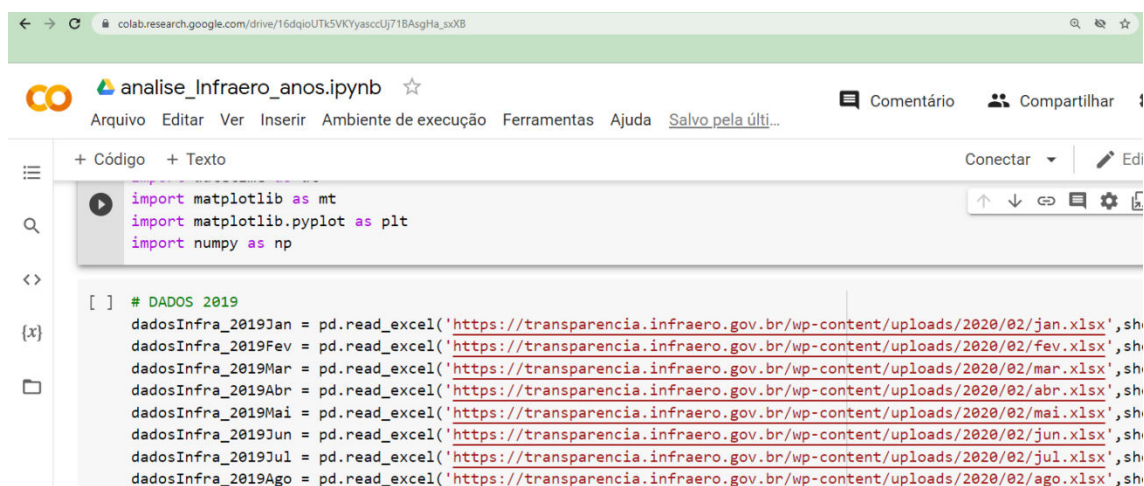
unidade INFRAERO, pois o valor já é a soma de pousos e decolagens de todas zonas aeroportuárias administradas pela INFRAERO.

4.3. Google Colaboratory, Pandas e Matplotlib

Para a codificação foi utilizada a linguagem de programação Python e o ambiente Google Colaboratory, também conhecido como Colab, como é apresentado na Figura 9. Tal ambiente é uma ferramenta desenvolvida pela Google que permite escrever e executar códigos na linguagem Python utilizando somente um navegador de internet, não necessita de instalação de ambientes (como Anaconda) ou de configurações de máquina local. O Colab ainda facilita o compartilhamento de código entre múltiplos usuários.

Encara-se o Colab como uma versão no navegador dos *notebooks* usados pelo Jupyter (jupyter.org), e os arquivos com extensão padrão *.ipynb* são intercambiáveis entre o Jupyter e o Google Colab. Tais ferramentas são bastante utilizadas para estudos em ciências de dados. Para a manipulação de dados foi utilizada a biblioteca Pandas, e para a confecção, visualização e exibição de dos gráficos e dados foi utilizada majoritariamente a biblioteca Matplotlib.

Figura 9: Tecnologia do Google, Google Colab



Fonte: Autoria própria (2021)

5. RESULTADOS OBTIDOS

Neste capítulo são apresentados os resultados obtidos durante a realização deste trabalho. Na seção 5.1 são apresentados os resultados do que ocorreu com as zonas aeroportuárias no início da pandemia até a finalização deste trabalho. Na seção 5.2 é apresentada uma análise geral dos resultados obtidos, congregando informações sobre a ANAC e a INFRAERO conforme proposto nos objetivos.

Nesta seção, os dados foram obtidos diretamente dos *sites* dos órgãos ANAC e INFRAERO. Para a ANAC foram analisados o aeroporto internacional de Recife (REC) e o aeroporto internacional de Guarulhos (GRU). Já os dados fornecidos pela INFRAERO contemplam todos os aeroportos administrados pela empresa.

5.1. Extração e Análise de Dados

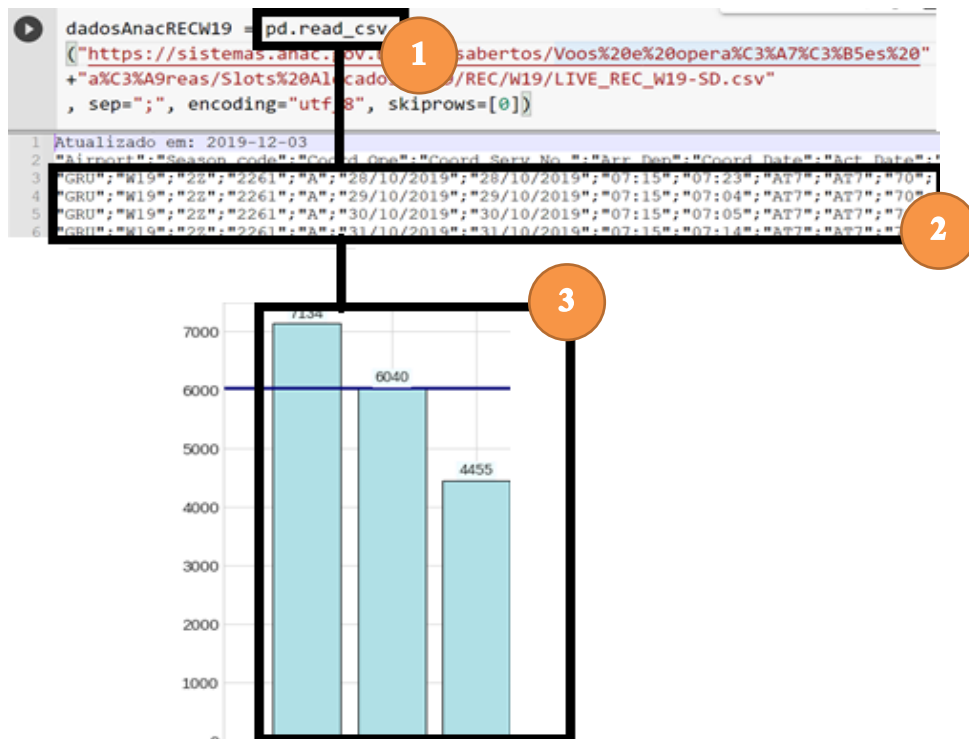
Na Figura 10 resumem-se 3 passos chave que foram seguidos para a extração dos dados dos arquivos disponibilizados pela ANAC e pela INFRAERO:

1 – O primeiro passo é por meio da função “`read_csv`” da biblioteca *pandas*, é possível obter os arquivos das instituições. A função necessita de argumentos para determinar qual é o separador dos dados, o tipo de encoding e o número de linhas que precisam ser ignoradas.

2 – O segundo passo, após a extração dos dados dos arquivos, é somar as linhas que irão gerar os dados sobre os movimentos nos aeroportos.

3 – O terceiro passo é gerar o gráfico somando as linhas que representam as operações de pousos e decolagens nos aeroportos.

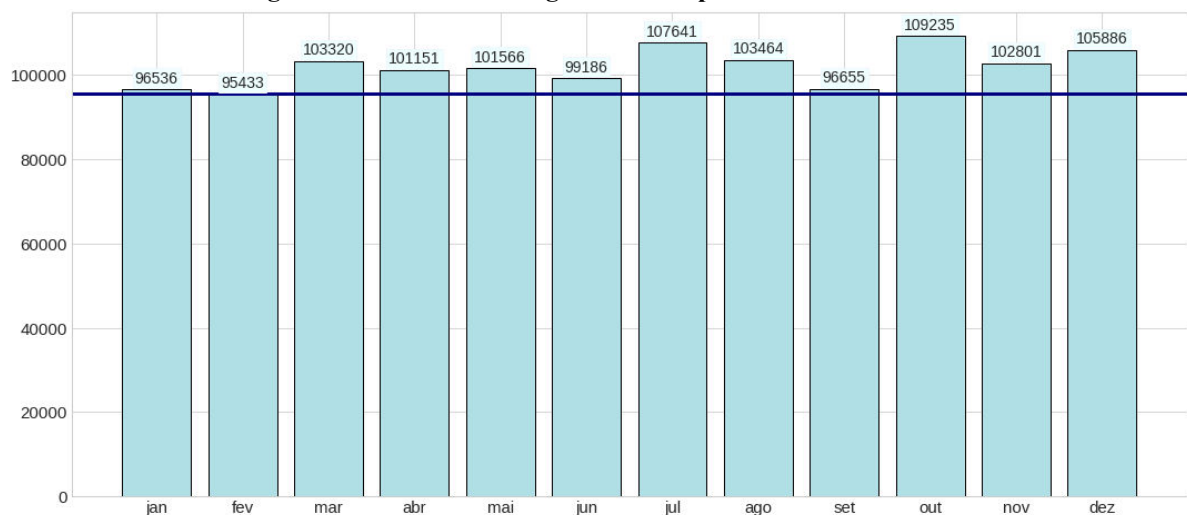
Figura 10: Passo a passo para extração dos dados



Fonte: Autoria própria (2021)

Em 2019, um ano antes da pandemia, não foi possível analisar dados em REC e GRU, pois a base de dados apresentou-se incompleta. A INFRAERO disponibiliza os dados conforme a Figura 11, onde são mostrados os pousos e decolagens ocorridos durante o ano de 2019. A linha azul representa a média das operações no período.

Figura 11: Pousos e decolagens nos aeroportos da INFRAERO

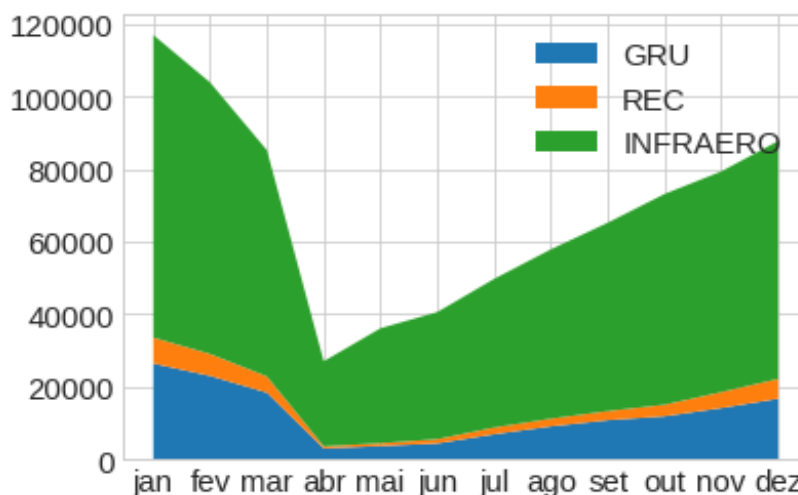


Fonte: Autoria própria (2021)

Para o ano de 2020 é possível extrair dados de REC e GRU, sendo possível verificar

que nos meses de fevereiro e março houve em redução no número de operações, no mês de abril foi o mês com o menor número de operações dos aeroportos, seguindo um aumento gradual das operações a partir de maio. Na Figura 12 é mostrado como estava antes e após a declaração de pandemia pela OMS.

Figura 12: Operações, pouso e decolagem, realizadas nos aeroportos analisados no ano de 2020



Fonte: Autoria própria (2021)

As escolhas das unidades aeroportuárias foram feitas através de pesquisas relacionadas ao movimento dos aeroportos no Brasil, de acordo com a revista AeroFlap⁷ em uma matéria publicada em 1 de maio de 2021, o aeroporto de Guarulhos está em primeiro lugar, como o mais movimentado do Brasil, enquanto na região Nordeste o destaque é para o aeroporto de Recife. Quanto ao número de operações, por meio da análise das bases de dados verifica-se na Tabela 2:

Tabela 2: Número de operações

	Março de 2020	Abril de 2020
ANAC/GRU	18416	3096
ANAC/REC	4455	601
INFRAERO	62271	23398

Fonte: Elaborada pelo autor com base na ANAC e INFRAERO, 2021

5.2. Processamento dos Dados da ANAC

Para a obtenção dos resultados, primeiro foi preciso realizar tratamento no arquivo .csv disponibilizado pela ANAC, pois existem linhas no arquivo que são desnecessárias. Na Figura

⁷ AeroFlap: <https://www.aeroflap.com.br/os-10-aeroportos-mais-movimentados-do-brasil/>

13 e Figura 14 há correspondência entre as cores para melhor entendimento.

Figura 13: Trecho do código responsável por fazer primeiro tratamento

```
dadosAnacRECW19 = pd.read_csv  
("https://sistemas.anac.gov.br/dadosabertos/Voos%20e%20opera%C3%A7%C3%B5es%20"  
+ "a%C3%A9reas/Slots%20Alocados/2019/REC/W19/LIVE_REC_W19-SD.csv"  
, sep=";", encoding="utf_8", skiprows=[0])
```

Fonte: Autoria própria (2021)

Na Figura 14, o trecho destacado representa a linha que precisa ser retirada, que é a linha de atualização do arquivo, pois ela não é importante para este trabalho. O argumento na função, *skiprows*, na Figura 13 é responsável por retirar a linha. O argumento no código, *sep* na Figura 13, é responsável por atribuir o caractere “;” (ponto-e-vírgula) como separador de coluna. O argumento, *encoding* na Figura 13, é responsável por forçar a codificação dos caracteres.

Figura 14: Tabela da ANAC

```
1 Atualizado em: 2019-12-03  
2 "Airport";"Season code";"Coord Ope";"Coord Serv No "; "Arr Dep";"Coord Date";"Act Date";"Coord Time";"Act_Tin"  
3 "GRU";"W19";"2Z";"2261";"A";"28/10/2019";"28/10/2019";"07:15";"07:23";"AT7";"AT7";"70";"68";"RAO";"8"  
4 "GRU";"W19";"2Z";"2261";"A";"29/10/2019";"29/10/2019";"07:15";"07:04";"AT7";"AT7";"70";"68";"RAO";"11"  
5 "GRU";"W19";"2Z";"2261";"A";"30/10/2019";"30/10/2019";"07:15";"07:05";"AT7";"AT7";"70";"68";"RAO";"10"  
6 "GRU";"W19";"2Z";"2261";"A";"31/10/2019";"31/10/2019";"07:15";"07:14";"AT7";"AT7";"70";"68";"RAO";"1"
```

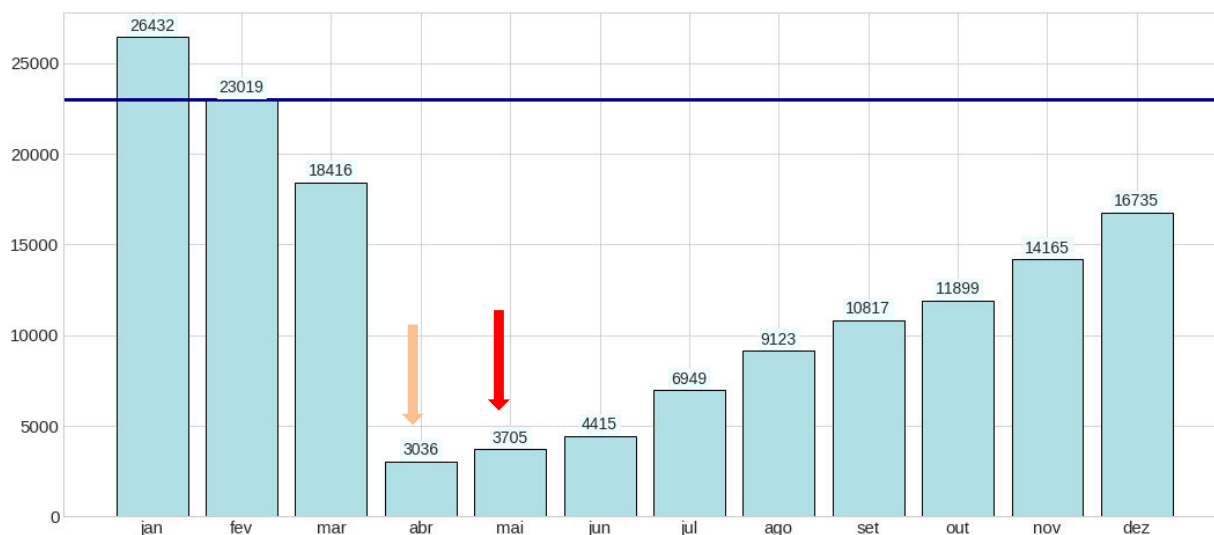
Fonte: Autoria própria (2021)

O *encoding* é necessário para que o programa consiga ler (input) corretamente que tipo de caractere está naquela posição, o idioma português possui caracteres que não existem no idioma inglês como é o exemplo do “ç”, que são considerados como caracteres especiais assim como caracteres acentuados (é, ã) presentes em nosso idioma e nos de origem latina. Após pesquisa na documentação⁹ do Python, foi possível analisar qual é o argumento utilizado na variável *encoding* para essa tradução, que é o “utf_8” ilustrado na Figura 13.

O primeiro resultado consta dos dados obtidos da ANAC sobre as operações de pouso e decolagem no aeroporto de Guarulhos no ano de 2020, antes da declaração da OMS da pandemia, não tendo sido a obtenção dos dados do ano de 2019 de forma completa. No ano de 2020, a partir do mês de abril indicado com setas na Figura 15, houve uma redução significativa nos números de operações em relação ao mês anterior.

⁹ Documentação Python: <https://docs.python.org/3/library/codecs.html>

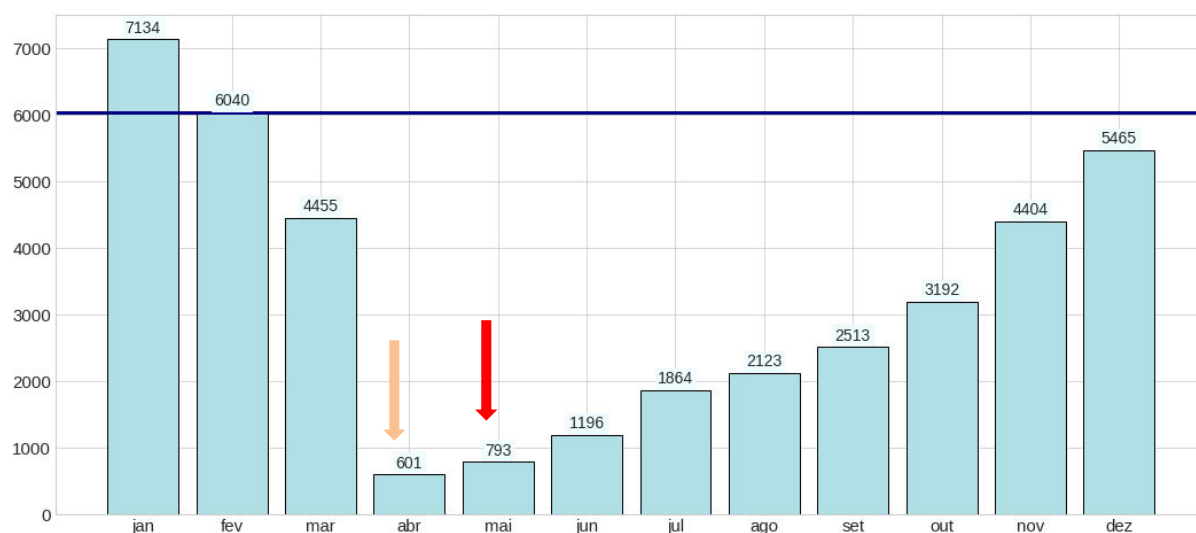
Figura 15: Operações no aeroporto de Guarulhos (GRU) em 2020



Fonte: Autoria própria (2021)

De forma análoga ao que aconteceu no aeroporto de Guarulhos, no aeroporto de Recife no mesmo período também se observa que a partir do mês de abril houve queda no número de operações, conforme indicado na Figura 16.

Figura 16: Operações no aeroporto de Recife de 2020



Fonte: Autoria própria (2021)

5.3. Processamento dos Dados da INFRAERO

Para a obtenção dos resultados na base de dados da INFRAERO, foi preciso realizar tratamento no arquivo *.xls*. Na Figura 17 está o trecho do código responsável por este tratamento, com destaque em cores diferentes.

Figura 17: Trecho do código que trata o arquivo da INFRAERO

```
dadosInfra_2020_Jan = pd.read_excel('https://transparencia.infraero.gov.br/wp-content/uploads/2021/02/jan.xlsx',
sheet_name=0, skiprows=4, index_col=[1])
dadosInfra_2020_Jan
```

Fonte: Autoria própria

Na Figura 17 o destaque no argumento, *skiprows*, é responsável por desconsiderar as 4 primeiras linhas do arquivo, destacado na Figura 18, que não é importante para o trabalho e dificulta a análise.

Figura 18: Arquivo .xlsl da INFRAERO sem tratamentos

 SUPERINTENDÊNCIA DE GESTÃO DA OPERAÇÃO - DOGP Movimento de Aeronaves da REDE INFRAERO em janeiro de 2021 Planilha elaborada em: 18/02/2021							
Discriminação	Pousos no mês	Pousos no Ano	Decolagens no mês	Decolagens no Ano	Pou + Dec. no mês	Pou + Dec. no Ano	
INFRAERO	34.370	34.370	34.248	34.248	68.618	68.618	
TRANSPORTE REGULAR	14.863	14.863	14.863	14.863	29.726	29.726	
VOO DOMÉSTICO	14.775	14.775	14.771	14.771	29.546	29.546	
NACIONAL	14.577	14.577	14.573	14.573	29.150	29.150	
REGIONAL	198	198	198	198	396	396	
VOO INTERNACIONAL	88	88	92	92	180	180	
TRANSPORTE NÃO REGULAR	19.507	19.507	19.385	19.385	38.892	38.892	
VOO DOMÉSTICO	19.366	19.366	19.251	19.251	38.617	38.617	
VOO INTERNACIONAL	141	141	134	134	275	275	
SBBE	1.637	1.637	1.629	1.629	3.266	3.266	

Fonte: Autoria própria (2021)

Durante a análise dos dados com o Google Colab foi observado que a tabela possuía uma coluna sem nome e valores que não apareciam (não estavam visíveis) no arquivo .xlsl. Essa coluna fica visível após a importação usando a biblioteca Pandas, destacada na Tabela 3. **Erro! Fonte de referência não encontrada..** O destaque no “*indexcol*”, o argumento é responsável por desconsiderar essa coluna e por transformar a coluna “Discriminação” na coluna de índice/chave primária do datasource.

Tabela 3: Dataframe após a importação dos dados da INFRAERO no Google Colab

	Unnamed: 0	Discriminação	Pousos no Mês	Pousos no Ano	Decolagens no Mês	Decolagens no Ano	Pou + Dec. no Mês	Pou + Dec. no Ano
0	NaN	INFRAERO	41685	41685	41609	41609	83294	83294
1	NaN	TRANSPORTE REGULAR	25031	25031	25118	25118	50149	50149
2	NaN	VOO DOMÉSTICO	24736	24736	24818	24818	49554	49554
3	NaN	NACIONAL	24515	24515	24606	24606	49121	49121
4	NaN	REGIONAL	221	221	212	212	433	433

Fonte: Autoria própria (2021)

O arquivo da INFRAERO possui quatro *sheetnames* (ou pastas, quando aberto no Microsoft Excel) como está destacado na Figura 19. Na Figura 17 o trecho destacado com o

argumento *sheetname*, é responsável por selecionar qual pasta deve ser utilizada na importação e na análise. Codificando *sheet_name=0*, há a seleção da primeira pasta, que é Aeronaves.

Figura 19: Sheetname do arquivo da INFRAERO.

VOO DOMÉSTICO	11	11
VOO INTERNACIONAL	0	0
Aeronaves	Carga Aérea	Correios
Passageiros		

Fonte: Autoria própria (2021)

Para extração e estudo dos dados da INFRAERO foi preciso selecionar linhas e colunas específicas, pois na base dados há colunas e linhas que não são interessantes para as análises, visto que há colunas que já são o somatório das duas operações de pouso e decolagens, há somatórios do mês ou do ano, e para o trabalho as análises foram feitas mês a mês, sendo selecionado a coluna “Pou+Dec.no Mês” da Figura 20.

Figura 20: Colunas da base de dados da INFRAERO

Planilha elaborada em: 18/02/2021						
	Pousos no Mês	Pousos no Ano	Decolagens no Mês	Decolagens no Ano	Pou + Dec. no Mês	Pou + Dec. no Ano
	34.378	34.378	34.248	34.248	68.648	68.648

Fonte: Autoria própria (2021)

O código que realiza a seleção da coluna está na Figura 21. Neste código há a seleção da coluna necessária para a realização do trabalho e seleciona parte da tabela que corresponde a somatória dos pousos e decolagens do mês correspondente.

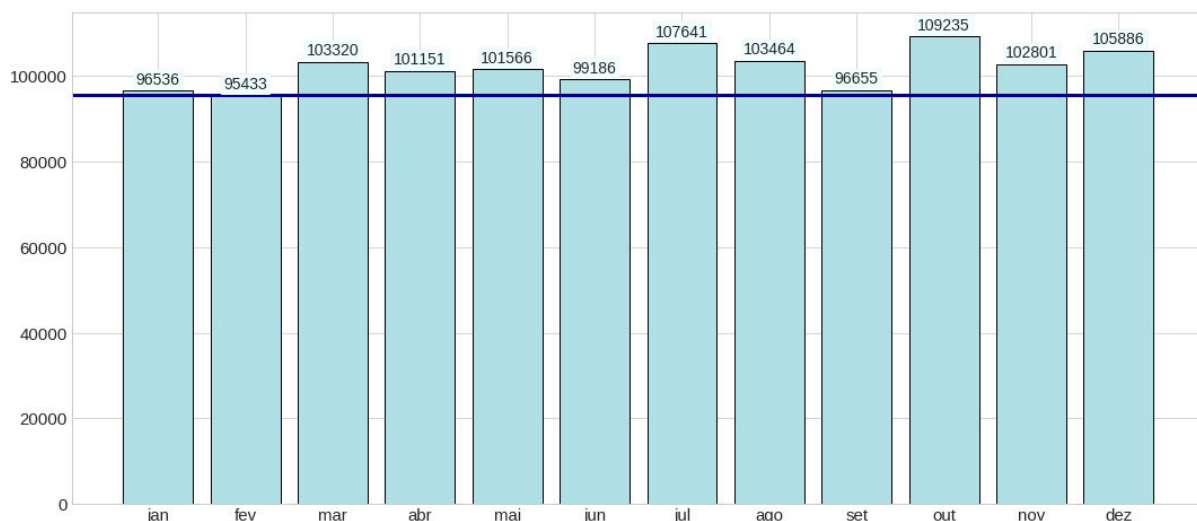
Figura 21: Trecho do código responsável por selecionar coluna da INFRAERO

```
colunasSelecionados = ['Pou + Dec. no Mês']
dadosInfra_2020_Jan = dadosInfra_2020_Jan.filter(items=colunasSelecionados)
dadosInfra_2020_Jan = dadosInfra_2020_Jan.iloc[0:1]
```

Fonte: Autoria própria (2021)

Após a codificação foi obtido o resultado apresentado na Figura 22, por meio de uma análise da base da INFRAERO sobre os as operações de pousos e decolagens nas zonas aeroportuária administradas pela empresa no ano de 2019. A linha azul representa a média das operações no período.

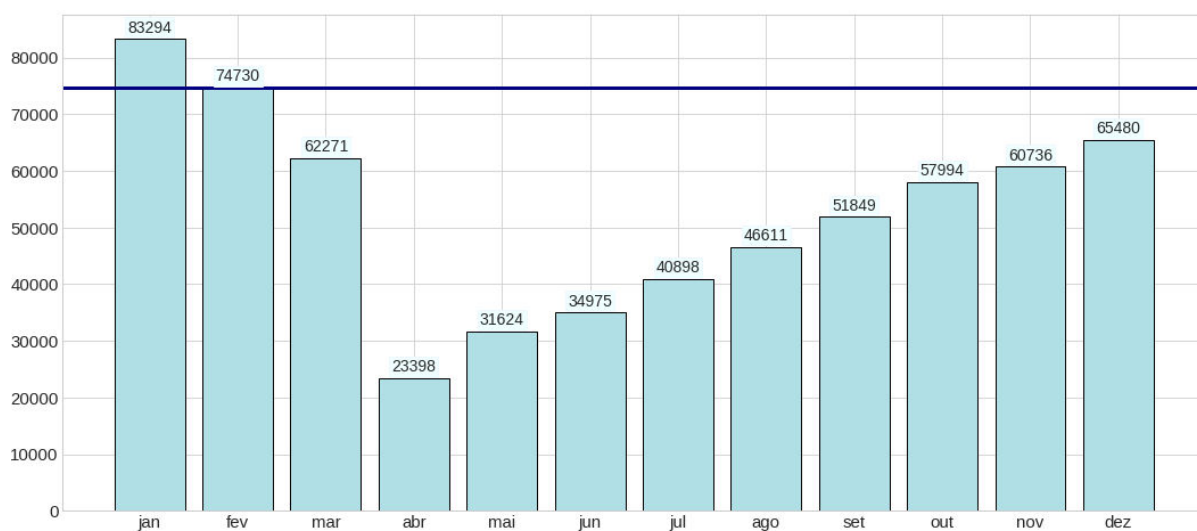
Figura 22: Operações na INFRAERO no ano de 2019



Fonte: Autoria própria (2021)

No ano de 2020, nos meses de janeiro até março, antes da OMS declarar pandemia, observa-se na Figura 23, que já há uma queda no número de operações realizadas nos aeroportos em relação ao mesmo período do ano anterior.

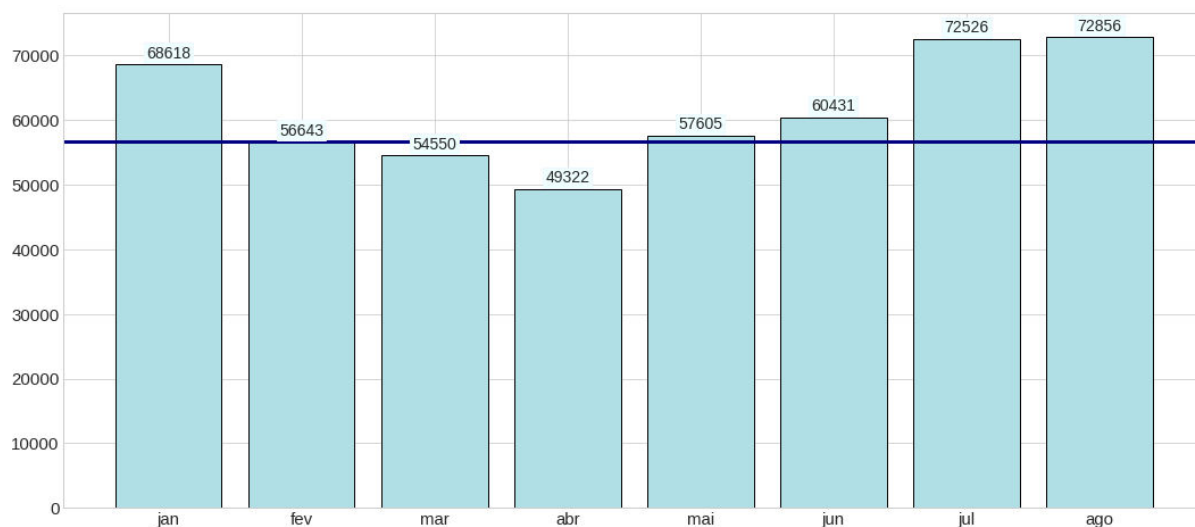
Figura 23: Operações na INFRAERO no ano de 2020



Fonte: Autoria própria (2021)

No ano de 2021 é possível obter dados até o mês de agosto, onde pode-se observar na Figura 24, o período de janeiro a agosto de 2021.

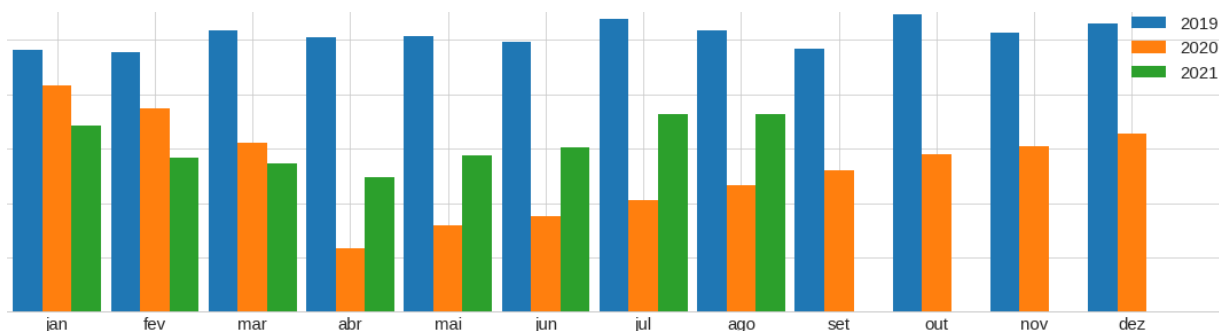
Figura 24: Operações na INFRAERO no ano de 2021



Fonte: Autoria própria (2021)

Na Figura 25, os anos de 2019, 2020 e 2021 são postos no mesmo gráfico mês a mês, observando que no início do ano de 2020 já se tinha uma queda nos pousos e decolagens mesmo antes da declaração da pandemia da COVID-19. Também se observa que nos períodos de março, abril e maio são os períodos de menores números de pousos e decolagens nos aeroportos.

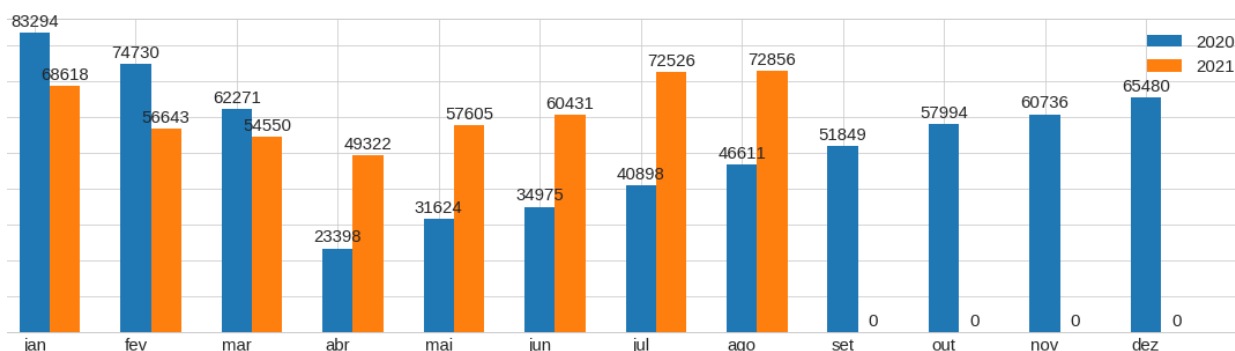
Figura 25: Operações mês a mês na INFRAERO



Fonte: Autoria própria (2021)

Na Figura 26, correspondentes aos anos de 2020 e 2021, é possível observar que no período dos meses de março, abril e maio do ano de 2020 teve os menores números de pousos e decolagens em relação ao ano de 2021.

Figura 26: Pousos e decolagens nos anos de 2020 e 2021



Fonte: Autoria própria (2021)

5.4. Análise dos Resultados Obtidos

Os voos dos aeroportos de GRU e REC são separados em tipos de serviços, onde cada tipo corresponde o tipo de voo que uma determinada empresa aérea está realizando desde a sua origem até o seu destino. Os tipos de serviços são identificados no dicionário de dados por letras maiúsculas (e.g., P - Voo de posicionamento, T - Voo de teste de aeronave, K - Voo de treinamento). Foram analisados apenas os voos de passageiros, correspondente a letra “J”, e os voos de carga e mala postal, sendo representado pelas letras “F” e “M” de acordo com o código na

Figura 27, onde é feito a triagem, para selecionar os tipos de voos específicos.

Figura 27: Trecho do código que seleciona os tipos de serviços

```
servicoS20 = dadosTipoServiS20[(dadosTipoServiS20['Tipodeserviço'] == 'J')
| (dadosTipoServiS20['Tipodeserviço'] == 'F') | (dadosTipoServiS20['Tipodeserviço'] == 'M')]
servicoS20.value_counts()
```

Fonte: Autoria própria (2021)

Nos dados disponibilizados pela INFRAERO não há separação dos tipos de voos, para realizar uma análise da quantidade de número de voos e carga movimentada é necessário modificar o argumento “*sheet_name*” no código, pois o arquivo da INFRAERO é do tipo .xlsx e possui pastas que diferenciam passageiros e cargas conforme a Figura 19.

Na base de dados da INFRAERO, a coluna que representa a operação de pouso e decolagem é “Pousos + Decolagens”, não especificando qual o tipo voo como é especificado na ANAC, distinguindo entre passageiros, cargas e malas postais, mas é possível analisar a quantidade de carga movimentada nos aeroportos administrado pela INFRAERO, a coluna que representa a carga movimentada é “Carga + Descarga”. O ano de 2020 teve uma redução de

48,16% no número de operações de pousos e decolagens em relação ao ano de 2019 e o movimento de cargas teve uma redução de 57,28% no mesmo período. Os valores quanto ao número de “Pousos+ Decolagens” e “Carga + Descarga” nos anos de 2019 e 2020 verifica-se na Tabela 4.

Tabela 4: Operações nos anos 2019 e 2020

	2019	2020
Pousos + Decolagens	1.235.775	633.860
Carga + Descarga	365.296.040 Kg	156.020.278 Kg

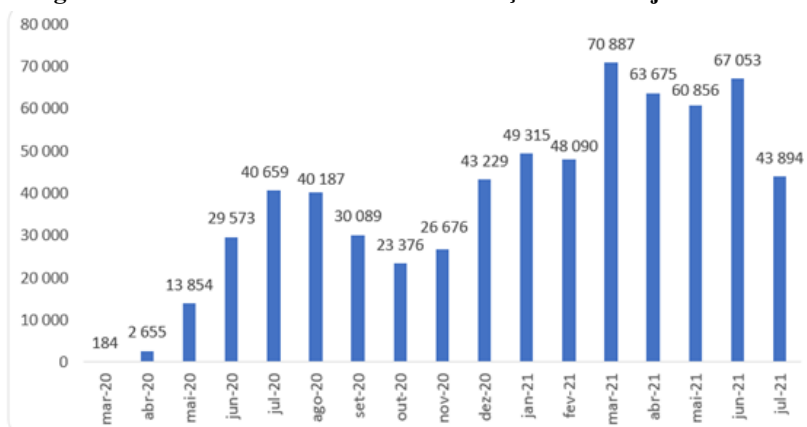
Fonte: Elaborada pelo autor com base na INFRAERO (2021)

Analisando o período de 2019, 2020 e 2021 na base de dados da INAFRAERO. O ano de 2019 o mês foi outubro obteve o maior número de operações de pousos e decolagens com 109.235 operações registrados e no de fevereiro registrou-se o menor número com 95.433 operações, para o mês de fevereiro e outubro houve um aumento de 14,46% no número de operações.

O ano de 2020, no mês de janeiro, foi registrado um pico de 83.294 operações registradas enquanto o mês de abril, já declarado a pandemia do covid-19, houve um número de 23.398 operações, uma queda de 71,90% nos números de operações e no mês seguinte, o mês de maio registrou 31.624 operações, um aumento de 35,15% no número de operações em contrapartida, no Brasil, de acordo com gráfico na

Figura 28, o mês de março de 2020 teve 184 casos registrados, o mês de abril 2.655 casos de covid-19 e o mês de maio 13.854 casos de covid-19, um acréscimo de 134,29% e um acréscimo de 421% no número de casos da doença entre março e abril e abril e maio, respectivamente.

Figura 28: Casos de COVID-19 entre março de 2020 e julho de 2021



Fonte: Correio de Minas¹⁰ (2021)

O ano de 2021 (dados obtidos até o mês de agosto) o mês que teve o maior número de operações, foi o mês de agosto com 72.856 números de operações e o menor mês, abril considerado o mês da segunda onda da covid-19, teve 49.322 operações, um aumento de 47,71%. No período da segunda onda houve um pico de infectados no mês de março de 2021 registrando 70.887 novos casos e o mês de abril 63.675 novos casos, uma redução de 10,17%, observou-se uma redução no movimento de operações nos aeroportos administrados pela INFRAERO a diferença entre os meses de abril e março que teve 54.550 operações teve uma redução de 9,58% no número de voos.

Analisando a base de dados da ANAC o ano de 2019 não é possível fazer uma análise completa do ano tanto dos aeroportos GRU e REC, pois a base de dados não possui registros do ano. O ano de 2020 o número de operações de pousos e decolagens no aeroporto de GRU e REC foi de 148.711 e 39.780 respectivamente, sendo os voos distribuídos em voos regulares de passageiros e voos responsáveis por operações de carga e mala postais. A distribuição está de acordo a Tabela 5.

Tabela 5: Distribuição dos voos em GRU e REC em 2020

	GRU	REC
Voos regulares de passageiros	134.347	35.868
Voos de carga e mala postais	6.417	1.681

Fonte: Elaborada pelo autor com base na ANAC, 2021

Nos aeroportos de REC e GRU, dados encontrados na base de dados da ANAC, o ano de 2020 o mês de pico foi registrado em janeiro, antes da declaração da OMS da pandemia, e o menor número registrado no mês de abril, após a declaração da pandemia, e pode-se ser observado de acordo Tabela 6.

Tabela 6: Melhor e pior mês de 2020 em GRU e REC

	GRU	REC
Melhor mês	26.432	7.134
Pior mês	3.036	601

Fonte: Elaborada pelo autor com base na ANAC, 2021

¹⁰ Correio de Minas: <https://correiodeminas.com.br/2021/08/02/diario-da-covid-19-em-julho-brasil-registra-as-menores-medias-de-2021/156100/>

Em REC a queda entre os picos de maior e menor volume foi de 91,57% e em GRU foi 88,51%. O mês de março foi registrado 4.455 operações e o de abril 601 em REC, uma queda de 86,50% e entre abril e maio, que teve 793, um aumento de 31,94%. Em GRU o mês de março foi registrado e 18.416 e o mês de abril 3.036 operações uma queda de 83,51% e entre o mês de abril e maio, que obteve 3.705 operações, um aumento de 22,03%.

Entre os meses de março e abril houve a maior queda do número de decolagens no ano de 2020, 86,13%, causada pelas medidas de contenção ao avanço da pandemia. Em relação ao mês de abril do ano anterior, houve queda de praticamente 90% no número de decolagens de acordo com a ANAC (2019, 2020).

O código fonte utilizado para a obtenção de todos os resultados apresentados nesse capítulo está acessível a qualquer usuário no repositório do GitHub¹¹.

¹¹ Github: <https://github.com/victorBezerra/Minera-o-de-dados-ANAC-e-INFRAERO>

6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Com o término do projeto, é observado que a pandemia causada pela covid-19 tem afetado os movimentos de passageiros entre diversas regiões do país e do mundo, tendo em vista que uma das principais formas de prevenção e controle da doença é por meio do isolamento. Consequentemente os aeroportos e companhias aéreas são diretamente impactados pelo aumento ou redução do número de infectados, como por exemplo, o período da segunda onda, onde o mês de março teve um pico no número de infectado e no próximo mês de abril houve uma redução no número de infectados, e nos aeroportos da INFRAERO no mesmo período também teve uma redução no número de voos.

Trabalhos futuros:

- Com o início da fase de vacinação contra a covid-19, pode-se verificar se a introdução da vacina contra a covid-19 afetou o número de operações realizadas nos aeroportos.
- Sem um prazo para o fim da pandemia da covid-19, verificar se em algum momento os números de operações nos aeroportos voltaram aos números pré-pandemia
- Com a descoberta de novas variantes da covid-19 e a possibilidade de nova onda da doença com o aumento no número de casos, pode-se verificar se a aparição de novas ondas da doença afetou o número de operações realizadas nos aeroportos.

REFERÊNCIAS

GOVERNO FEDERAL. **Governo Federal anuncia medidas para minimizar impactos na aviação civil**. Governo Federal. Disponível em: <<https://www.gov.br/infraestrutura/pt-br/assuntos/noticias/ultimas-noticias/governo-federal-anuncia-medidas-para-minimizar-impactos-na-aviacao-civil>>. Acesso em: 10 Nov. 2021.

INFRAERO. **AEROPORTOS DA INFRAERO MOVIMENTAM 83,9 MILHÕES DE PASSAGEIROS EM 2019**. Infraero. Disponível em: <https://www4.infraero.gov.br/imprensa/noticias/aeroportos-da-infraero-movimentam-83-9-milhoes-de-passageiros-em-2019/>. Acesso em: 10 Nov. 2021.

ROJAS, Paulo Ferreira. **Impactos do Covid-19 nas principais companhias aéreas brasileiras**. 2021. Universidade do Sul de Santa Catarina, Palhoça, 2021.

ICAO. **Economic Impacts of COVID-19 on Civil Aviation**. *International Civil Aviation Organization* (ICAO). Disponível em: <<https://www.icao.int/sustainability/Pages/Economic-Impacts-of-COVID-19.aspx>>. Acesso em: 10 Nov. 2021.

CHAO, Wang. **A quarta revolução**. Unesco. 2018. Disponível em: <<https://pt.unesco.org/courier/2018-3/quarta-revolucao>> . Acesso em: 10 Nov. 2021.

BRASIL. Lei nº 11.182, de 27 de Setembro de 2005. CRIA A AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL - ANAC, E DÁ OUTRAS PROVIDÊNCIAS.

BRASIL. Lei nº 5.862, de 12 de Dezembro de 1972. AUTORIZA O PODER EXECUTIVO A CONSTITUIR A EMPRESA PÚBLICA DENOMINADA EMPRESA BRASILEIRA DE INFRA-ESTRUTURA AEROPORTUÁRIA - INFRAERO, E DÁ OUTRAS PROVIDÊNCIAS.

CABENA, P; HADJINIAN, P; STADLER, R; JAAPVERHEES; ZANASI, A. *Discovering Data Mining: From Concept to Implementation*. Prentice Hall, 1998.

CHEN, Wei; GUO, Fangzhou; WANG, Fei-Yue. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, v. 16, n. 6, p. 2970-2984, 2015.

DE OLIVEIRA, Bruno Carneiro; CASTILHO, Denis. **ENTRE FLUXOS E ESCALAS: COVID-19 E SUAS IMPLICAÇÕES NA REDE DE TRANSPORTE AÉREO NO PRIMEIRO SEMESTRE DE 2020**. *GEOgraphia*, v. 22, n. 49, 2020. Disponível em: <<https://periodicos.uff.br/geographia/article/view/45882>> . Acesso em: 10 Nov. 2021.

TEIXEIRA, Vinicius Modolo. O impacto da crise da covid-19 no setor aeronáutico. **Boletim de Conjuntura (BOCA)**, v. 2, n. 4, p. 40-45, 2020. Disponível em: <<https://revista.ufrir.br/boca/issue/view/281>>. Acesso em: 10 Nov. 2021

Modeling Impacts of Aviation Mitigations and ATC Delay on Passenger COVID-19 Infection Risk

CRAVO, Beatriz Malerba. A alocação de slots e a concorrência no setor de transporte aéreo. *Journal of Transport Literature*, v. 8, p. 159-177, 2014. Disponível em: <

<https://www.scielo.br/j/jtl/a/c6gkYZSDmFPYpt8kvWcTRdf/abstract/?lang=pt&format=html>
> Acesso em: 10 Nov. 2021.

CAMILO, Cássio Oliveira; SILVA, João Carlos da. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. Universidade Federal de Goiás (UFG), v. 1, n. 1, p. 1-29, 2009.

TEIXEIRA, Fabio Pereira et al. COVID 19 Impact on Aircraft Ground Time at Congonhas Airport (CGH). 2020.

PETROVIĆ, Nenad; ROBLEK, Vasja; PAPACHASHVILI, Nino. Decision Support Based on Data Mining for Post COVID-19 Tourism Industry.

CARNEIRO, Juliana; ALLIS, Thiago. Como se move o turismo durante a pandemia da COVID-19?. Revista Brasileira de Pesquisa em Turismo, v. 15, 2021. Disponível em:<<https://www.scielo.br/j/rbtur/a/rmSVQFWCd3CXdDJvqcYBRpt/abstract/?lang=pt>>. Acesso em: 10 Nov 2021.