

Embraer Social Tech III
Projeto Integrador - Grupo 6
Tema: Análise dos Custos de Transporte na Embraer

Índice

Índice	1
Desafio	2
Contextualização	2
Sobre a Embraer	2
Fabricação de Aeronaves	3
Aeronaves Agrícolas	4
Exportação Embraer	5
Regime Aduaneiro Embraer	5
Projeto Integrador	6
Kanban	6
Construção do banco de dados MySQL	6
Desenvolvimento Power BI	7
Documentação de Arquitetura de Dados	8
Desenvolvimento Excel	19
Confiabilidade dos dados	20
Análises e Plano de Ação	21
Previsão do preço do combustível	21
Custo do Combustível	21
Modelo de previsão: ARIMA	22
Análise de Indicadores	24
Entregas de Aeronaves	24
Aumento dos custos de combustível	25
Modelos mais vendidos de Aeronaves	25
Autonomia das Aeronaves	25
Atuação no mercado	25
Previsão de demanda + custo de frete	26
Eficiência das entregas	26
Plano de Ação	26
Referências // Anexos	28

Desafio

Avaliar os *custos de transporte* (custos associados ao transporte de mercadorias, que podem incluir transporte terrestre, marítimo ou aéreo) na área de logística e supply chain da Embraer e, com base nos resultados encontrados, fazer um mapeamento do cenário e sugerir um plano de atuação.

Optamos por esse tema porque vimos muito potencial no desenvolvimento de uma análise aprofundada e real, buscando não apenas compreender os desafios logísticos enfrentados pela Embraer, mas também buscando oferecer insights e projeções valiosas referente ao custo de transporte pré e pós pandemia, garantindo eficiência e principalmente, planejamento para os próximos meses voltado para esse setor dinâmico e estratégico.

Contextualização

Sobre a Embraer

A Embraer é uma empresa brasileira que se destaca na fabricação e comercialização de aeronaves. Fundada em 1969, iniciou as suas atividades como uma empresa estatal e, posteriormente, em 1994, passou por um processo de privatização - [ampliando a sua atuação para mais de 60 países](#). Atualmente, trabalha nos segmentos Aviação Comercial, Aviação Executiva, Defesa e Segurança e Aviação Agrícola.

De acordo com o site de Relações com Investidores da Embraer, em média, a cada 10 segundos uma aeronave fabricada pela Embraer decola de algum lugar do mundo, transportando anualmente mais de 145 milhões de passageiros. Além disso, a empresa é líder na fabricação de jatos comerciais de até 150 assentos e a principal exportadora de bens de alto valor agregado do Brasil.

A Embraer começou a ganhar reconhecimento internacional na década de 1990, quando lançou a família de aeronaves ERJ (Embraer Regional Jet) e, depois, os jatos E-Jet, que são amplamente utilizados nos voos regionais ao redor do mundo. Além dos setores civis, a Embraer também se destacou na produção de aeronaves militares, como o avião militar de treinamento e ataque ao solo A-29 Super Tucano.

Ao longo de sua história, a Embraer tem desempenhado um papel significativo no desenvolvimento da indústria aeroespacial brasileira e se consolidou como uma das principais fabricantes de aeronaves do mundo.

Fabricação de Aeronaves

A fabricação de aviões envolve uma variedade de componentes e materiais para garantir segurança, eficiência e desempenho operacional. A escolha desses materiais, muitas vezes, depende da função específica e das características desejadas para a aeronave. Veja alguns deles:

Estrutura Principal

- Alumínio: tradicionalmente, as estruturas principais dos aviões eram feitas, em sua grande maioria, de alumínio. O material é leve e possui boa resistência estrutural.
- Compostos de fibras de carbono: cada vez mais, a indústria aeronáutica tem adotado materiais compósitos, como fibra de carbono reforçada com resina epóxi. Esses materiais oferecem alta resistência, com peso reduzido.

Asas

- Compostos: as asas costumam incorporar materiais compósitos, como fibras de carbono e materiais compostos de matriz metálica ou polimérica, para otimizar a relação peso-resistência.

Motores

- Ligas de titânio: partes dos motores podem ser feitas com ligas de titânio, devido à sua resistência a altas temperaturas e, também, à sua leveza.
- Superligas: em áreas sujeitas a altas temperaturas, como as seções quentes dos motores, são utilizadas superligas à base de níquel ou cobalto.

Revestimento externo

- Alumínio e titânio: o revestimento externo da aeronave geralmente é feito de painéis de alumínio ou titânio, proporcionando uma superfície aerodinâmica e resistência estrutural.
- Compósitos: em algumas aeronaves mais modernas, painéis de compósitos também são utilizados no revestimento externo.

Sistemas elétricos e eletrônicos

- Fios de cobre: os sistemas elétricos são interligados por fios de cobre.
- Fibra óptica: usada para comunicações de dados de alta velocidade.

Interior

- Materiais leves: os interiores normalmente incorporam materiais mais leves, como plásticos, com a finalidade de reduzir o peso total da aeronave.

Vidro e materiais transparentes

- Acrílico e policarbonato: as janelas e partes transparentes da cabine podem ser feitas tanto de acrílico, quanto de policarbonato.

A fabricação deste tipo de produto conta com uma combinação complexa de materiais tradicionais e avançados, visando otimizar o desempenho, a eficiência e a segurança deste

meio de transporte. Avanços contínuos têm permitido a criação de modelos mais leves, eficientes e resistentes.

Aeronaves Agrícolas

Como complemento ao nosso trabalho, resolvemos trazer aqui algumas informações sobre um dos modelos Agrícolas da Embraer, uma vez que o foco do trabalho acabará sendo mais voltado para os aviões comerciais e executivos.

Dessa forma, falaremos sobre o Ipanema (EMB 200 / 201 / 202 / 203), que é um dos modelos que possui uma média de venda anual de 75 unidades.

O Ipanema é um avião agrícola leve e de extrema durabilidade, muito popular nas propriedades agrícolas do Brasil. Mais tarde se tornaria **o único avião do mundo movido a etanol**, o que lhe valeu o título de uma das 50 melhores invenções de 2005 pela revista “Scientific American”. Atualmente, é fabricado em três variações do modelo, domina 60% do mercado e já contabiliza mais de 1.500 unidades produzidas desde o seu lançamento. No ano de 2021 a frota de aviões agrícolas foi de 2432 aeronaves e os aviões da Embraer correspondem a 1349 unidades. O Mato Grosso do Sul foi o Estado com maior número de vendas.

Idealizado para pulverizar plantações com fertilizantes e pesticidas, também pode ser utilizado para espalhar sementes, combater incêndios e criar chuva artificial. Para proteger o piloto do contato com os produtos químicos, a cabine do Ipanema conta com um sistema de vedação e a dispersão dos defensivos químicos ocorre na parte posterior das asas. Por voar em baixas altitudes, conta com um equipamento corta-fios. Ao longo dos anos diversas variantes do Ipanema foram certificadas e postas à disposição.

Sustentabilidade

O Ipanema foi a aeronave escolhida para servir como o Demonstrador Elétrico da Embraer, que realizou seu primeiro voo em agosto de 2021. A propulsão elétrica está se tornando cada vez mais disponível e confiável para uso aeronáutico, o que pode representar um grande avanço na transição para um modelo de aviação de emissão zero.

Exportação Embraer

De acordo com uma [reportagem da Exame](#), a Embraer já atingiu a marca de mais de 8.000 aeronaves entregues desde sua fundação em 1969, com mais de 90% de sua produção exportada, sendo os Estados Unidos e a Europa os principais destinos. Especialistas

consultados pela CNN atribuem o sucesso internacional da Embraer ao fato de o Brasil não possuir um mercado de aviação regional robusto, ao contrário de países europeus e dos EUA.

No Brasil, a Azul é a principal compradora das aeronaves da Embraer, recebendo 60 delas no ano anterior (2022). As companhias aéreas GOL e Latam, no entanto, não possuem aeronaves da Embraer em suas frotas, devido a rotas maiores que demandam aeronaves maiores, como as da Boeing e Airbus.

Regime Aduaneiro Embraer

O Regime Aduaneiro, basicamente, compreende o conjunto de normas e procedimentos fiscais relacionados às operações de importação e exportação. Para empresas envolvidas no comércio internacional, como a Embraer, seguir essas regras específicas é crucial. O regime aduaneiro define como essas operações são tratadas em termos de impostos, tarifas e documentação.

No caso da Embraer, o regime adotado é o RECOF (Regime de Entrepasto Industrial sob Controle Aduaneiro Informatizado), e o artigo *“Análise logística de regime aduaneiro utilizado em uma empresa nacional de produção aviônica”* detalha o fluxo do processo de liberação alfandegária através desse regime. Um ponto interessante, ressaltado anteriormente, é que mais de 90% da produção da Embraer é exportada, enquanto a empresa importa 95% das matérias-primas, partes e peças para a produção, incluindo peças de reposição para as aeronaves fabricadas.

O artigo também menciona que as autoridades aduaneiras, referentes ao RECOF, procuram evitar a retenção dos produtos por longos períodos. Além disso, empresas credenciadas, como a Embraer, têm a vantagem de manter baixos níveis de estoque, possibilitando um abastecimento *just-in-time*. Isso não apenas reduz os custos com armazenagem, mas também aumenta o giro das mercadorias estocadas e agiliza a distribuição dos produtos no mercado.

Projeto Integrador

Kanban

No contexto da gestão eficiente do Projeto Integrador, optamos por implementar um sistema Kanban, dentro do Trello, para organizar e monitorar todas as atividades do time. Essa abordagem, inspirada nas melhores práticas da metodologia ágil, visa criar um

ambiente transparente do fluxo de trabalho, assim como organizar as tarefas para um acompanhamento contínuo e colaboração do time.

Inicialmente, definimos as etapas principais do projeto, alinhadas aos prazos de entrega de cada uma delas e aos seus respectivos responsáveis. A partir disso, seguimos para o preenchimento de cada uma das subtarefas relacionadas às macro estabelecidas. Dentro do nosso quadro, dividimos o fluxo em 4 principais etapas: BACKLOG (a fazer), DOING (em andamento), REVIEW (revisão) e DONE (concluído).

Além do acompanhamento por meio do Trello, nosso grupo criou uma rotina de reunião semanal (às quintas-feiras, 21h00), para falarmos sobre as tarefas em andamento, necessidade de apoio nas demandas, revisão dos prazos de entrega e criação dos dashboards.

Link do Trello: <https://trello.com/w/embraersocialtechprojetointegrador>

Quadro Trello: <https://trello.com/b/qfxHvCIA/projeto-integrador>

Construção do banco de dados MySQL

Parte I: Levantamento de Dados

Link Banco de Dados no Excel:  Banco de dados - Projeto Integrador

Para começarmos a construir o banco de dados do Projeto Integrador, realizamos inicialmente um mapeamento aprofundado sobre a Embraer e o cenário/segmento em que a empresa está inserida. Esse processo envolveu o levantamento de diversas referências bibliográficas, cujas fontes foram devidamente documentadas ao final deste relatório, e a coleta precisa de dados retirados do [site de Relações com Investidores da Embraer](#). Essa investigação inicial trouxe uma compreensão mais abrangente sobre a estrutura organizacional da empresa e os fatores que impactam diretamente os custos de transporte das mercadorias.

Para enriquecer a estrutura do banco de dados, dentro da temática escolhida pelo grupo, pesquisamos sobre informações adicionais que julgamos serem necessárias para darmos continuidade a nossa análise. Entre elas, podemos destacar: tipos de combustíveis utilizados pela frota, custos associados por litro e quilometragem, principais compradores de aeronaves da Embraer, modelos de aviões fabricados pela empresa, dados de aviões elétricos/híbridos/movidos a hidrogênio, capacidade do tanque de combustível e plantas da Embraer.

A partir de todo esse levantamento, começamos a unificar as informações para construir uma base sólida, que servirá, em um segundo momento, para o desenvolvimento do dashboard e de todas as análises que virão a ser feitas no decorrer do projeto.

Parte II: Construção da Base

Para fazer a construção da base de dados, utilizamos os dados reais coletados da Embraer, somados às pesquisas que realizamos online, e geramos dados aleatórios para preencher colunas que ficaram faltando na análise que gostaríamos de entregar. Os dados foram:

- **Ano de entrega de aviões comerciais:** para a definição dos anos, consideramos o período de produção de cada modelo de aeronave, a fim de termos dados mais próximos da realidade;
- **Países e companhias aéreas que adquiriram aviões comerciais e executivos:** pegamos os dados existentes de clientes da Embraer e geramos uma base aleatória para complementar os dados faltantes.
- **Países que podem ser potenciais clientes de aviões “sustentáveis”:** geramos dados aleatórios, com base no Environmental Performance Index, para mapear o ranking de países mais sustentáveis e que poderiam ter potencial interesse em adquirir aviões híbridos, elétricos e/ou movidos a hidrogênio.

Além disso, a fim de agregar mais dados à base de dados, fizemos um tipo de aumento de dados, baseado na quantidade de aviões entregue sendo ela dividida aleatoriamente entre os dados faltantes, ora entre ano de entrega, ora entre países e companhias aéreas.

Os dados aleatórios foram gerados pelo Notebook de Python Google Colab. Veja a seguir:

 `Base_de_dados.ipynb` .

Desenvolvimento Power BI

Listagem dos indicadores e análises que gostaríamos de trazer para o trabalho:

 PROJETO INTEGRADOR - EMBRAER

Documentação de Arquitetura de Dados

1. Introdução

Esta documentação tem como propósito proporcionar uma compreensão clara da arquitetura dos dados, fornecendo recomendações práticas para aprimorar a confiabilidade e qualidade dos dados ao longo do tempo. As orientações visam

comunicação entre desenvolvedores, designers de banco de dados e outros stakeholders envolvidos no projeto.

Exemplos de consultas SQL frequentemente utilizadas para melhor compreensão:

```
1 • use projetoEmbraer;
2 • select * from BD_aumentado_comercial;
3
4
5 -- Conferindo Registros da Coluna 'Ano de Fabricação'
6 • SELECT DISTINCT `Ano de Fabricação` FROM BD_aumentado_comercial;
7
8 -- Filtrar por Segmento e Ano de Fabricação:
9 • SELECT * FROM BD_aumentado_comercial WHERE `Segmento` = 'Aviação Comercial' AND `Ano de fabricação` = 2022;
10
11 -- Filtrando por Ano de Fabricação:
12 • SELECT * FROM BD_aumentado_comercial WHERE `Ano de Fabricação` = 2018;
13
14 -- Filtrando por Produtos com Quantidade Entregue Maior que um Valor:
15 • SELECT * FROM BD_aumentado_comercial WHERE `Quantidade entregue` > 10;
16
17 -- Filtrar por Origem e Destino:
18 • SELECT * FROM BD_aumentado_comercial WHERE `Origem` = 'São José dos Campos' AND `Destino` = 'Egito' LIMIT 0, 1000;
19
20 -- Filtrar por Produtos com Pedidos Firmes em Carteira:
21 • SELECT * FROM BD_aumentado_comercial WHERE `Pedidos Firmes em Carteira` > 0;
22
```

Figura 2 - Parte de código que descreve rotinas utilizadas em MySQL.

Realizamos as conexões entre as tabelas e a atribuição das chaves, bem como a definição dos tipos de dados (Text, Varchar, etc.), por meio de linhas de código SQL. Esse processo garante a integridade e eficiência da estrutura do banco de dados, permitindo a gestão precisa das relações entre as diferentes entidades. Através do SQL, conseguimos não apenas visualizar, mas também implementar de maneira prática e eficaz os relacionamentos e características fundamentais de nossas tabelas no MySQL.

```

use projetoembraer;
-- 1. Supondo que a coluna 'Destino' seja do tipo TEXT
-- Altere o tipo de dados para VARCHAR com um comprimento máximo adequado
> Execute|JSON
ALTER TABLE BD_aumentado_comercial
MODIFY Destino VARCHAR(255); -- Altere para o comprimento desejado

-- 2. Verificação de Dados (certifique-se de que não há valores nulos na coluna 'Destino')
> Execute|JSON
SELECT *
FROM BD_aumentado_comercial
WHERE Destino IS NULL;

-- 3. Supondo que a coluna 'Destino' seja do tipo TEXT
-- Altere o tipo de dados para VARCHAR com um comprimento máximo adequado
> Execute|JSON
ALTER TABLE BD_aumentado_comercial_exec
MODIFY Destino VARCHAR(255); -- Altere para o comprimento desejado

-- 4. Verificação de Dados (certifique-se de que não há valores nulos na coluna 'Destino' em 'BD_aumentado_comercial_exec')
> Execute|JSON
SELECT *
FROM BD_aumentado_comercial_exec
WHERE Destino IS NULL;

```

Figura 3 - Parte de código que modifica variáveis em uma tabela do MySQL.

```

-- 5. Agora, adicione o índice à coluna 'Destino' em 'BD_aumentado_comercial_exec'
> Execute|JSON
CREATE INDEX idx_destino_aumentado ON BD_aumentado_comercial_exec(Destino);

-- 6. Verifique novamente se não há valores nulos na coluna 'Destino' em 'BD_aumentado_comercial'
> Execute|JSON
SELECT *
FROM BD_aumentado_comercial
WHERE Destino IS NULL;

-- 7. Certifique-se de que todos os valores em Destino já existem em Destino
> Execute|JSON
SELECT DISTINCT Destino
FROM BD_aumentado_comercial
WHERE Destino NOT IN (SELECT Destino FROM BD_aumentado_comercial_exec);

-- 8. Desative temporariamente as verificações de chave estrangeira
> Execute|JSON
SET foreign_key_checks = 0;

-- 9. Se o resultado for vazio, todos os valores já existem, então você pode adicionar a chave estrangeira
> Execute|JSON
ALTER TABLE bd_aumentado_comercial
ADD CONSTRAINT fk_destino_aumentado
FOREIGN KEY (Destino) REFERENCES BD_aumentado_comercial_exec(Destino);

```

Figura 4 - Parte de código que insere uma chave estrangeira, criando o relacionamento entre tabelas no MySQL.

3. Modelos de Dados no Power BI

As imagens a seguir representam medidas importantes realizadas com o auxílio da ferramenta do Power BI.

- Soma do Custo de Total de Entrega x Ano utilizando tanto modelos de aviação comercial quanto os modelos de aviação executiva.



Figura 5 - Comparação da Soma do Custo Total de Frete ao longo dos anos na Aviação Comercial.



Figura 6 - Comparação da Soma do Custo Total de Frete ao longo dos anos na Aviação Executiva.

- Média do Custo do Combustível (QAV) por período, sendo o período caracterizado por pré-pandemia (antes de 2020) e pós-pandemia (2020 em diante), e o impacto no frete para os mesmos períodos

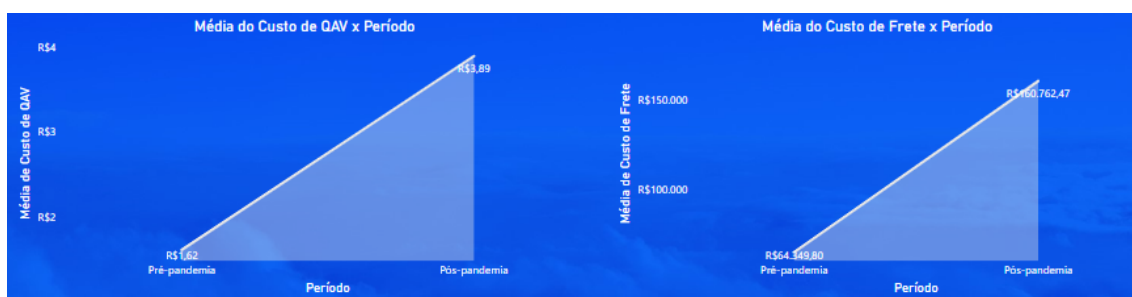


Figura 7 - A Figura a direita representa a média do custo do combustível (QAV) por período (pré e pós pandemia). A Figura a esquerda representa a média do custo de frete por período.

- A quantidade de aeronaves entregue por ano e por modelo, tanto para o segmento de Aviação Comercial (Figura 8), quanto para o segmento de Aviação Executiva (Figura 9).

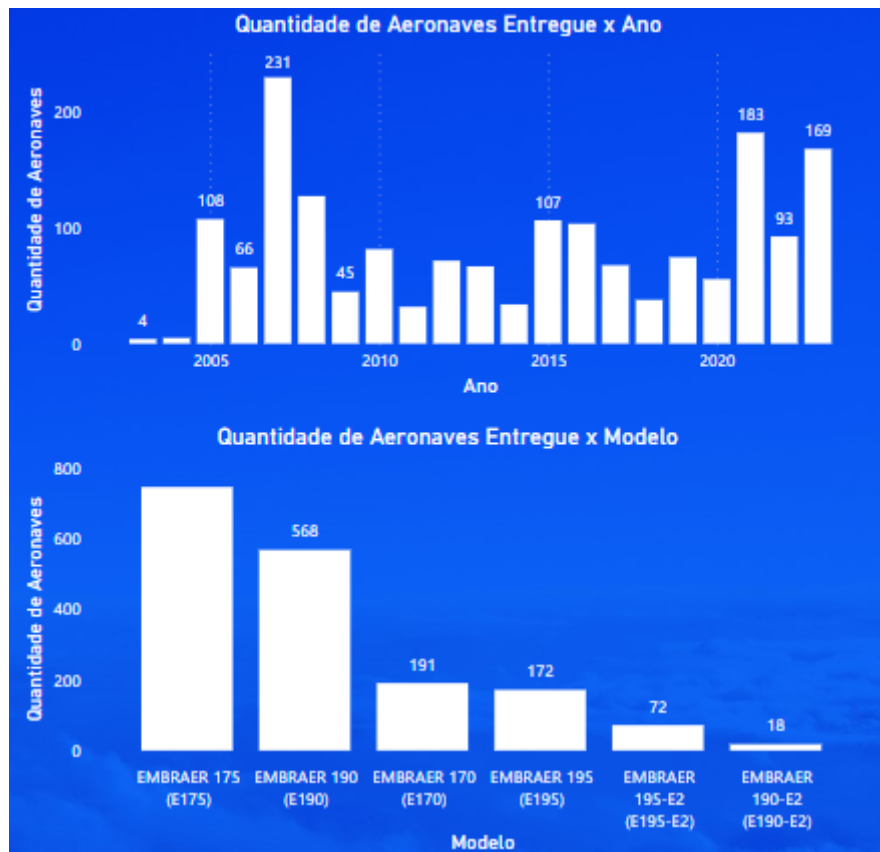


Figura 8 - A primeira análise (acima) representa a quantidade de aeronaves entregues pela Embraer por ano. A Segunda análise (abaixo) representa a quantidade de aeronaves entregue por modelo na Aviação Comercial.

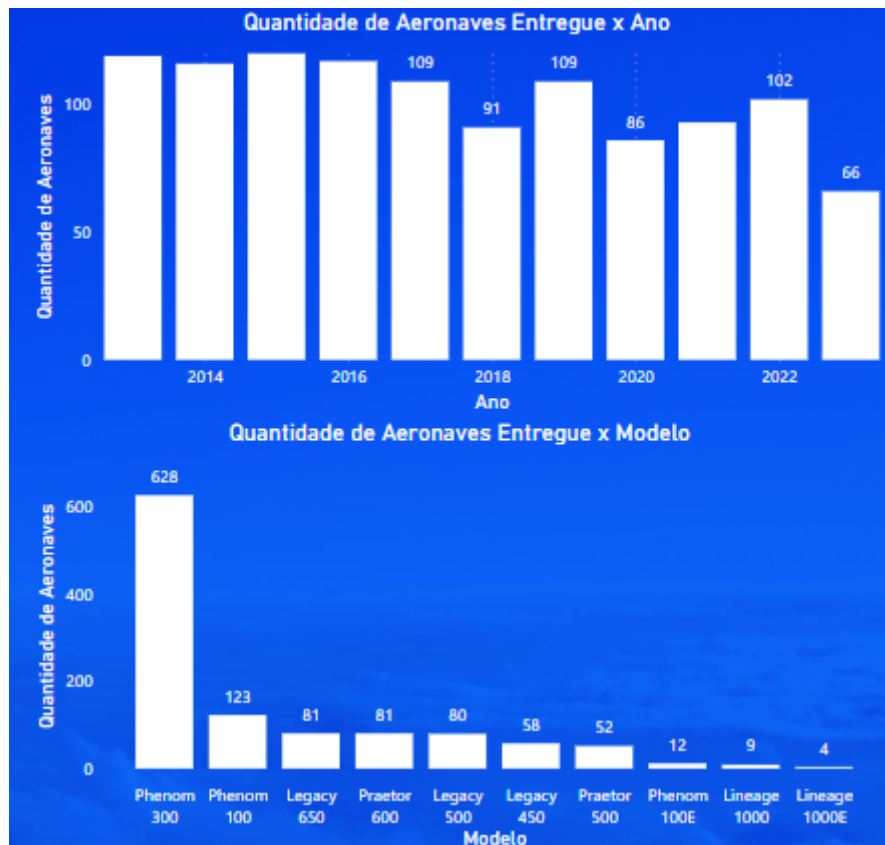


Figura 9 - A primeira análise (acima) representa a quantidade de aeronaves entregues pela Embraer por ano. A Segunda análise (abaixo) representa a quantidade de aeronaves entregue por modelo na Aviação Executiva.

- Uma análise que representa a eficiência de entregas da Embraer relacionadas a Aviação Comercial para os anos de 2003 a 2023 (85,8% dos pedidos realizados já foram entregues).

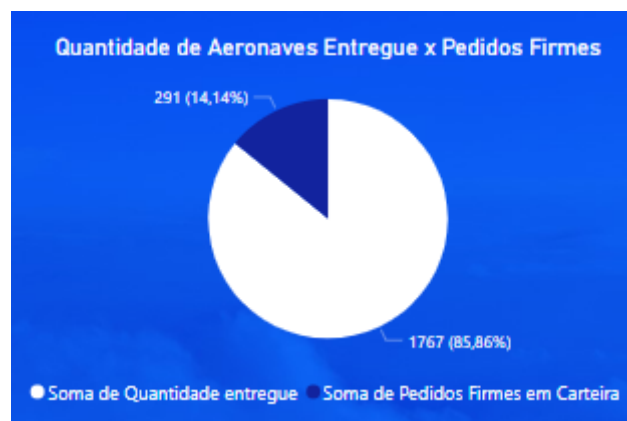


Figura 10 - Gráfico de pizza que mostra a quantidade de aeronaves entregue por pedidos firmes em carteira do segmento de Aviação Comercial.

Lógica das Transformações:

As transformações realizadas foram limpeza, normalização e aumento de dados.

A limpeza inclui retirada de informações não usuais (como, por exemplo, 'Clientes não-divulgados') que foram substituídas por valores gerados aleatoriamente entre os valores já existentes. Além da padronização dos valores caracterizados por *string*.

A normalização entre o preço de custo de frete de 1 (um) avião para a quantidade de aviões entregue para um cliente.

$$\text{Custo de frete: } Tam_{tanque} * Val_{comb} + \left(\frac{Au}{Dest} \right) * Tam_{tanque} * Val_{comb},$$

onde Tam_{tanque} = tamanho do tanque de combustível da aeronave, Val_{comb} = valor do combustível para determinado ano, Au = Autonomia da aeronave e $Dest$ = Distância ao país de destino a partir de São José dos Campos.

Custo total de frete:

$$Qtd * (Tam_{tanque} * Val_{comb} + \left(\frac{Au}{Dest} \right) * Tam_{tanque} * Val_{comb}),$$

onde Qtd é a quantidade de aeronaves entregue a determinado país.

O aumento de dados foi realizado com base na quantidade de aviões entregue relacionado aos dados faltantes de cada tabela. Se a quantidade de aviões entregue é x maior do 1, dividimos x em valores inteiros aleatórios com a condição de que a soma seja igual a x . Por exemplo, $x = 3$, geramos aleatoriamente y e z tal que $y + z = x$, ou a, b e c , tal que $a + b + c = x$, nestes casos y seria igual a 2 e z igual a 1, e $a = 1, b = 1$ e $c = 1$.

Então, em ambos os casos nós aumentamos a quantidade entregue, mas deixamos fixas as informações já existentes. Por exemplo, na base de Aviação Comercial, nós não tínhamos acesso ao ano de entrega dos aviões, mas tínhamos acesso ao país de

destino. Portanto, nós repetimos os países de destino para cada quantidade aumentada, e geramos aleatoriamente o ano de entrega dos aviões (prezando o período em que o modelo de aeronave foi fabricado). Exemplo: $x = 3$; $a, b, c = 1$.
Geramos:

Quantidade entregue	Ano de entrega	País de destino
1	2006	Austrália
1	2007	Austrália
1	2008	Austrália

Tabela 1 - Exemplo de como os dados faltantes são gerados.

O mesmo procedimento foi realizado para a base de dados de Aviação Executiva, sendo os dados faltantes o país de destino e os dados preexistentes que foram repetidos foram os anos de entrega.

Fluxos de Dados:

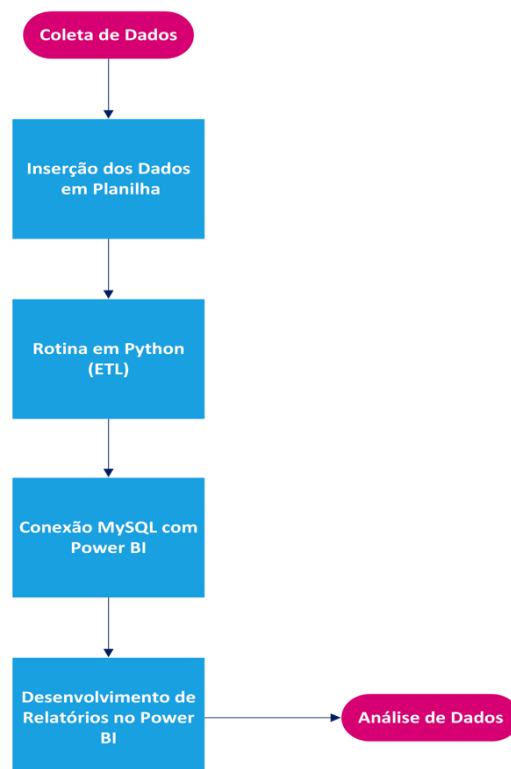


Figura 11 - Fluxograma da sequência de dados.

3.1 Descrição do Fluxograma:

Coleta de Dados:

- Inicia o processo de coleta de dados de diversas fontes na internet.

Inserção em Planilha:

- Os dados coletados são inseridos em uma planilha para organização.

Rotina em Python (ETL):

- Utiliza-se uma rotina em Python para realizar as operações de Extração, Transformação e Carga (ETL) nos dados armazenados na planilha. Operações ETL incluem:
 - Extração: Os dados são lidos da planilha.
 - Transformação: Aplicam-se transformações conforme necessário (limpeza, manipulação de dados).
 - Carga: Os dados transformados são carregados nas tabelas do MySQL.

Conexão MySQL com Power BI:

- Realiza-se a conexão entre o MySQL e o Power BI para possibilitar o desenvolvimento de relatórios.

Desenvolvimento de Relatórios no Power BI:

- Utiliza-se o Power BI para criar relatórios interativos com base nos dados carregados no MySQL.

Análise de Dados:

- Exploração e interpretação dos dados para obter insights valiosos.

3.2 Ferramentas Específicas Utilizadas durante o Processo:

A seguinte rotina em Python utiliza a biblioteca SQLAlchemy para realizar operações de ETL (Extração, Transformação e Carga) em dados armazenados em arquivos CSV para tabelas MySQL.

Importação de Bibliotecas:

- `create_engine` e `pd` (Pandas): Funções do SQLAlchemy e Pandas, respectivamente, para criação de engine de banco de dados e manipulação de dados tabulares.

```
from sqlalchemy import create_engine
import pandas as pd
```

Figura 12 -Bibliotecas utilizadas.

Configurações do Banco de Dados:

- `db_config`: Dicionário contendo informações como host, usuário, senha e nome do banco de dados.

```
# Configurações do banco de dados
db_config = {
    'host': 'localhost',
    'user': 'root',
    'password': 'natasha57',
    'database': 'projetoEmbraer'
}
```

Figura 13 - Configurações utilizadas no banco de dados.

Conexão ao MySQL usando SQLAlchemy:

- `engine`: Cria uma instância de engine para gerenciar a conexão com o banco de dados.

Mapeamento de Arquivos CSV para Tabelas:

- `mapeamento_arquivos_tabelas`: Dicionário que associa nomes de arquivos CSV a nomes correspondentes de tabelas MySQL.

```
engine = create_engine(f"mysql+mysqlconnector://{db_config['user']}:{db_config['password']}@{db_config['host']}/{db_config['database']}")
```

Figura 14 - Mapeamento de arquivos CSV para Tabelas do MySQL.

Loop sobre Arquivos e Tabelas:

A rotina percorre cada arquivo e tabela correspondente, lê os dados do arquivo CSV usando Pandas e os insere no MySQL. Essa rotina realiza a ETL, garantindo a integridade e consistência dos dados no MySQL, além de fornecer informações detalhadas sobre eventuais erros encontrados durante o processo.

```
for nome_arquivo_csv, nome_tabela in mapeamento_arquivos_tabelas.items():
    try:
        # Ler dados do arquivo CSV usando pandas
        dados_csv = pd.read_csv(nome_arquivo_csv)

        # Remover espaços extras no nome das colunas
        dados_csv.columns = dados_csv.columns.str.strip()

        # Criar a tabela no MySQL (se ela não existir)
        dados_csv.head(0).to_sql(nome_tabela, con=engine, if_exists='replace', index=False)

        # Inserir dados no MySQL
        dados_csv.to_sql(nome_tabela, con=engine, if_exists='append', index=False)

        # Exibir os dados
        print(f'Dados do arquivo {nome_arquivo_csv} inseridos com sucesso na tabela {nome_tabela}')

    # Tratamento de exceções específicas.
    # Se o arquivo CSV não for encontrado, a exceção FileNotFoundError será capturada.
    # A variável 'e' armazena a exceção capturada, permitindo o acesso às informações sobre a exceção.
    except FileNotFoundError as e:
        print(f'Erro: Arquivo {nome_arquivo_csv} não encontrado. Verifique o nome do arquivo. Detalhes do erro: {e}')

    # Tratamento de exceções genéricas.
    # Se ocorrer qualquer outra exceção, ela será capturada pelo bloco except Exception.
    except Exception as e:
        print(f'Erro ao inserir dados do arquivo {nome_arquivo_csv} no MySQL: {e}')

# Bloco finally.
# Independentemente de ocorrer uma exceção ou não, o bloco finally é executado.
# No caso, é usado para garantir que a conexão com o MySQL seja fechada.
finally:
    engine.dispose() # Fechar a conexão
```

Figura 15 -Código para inserir as tabelas CSV no MySQL através do Python.

4. Recomendações para Aprimorar a Confiabilidade e Qualidade dos Dados:

Adotar Boas Práticas de Coleta:

- Implementar validação de formato e fazer verificações de consistência durante a entrada de dados;
- Considerar a adoção de técnicas de aprendizado de máquina para detecção automática de anomalias;
- Armazenar Dados:
 - Normalizar o banco de dados para reduzir redundâncias e melhorar a consistência.
 - Manter um plano de backup regular e testar procedimentos de recuperação.
- Limpar Dados:
 - Utilizar expressões regulares ou ferramentas específicas para corrigir problemas de formatação.
 - Automatizar processos de limpeza sempre que possível para eficiência a longo prazo.
- Monitorar Continuamente:
 - Implementar alertas automáticos para notificar sobre possíveis problemas nos dados.
 - Realizar auditorias periódicas para garantir a aderência contínua às boas práticas.

Desenvolvimento Excel

Para iniciar as análises das informações coletadas, fizemos dois dashboards no Excel: um voltado para aviação comercial, entre os anos de 2003 e 2023, e outro para aviação comercial e executiva entre 2013 e 2023.

Com base no banco de dados levantado, criamos algumas tabelas dinâmicas para consolidar as informações e cruzar variáveis que fizessem sentido entre si. Para cada uma

das dinâmicas, geramos gráficos, para uma análise mais visual, e inserimos segmentações de dados (filtros) para agregar na investigação e ajudar a extrair insights mais detalhados.

Para o layout do arquivo, utilizamos a paleta de cores da Embraer. Extraímos o código hexadecimal do próprio site da empresa e trouxemos o logo para atribuir o projeto à marca.

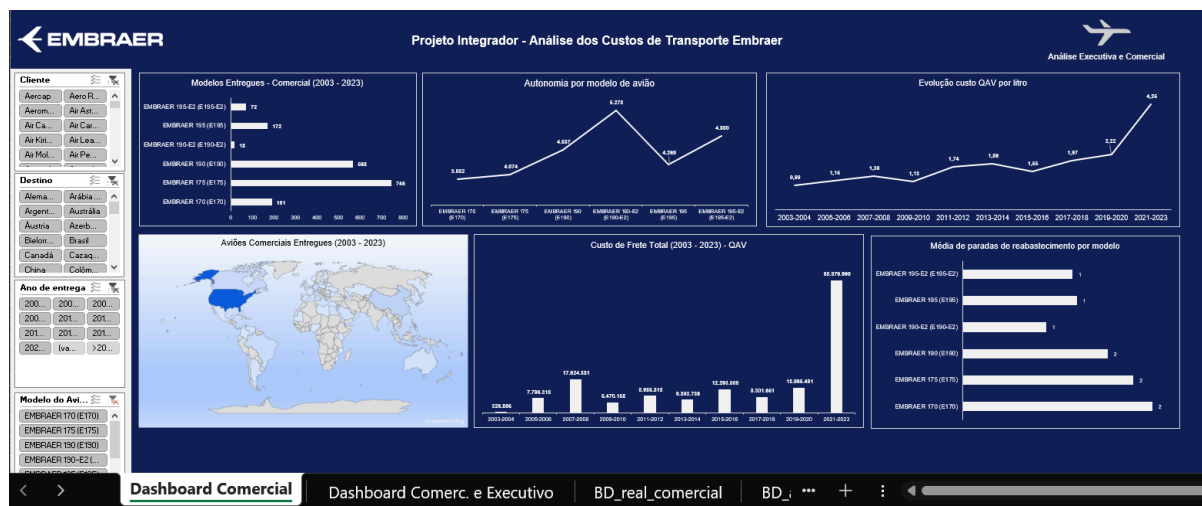


Figura 16 - Dashboard desenvolvido no Excel sobre Custo de Transporte da Aviação Comercial.

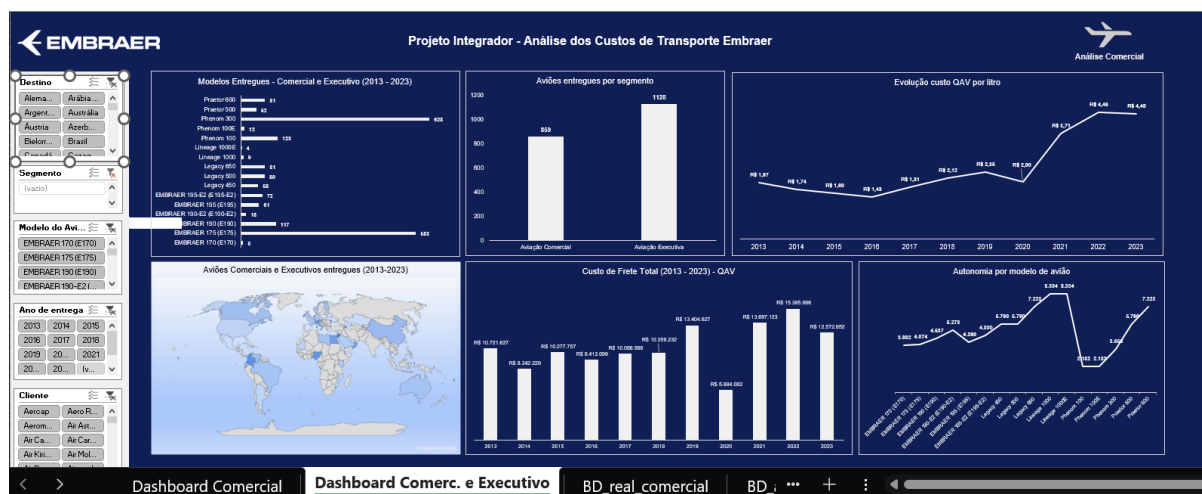


Figura 17 - Dashboard desenvolvido no Excel sobre Custo de Transporte da Aviação Comercial e Executiva.

Observação: todos os filtros funcionam para todos os gráficos apresentados no dashboard.

Conseguimos, também, percorrer ambos os dashboards clicando no avião que está no canto superior direito do dash. O botão facilita o acesso, principalmente quando “escondemos” as abas em uma apresentação comercial.

Confiabilidade dos dados

Para coleta de dados usamos fontes confiáveis, como o site de relacionamento com investidores da Embraer > olhamos apresentações de resultados, releases de entregas e

backlogs e puxamos as especificações de cada um dos modelos de aviões. Para informações adicionais, acessamos artigos online, grandes sites de notícias e relatórios publicados por universidades.

Normalização dos dados:

- Ajustamos tudo o que estava como “jatos leves” e “jatos grandes” na tabela, colocando os modelos respectivos entregues por trimestre na Embraer.
- Criação de novos campos para ajudar na análise dos dados: tamanho do tanque (litro) por modelo, peso do avião vazio por modelo, autonomia por modelo, paradas para abastecer, custo de frete pré e pós-pandemia, custo total de transporte, etc.

Análises e Plano de Ação

Previsão do preço do combustível

Custo do Combustível

Aprofundamos a análise do preço do Querosene de Aviação com o intuito de avaliar as entregas futuras (considerando os pedidos firmes em carteira).

Criamos um [banco de dados](#) com o valor do Querosene de Aviação com todas as semanas desde Janeiro de 2002 a Janeiro de 2004, totalizando 1148 semanas. Os dados foram retirados do site da [ANP - Preços de combustíveis, derivados de petróleo e biodiesel](#).

A Figura 18 mostra a distribuição do preço do QAV.

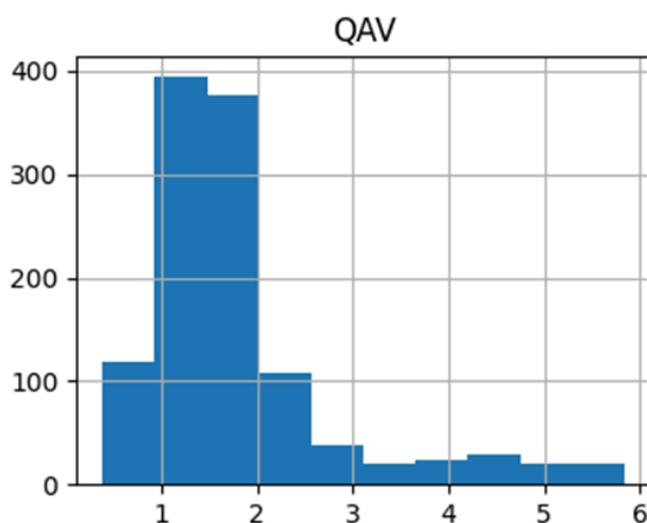


Figura 18 - Distribuição do preço do QAV.

Nota-se que a maior parte dos dados se encontra entre R\$0,80 e R\$2,00.

Na Figura 19, podemos observar a variação do preço do QAV durante os anos.



Figura 19 - Variação do preço do QAV por Ano.

Notamos que os valores aumentam consideravelmente no período pós-pandemia. A fim de considerar uma análise preditiva do preço do QAV, nós consideramos o modelo ARIMA.

Modelo de previsão: ARIMA

O modelo *AutoRegressive Integrated Moving Average* (ARIMA) se destaca como um modelo estatístico poderoso para análise e previsão de dados de séries temporais, e justamente por isso que o escolhemos.

A palavra ARIMA refere-se a composição de uma série temporal. Uma série temporal pode ter um componente autorregressivo (AR) e um componente de médias móveis (MA); e esses componentes podem estar integrados (I).

- Autorregressão (AR): É um modelo que utiliza a correlação entre a observação atual e as observações anteriores (no período de tempo).
- Integrados (I): O uso de diferenciação de observações brutas para tornar a série temporal estacionária.
- Média Móvel (MA): Um modelo leva em consideração a relação entre a observação atual e os erros residuais de um modelo de média móvel aplicado a observações passadas.

Cada um desses componentes é especificado explicitamente no modelo como um parâmetro. Uma notação padrão é usada para ARIMA(p,d,q) onde os parâmetros são substituídos por valores inteiros para indicar rapidamente o modelo ARIMA específico que está sendo usado.

Os parâmetros do modelo ARIMA são definidos da seguinte forma:

- p: A ordem de defasagem, representando o número de observações de defasagem (valores anteriores no tempo) incorporadas no modelo.
- d: Grau de diferenciação, denotando o número de vezes que as observações brutas sofrem diferenciação.
- q: Ordem da média móvel, indicando o tamanho da janela de média móvel.

Quando utilizamos o ARIMA, um modelo de regressão linear é construído incluindo o número e tipo especificados de termos, e os dados são preparados por um grau de diferenciação para torná-los estacionários, ou seja, para remover tendências e estruturas sazonais que afetam negativamente o modelo de regressão.

O modelo ARIMA *default* (padrão) da biblioteca *statsmodel* tem como parâmetros ($p=0$, $d=0$, $q=0$), nós escolhemos os parâmetros ($p=5$, $d=1$, $q=0$), que dentre os analisados resultou menor erro nas avaliação que fizemos.

Então, utilizamos a base com valores de querosene, [iniciamos o modelo ARIMA](#), com o auxílio da biblioteca *statsmodel* e invocando *ARIMA()* e especificamos os parâmetros p , d e q .

Para treinar e testar o modelo, nós separamos a base de dados em dois conjuntos, um de treino - contendo 66% dos dados, e um de teste - contendo o restante dos dados. Para o teste, nós utilizamos valores de 07/2016 até 01/2014.

Geramos as previsões utilizando o conjunto de teste, e comparamos com os valores reais. Essa comparação pode ser observada na Figura 20.

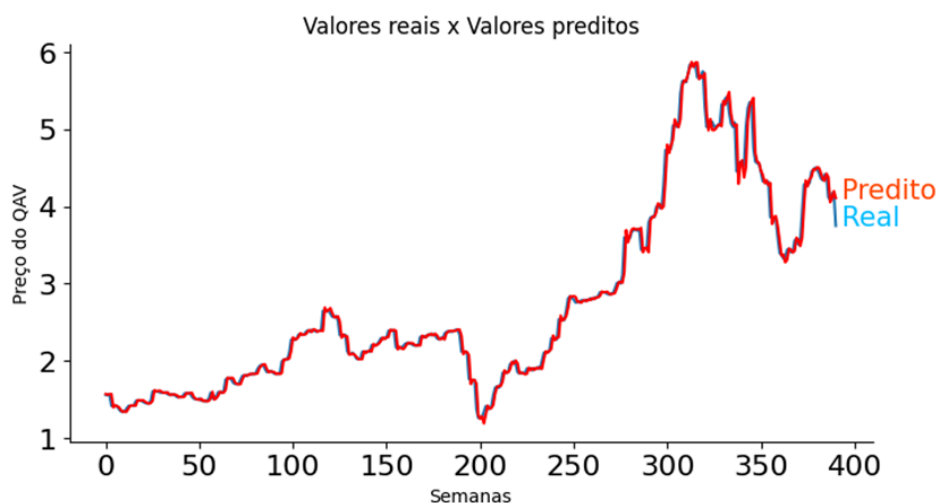


Figura 20 - Comparação entre valores reais e valores preditos pelo modelo ARIMA.

Como podemos notar, as curvas são semelhantes uma das outras, nós obtivemos os valores preditos com erro quadrático médio de 0,1.

Com o treinamento do modelo, nós podemos prever os valores futuros. Na figura 21, nós conseguimos prever o preço do combustível para as próximas 24 semanas.

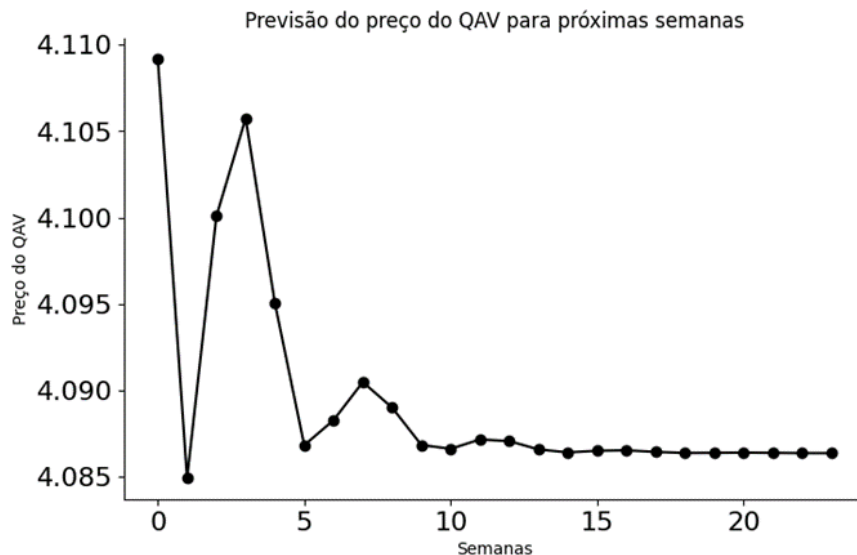


Figura 21 - Previsão do preço do combustível utilizando o modelo ARIMA.

Nós podemos observar que a partir de um certo valor de semanas, as previsões começam a estabilizar e mostrar valores cada vez mais constantes, isso significa que a partir desse valor (aproximadamente 15 semanas) já não é mais possível prever o valor do combustível. Isto se deve a falta de dados referente a este período.

Nós utilizamos essa projeção para estimar o custo de transporte das entregas futuras, considerando pedidos firmes em carteira.

Análise de Indicadores

Entregas de Aeronaves

O ano com maior número de entregas de aeronaves comerciais foi em 2007, com 231 modelos entregues, porém 2021 e 2023 foram anos com uma entrega também positiva, com 183 e 169, respectivamente. Lembrando que estamos olhando números de 2023 consolidados até o 3 trimestre do ano, então temos potencial de aumentarmos esse número.

Com relação às aeronaves executivas, estávamos tendo um período positivo de entrega entre 2013 e 2019, com uma média de 111 jatos por mês. A partir de 2020 esse número

passou para uma média de 86/mês. Apesar disso, a [Embraer afirmou que a demanda de jatos pós pandemia continuou](#), principalmente pela necessidade dos clientes de minimizar a exposição ao coronavírus. A previsão para 2023 era de fechar com a entrega de 130 jatos particulares, atingindo o maior número desde 2010.

Aumento dos custos de combustível

O aumento do custo de QAV pós-pandemia impactou no aumento do valor do frete para a Embraer em 90% (considerando uma média do período 2016 - 2023).

Modelos mais vendidos de Aeronaves

Das Aeronaves Comerciais, os modelos E175 e E190 foram os modelos mais demandados entre os anos de 2003 e 2023. Já das Aeronaves Executivas, o modelo Phenom 300 acabou sendo o principal modelo adquirido, entre os anos de 2013 e 2023. Importante destacar que o Phenom 300E da Embraer é o jato executivo mais vendido do mundo.

Autonomia das Aeronaves

O modelo E170 da Embraer é o que tem menor autonomia entre todos os comerciais analisados e, conseqüentemente, quando olhamos para a média de paradas, acaba tendo a maior quantidade. São 2,28 paradas, em comparação a 1 parada do E190-E2, a aeronave com maior autonomia.

Olhando para os modelos executivos, temos essa mesma visão considerando o Phenom 100 (menor autonomia e 4,26 paradas) vs. o Lineage 1000E (maior autonomia e 0,67 paradas). Aqui entendemos que o tamanho da aeronave impacta bastante no tamanho do tanque, uma vez que consideramos na mesma análise jatos leves e jatos grandes.

Atuação no mercado

Os Estados Unidos são o principal comprador de aviões comerciais da Embraer, com 928 entregas realizadas nos últimos 20 anos, um número 12,5x maior do que o número de aeronaves entregues no Brasil. Quando olhamos o custo médio de frete para cada uma dessas regiões, vemos que os EUA requerem o dobro deste custo (70k) vs. os 35k do Brasil. Apesar de ser a maior fabricante de aeronaves no Brasil, a Embraer exporta 90% de sua mercadoria.

Seria interessante uma análise de estratégia do negócio para ver a viabilidade de investimento comercial e promocional no Brasil. Tanto Gol, quanto Latam são potenciais clientes para os modelos de aeronaves de média distância, e aumentando os clientes nacionais, conseguiríamos ter um impacto positivo no custo total de frete.

Além disso, com a recente [suspensão dos voos da Boeing 747 Max](#) pela ANAC e pela agência nacional americana, pode ser interessante uma abordagem mais agressiva à Copa Airlines (empresa que trabalha com esse modelo de avião em ambos os países) e que ainda não é cliente Embraer. Sugerimos uma aproximação para uma oportunidade de negociação.

Previsão de demanda + custo de frete

Estima-se que a vida útil de um avião comercial é de 30 anos. Por esse motivo, fizemos uma previsão de demanda, para 2035, de aviões comerciais que foram entregues entre os anos de 2003 e 2005, com a intenção de fornecer sempre modelos mais atualizados e sustentáveis para o mercado, garantindo segurança nos voos e menor demanda de manutenção. Serão 117 aeronaves e olhando os maiores clientes, teríamos 22% para a Air Canadá, 22% para a JetBlue, 15% para a US Airways e 13% para a Saudi Arabian Airlines. O custo de frete para essas aeronaves ficará na casa de R\$ 6.376.159,73 considerando uma análise feita pela exame que prevê um aumento de 4% no valor do QAV para cada ponto percentual na redução das emissões. Em 2035 a previsão é de 8% das emissões e, por isso, 32% de aumento no preço do combustível.

Eficiência das entregas

Fazendo uma análise de eficiência das entregas de aeronaves comerciais (pedidos em carteira vs. os pedidos entregues) a Embraer está em um patamar bastante positivo, com 86% (~1.770) dos pedidos entregues até o Q3 de 2023, considerando um total analisado de 2.058 aeronaves. Para os pedidos em carteira, teremos um custo de frete de R\$ 8.704.858,30 (com base no método de regressão ARIMA para os próximos 6 meses que explicamos anteriormente). Esse valor representa 0,01% do valor de venda dos aviões em carteira (78,2 bi).

Modelo de PDCA: A3

O A3 é um template de PDCA (PLAN/DO/CHECK/ACT). Acaba sendo uma maneira visual de mostrar um Kaizen, para organizar todo um projeto, desde o planejamento até um plano de ação. Qualquer pessoa que pegue esse documento deverá entender o que foi feito, quando e como.

Deixaremos uma imagem abaixo representando o A3 e, também, o link do arquivo caso seja necessário acessar.

A3 Problem Solving - SOCIAL TECH CAREERS - EMBRAER		
Título / Tema:	Análise dos Custos de Transporte na Embraer	Análise e Confirmação da Causa-Raiz:
Contexto do Projeto:		Após análise dos indicadores, observamos os seguintes dados: O aumento do QVA pós pandemia gerou um impacto de 90% de aumento nos custos de frete; Entre 2013 -2019 a média de entrega foram 111 Aeronaves Executivas; Em 2007,2021 e 2023 tivemos o maior número de entrega em modelos comerciais; 130 Jatos particulares foram vendidos em 2023, atingindo o maior número desde 2010.
Avaliar os custos de transporte (custos associados ao transporte de mercadorias, que podem incluir transporte terrestre, marítimo ou aéreo) na área de logística e supply chain da Embraer e, com base nos resultados encontrados, fazer um mapeamento do cenário e sugerir um plano de atuação.		
Situação Atual		Situação Futura:
Nossa análises tem como base o cenário geral da Embraer, foco no cenário pós-pandemia e previsões de demanda a partir dos dados. A construção do nosso banco de dados contém uma pesquisa criteriosa, onde buscamos informações reais.		Prospecção de novos clientes, como Gol, Latam e Copa Airlines; Lançamentos de novos modelos aproveitando a nova previsão de demanda; Acompanhamento do custo de frete, projeto de Melhoria Continua na área logística; Buscar oportunidades em financiamentos governamentais, como BNDS para auxiliar na produção, exportação e melhora nos fretes.
Meta / Objetivo:		Plano de Ação:
Analisar Custos de Transporte na Embraer pós-pandemia.		Direcionamento da estratégia comercial para uma aproximação mais agressiva das companhias brasileiras; Focar a venda de aviões com menor autonomia para distâncias mais curtas; Previsão de demanda para 2035.
Identificação de Possíveis Causas / Análises:		Acompanhamento / Validação / Testes:
O cenário pós-pandemia mudou o rumo do mercado no mundo. Os dados apontam os impactos com o aumento do QVA, mas também tivemos uma alta demanda em jatos particulares, por exemplo. Pesquisamos pedidos, clientes, aeronaves, custos de fretes, entre outras informações intrínsecas ao projeto. Trabalhamos com Excel, Power BI, MySql e Python.		Acompanhamento dos dados modelo de previsão de séries temporais ARIMA; Projetos de Melhoria Continua; Projetos focados em captação de novos clientes.
		Stakeholders:
		Giovanna de Luca Isabella Qualharello Natasha Brandão Tahine Santos Thamiris Neves

Link do arquivo: [X A3 - EMBRAER SOCIAL TECH CAREERS.xlsx](#)

Plano de Ação

Dividimos o nosso plano de ação em 3 principais estratégias:

1. Direcionamento da atuação comercial para uma aproximação mais agressiva das companhias brasileiras (Latam e Gol), assim como aproximação com a Copa Airlines, aproveitando o momento de vulnerabilidade do mercado de aviação.
2. Focar a venda de aviões com menor autonomia para distâncias mais curtas (diminuindo a quantidade de paradas). Um exemplo: O E170 é o avião que possui menor autonomia e, conseqüentemente, uma maior quantidade de paradas para abastecimento. Seria um modelo interessante para direcionarmos as vendas para América do Sul. Já o E190-E2, o modelo que possui maior autonomia, poderia ser direcionado para a Europa. Essa ação conseqüente impactaria na diminuição dos custos de frete.
3. Fizemos uma previsão de demanda para 2035, com base nos 30 anos de vida útil do avião. A Embraer poderia aproveitar a previsão para lançar os modelos de aviões da família de energia elétrica e aviões híbridos em 2030 e oferecer modelos mais sustentáveis e tecnológicos. O intuito dessa ação seria de introduzir o novo portfólio para o mercado e iniciar essas vendas por clientes “fiéis”, que já são clientes da empresa há décadas. Podemos até realizar essa ação aplicando gatilhos mentais de exclusividade e gerando um certo “burburinho” no mercado.

Referências // Anexos

[2022 EPI Results](#)

[Aeronaves poderão ter motores à eletricidade e biocombustíveis](#)

[Análise logística de regime aduaneiro utilizado em uma empresa nacional de produção
aviônica - Everton Alves de Araujo \(UFAM\)](#)

[ANP - Preços de combustíveis, derivados de petróleo e biodiesel](#)

[Aviões a hidrogênio podem ser os mais baratos já em 2035](#)

[Azul Conecta apresenta motor híbrido na LABACE](#)

[Cálculo de carreta](#)

['Carros voadores' com 12 milhões de passageiros por ano: as metas para o eVTOL da Embraer](#)

[Como calcular custo de transporte de cargas na prática?](#)

[Como funciona a tabela de preços do transporte de carga aérea no Brasil](#)

[Como uma aeronave é fabricada? - Tecmundo](#)

[Conheça os aviões mais antigos que estão na frota de Azul, Gol e Latam](#)

[Conheça os bastidores da produção de um avião na Embraer](#)

[De São Paulo a Amsterdã em um E175: como é feita a entrega de um novo avião da Embraer](#)

[Documento Projeto Integrador - Social Tech Embraer](#)

[Embraer anuncia nova fábrica de carro voador no interior de São Paulo - IPEA](#)

[Embraer Day: Embraer Investor Relations](#)

[Embraer: Relações com Investidores](#)

[Embraer: Todas as aeronaves da Embraer](#)

[Embraer: um caso de sucesso com o apoio do BNDES - Paulus Vinícius da Rocha Fonseca](#)

[Entenda por que a Embraer vende mais avião no exterior do que no Brasil - Exame](#)

[Estrutura de aeronaves: análise do aumento da utilização de materiais compósitos - Denis Rodrigues da Rocha](#)

[EPE - Projeção de preços internacionais de petróleo e derivados: 2020 - 2030](#)

[Frete Marítimo](#)

[Frota de aviões elétricos deve suprir toda a Noruega até 2040](#)

[iNFRADebate: Sustentabilidade na aviação civil – rumo a um futuro mais verde](#)

[Operação da Embraer exige logística complexa - Investe SP](#)

[O que é SAF? Conheça as diferentes rotas de combustível sustentável de aviação](#)
[Perspectivas e Desafios para Projetos de Exportação de Hidrogênio Verde](#)

[Petrobras aumenta preço do querosene de aviação \(QAV\) em 17,1%](#)

[Petrobras elevou preço de querosene de aviação em 17,1% nas refinarias, diz Abear](#)

[Piauí: Passagens nas alturas](#)

[Poder 360: Petrobras aumenta preço do querosene de aviação em 17%](#)

[Preço do Etanol no Brasil 2023: veja os valores em cada estado!](#)

[Preços do diesel e do querosene de aviação podem aumentar com projeto de incentivo aos biocombustíveis, diz governo](#)

[Quanto custa o litro do Avgas?](#)

[Site Institucional Embraer](#)

[Tabela de Preços Viracopos](#)

[ARIMA - Modelo de Previsão](#)