

**Explorando Padrões do Comércio Exterior da Baixa Mogiana: Uma análise
Baseada em Dados.**

Juliano Alessandro dos Santos

Mateus Guilherme Fuini

Fatec Itapira - "Dr. Ogari de Castro Pacheco"

Agosto/2023 até Dezembro/2023

*Dedico esta conquista acadêmica a
minha mãe, Neuza Francisca dos Santos.
Você foi uma pessoa fundamental nesta
jornada de aprendizado.*

*Agradeço pela compreensão e
carinho.*

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a Deus, por me guiar e me apoiar nesta trajetória de aprendizado.

À minha querida mãe, Neuza Francisca dos Santos, no qual eu dedico este trabalho acadêmico por ser uma pessoa forte e trabalhadora que abdicou de alguns de seus sonhos para poder cuidar de mim e me dar uma educação de qualidade e por me ensinar a nunca desistir e persistir nos meus sonhos.

Ao meu Orientador Mateus Guilherme Fuini, por me proporcionar esta oportunidade de extensão acadêmica, apresentando para mim esse universo de pesquisa acadêmica através desta monografia.

À professora Naely Iamarino Pizzi Cazarin, por me auxiliar na correção do presente relatório verificando possíveis erros de ortografia, concordância e informações que fogem da regra da ABNT.

A instituição acadêmica Fatec Itapira - "Dr. Ogari de Castro Pacheco", por me proporcionar essa oportunidade de me aprofundar em uma área de pesquisa de meu interesse.

*“Se eu vi mais longe, foi por estar sobre
ombros de gigantes.”*

(Sir Isaac Newton-1675)

Explorando Padrões do Comércio Exterior da Baixa Mogiana: Uma análise Baseada em Dados.

Juliano Alessandro dos Santos
Mateus Guilherme Fuini
Fatec Itapira - "Dr. Ogari de Castro Pacheco"

Resumo: Este estudo, referente à área de engenharia de dados, tem como objetivo realizar a extração, processamento e análise dos dados da Baixa Mogiana, com isso permitindo auxiliar na compreensão de tendências e padrões. Em virtude disso, é notório citar a relevância que esses dados terão, haja visto que uma empresa pode compreender de maneira mais clara e dinâmica o comportamento das exportações de determinada cidade ou estado ao longo do tempo. Para poder atingir tal objetivo, todo o estudo será contemplado com uma metodologia descritiva dos dados, ou seja, iremos recolher as informações e apresentar todos esses dados não se preocupando do porquê. Em virtude disso, serão apresentados os resultados de como foram feitos os processos de extração e limpeza, além de apresentar um painel de dados no qual, demonstra todos os dados cruzados organizados e em formato de gráfico, com filtro de tempo para compreender ainda melhor essa dinâmica. Portanto, podemos inferir que os resultados foram obtidos e que os limites dos meus estudos são pautados pelas limitações dos dados, no qual não foi possível cruzar as informações de transporte, além disso, informações de importação não foram trazidas devido ao limite de tempo estabelecido pelo projeto.

Palavras-chave: exportação; processamento; limpeza; análise; padrões.

LISTA DE FIGURAS

Figura 1- Ilustração do Processo de ETL	3
Figura 2 - Ilustração de um processo de Data Warehouse	4
Figura 3 - Apresentação Didática do Modelo Estrela	5
Figura 4 - Resumo do Projeto desenvolvido	6
Figura 5 - Ilustração dos Códigos Necessários para a extração dos arquivos CSV	8
Figura 6 - Ilustração do Banco de Dados no formato Estrela	9
Figura 7 - Ilustração do Processo de junção e seleção de colunas específicas	10
Figura 8 - Demonstração do processo de Extração de Duplicidade de dados	10
Figura 9 - Comando para a criação de Chave Primária	11
Figura 10 - Comandos para a criação da chave estrangeira e mudança de tipagem.	11
Figura 11 - Ilustração do Processo de Junção e Adição de Colunas no Pentaho	12
Figura 12 - Demonstração dos Comandos SQL para a granularização dos campos	13
Figura 13- Ilustração do Processo de Limpagem de dados no Pentaho	13
Figura 14 - Escolha das cores dos painéis de visualização	14
Figura 15 - Desempenho Geral das Exportações de mercadoria	15
Figura 16 - Painel de Análise Estatística para Engenheiro Coelho	19
Figura 17 - Painel de Análise Estatística para a cidade de Estiva Gerbi	20
Figura 19 - Painel de Análise Estatística para a cidade de Itapira	21
Figura 20 - Painel de Análise Estatística para a cidade de Mogi Guaçu	21
Figura 21 - Painel de Análise Estatística para a cidade de Mogi Mirim	22
Figura 22 - Painel de Análise Estatística para Santo Antônio de Posse	23

SUMÁRIO

1 INTRODUÇÃO	1
2 FUNDAMENTAÇÃO TEÓRICA.....	2
2.1 COMPREENSÃO DOS DADOS	2
2.2 ENGENHARIA DE DADOS	2
2.2.1 PIPELINES DE DADOS(ETL)	3
2.3 DATA WAREHOUSE.....	4
2.4 ANÁLISE DE DADOS.....	5
2.5 PENTAHO	5
3 METODOLOGIA.....	5
4 RESULTADOS E DISCUSSÕES	7
4.1 EXTRAÇÃO DE DADOS	7
4.2 TRANSFORMAÇÃO DOS DADOS	8
4.2.1 SELEÇÃO DAS COLUNAS DESEJADAS.....	9
4.2.2 REMOÇÃO DE VALORES DUPLICADOS E CRIAÇÃO DE CHAVES PRIMÁRIAS	10
4.2.3 MUDANÇA DA TIPAGEM DOS DADOS	11
4.2.4 ADIÇÃO DE COLUNA E JUNÇÃO DE ARQUIVOS	11
4.2.5 ALTERAÇÃO DOS NOMES DAS TABELAS	12
4.2.6 GRANULARIZAÇÃO DOS CAMPOS - DIMENSÃO DE DATAS	12
4.2.7 REMOÇÃO DE CARACTERES ESPECIAIS	13
4.3 VISUALIZAÇÃO DOS DASHBOARDS	14
5 CONCLUSÃO	15
6 REFERÊNCIAS BIBLIOGRÁFICAS	16

APÊNDICE A - ANÁLISE DE DADOS	19
A.1 ENGENHEIRO COELHO	19
A.3 ITAPIRA	21
A.5 MOGI MIRIM.....	22
A.6 SANTO ANTÔNIO DE POSSE	23

1 INTRODUÇÃO

A consolidação do Brasil no âmbito mundial de exportação agrícola ocorreu com a eliminação das barreiras não tarifárias nos anos 1990(LIMA; RIBEIRO, 2016,p.1). Nesse Contexto, conforme mencionado por Garcia(2019, p. 20), no, período pré-crise de 2008, as exportações totais do Brasil Somavam US\$ 197 Bilhões, no ano seguinte pós-crise econômica de 2008, este valor teve uma queda passando a corresponder US\$ 152 Bilhões, além disso, as importações tiveram uma variação nesse período registrando US\$ 172 Bilhões; em 2008 e em 2009, US\$ 127 Bilhões. Portanto, pelos dados trazidos pelo autor é perceptível o protagonismo da Exportação no mercado Mundial.

Desse modo, conforme destacado por Abreu(2015, p. 1) a exportação é responsável por impulsionar o ramo agrícola a se desenvolver cada vez mais. Nesse contexto, em 2014 o agronegócio foi responsável por 44% das exportações(ABREU,2015,p.3). Além disso, o número de parceiros comerciais do Brasil cresceu ainda mais chegando a 150 países, sendo que os produtos mais exportados são: café, grãos, carnes, sucos, celulose, madeiras, frutas e hortaliças in natura de acordo com Abreu(2015,p. 3), demonstrando dessa forma a relevância que o país tem no setor agrícola.

Em suma, destaca-se a relevância de se ter Dashboards Interativos, para a visualização das inúmeras informações que englobam o comércio exterior como: Receita obtida por exportação, países que mais recebem exportações, e quais produtos que são mais exportados. Desse modo, a ideia central do projeto é reunir os diversos arquivos CSVs presentes no site do Governo, permitindo criar os relacionamentos entre esses arquivos, que antes não existiam para que dessa forma possa gerar percepções que antes não eram tão visíveis em virtude de os arquivos estarem separados entre si.

Na trajetória de desenvolvimento deste trabalho, foi crucial identificar as perguntas fundamentais que orientaram a construção do Objeto de Estudo proposto nesta monografia. Nesse contexto, as seguintes questões desempenharam um papel central na elaboração do projeto.

- Os Dashboards Desenvolvidos permitiram uma visualização clara e objetiva dos diversos arquivos CSV presentes no site do governo?

- Foi possível identificar tendências e padrões nos dados observados?

O objetivo deste projeto é realizar a extração, processamento, armazenamento, análise e visualização de dados do comércio exterior da região da Baixada Mogiana, com o intuito de fornecer identificações de padrões, tendências e insights relevantes para auxiliar na compreensão e tomada de decisões estratégicas, no âmbito do comércio internacional, para empresas de todos os segmentos econômicos.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão abordados os principais conceitos que abrangem a área de Ciência de Dados, que foram de suma importância para a elaboração do Produto final, ou seja, os dashboards interativos dos dados.

2.1 COMPREENSÃO DOS DADOS

Os dados isolados sem nenhum contexto associado, não possuem nenhum valor agregado os mesmos só ganham valor quando se torna informação, ou seja, quando são colocados em algum contexto e tratados pelo profissional específico da área (SANTOS D'AMORIM et al., 2020 apud NUNES, 2022, p. 9).

2.2 ENGENHARIA DE DADOS

Em conformidade com o parágrafo anterior podemos definir que a área de engenharia de dados está preocupada em: extrair, transformar, limpar e analisar esses dados, para que dessa forma possa-se obter o produto final que é a informação (RAUDEMBERG; CARMO, 2019 apud NUNES, 2022, p. 12).

Dessa forma, a área de dados está preocupada em resolver um problema a partir de uma imensa quantidade de informações, com isso o profissional dessa área fica responsável por minimizar perdas em uma empresa, auxiliar na tomada de decisão a partir de dados e sugerir propostas e melhorias de acordo com o cenário vislumbrado pelos dados tratados (AMARAL, 2016 apud SOUZA et al., 2021, p. 15).

2.2.1 PIPELINES DE DADOS (ETL)

A função primordial de um engenheiro de dados é tornar os dados úteis para posteriormente serem utilizados por um analista de dados (DHAR, 2013 apud ARRAIS, 2022, p. 24).

Nesse contexto, construir pipeline de dados eficientes se torna essencial na vida de um engenheiro de dados, ademais podemos entender pipelines de dados, como o transporte de dados de uma origem, que pode ser diverso, até um determinado destino comum, no qual todos os dados sofreram processos de ETL (RAJ, et al., 2020 apud ARRAIS, 2022, p. 25).

O termo ETL faz referência aos processos de Extração, Transformação e Carregamento dos dados, conforme exemplificado na figura 1, no qual a extração significa a retirada dos dados de localidades variadas, enquanto a Transformação são os processos de limpeza como: retirada da duplicidade de dados, eliminação de caracteres indesejados, exclusão de colunas que não precisam ser usadas. Por último temos o carregamento, que significa armazenar esses dados tratados em um banco de dados (JO; LEE, 2019 apud ARRAIS, 2022, p. 25).

Figura 1- Ilustração do Processo de ETL



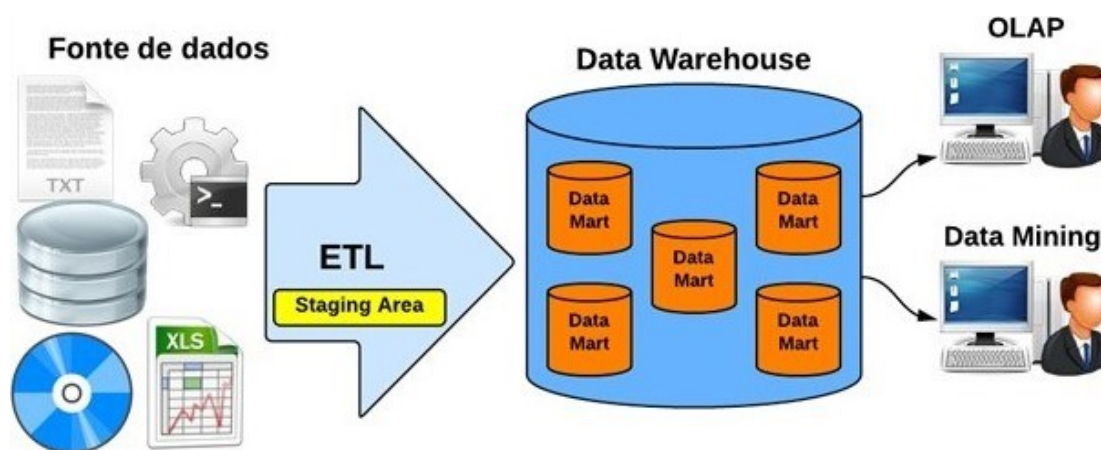
Fonte: NABEREZNY (2022)

2.3 DATA WAREHOUSE

Em virtude da imensa quantidade de dados que são gerados diariamente por empresas, torna-se necessário um banco de dados capaz de ser rápido, flexível e intuitivo na hora de realizar o retorno das consultas, em virtude disso surgem os bancos de dados orientados a processamento analítico online (OLAP) (VAISMAN; ZIMANYI, 2016 apud ARRAIS, 2022,p. 27).

Para conseguir atender este novo Paradigma, emerge os *Data Warehouse*, um banco de dados que armazena diversos tipos de fontes em um único local, para posteriormente serem utilizado (MOREIRA, et al., 2006, p. 2), na figura 2 é apresentado a ilustração de um DW, no qual apresenta a ilustração do processo mencionado anteriormente, além disso é mostrado que esses dados tratados podem ter dois destinos: um OLAP, citado anteriormente ou um Data mining.

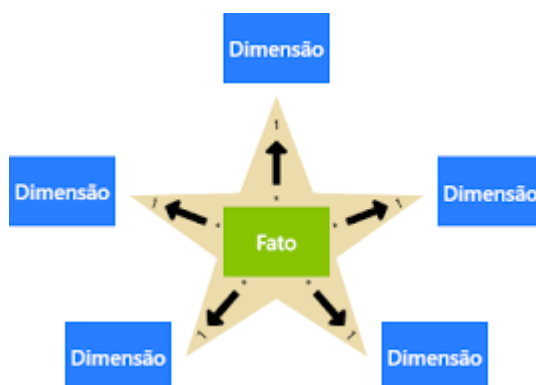
Figura 2 - Ilustração de um processo de Data Warehouse



Fonte: ELIAS (2014)

Nessa linha de pensamento, para a construção de um DW, é necessário padronizar o mesmo em um modelo multidimensional, no qual as tabelas presentes no banco de dados devem ser chamadas de Dimensão, quando se refere a algum assunto específico, ou seja é a identificação dos elementos presentes naquela tabela, enquanto a tabela fato, se refere a coleção de eventos que ocorrem ao longo do tempo, no qual as tabelas dimensões participam (MACHADO, 2016, p.79 apud SOUZA et al., 2021, p. 36). Na figura 3, está sendo ilustrado a explicação anterior

Figura 3 - Apresentação Didática do Modelo Estrela



Fonte: ANTONCHUK(2020)

2.4 ANÁLISE DE DADOS

Na sociedade atual, o dado é o novo petróleo da humanidade, por isso torna-se necessário um profissional capaz de analisar essas informações, realizar planejamentos estratégicos e tomar decisões estratégicas com base nos dados analisados, tal prática é chamada de data driven (FÁVERO; BELFIORE, 2017 apud ARRAIS, 2022, p. 44).

2.5 PENTAHO

Ferramenta de inteligência empresarial de código aberto escrita em linguagem JAVA, destinado a otimizar e facilitar processos de ETL (ROCHA, 2015, p. 22). Ademais, é uma tecnologia orientada a metadados que cria telas responsáveis por realizar determinadas regras de negócios (ROCHA, 2015, p. 18 apud PENTAHO, 2012).

3 METODOLOGIA

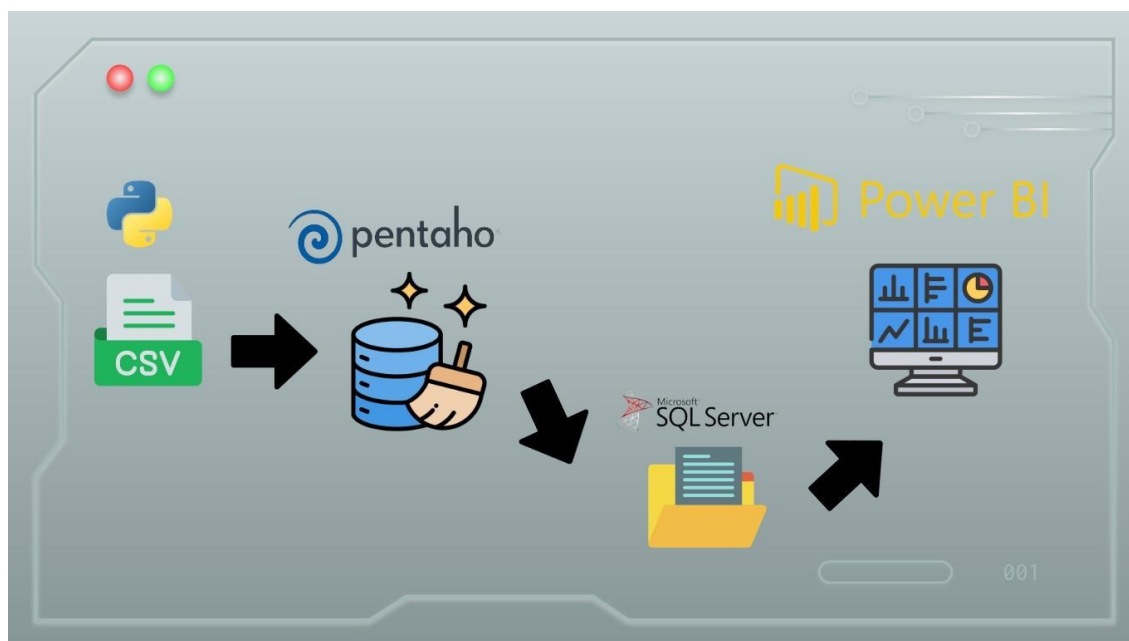
A metodologia que contempla essa iniciação é a quantitativa, visto que toda a análise deste projeto será pautada em um estudo descritivo dos dados, em outras palavras utilizaremos todos os dados apresentados no site: Estatísticas de Comércio Exterior em Dados Abertos para realizar ações com base em técnicas estatísticas. Desse modo, teremos as receitas obtidas através dos produtos exportados para os diversos países.

Toda a análise de dados desse Projeto foi concentrada na região da baixa Mogiana, abordando dessa maneira as seguintes cidades: Itapira, Mogi Guaçu, Mogi Mirim, Estiva Gerbi, Engenheiro Coelho e Santo Antônio de Posse.

O projeto acadêmico foi conduzido nas dependências da instituição Fatec de Itapira - Ogari de Castro Pacheco, nos seguintes dias: Segunda e quarta nos horários das 14h00 às 18h00.

Para obter os dados que foram analisados, foi necessário coletar os arquivos CSV, apresentados no site: Estatísticas de Comércio Exterior em Dados Abertos, utilizando-se a linguagem de Programação Python (EXTRAÇÃO) , com os dados obtidos foi necessário realizar o tratamento dos dados através de uma ferramenta de Limpeza chamada Pentaho (TRANSFORMAÇÃO), para que dessa forma pudéssemos armazená-los em um banco de dados relacional, seguindo o modelo estrela, em seguida realizamos a importação dessa base de dados em um Power BI (CARREGAMENTO), para assim realizar a análise dos dados, na figura 4 é apresentado o resumo de todas as etapas citadas anteriormente no texto.

Figura 4 - Resumo do Projeto desenvolvido



Fonte: Próprio Autor (2023)

4 RESULTADOS E DISCUSSÕES

Esta seção será destinada à apresentação dos resultados obtidos. Em virtude disso, será lembrado que o intuito desse projeto é realizar a extração, processamento, armazenamento e visualização de dados do comércio exterior da região de Itapira, com o intuito de permitir que profissionais da área de Comércio Exterior possam identificar padrões, tendências e insights relevantes para auxiliar na compreensão e tomada de decisões, tendo como foco da pesquisa a Baixa Mogiana (Itapira, Mogi Guaçu, Mogi Mirim, Estiva Gerbi, Engenheiro Coelho e Santo Antônio de Posse.). Ademais, utilizaremos a metodologia proposta na seção anterior desse projeto, para poder apresentar e discutir os principais aspectos da exportação de produtos.

4.1 EXTRAÇÃO DE DADOS

A primeira etapa deste projeto foi realizar a coleta de dados dos arquivos CSVs, presentes no site: Estatísticas de Comércio Exterior em Dados Abertos, para fazer tal etapa foi necessário importar algumas bibliotecas do Python. Desse modo, foi utilizado as seguintes Bibliotecas:

- **Requests:** responsável por realizar a requisição no site para assim podermos extrair os dados.
- **Pandas:** utilizado para poder visualizar os arquivos de maneira mais rápida e agradável no formato tabular.
- **Io:** Utilizado para lidar com diversos tipos de arquivos de texto, binários e dados brutos.
- **urllib3:** Essa biblioteca foi de utilidade fundamental no decorrer das atividades. Sempre que fizemos uma requisição ao servidor, algum tipo de erro pode acontecer no caminho impedindo assim nossa requisição. Nessa linha de pensamento, toda vez que eu fazia uma requisição ao servidor, eu era impedido de acessar o conteúdo por motivos de SSL, que tem o intuito de criptografar os dados e permitir segurança dos dados. Então, ele me impediu de pegar esses dados por segurança. No entanto, utilizando a referida biblioteca, pude ignorar essa chamada de atenção, conseguindo trazer os dados para o ambiente de

desenvolvimento. Na figura 5 estão ilustrados os códigos utilizados para a obtenção dos arquivos necessários para a próxima etapa do projeto, ou seja, a transformação desses dados.

Figura 5 - Ilustração dos Códigos Necessários para a extração dos arquivos CSV

```
import requests
import pandas as pd
import io
import urllib3

http = urllib3.PoolManager(maxsize=10)
urllib3.disable_warnings()

dfs = []
anos = range(2018,2024)

for ano in anos:
    print(f"\nDados de Exportação do ano de {ano}")

    url = f'https://balanca.economia.gov.br/balanca/bd/comexstat-bd/mun/EXP_{ano}.MUN.csv'
    arquivos = requests.get(url, verify=False)
    arquivos.encoding = 'utf-8'

    csvio = io.StringIO(arquivos.text, newline='')
    df = pd.read_csv(csvio, delimiter=',')
    dfs.append(df)
    print(df.head(3))

combined_df = pd.concat(dfs, ignore_index=True)

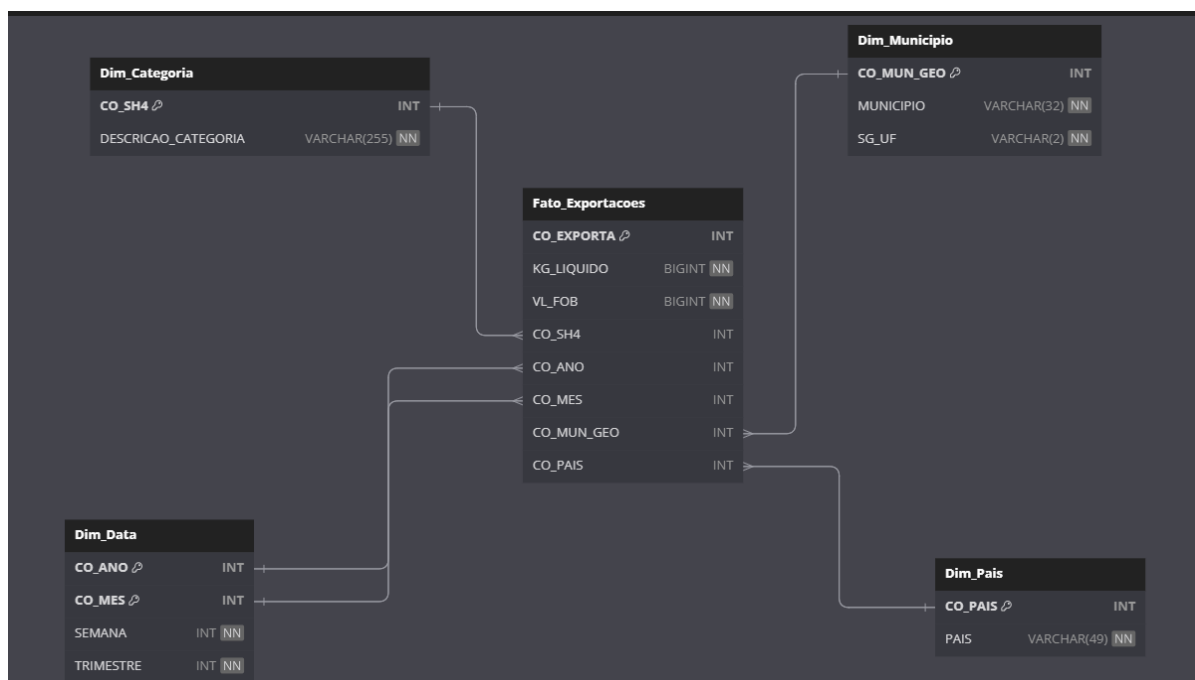
combined_df.to_csv('C:\\Users\\juliano\\Documents\\ExtracaoDados\\Exportacoes20182023.csv', index=False)
```

Fonte: Próprio Autor (2023)

4.2 TRANSFORMAÇÃO DOS DADOS

Nesta segunda etapa do projeto foi realizado o tratamento dos dados, para logo em seguida o mesmo pudesse ser armazenado em um data warehouse. Para realizar o processo de transformação dos dados foi utilizado o Pentaho, uma ferramenta de suma importância para realizar o processo de ETL, mencionado na Fundamentação teórica do referido projeto, em conjunto com o banco de dados SQL Server. Portanto, todas as limpezas que serão mencionadas neste capítulo têm como intuito, construir um banco de dados multidimensional no modelo estrela, cujo resultado é demonstrado na figura 6.

Figura 6 - Ilustração do Banco de Dados no formato Estrela

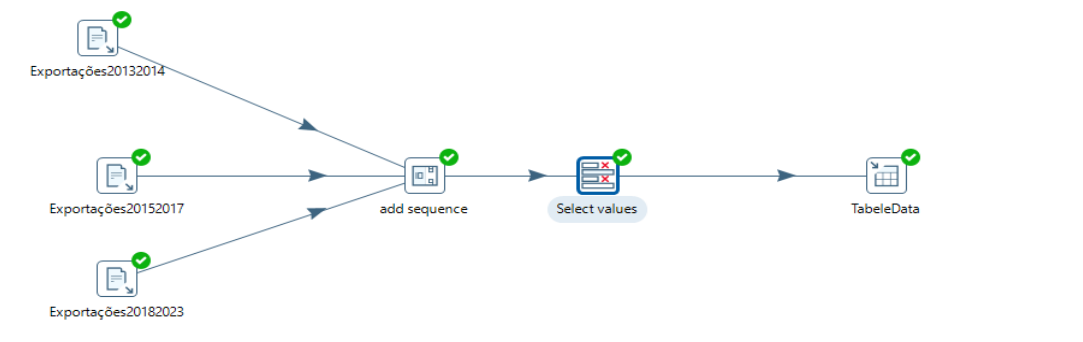


Fonte: Próprio Autor (2023)

4.2.1 SELEÇÃO DAS COLUNAS DESEJADAS

Para a criação da tabela **Dim_Data**, foi necessário selecionar algumas colunas desejadas, como "CO ANO" e "CO_MES", de um arquivo CSV que possui mais de 7 campos. Logo, foi necessário excluir alguns campos e priorizar os dois campos mencionados utilizando a função **SELECT VALUES** do Pentaho. Portanto, a figura 7 apresenta o resultado obtido.

Figura 7 - Ilustração do Processo de junção e seleção de colunas específicas

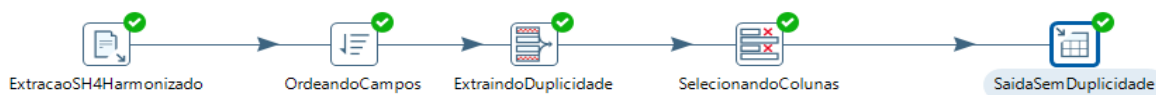


Fonte: Próprio Autor (2023)

4.2.2 REMOÇÃO DE VALORES DUPLICADOS E CRIAÇÃO DE CHAVES PRIMÁRIAS

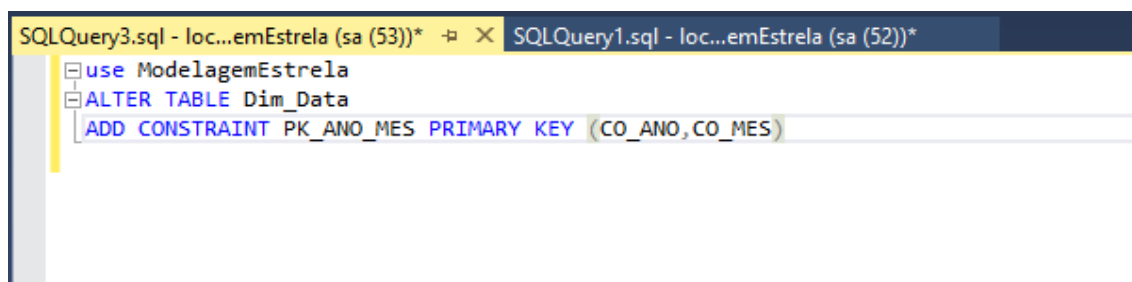
Como estamos normalizando os dados para o Padrão Relacional, torna-se necessário um campo do tipo PK em nossa base de dados. Para garantir que um campo seja considerado PK, o mesmo tem algumas restrições que devem ser seguidas para garantir a integridade dos relacionamentos. Uma dessas restrições é que a chave não tenha valor duplicado, portanto, foi necessário fazer a retirada de valores duplicados ao longo desta tabela. Em seguida, no SQL Server, alterar o tipo do campo para PK. Na figura 8 e 9 serão ilustrados tais processos.

Figura 8 - Demonstração do processo de Extração de Duplicidade de dados



Fonte: Próprio Autor (2023)

Figura 9 - Comando para a criação de Chave Primária



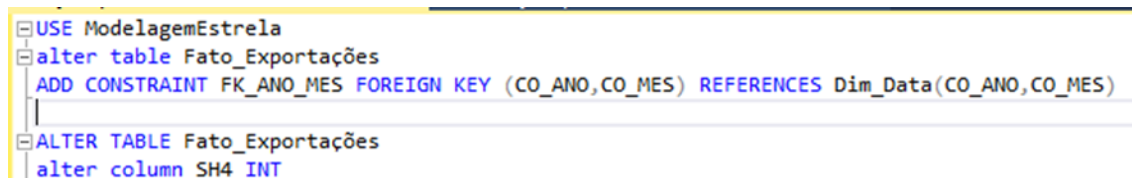
```
SQLQuery3.sql - loc...emEstrela (sa (53))* X SQLQuery1.sql - loc...emEstrela (sa (52))*
use ModelagemEstrela
ALTER TABLE Dim_Data
ADD CONSTRAINT PK_ANO_MES PRIMARY KEY (CO_ANO,CO_MES)
```

Fonte: Próprio Autor (2023)

4.2.3 MUDANÇA DA TIPAGEM DOS DADOS

Ao criar uma chave estrangeira para assegurar relacionamentos com a tabela fato, é crucial que os dados sejam compatíveis a fim de garantir essa conexão. Em algumas situações, foi necessário ajustar a tipagem dos dados nas tabelas de dimensões para assegurar essa correspondência, com isso na Figura 10 serão apresentados os comandos realizados para permitir esses relacionamentos.

Figura 10 - Comandos para a criação da chave estrangeira e mudança de tipagem.



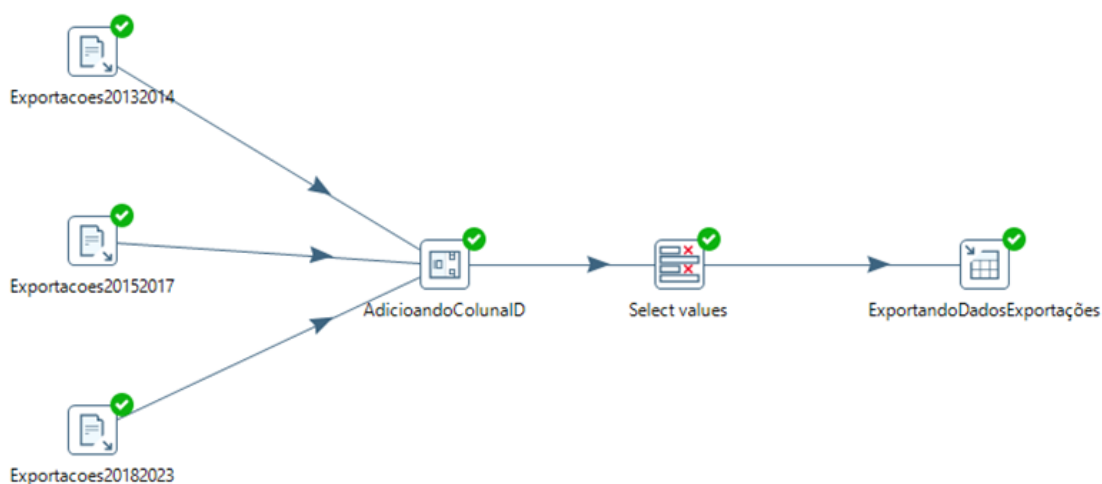
```
USE ModelagemEstrela
alter table Fato_Exportações
ADD CONSTRAINT FK_ANO_MES FOREIGN KEY (CO_ANO,CO_MES) REFERENCES Dim_Data(CO_ANO,CO_MES)
|
ALTER TABLE Fato_Exportações
alter column SH4 INT
```

Fonte: Próprio Autor (2023)

4.2.4 ADIÇÃO DE COLUNA E JUNÇÃO DE ARQUIVOS

Na construção da tabela fato, foi necessário utilizar um identificador desse campo, já que o mesmo não possuía um campo que realizasse determinada ação antes de ser tratado. Com isso, o campo ID EXPORTAÇÕES foi criado. Além disso, como realizamos a extração de diversos arquivos, tivemos que juntar esses dados de exportação de 2013 até 2023 em um único arquivo, como ilustrado na figura 11.

Figura 11 - Ilustração do Processo de Junção e Adição de Colunas no Pentaho



Fonte: Próprio Autor (2023)

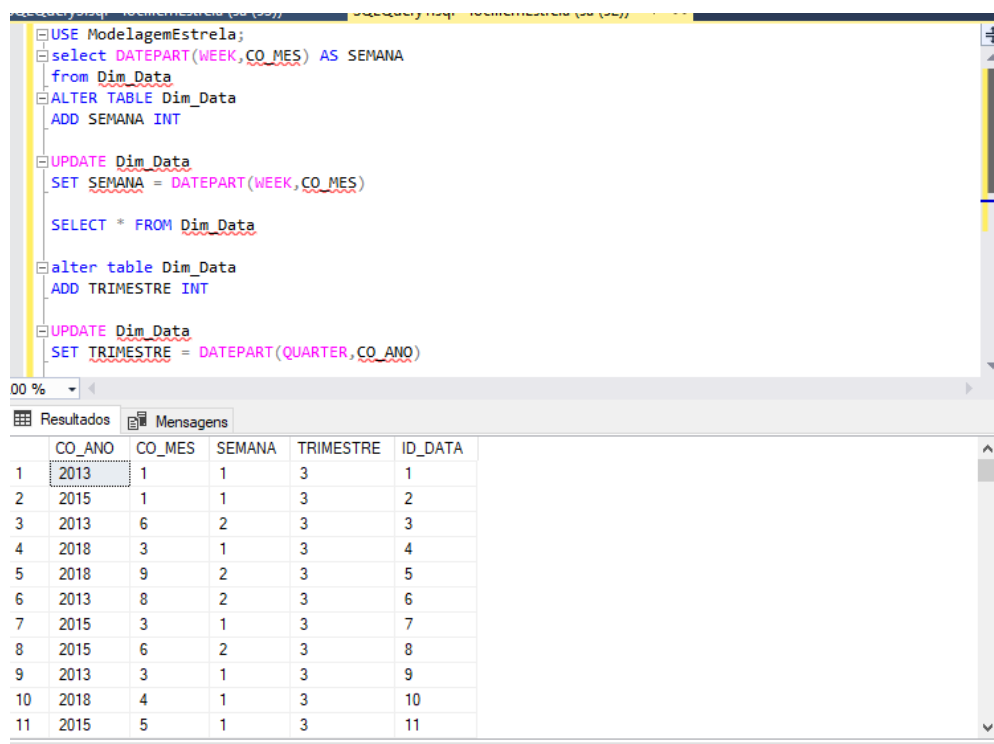
4.2.5 ALTERAÇÃO DOS NOMES DAS TABELAS

Para seguir a padronização do modelo Estrela, que consiste em ter tabelas com o sufixo "Dim" e "Fato", foi necessário alterar o nome das tabelas extraídas do site do governo. Isso foi realizado pelo Pentaho, permitindo assim enviar tal alteração para o SQL SERVER.

4.2.6 GRANULARIZAÇÃO DOS CAMPOS - DIMENSÃO DE DATAS

O processo de granularização envolve o aumento do nível de detalhamento de uma tabela, expandindo o número de campos. Esse procedimento adquire extrema relevância em cenários de larga escala, onde auxilia as empresas na tomada de decisões fundamentais. Nesse contexto, os campos CO_ANO e CO_MES passaram por essa operação, conforme ilustrado na figura 12.

Figura 12 - Demonstração dos Comandos SQL para a granularização dos campos



Fonte: Próprio Autor (2023)

4.2.7 REMOÇÃO DE CARACTERES ESPECIAIS

O arquivo Categoria após ser extraído apresentou nos campos um conjunto de aspas antes do texto, logo tornou-se necessário utilizar a função do Pentaho que limpa esses caracteres indesejados. Na figura 13 é possível visualizar como foi feito esse procedimento.

Figura 13- Ilustração do Processo de Limpagem de dados no Pentaho



Fonte: Próprio Autor (2023)

4.3 VISUALIZAÇÃO DOS DASHBOARDS

As escolhas das cores para o Dashboard foram pautadas nas cores da Fatec, através do site: Wibushi no qual o usuário insere uma imagem, retornando dessa maneira as cores presentes na figura. A figura 14, apresenta a paleta de cores que o site retornou para as cores do logo da instituição.

Figura 14 - Escolha das cores dos painéis de visualização



Fonte: Próprio Autor (2023)

Na figura 15, é apresentado o Dashboard produzido, nele é possível observar a divisão da tela em cards que apresentam dados estatísticos das exportações como: Maior valor obtido por exportação, média das exportações e soma total das exportações geradas. Além disso, o painel é dividido em 4 gráficos sendo eles: cidades, no qual apresenta as cidades que mais exportaram, é válido reforçar que esta monografia aborda apenas as cidades da baixa Mogiana, por isso através desse gráfico, podemos visualizar quais destas mais exportam.

O segundo gráfico apresentado demonstra quais produtos são mais exportados nesse grupo de 6 cidades analisadas.

O terceiro gráfico elucida quais são os países que mais contribuem para as exportações, e por fim temos o quarto e último gráfico que apresenta a evolução dos lucros obtidos de exportação ao longo do tempo.

O Dashboard apresenta um filtro de anos, no qual podemos verificar se essa tendência de determinada cidade exportar mais persiste ao longo de um determinado intervalo de tempo, também é possível mensurar se essa tendência vai se manter para os produtos e

parceiros comerciais. Portanto, podemos inferir que tal análise da figura 15, corrobora para responder tanto a primeira quanto a segunda questão trazida na introdução.

Figura 15 - Desempenho Geral das Exportações de mercadoria



Fonte: Próprio Autor (2023)

Além disso, é possível perceber uma barra lateral no painel, no qual serve de navegabilidade para o usuário, ao clicar em uma das figuras tanto a imagem da cidade quanto na figura do Brasil é possível acessar informações específicas para cada cidade mencionada na introdução, ademais clicando na figura do Brasil, o usuário será capaz de analisar o desempenho dos países presentes no Mercosul para cada cidade, portanto ele será capaz de visualizar qual País do Mercosul mais contribui por cidade, qual o produto mais importado para esse bloco econômico. Portanto, o dashboard apresentado cruza diversas informações e apresenta de maneira mais ilustrativa possível, com isso a empresa que tiver essas informações em mão será capaz de averiguar tendências e padrões, através de uma análise profunda.

5 CONCLUSÃO

Ao final desta monografia, é perceptível a relevância desses dados contidos no site do governo para a gestão de atividades comerciais de uma empresa. Nessa mesma linha de pensamento, é notório perceber a importância de uma empresa ter todas as informações centralizadas e cruzadas entre si, pois tomar decisões com base em dados confiáveis e tratados, se torna crucial, pois evita perdas econômicas, ajuda na criação de insights,

tornando dessa maneira a empresa competitiva e estável no mercado mundial e econômico. Da mesma forma que muitos objetivos foram cumpridos, como: conseguir extrair os arquivos do site, realizar os devidos tratamentos e visualizar esses dados de forma dinâmica e ilustrativa em um dashboard, alguns pontos deixaram a desejar durante esse processo de aprendizado, uma das limitações dos dados foi não ter conseguido traçar quais são os meios de transportes que esses produtos são levados da origem até o destino.

Dados sobre as importações faltaram para completar os estudos, haja visto que o tratamento dos dados de exportação consumiu um tempo considerável de trabalho. Paralelamente, ao que foi mencionado no parágrafo anterior, esse estudo sobre comércio exterior abre margem para novas pesquisas, que também possuem muita relevância no cenário econômico algumas dessas pesquisas poderiam ser pautadas nos seguintes tópicos: Quais blocos econômicos mais exportam e importam no cenário mundial, quais cidades da baixa Mogiana mais exportam e importam produtos, quais são os maiores estados importadores e exportadores de produtos, e por fim quais são os meios de transportes mais usados nas exportações e importação. Considero de suma importância que um estudante tenha um projeto que o mesmo possa chamar de seu, ao final deste projeto me sinto realizado, pois estive em contato com uma área no qual me interessei desde o segundo semestre da faculdade, e através da iniciação científica, pude me aprofundar nos principais desafios da área de ciência de dados, desenvolvi muitas habilidades que tenho certeza que futuramente serão relevantes nessa área que pretendo seguir.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, Kátia. **A importância das Exportações**. Revista de Política Agrícola, Brasília, v. 24, n. 3, p. 1-3, Jul./Ago./Set. 2015.

ANTONCHUK, Yuri. **Conceitos de Tabelas de Dimensão e Fato na Construção do Data Warehouse**. MentorsTec, 2020. Disponível em: <https://www.mentorstec.com.br/post/conceitos-de-tabelas-dimens%C3%A3o-e-fato-na-constru%C3%A7%C3%A3o-do-data-warehouse>. Acesso em: 28 Set. 2023.

ARRAIS, Karolayne. **Construção de um pipeline de dados utilizando serviços da nuvem.** Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Computação). Universidade Federal de São Carlos Centro de Ciência Exatas e de Tecnologia, São Carlos. P. 91. 2022.

CARVALHO, Ulysses. **Reestruturação Econômica e Urbana no Estado de São Paulo: O papel de Mogi Guaçu e Mogi Mirim na Complexa Rede Urbana Paulista..** 100-104p. Dissertação de Mestrado (Pós-Graduação em Geografia). Universidade Federal do Rio Grande do Sul Instituto de Geociências, Porto Alegre. 2013.

ELIAS, Diego. **Conhecendo a arquitetura de Data Warehouse.** Canaltech, 2014. Disponível em: <https://arquivo.canaltech.com.br/business-intelligence/conhecendo-a-arquitetura-de-data-warehouse-19266/>. Acesso em: 07 Out. 2023.

FERNANDES, Pedro. **O comércio bilateral do Complexo de Soja entre Brasil e China: Uma Análise Desagregada usando indicadores do Comércio Exterior.** Monografia (Bacharelado em Ciências Econômicas). Universidade Federal da Paraíba Centro de Ciências Sociais Aplicadas, João Pessoa, p. 14. 2016.

GARCIA, Ana Laura. **As Exportações Brasileiras entre 1998 e 2018: Uma análise sobre a reprimarização.** Monografia (Bacharel em Ciências Econômicas). Universidade Federal de Uberlândia Instituto de Economia e Relações Internacionais, Uberlândia. p. 20. 2019.

LIMA, Rebecca; RIBEIRO, José. **A Dinâmica do crescimento das exportações do Agronegócio Brasileiro.** Número 2249. Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília, 2016.

MOREIRA, Lucyanno; ASSIS, Isabela; ROCHA, Ricardo; BARRETO, David. **Data webhouse: A Evolução do Data Warehouse para Web e suas Contribuições para o Aperfeiçoamento do Relacionamento com Clientes.** XIII SIMPEP, Nov. 2006.

NABEREZNY, Thais. **O que é ETL e por que é importante a sua utilização?** Rox Partner, 2022. Disponível em: <https://roxpartner.com/o-que-e-etl-e-por-que-e-importante-a-sua-utilizacao/>. Acesso em: 22 Set. 2023.

NETO, Onofre. **O Brasil no Mercado Mundial de Carne Bovina: Análise da Competitividade da Produção e da Logística de Exportação Brasileira.** Ateliê Geográfico, v. 12, n. 2, p. 4-5, Ago. 2018.

NUNES, Leidiane. **Ciência de Dados como método de transformação de dados em informação.** Trabalho de Conclusão de Curso (Tecnólogo em Sistemas Para Internet.). Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano, Pernambuco. p. 24. 2022.

ROCHA, Rafael. **Analisando o Pentaho como Ferramenta de Business Intelligence.** Trabalho de Conclusão de Curso (Bacharel em Sistemas de Informação). Fundação de Ensino “Eurípides Soares da Rocha” Centro Universitário Eurípides de Marília - Univem, Marília. P. 18-22. 2015.

SANJUAN, Adriana. **Análise do Processo de Configuração Desigual do Setor de Celulose e Papel no Brasil.** Pesquisa & Debate, São Paulo, v. 23, n. 2, p. 24, jul-dez. 2013.

SANTOS, Angela; GONÇALVES, João. **Evolução do Comércio Exterior do Complexo Automotivo.** BNDES Setorial, Brasília, n. 13, p. 9, mar. 2001.

SOUZA, Amaranta; SALAZAR, Eder; SOARES, Liriane; BRUSCHI, Paula; ALMEIDA, Tayná; RIBEIRO, Verônica. **Ciência de Dados, Business Intelligence e Big Data: conceitos e aplicações.** 1 ed. Curitiba: Appris, 2021.

APÊNDICE A - ANÁLISE DE DADOS

Nesta seção serão discutidas as análises referentes aos dados previamente limpos e filtrados com a criação dos painéis de visualização.

A.1 ENGENHEIRO COELHO

Conforme apontado por Souza e Neto (2016 apud Aurélio, 2012, p. 42) um dos fatores que levam ao crescente aumento das exportações de frutas, são a qualidade destes produtos, assim como a cotação da nossa moeda que convertida para outros países se torna mais barata, logo a mercadoria se torna mais acessível naquele país, fato que é vislumbrado na figura 16 ao se observar que a Bélgica é o maior parceiro comercial de Engenheiro Coelho. Portanto, podemos deduzir que a exportação, são em sua maioria destinada para os países Europeus, haja visto que a taxa do câmbio nominal, será menor.

Figura 16 - Painel de Análise Estatística para Engenheiro Coelho



Fonte: Próprio Autor, 2023

A.2 ESTIVA GERBI

Figura 17 - Painel de Análise Estatística para a cidade de Estiva Gerbi



Fonte: Próprio Autor (2023)

A ascensão econômica de Estiva Gerbi, ocorre quando a mesma, conquista sua emancipação de Mogi Guaçu, tornando-se independente economicamente, recebendo nesse momento o status Município (CARVALHO, 2013, p. 104).

Deve-se abordar, ainda que por ter uma forte dependência econômica de Mogi Guaçu, a mesma começou a se enfraquecer no setor cerâmico, haja visto que Mogi Guaçu era o maior produtor de cerâmica, porém começou a diversificar seu setor econômico não ficando somente em uma segmentação (CARVALHO, 2013, p. 100). Entretanto, se observarmos o gráfico da figura 17, de produtos exportados, notamos uma dualidade entre a citação trazida pelo autor e os dados coletados do site do Comércio do Governo, que demonstra que a cerâmica continua sendo ainda o carro chefe da cidade, sendo o produto mais exportado.

A.3 ITAPIRA

Figura 18 - Painel de Análise Estatística para a cidade de Itapira



Fonte: Próprio Autor (2023)

Observa-se na figura 18, que a cidade de Itapira tem como produto mais exportado a soja, tal fato é sustentado por Fernandes(2016 apud EMBRAPA, 2014, p. 14), que afirma que a soja corresponde a 49% das áreas plantadas em todo Brasil. Em conformidade com Fernandes (2016 apud LAZZAROTTE; HIROSHI et al., 2009, p. 14) o mesmo afirma que a grande parte das exportações de soja no Brasil são direcionadas para países asiáticos, em especial na China. Portanto, ao se analisar a imagem, percebemos que o maior parceiro comercial de Itapira é a China.

A.4 MOGI GUAÇU

Figura 19 - Painel de Análise Estatística para a cidade de Mogi Guaçu



Fonte: Próprio Autor (2023)

Na figura 19, é notório perceber que o setor da celulose, com seus diversos derivados predomina no âmbito de exportação, tal fato é confirmado pela autor Sanjuan (2013 apud ROCHA, 2006, p. 24) ao inferir que o crescente interesse dos

estrangeiros por celulose na América Latina, deve-se ao fato da preferência do país em produzir fibras curtas, ao invés das longas que são mais caras. Ademais, Sanjuan (2013 apud ROCHA, 2005, p. 24), complementa a discussão ao reforçar as vantagens geográficas do País como: localização, abundância de água e terras, com rápido crescimento da árvore de eucalipto.

A.5 MOGI MIRIM

Figura 20 - Painel de Análise Estatística para a cidade de Mogi Mirim



Fonte: Próprio Autor (2023)

Através da figura 20, é perceptível notar que o maior produto exportado pela cidade de Mogi Mirim é partes e acessórios dos veículos automóveis, sendo que os países que mais recebem são: Estados Unidos em primeiro lugar, e ocupando a segunda posição vem a Argentina, tal fato é comprovado quando Santos e Gonçalves (2001,p. 9), afirmam que os EUA é o principal consumidor de autopeças absorvendo entre 35% até 40% das autopeças do Brasil. Ademais, o mesmo afirma que na segunda colocação vem a Argentina absorvendo entre 20% até 28% do mercado de autopeças. Portanto, os dados trazidos por esses autores validam a análise estatística trazida por esse dashboard.

A.6 SANTO ANTÔNIO DE POSSE

Em estudos realizados por Neto (2018 apud OLIVIO, 2008, p. 5), o mesmo aponta que os EUA, apresenta uma tecnologia sofisticada capaz de aproveitar uma maior produção de arrobas, além de conseguir aproveitar os restos dos animais. Dessa forma, o autor corrobora, com os resultados obtidos nesse projeto, que apontam que Santo Antônio de Posse tem como maior Parceiro Comercial os Estados Unidos, devido ao fato do mesmo ser o maior importador de carne.

Da mesma forma, Neto (2018 apud USDA, 2016, p. 4) aponta que em 2016 o Brasil apresentava o segundo maior rebanho do mundo. Nessa linha de pensamento, observa-se na figura 21, no qual possui um gráfico de evolução de exportações ao longo dos anos, que no referido ano Santo Antônio de Posse apresentou uma evolução nas exportações e nos próximos anos percebemos uma permanência na evolução e em 2019 a evolução continuou.

Figura 21 - Painel de Análise Estatística para Santo Antônio de Posse



Fonte: Próprio Autor (2023)