

**FACULDADE DE TECNOLOGIA DE SÃO JOSÉ DOS CAMPOS
FATEC PROFESSOR JESSEN VIDAL**

**ALENDRE ESCOBAR GARCIA SILVEIRA
ANA JULIA DO COUTO BRANDÃO
HELLEN DE SOUSA SANTOS
STEFFANY CAROLINE VIEIRA SANTO
THAYSSA ANDRADE SANTOS
YURI GONÇALVES ALEIXO
PEDRO LOUIS LOURENÇO**

PROJETO INTEGRADOR SPRINT 2

São José dos Campos
2024

**ALENDRE ESCOBAR GARCIA SILVEIRA
ANA JULIA DO COUTO BRANDÃO
HELLEN DE SOUSA SANTOS
STEFFANY SANTO
THAYSSA ANDRADE SANTOS
YURI GONÇALVES ALEIXO**

PROJETO INTEGRADOR SPRINT 2

Trabalho avaliativo apresentado à Faculdade de Tecnologia de São José dos Campos, como parte dos requisitos necessários para obtenção da nota semestral do Projeto Integrador – PI no curso de logística.

Orientador: Newton Eizo Yamada
Coorientador: Professor Mestre Marcus Vinicius do Nascimento

São José dos Campos
2024

Os autores oferecem a obra à familiares e amigos.

AGRADECIMENTOS

Aos professores, pelas aulas e instruções dadas para que assim fosse possível a realização deste trabalho.

“Não se deve ir atrás de objetivos fáceis,
é preciso buscar o que só pode ser
alcançado por meio dos maiores
esforços.”

Albert Einstein

RESUMO

Nessa sprint, foi utilizado o Jira Software para formar o backlog do produto, estruturando para a próxima sprint, atribuindo cada tarefa a um integrante do grupo.

Este estudo foca na construção e tratamento de uma base de dados da ANTAQ, sendo assim, após exercer as tarefas atribuídas buscamos tratar e analisar os dados relacionados a cargas e atracação de navios, abrangendo técnicas de limpeza de dados, filtragem e integração de informações.

A análise de dados realizada envolveu três etapas principais: a filtragem e carga exclusiva, o cálculo da prancha média e a criação de um dashboard para visualização dos resultados.

O código começa importando as bibliotecas necessárias e carregando um arquivo CSV do Google Drive para um DataFrame. Após isso, realiza um filtro nos dados, mantendo apenas as linhas em que a coluna `snaturezacarga` é igual a 'Exclusivo', e salva o DataFrame filtrado em um novo arquivo.

Em seguida, é feito o cálculo da prancha média para avaliar a produtividade dos portos, com a fórmula: $\text{Prancha Média} = \text{Quantidade total de carga} / \text{Total de horas trabalhadas} * \text{Pontos de operação}$.

Por fim, é desenvolvido um dashboard para exibir os 10 portos com melhor prancha média, apresentando informações sobre a quantidade de berços e a movimentação por peso de carga.

Palavras-Chave: Antaq, Dataframe, CSV, Atracação, Filtragem, Prancha média, Base de dados

ABSTRACT

In this sprint, Jira Software was used to create the product backlog, structuring tasks for the next sprint and assigning each task to a team member.

This study focuses on the construction and processing of a database from ANTAQ. After completing the assigned tasks, we worked on cleaning and analyzing data related to cargo and ship docking, applying data cleaning techniques, filtering, and information integration.

The data analysis consisted of three main steps: filtering and exclusive loading, calculating the average dock, and creating a dashboard to visualize the results.

The code begins by importing the necessary libraries and loading a CSV file from Google Drive into a DataFrame. It then filters the data, keeping only the rows where the column `stnaturezacarga` equals 'Exclusive', and saves the filtered DataFrame into a new file.

Next, the average dock is calculated to evaluate port productivity, using the formula: $\text{Average Dock} = \text{Total amount of cargo} / \text{Total hours worked} * \text{Operation points}$.

Finally, a dashboard is developed to display the top 10 ports with the best average dock, showing information on the number of docks available and cargo movement by weight.

Keywords: ANTAQ, DataFrame, CSV, Docking, Filtering, Average Dock, Database

1. INTRODUÇÃO SPRINT 1

1.1. Objetivo do Trabalho

Este estudo está situado na área de Logística, tendo como tema a importância a construção e tratamento de uma base de dados da ANTAQ (Agência Nacional de Transporte Aquaviário) de movimentação portuária de 2014 a 2024.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. O que é logística

Segundo Ballou (2001) a logística inclui todas as atividades importantes para a disponibilização de bens e serviços aos consumidores quando e onde estes quiserem adquiri-los. Essas atividades incluem planejamento, transporte, armazenagem e etc.

Segundo Daskin (1995), define a logística como sendo o planejamento e a operação de sistemas físicos (veículos, armazéns, redes de transporte, etc), informacionais e gerenciais (processamento de dados, teleinformática, processos de controle gerenciais, etc) necessários para que insumos e produtos vençam condicionantes físicas e temporais de forma econômica.

2.2. Logística Portuária

Trata-se de todos os processos de transporte relacionados ao uso, armazenamento e recebimento de produtos que chegam em um porto. Ela faz parte dos vários tipos de logísticas que existem e conta com algumas particularidades. Além disso, inclui toda a parte de documentação e procedimentos alfandegários, questões fiscais e legais, sem falar no planejamento e na gestão das rotas de transporte. Por isso, ela requer um nível de organização gigantesco e uma coordenação minuciosa para que tudo saia conforme o planejado.

O objetivo é otimizar o tempo, poupar recursos e, claro, evitar dores de cabeça para todos os envolvidos, desde as transportadoras até as empresas de logística e os consumidores finais.

Segundo Neto (2019), a logística portuária abrange a movimentação e a gestão marítima, como o embarque, desembarque, estocagem, armazenamento e escoamento das mercadorias e a gerência dos modais. A estrutura da logística portuária pode ser dividida

em três tópicos: Complexo Fixo, Administração e Operação.

2.3. Python

Python é uma linguagem de programação de alto nível, ou seja, com sintaxe mais simplificada e próxima da linguagem humana, utilizada nas mais diversas aplicações, como desktop, web, servidores e ciência de dados.

2.4. Colab

É uma ferramenta em nuvem que permite criar e executar códigos na linguagem Python. Com ele, você pode rodar os programas diretamente do seu navegador, de forma simples e rápida.

2.5. Power BI

O Power BI é uma coleção de serviços de software, aplicativos e conectores que trabalham juntos para transformar suas fontes de dados não relacionadas em informações coerentes, visualmente envolventes e interativas.

2.6. Github

O GitHub é um repositório de dados compartilhado. É uma plataforma que permite uma hospedagem prática de código-fonte e arquivos em nuvem.

2.7. Jira Software

O Jira Software é um software pensado para o uso em metodologias ágeis, como o Scrum ou o Kanban. O software permite o gerenciamento de projetos e monitoramento de tarefas em um único lugar, garantindo o alinhamento entre a equipe e a execução das demandas, assim permitindo a flexibilidade nos processos de acordo com as necessidades de cada equipe.

2.8. Slack

O Slack é um app de mensagens para empresas que conecta as pessoas às informações de que elas precisam. Reunindo pessoas para trabalhar como uma equipe unificada, o Slack transforma a forma como as organizações se comunicam.

2.9. Prancha Média

A prancha média é um cálculo utilizado para medir a eficiência nas operações de carga e descarga em portos. Basicamente, ela mostra quantas toneladas de carga são movimentadas, em média, por hora em cada ponto de operação. Esse cálculo ajuda a entender a velocidade com que uma carga é transferida de um navio para o porto, e permite ter uma visão clara da produtividade durante o processo.

4. BANCO DE DADOS ESTRUTURADO

4.1. Desenvolvimento

Para avaliação desta segunda parte do projeto, houve-se requisitos para serem atendidos, estes são:

- o uso do GitHub.
- Planilhas com dados extraídos.
- o uso do Power BI.
- Uso do Colab/Python.

Esses itens auxiliaram para a visualização e demonstração dos dados tratados, como também, do processo para a conclusão do projeto. Assim, a partir dos dados pré tratados deu-se o início ao processo de desenvolvimento do projeto, pois observa-se que o banco de dados incluídos dentro deste site, podem ser filtrados por região e municípios, uma vez que filtrados, esses dados podem ser baixados em planilhas do Excel, sendo fundamental para a construção do projeto, tendo em vista que essas filtragens e planilhas serão usadas diretamente no software do Power BI. Uma vez que os arquivos retirados do site são abertos dentro do Excel, serão transformados em tabelas usados por essa ferramenta.

Logo, o objetivo é transformá-los em dados visíveis como, por exemplo, visualização em gráficos, tabelas, indicadores e, também, com os dados filtrados por região é possível a visualização dos dados através de um mapa interativo, sendo incluído no arquivo e no requisito do projeto.

Após a formatação dos arquivos no software BI, são então transferidos para o Colab/Python/Linguagem R, onde é possível conectar o todo do projeto, ou seja, as planilhas e os dados ao Power BI, para que, sejam demonstrados de uma forma ágil e prática.

5. ANÁLISE DE DADOS

5.1. Código de filtragem e carga exclusiva

O código começa importando as bibliotecas necessárias, como o `pandas` para manipulação de dados e o `drive` do Google Colab para acessar arquivos no Google Drive. Em seguida, monta o Google Drive, permitindo o acesso aos arquivos armazenados. Define o caminho para um arquivo CSV e carrega esse arquivo em um DataFrame, especificando que o delimitador é o ponto e vírgula.

Após isso, o código exibe as colunas do DataFrame para verificar se os dados foram carregados corretamente. Em seguida, filtra o DataFrame, mantendo apenas as linhas em que a coluna `stnaturezacarga` é igual a 'Exclusivo'. Finalmente, o DataFrame filtrado é exibido e salvo em um novo arquivo CSV no Google Drive.

```
# Importar bibliotecas necessárias
import pandas as pd
from google.colab import drive

# Montar o Google Drive para acessar os arquivos
drive.mount('/content/drive')

# Definir o caminho do arquivo CSV no Google Drive
caminho_arquivo = '/content/drive/My Drive/Base/Resultado_merge.csv' # Atualize com o caminho correto

# Carregar o arquivo CSV em um DataFrame
df = pd.read_csv(caminho_arquivo)

# Carregar o arquivo CSV em um DataFrame, especificando o delimitador como ponto e vírgula
df = pd.read_csv(caminho_arquivo, sep=';')

# Exibir as colunas para verificar o nome correto
print(df.columns)

# Filtrar os dados que possuem "Exclusivo" na coluna 'stnaturezacarga'
df_exclusivo = df[df['stnaturezacarga'] == 'Exclusivo']

# Exibir ou salvar o novo DataFrame filtrado
print(df_exclusivo)

# Salvar o DataFrame filtrado em um novo arquivo CSV no Google Drive
df_exclusivo.to_csv('/content/drive/My Drive/Base/Resultado_merge2.csv', index=False)
```

Fonte: Autores Próprios

5.2. Prancha Média

Foi usado um cálculo que ajuda a entender a velocidade com que uma carga é transferida de um navio para o porto, e permite ter uma visão clara da produtividade durante o processo. Para obtê-la, basta dividir a quantidade total de carga movimentada pelo tempo total de operação, ou seja:

Prancha Média = Quantidade total de carga/Total de horas trabalhadas * Pontos de operação.

O código executado no Power BI utiliza a função DIVIDE para dividir os valores da coluna vlpesocargabruta, que indica o peso bruto da carga, pela coluna tempo_operacao, que representa o tempo de operação. A função garante que, se o tempo de operação for zero, o resultado será zero, evitando assim erros de divisão por zero. O resultado dessa operação é armazenado na variável Prancha_media, refletindo a relação entre a carga e o tempo de operação.

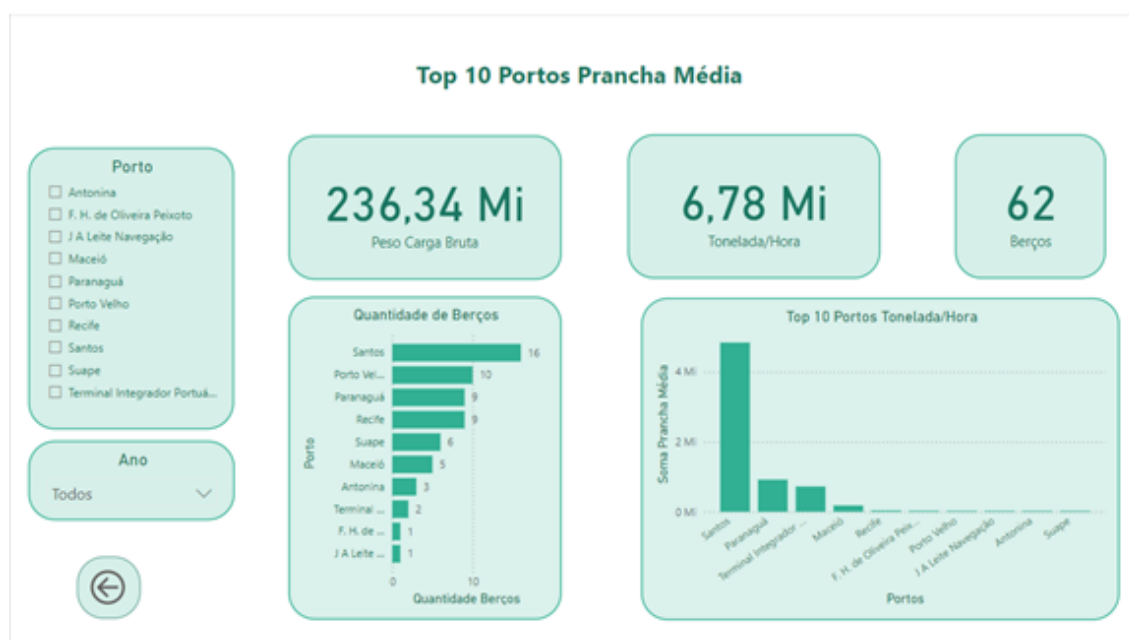
Prancha_media = DIVIDE(Resultado_merge2[vlpesocargabruta], Resultado_merge2[tempo_operacao],0)

stnaturezacarga	stshz	stsh4	natureza da carga	sentido	tea	qtz carga	vlpesocargabruta	Prancha_media
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	19670.485	1035.28868421053
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	15149.79	552.575927051672
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	4914.214	144.89083046683
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	3522.594	103.860265356265
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	3000	88.4520884520885
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	5000	147.420147420147
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	26294.46	775.266633906634
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	618.495	18.2357248157248
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	747.853	22.0497199017199
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	2667.418	78.646230958231
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	500	14.7420147420147
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1921.234	56.6457199017199
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	7625.152	224.820206388206
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1133.852	33.4305257985258
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	2547.769	75.1184963144963
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	587.038	17.3082457002457
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	826.057	24.3554889434889
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1503.694	44.3349582309582
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	140.728	4.1492285012285
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	2304.637	67.9499852579853
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1561.82	46.0487469287469
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	500	14.7420147420147
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1083.943	31.9590073710074
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	391.81	11.5521375921376
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	1278.767	37.7032039312039
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	4884.739	69.6162327790974
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	4500	64.1330166270784
Exclusivo	Exclusivo	Exclusivo	Granel Sólido	Embarcados	0	0	9017.083	128.509496437055

Fonte: Autores Próprios

5.3. Dashboard prancha média

No Dashboard, apresentamos o Top 10 Portos com base na prancha média. Para cada porto listado, incluímos a informação sobre a quantidade de berços disponíveis. Além disso, é possível visualizar a movimentação por peso de carga bruta. Os dados podem ser filtrados por ano, de 2014 a 2023, e também por porto específico dentro da lista.



Fonte: Autores Próprios

5.3. Dashboard prancha média

Os 10 principais portos foram definidos com base na prancha média. Essa lista inclui: Antonina, F. H. de Oliveira Peixoto, J A Leite Navegação, Maceió, Paranaguá, Porto Velho, Recife, Santos, Suape e Terminal Integrador Portuário. De acordo com o gráfico do Dashboard, o porto de Santos apresenta a maior prancha média, enquanto Suape possui a menor entre os dez listados.



Fonte: Autores Próprios

