



INSTITUTO POLITÉCNICO DE BEJA
Escola Superior de Tecnologia e Gestão
Licenciatura em Engenharia Informática
Sistemas de Informação

Trabalho de Grupo 1

Martinho José Novo Caeiro
Paulo António Tavares Abade



Beja, novembro de 2024

INSTITUTO POLITÉCNICO DE BEJA
Escola Superior de Tecnologia e Gestão
Licenciatura em Engenharia Informática
Sistemas de Informação

Trabalho de Grupo 1

Martinho José Novo Caeiro
Paulo António Tavares Abade

Beja, novembro de 2024

Conteúdo

1	Introdução	1
2	Metodologia de Trabalho	1
3	ETL	1
3.1	Extração de Dados	1
3.2	Transformação de Dados	2
3.3	Inserção de Dados	2
4	Criação da Dataware House	3
5	Contexto Global	3
6	Conclusões e Perspectivas de Trabalho Futuro	3
	Bibliografia	4

1 Introdução

Neste trabalho irá ser abordado com a Crise do Petróleo influenciou a produção e desenvolvimento dos carros entre os anos 1970 e 1980, alterando o motor e por consequência a sua aceleração, eficiência e consumo.

Para analisar esta situação foi utilizada uma base de dados com 393 entradas, com 9 atributos iniciais, sendo estes: Milha por Galão, Cilindros, Cilindrada, Cavalagem, Peso, Aceleração, Ano de Fabrico, País de Origem e o Nome Completo do Carro.

Ainda foi utilizada outra base de dados que indica o preço médio de combustível por ano, entre 1970 e 1982, nos Estados Unidos da América.

Por fim, a nossa teoria é que a Crise do Petróleo influenciou a produção e desenvolvimento dos carros entre os anos 1970 e 1980, alterando o motor e por consequência a sua aceleração, eficiência e consumo.

2 Metodologia de Trabalho

Para o desenvolvimento deste trabalho foi utilizado o *Github* para a partilha de código e documentação, o *Visual Studio Code* como IDE de desenvolvimento, o *Excel - PowerQuery* ferramenta principal de tratamento e análise de dados e por fim, o *SQL Server Managment* para a criação da Dataware House.

3 ETL

O trabalho de ETL foi realizado em 3 fases, a primeira foi a extração dos dados, a segunda a transformação e por fim a inserção dos dados em tabelas dinâmicas.

3.1 Extração de Dados

Os dados foram extraídos de um ficheiro *.csv* que foi encontrado no Kaggle (Elmetwally, 2023) e importados para o *Excel* utilizando o PowerQuery.

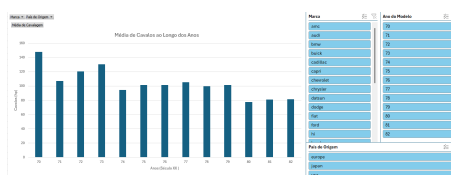
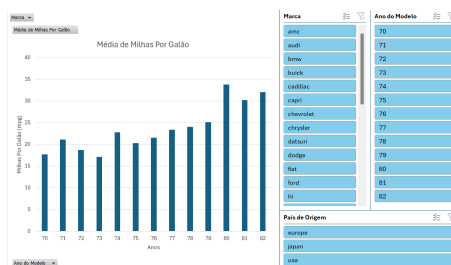
3.2 Transformação de Dados

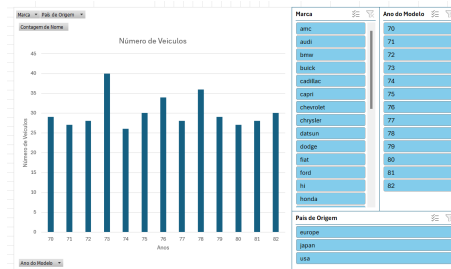
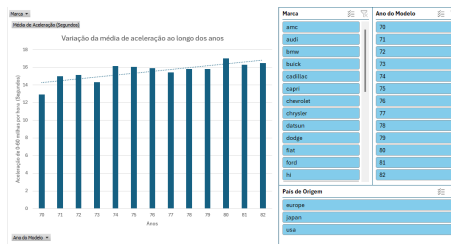
Com o auxílio do PowerQuery foi possível transformar os dados, e foram encontrados diversos problemas, como por exemplo, a presença de valores nulos, duplicados e a necessidade de alterar o tipo de dados de algumas colunas. Ainda foram encontrados problemas com a formatação dos dados, como por exemplo, a presença de vírgulas em vez de pontos. O nome das colunas também foi alterado para facilitar a sua identificação e foram adicionadas novas colunas para facilitar a análise e permitir a criação de tabelas dinâmicas, sendo que as novas colunas são: Marca e Modelo.

3.3 Inserção de Dados

Utilizando as tabelas dinâmicas do Excel, foi possível visualizar os dados de forma mais clara e objetiva, para isso, foram utilizados gráficos dinâmicos onde é possível visualizar a evolução dos carros ao longo dos anos, a evolução da eficiência dos carros ao longo dos anos e ainda a evolução da eficiência dos carros por país.

Para isso, o Excel possui opção de selecionar a visualização dos dados por ano, por país e por marca. Isso foi feito pensando na segmentação dos dados, para que seja possível visualizar os dados de forma mais clara e objetiva.





4 Criação da Dataware House

5 Contexto Global

6 Conclusões e Perspectivas de Trabalho Futuro

Texto das conclusões.

Bibliografia

Elmetwally, T. (2023). *Car information dataset* [Base De Dados Kaggle]. Obtido outubro 14, 2024, de <https://www.kaggle.com/datasets/tawfikelmetwally/automobile-dataset/data>