

## SEÇÃO 1 - RESOLUÇÃO DAS QUESTÕES

A resolução das questões foi realizada tanto com a linguagem python, tópico 1.1, quanto com a linguagem SQL, tópico 1.2. A seguir segue a apresentação das consultas com os respectivos resultados.

### 1.1 Resolução das questões com python

#### Extração dos dados

```
qDataProduct = "SELECT * FROM data_product;"
qDataProductSales = "SELECT * FROM data_product_sales;"
qDataStoreCad = "SELECT * FROM data_store_cad;"
qDataStoreSales = "SELECT * FROM data_store_sales;"

df_dataProduct = pd.read_sql(qDataProduct, db_connection)
df_dataProductSales = pd.read_sql(qDataProductSales, db_connection)
df_dataStoreCad = pd.read_sql(qDataStoreCad, db_connection)
df_dataStoreSales = pd.read_sql(qDataStoreSales, db_connection)
```

#### Questão 1.How many products does the company have?

```
df_dataProduct['PRODUCT_COD'].count()
```

Resultado: 9994

#### Questão 2.What are the 10 most expensive products in the company?

```
df_dataProduct.sort_values('PRODUCT_VAL',ascending=False).head(10)
```

Resultado:

	PRODUCT_COD	PRODUCT_NAME	PRODUCT_VAL	DEP_NAME	DEP_COD	SECTION_NAME	SECTION_COD
7942	301409	Whisky Escoces THE MACALLAN Ruby Garrafa 700ml...	741.99	BEBIDAS	2	BEBIDAS	4
5374	176185	Whisky Escoces JOHNNIE WALKER Blue Label Garra...	735.90	BEBIDAS	2	BEBIDAS	4
9032	315481	Cafeteira Expresso 3 CORACOES Tres Modo Vermelho	499.00	BEBIDAS	2	BEBIDAS	4
3057	100280	Vinho Portugues Tinto Vintage QUINTA DO CRASTO...	445.90	BEBIDAS	2	VINHOS	30
9237	320046	Escova Dental Eletrica ORAL B D34 Professional...	399.90	PERFUMARIA	8	HIGIENE BUCAL	39
5677	190817	Champagne Rose VEUVE CLICQUOT PONSARDIM Garraf...	366.90	MERCEARIA	6	ARTIGOS-PARA-O-LAR	2
4640	153795	Champagne Frances Brut Imperial MOET Rose Garr...	359.90	MERCEARIA	6	ARTIGOS-PARA-O-LAR	2
8647	311397	Conjunto de Panelas Allegra em Inox TRAMONTINA...	359.00	MERCEARIA	6	ARTIGOS-PARA-O-LAR	2
4449	147706	Whisky Escoces CHIVAS REGAL 18 Anos Garrafa 750ml	329.90	BEBIDAS	2	BEBIDAS	4
4669	154431	Champagne Frances Brut Imperial MOET & CHANDON...	315.90	MERCEARIA	6	ARTIGOS-PARA-O-LAR	2

#### Questão 3.What sections do the 'BEBIDAS' and 'PADARIA' departments have?

```
filtro = df_dataProduct['DEP_NAME'].isin(['BEBIDAS','PADARIA'])
dfResultado = df_dataProduct[filtro].groupby(['DEP_NAME','SECTION_NAME'])
dfResultado.describe()
```

Resultado

Bebidas: bebidas, cervejas, refrescos, vinhos

Padaria: doces-e-sobremesas, gestante, padaria, queijos-e-frios

#### Questão 4.Which store sold the most products in one day? Which day?

```
df_dataStoreSales.sort_values('SALES_VALUE',ascending=False).head(1)
```

Resultado:

	STORE_CODE	DATE	SALES_VALUE	SALES_QTY
29464	11	2019-02-20	479566.46	16072

#### Questão 5.What was the total sale of products (in \$) of each Business Area in the first quarter of 2019?

```
df_merge = pd.merge(df_dataStoreSales, df_dataStoreCad[['STORE_CODE','BUSINESS_NAME','BUSINESS_CODE']],
                    how = 'inner', on = 'STORE_CODE')
df_merge.dtypes
df_merge['DATE_CONVERT'] = pd.to_datetime(df_merge['DATE'])
filtro = (df_merge['DATE_CONVERT'] >= '2019-01-01') & (df_merge['DATE_CONVERT'] <= '2019-03-31')
```

```
df_merge[filtro].groupby(['BUSINESS_NAME','BUSINESS_CODE'])[['SALES_VALUE','SALES_QTY']].sum()
```

Resultado:

		SALES_VALUE	SALES_QTY
BUSINESS_NAME	BUSINESS_CODE		
Atacado	5	80384884.60	5221672
Farma	4	81776691.73	2854179
Posto	3	32072326.40	2347032
Proximidade	2	80171122.80	5209730
Varejo	1	81032347.65	5252352

## 1.2 Resolução das questões com SQL

Questão 1. How many products does the company have?

```
'SELECT COUNT(PRODUCT_COD) AS QuantidadeProdutos FROM data_product;'
```

Questão 2. What are the 10 most expensive products in the company?

```
'SELECT * FROM data_product ORDER BY PRODUCT_VAL DESC LIMIT 10;'
```

Questão 3. What sections do the 'BEBIDAS' and 'PADARIA' departments have?

```
'SELECT DISTINCT dep_name, section_name FROM data_product where (dep_name="BEBIDAS" OR dep_name="PADARIA")
order by dep_name, section_name;'
```

Questão 4. Which store sold the most products in one day? Which day?

```
'SELECT sales.STORE_CODE, cad.STORE_NAME, sales.DATE, sales.SALES_VALUE, sales.SALES_QTY from
data_store_cad AS cad INNER JOIN data_store_sales as sales ON sales.STORE_CODE=cad.STORE_CODE ORDER BY
SALES_QTY'
```

Questão 5. What was the total sale of products (in \$) of each Business Area in the first quarter of 2019?

```
'SELECT cad.BUSINESS_CODE, cad.BUSINESS_NAME, SUM(sales.SALES_VALUE) AS TOTAL_VENDAS FROM data_store_sales
AS sales JOIN data_store_cad AS cad ON sales.STORE_CODE=cad.STORE_CODE WHERE DATE BETWEEN "2019-01-01" AND
"2019-03-31" GROUP BY cad.BUSINESS_NAME, cad.BUSINESS_CODE'
```

## 1.3 Conexão com o banco de dados

O código a seguir foi utilizado para conexão com o banco de dados. Foi utilizado tanto nas consultas com python quanto com SQL, tópico 1.1 e tópico 1.2

```
import mysql.connector
import pandas as pd

db_connection = mysql.connector.connect(host="xxxx", database= "xxxx", user="xxxx", password="xxxx")

if db_connection.is_connected():
    db_info = db_connection.get_server_info()
    print("Conectado ao servidor MySQL versão", db_info)
else:
    print("nao conectado")

#Finalizando conexão
if db_connection.is_connected():
    db_connection.close()
    print("Conexão encerrada")
```