

Coffee Shop Sales

Bootcamp Data Analyst with SQL & Python using Google Platform

Iqbal Alfarirdzi Hakim



Contents

1. Introduction
2. Goals (Scope of Analysis)
3. Mengumpulkan Data
4. Memvalidasi Data
5. Menentukan Objek Data
6. Mengkontruksi Data
7. Mengintegrasikan Data
8. Membuat Busines Intelligence
9. Conclusion & Rekomendations

Introductions

Coffee Shop Sales merupakan datasets yang menunjukkan transaksi penjualan beberapa produk coffee shop yang di jual oleh tiga toko coffee shop yang berbeda selama enam bulan mulai dari 1 Januari 2023 hingga Juni 2023.



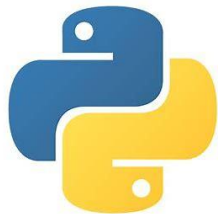
Goals (Scope of Analysis)

1. Menganalisis Tren Penjualan berdasarkan Lokasi Coffee Shop
2. Menganalisis Kategori dan tipe produk dengan total penjualan tertinggi
3. Menganalisis Total Penjualan Lokasi Coffee Shop Berdasarkan Jam Operasional Toko

Mengumpulkan Data

Tools Used

Analysis



Data Visualization



seaborn

matplotlib



Looker

Mengumpulkan Data

Upload datasets yang akan digunakan beserta library yang dibutuhkan untuk proses analisis data

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

df = pd.read_excel(r"C:\Users\lenovo\Documents\Bootcamp Data Analyst\DQLAB\Final Project\Coffee.xlsx")

df.head()
```

<https://www.kaggle.com/datasets/ahmedabbas757/coffee-sales>

Memvalidasi Data

Memvalidasi Data

transaction_id	transaction_date	transaction_time	transaction_qty	store_id	store_location	product_id	unit_price	product_category	product_type	product_detail
1	2023-01-01	07:06:11	2	5	Lower Manhattan	32	3.0	Coffee	Gourmet brewed coffee	Ethiopia Rg
2	2023-01-01	07:08:56	2	5	Lower Manhattan	57	3.1	Tea	Brewed Chai tea	Spicy Eye Opener Chai Lg
3	2023-01-01	07:14:04	2	5	Lower Manhattan	59	4.5	Drinking Chocolate	Hot chocolate	Dark chocolate Lg
4	2023-01-01	07:20:24	1	5	Lower Manhattan	22	2.0	Coffee	Drip coffee	Our Old Time Diner Blend Sm

- Terdapat 11 kolom yaitu: **transaction_id**, **transaction_date**, **transaction_time**, **transaction_qty**, **store_id**, **store_location**, **product_id**, **unit_price**, **product_category**, **product_type**, **product_detail**

Menentukan Objek Data

Menentukan Objek Data

```
# mencari total sales
df['total_sales'] = df['transaction_qty'] * df['unit_price']
sales_per_product = df.groupby('product_id')['total_sales'].sum().reset_index()

# Membuat variabel datetime
df['datetime'] = df['transaction_date'] + df['transaction_time']
```

Untuk kebutuhan proses analisis lebih lanjut. Dibuatlah object data baru yaitu **total_sales** dan juga **datetime**.

Mengkontruksi Data

Data Cleaning

melihat data yang hilang

```
df.isnull().sum()
```

```
df.info()
```

```
transaction_id      0
transaction_date     0
transaction_time     0
transaction_qty      0
store_id            0
store_location       0
product_id          0
unit_price           0
product_category     0
product_type         0
product_detail       0
total_sales          0
datetime            0
hour                0
dtype: int64
```

melihat data duplikasi

```
df.duplicated().sum()
```

```
0
```

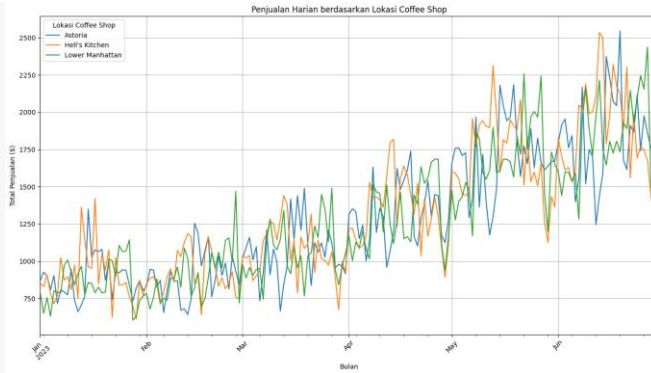
Dari sekian rangkaian data cleaning yang dilakukan seperti mencari missing values dan data yang terduplikat ternyata tidak ada kesalahan pada datasets ini

Mengintegrasikan Data

Menganalisis Tren Penjualan berdasarkan Lokasi Coffee Shop

```
# mencari penjualan harian berdasarkan coffee shop
```

```
daily_sales_by_location = df.groupby(['transaction_date', 'store_location'])['total_sales'].sum().unstack(  
daily_sales_by_location
```

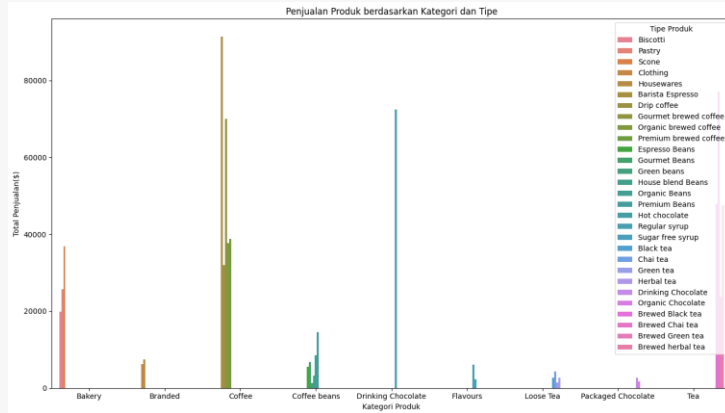


Dari hasil analisis ternyata didapat bahwa penjualan ketiga Lokasi coffee shop mengalami kenaikan seiring berjalannya waktu dan puncak penjualan terdapat pada bulan Juni yang dimana daerah **Astoria** mencapai puncak penjualan tertinggi diikuti oleh daerah **Hells Kitchen** dan daerah **Lower Manhattan**

Menganalisis Kategori produk apa dengan total penjualan tertinggi

Menampilkan penjualan produk berdasarkan kategori dan tipe produk sebagai gambaran

```
product_sales = df.groupby(['product_category', 'product_type'])['total_sales'].sum().reset_index()
product_sales
```



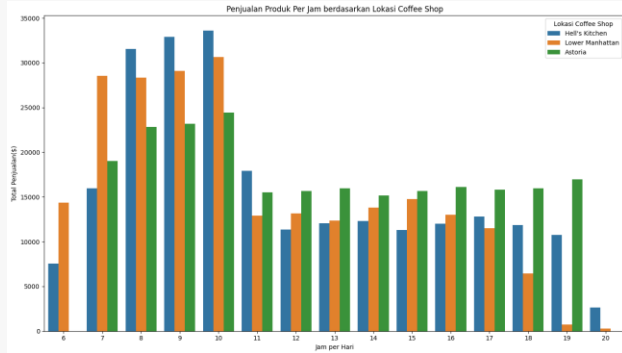
product_category	total_sales
Coffee	269952.45
Tea	196405.95
Bakery	82315.64
Drinking Chocolate	72416.00
Coffee beans	40085.25
Branded	13607.00
Loose Tea	11213.60
Flavours	8408.80
Packaged Chocolate	4407.64

Dari hasil analisis didapat bahwa kategori produk **Coffee** merupakan produk terlaris diikuti oleh **Tea** dan **Bakery**

Menganalisis Total Penjualan Lokasi Coffee Shop Berdasarkan Jam Operasional

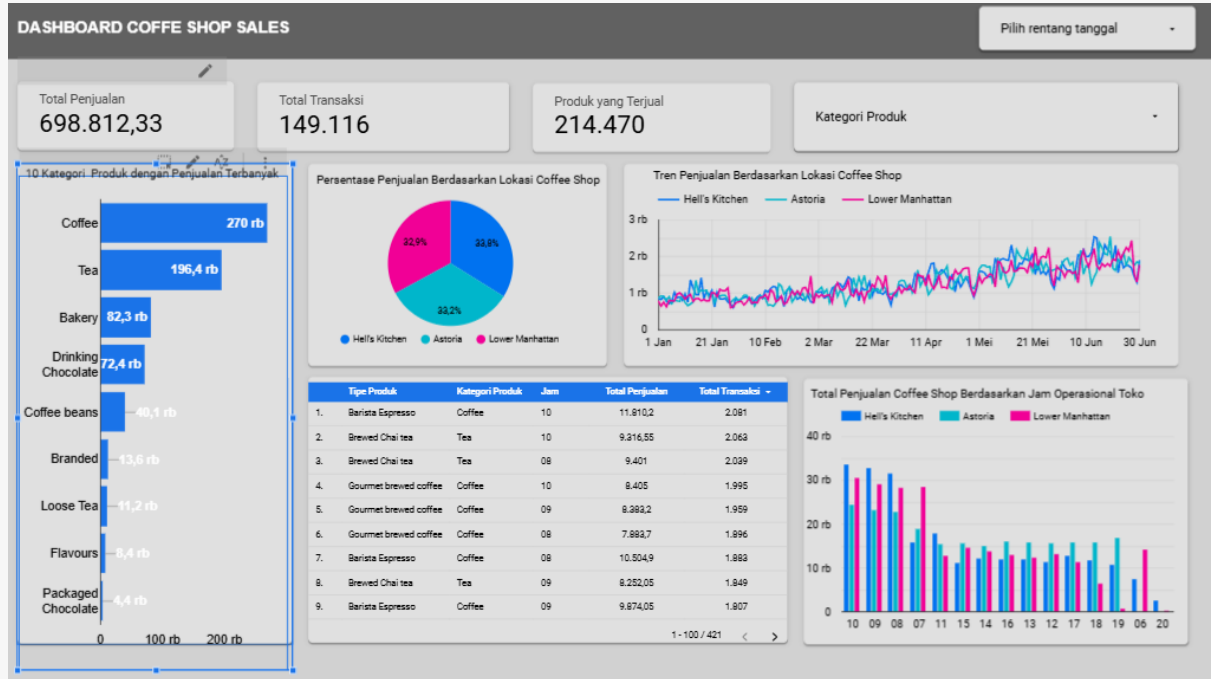
```
# Buat variabel baru untuk mengetahui penjualan berdasarkan jam operasional coffee shop
df['hour'] = df['datetime'].dt.hour
hourly_sales_by_product = df.groupby(['hour', 'product_category'])['total_sales'].sum().unstack()
```

```
# Buat variabel baru untuk mengetahui total penjualan setiap lokasi coffee shop berdasarkan jam operasional
hourly_sales_by_location = df.groupby(['store_location', 'hour'])['total_sales'].sum().reset_index()
hourly_sales_by_location
```



Berdasarkan total penjualan pada tiga toko coffee shop didapat bahwa penjualan tertinggi terdapat pada jam 7 hingga jam 10 pagi. Ini mengindikasikan bahwa kebanyakan pelanggan bertransaksi di coffee shop merupakan kalangan pekerja yang membeli produk pada jam-jam sebelum berangkat bekerja

Membuat Business Intelligence



<https://lookerstudio.google.com/reporting/a4eeb812-75dd-4ad5-912d-9beb05cffc9a>

Conclusion & Recomendations

Conclusion

Dari hasil analisis ini didapat bahwa penjualan ketiga coffee shop terus mengalami kenaikan dari awal bulan tahun 2023 hingga bulan Juni tahun 2023. Lalu penjualan tertinggi mencapai puncaknya pada bulan Juni yang dipimpin oleh coffee shop yang berasal dari daerah Astoria diikuti oleh daerah Hells Kitchen dan yang terakhir daerah Lower Manhattan.

Sementara itu produk yang disenangi oleh konsumen diantaranya adalah coffee, tea dan drinking chocolate yang masing-masing merupakan produk yang memiliki penjualan tertinggi.

Lalu berdasarkan hasil analisis penjualan tertinggi dari ketiga coffee shop ternyata penjualan tertinggi itu terdapat pada jam-jam para pekerja memulai aktivitasnya yaitu sekitar jam 7 hingga jam 10 pagi.

Recomendations

Dari hasil analisis data yang telah dilakukan, coffee shop dari ketiga daerah disarankan untuk terus meningkatkan penjualannya dengan cara memaksimalkan penjualan pada produk-produk yang disenangi oleh konsumen yaitu coffee, tea dan bakery. Lalu coffee shop dari ketiga daerah tersebut juga harus memaksimalkan peak hour yaitu sekitar jam 7 hingga jam 10 pagi dikarenakan dilihat dari tren penjualan jam tersebut lah yang memiliki tingkat penjualan paling tinggi.

TERIMA KASIH

Thanks!