

Portofolio

Rifcki Dwiyansyah

Data Explore Superstore

Project Overview: Project ini merupakan pembuatan sebuah dashboard superstore. Namun sebelum dashboard dibangun, kita perlu mencari tahu hal apa saja yang akan divisualisasikan. maka diperlukanlah data explorer menggunakan SQL

Tools: SQL

Code:

```
SELECT Category, SUM(Sales) AS TotalSales
FROM `capstoneteam16.Superstore.Superstore`
GROUP BY Category
ORDER BY TotalSales DESC;
```

Row	Category	TotalSales
1	Technology	836154.0329999...
2	Furniture	741999.7953000...
3	Office Supplies	719047.0320000...

```
SELECT Category, SUM(Profit) AS TotalProfit
FROM `capstoneteam16.Superstore.Superstore`
GROUP BY Category
ORDER BY TotalProfit DESC;
```

Row	Category	TotalProfit
1	Technology	145454.9481
2	Office Supplies	122490.8008000...
3	Furniture	18451.27280000...

```
SELECT Category, AVG(Discount) AS AVGDisc
FROM `capstoneteam16.Superstore.Superstore`
GROUP BY Category
Order BY AVGDisc ASC;
```

Row	Category ▼	AVGDisc ▼
1	Technology	0.132322685435...
2	Office Supplies	0.157285097909...
3	Furniture	0.173922677982...

```
SELECT State, SUM(Profit) AS TotalProfit
FROM `capstoneteam16.Superstore.Superstore`
Where Profit < 0
GROUP BY 1
ORDER BY 2 ASC
```

Row	State ▼	TotalProfit ▼
1	Texas	-36813.1875
2	Ohio	-21750.0001999...
3	Pennsylvania	-21602.8514999...
4	Illinois	-19501.6974999...

```
WITH Profit_segment_state AS (
  SELECT State, Segment, Profit, ROW_NUMBER() OVER (PARTITION BY State ORDER BY Profit DESC) AS profit_rank
  FROM `capstoneteam16.Superstore.Superstore`
)
SELECT State, Segment, Profit
FROM Profit_segment_state
WHERE profit_rank = 1
ORDER BY State ASC;
```

Row	State ▼	Segment ▼	Profit ▼
1	Alabama	Corporate	1459.2
2	Arizona	Consumer	211.4955
3	Arkansas	Consumer	843.1706
4	California	Corporate	1906.485

Pre-processing data sentiment

Project Overview: Project ini merupakan pembuatan Machine Learning untuk mengidentifikasi sebuah Headline berita mengenai saham. Namun sebelum diimplementasikan kedalam sebuah model, sebuah proses Pre-processing perlu dilakukan untuk mendapatkan akurasi yang baik.

Tools: Python

Code:

Pertama adalah menyimpan dataset kedalam sebuah variable DF

```
df = pd.read_csv('stock_senti_analysis.csv', encoding = 'Latin-1')
df.head(5)
```

	Date	Label	Top1	Top2	Top3	Top4	Top5
0	2000-01-03	0	A 'hindrance to operations': extracts from the...	Scorecard	Hughes' instant hit buoys Blues	Jack gets his skates on at ice-cold Alex	Chaos as Maracana builds up for United
1	2000-01-04	0	Scorecard	The best lake scene	Leader: German sleaze inquiry	Cheerio, boyo	The main recommendations

Selanjutnya adalah melihat apakah ada nilai NaN pada dataframe df

```
df.isna().any()

Date      False
Label     False
Top1      False
Top2      False
Top3      False
Top4      False
Top5      False
Top6      False
Top7      False
Top8      False
Top9      False
Top10     False
Top11     False
Top12     False
Top13     False
Top14     False
Top15     False
Top16     False
Top17     False
Top18     False
Top19     False
Top20     False
Top21     False
Top22     False
Top23     True
Top24     True
Top25     True
dtype: bool
```

Setelah mengetahui terdapat null, selanjutnya adalah menghapus row data yang kosong.

```
[ ] df.dropna(inplace=True)
    print(df.shape)
```

```
➡ (4098, 27)
```

```
▶ #lowercasing
for col in trainclean.columns:
    trainclean[col]=trainclean[col].str.lower()

for col in testclean.columns:
    testclean[col]=testclean[col].str.lower()
```

+ Code

+ Text

Penjelasan Code diatas : code diatas digunakan untuk melakukan lowercasing setiap kata kata yang ada di dalam trainclean dan testclean hal ini dibuat dengan tujuan agar isi datanya menjadi lebih konsisten.

```
[ ] #remove_punctuation
import string

def remove_punctuations(text): # Untuk menghapus tanda baca
    for punctuation in string.punctuation:
        text = text.replace(punctuation, '')
    return text # mengembalikan teks yang sudah dihapus tanda bacanya setelah proses
```

Penjelasan Code diatas : code diatas adalah code untuk menghapus tanda baca. hal ini dilakukan karena tanda baca tidak diperlukan pada sentiment analysis. selain untuk

```
▶ #Menggabungkan seluruh kata data

traindat = []
testdat = []

for row in range(0, trainclean.shape[0]):
    traindat.append(' '.join(str(x) for x in trainclean.iloc[row, 2:25]))

for row in range(0, testclean.shape[0]):
    testdat.append(' '.join(str(x) for x in testclean.iloc[row, 2:25]))

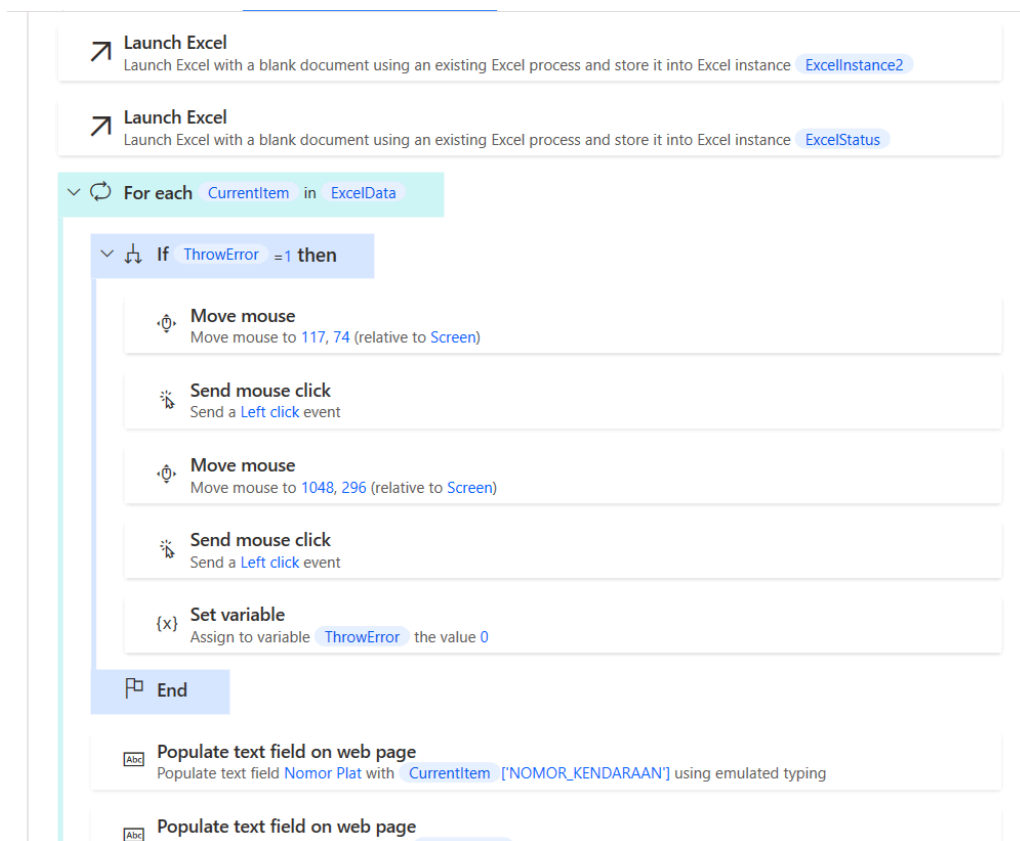
print(traindat)
```

Scrapping Data Samsat

Project Overview: Project ini dilakukan untuk melakukan scrapping detail data plat nomor dari website SAMSAT. Tujuannya adalah mengetahui jenis mobil yang digunakan oleh para pengunjung suatu event yang memiliki data plat nomor.

Tools: Tools yang digunakan adalah Microsoft Power Automate dan Google Spreadsheet

Step 1. Automate Scrapping Data



Step 2. Data Cleaning

Merubah data mentah yang di dapat dari scrapping

Informasi data kendaraan dan pajak
kendaraan bermotor
pemerintah provinsi DKI Jakarta
05 Agustus 2024 - 13:36:32
Nopol
#####
Kendaraan ke
1
Nama
XXXXX XXXXX
NIK
XXXXX
Alamat
XXXXX XXX XX
No. Rangka / Mesin
XXXXXX / XXXXXXXX
No BPKB
XXXXX
Merek / Type
WULING / E230REV30KW3LV24X2AT
Model / Pembuatan
MICRO/MINIBUS Tahun: 2022
Warna Kendaraan
PUTIH
Warna TNKB
HITAM
Bhn Bakar / Cylinder
LISTRIK / 30 kWh
Masa Berlaku: 5 TAHUN

Menjadi data yang siap digunakan. Data ini diubah menggunakan Microsoft Excel dengan memanfaatkan Split, IF Else dan lainnya. Setelah itu data siap digunakan sesuai kebutuhan.

D	E	F	G	H	I	J	K
NO PLAT	KODE_WILAYAH	MOR_KENDARA	ODE_BELAKAN	Wilayah	Lingkup	n Bakar / Cylin	Nilai Jual
				Kota Administr	JABODETABEK	LISTRIK / 30 kW	Rp. 185.000.00
				Kota Administr	JABODETABEK	LISTRIK / 61 kW	Rp. 347.000.00
				Kota Administr	JABODETABEK	BENSIN / 1498	Rp. 118.000.00
				Kota Administr	JABODETABEK	SOLAR / 2977 C	Rp. 0
				Kota Administr	JABODETABEK	BENSIN / 1499	Rp. 215.000.00
				Kota Administr	JABODETABEK	BENSIN / 1497	Rp. 202.000.00
				Kota Administr	JABODETABEK	BENSIN / 1998	Rp. 197.000.00
				Kota Administr	JABODETABEK	BENSIN / 1497	Rp. 154.000.00
				Kota Administr	JABODETABEK	BENSIN / 1298	Rp. 72.000.00
				Kota Administr	JABODETABEK	BENSIN / 1798	Rp. 132.000.00
				Kota Administr	JABODETABEK	BENSIN / 1495	Rp. 83.000.00
				Kota Administr	JABODETABEK	BENSIN / 1498	Rp. 112.000.00
				Kota Administr	JABODETABEK	BENSIN / 1997	Rp. 323.000.00
				Kota Administr	JABODETABEK	BENSIN / 1497	Rp. 223.000.00
				Kota Administr	JABODETABEK	BENSIN / 1197	Rp. 114.000.00
				Kota Administr	JABODETABEK	BENSIN / 1497	Rp. 243.000.00
				Kota Administr	JABODETABEK	BENSIN / 1197	Rp. 129.000.00
				Kota Administr	JABODETABEK	BENSIN / 1498	Rp. 302.000.00
				Kota Administr	JABODETABEK	LISTRIK / 70 kW	Rp. 244.000.00

Data Visualization

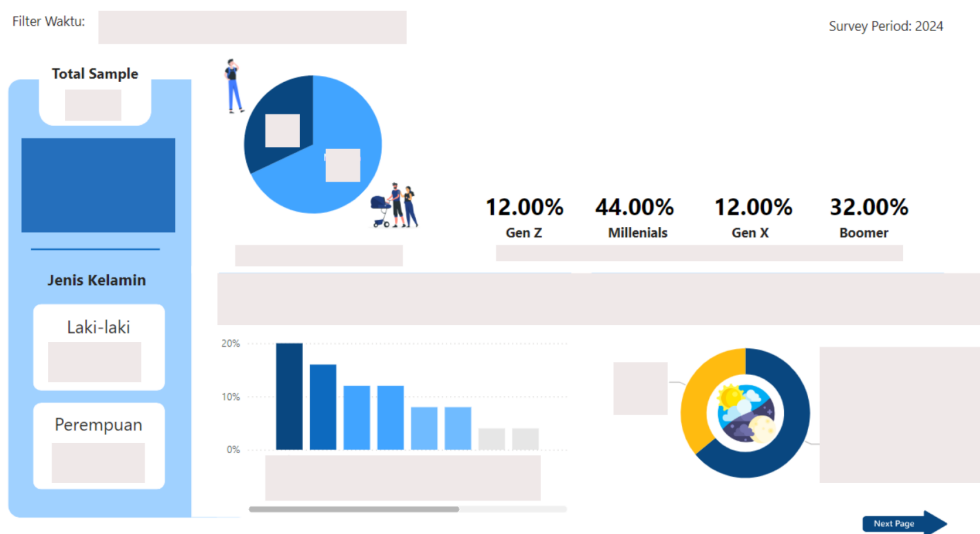
Dashboard Pengunjung

Project Overview: Project ini dibuat untuk melihat demografi pengunjung dari suatu tempat. Dashboard ini menampilkan personalisasi profile pengunjung dan minat properti dari para pengunjung. Project ini dibuat dengan menggunakan software Visualisasi Microsoft Power BI dan menggunakan data yang di buat sendiri.

Tools: Microsoft Power BI

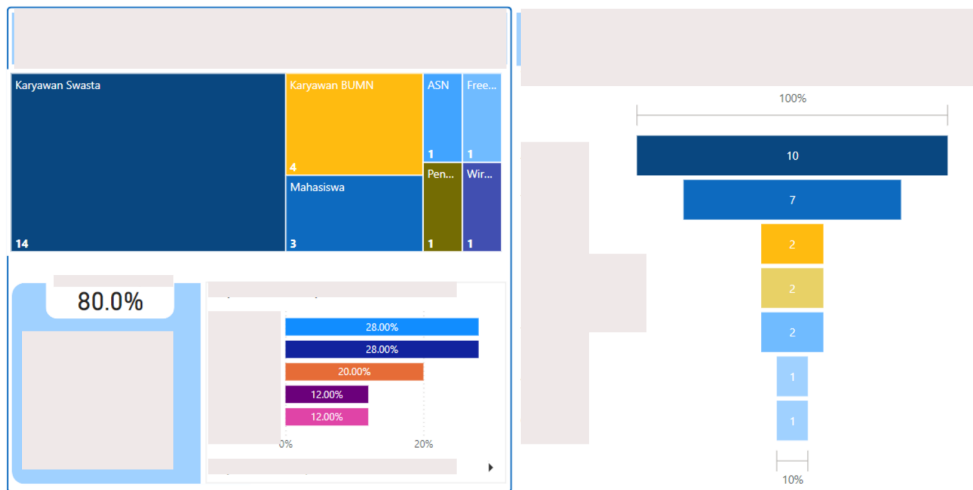
Skill: Data Understanding, Data Visualization

Project:



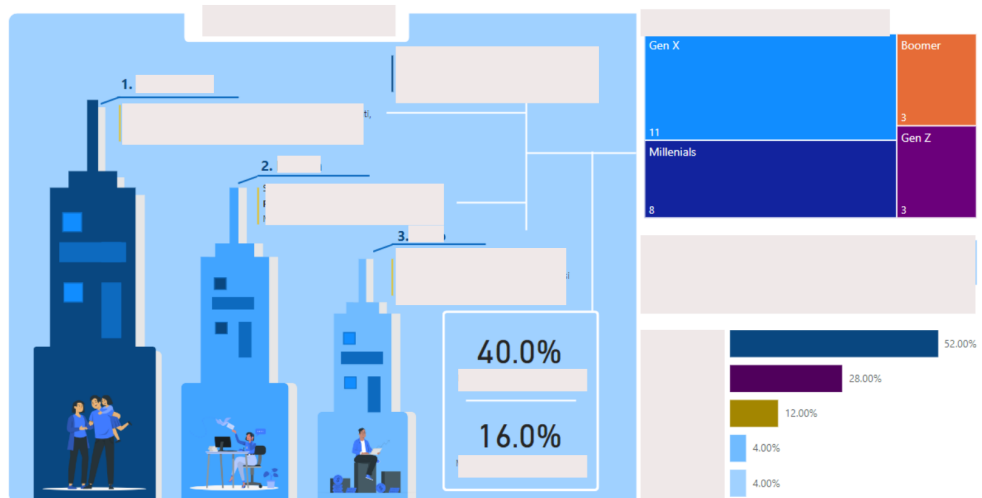
Filter Waktu:

Survey Period: 2024



Filter Waktu:

Survey Period: 2024



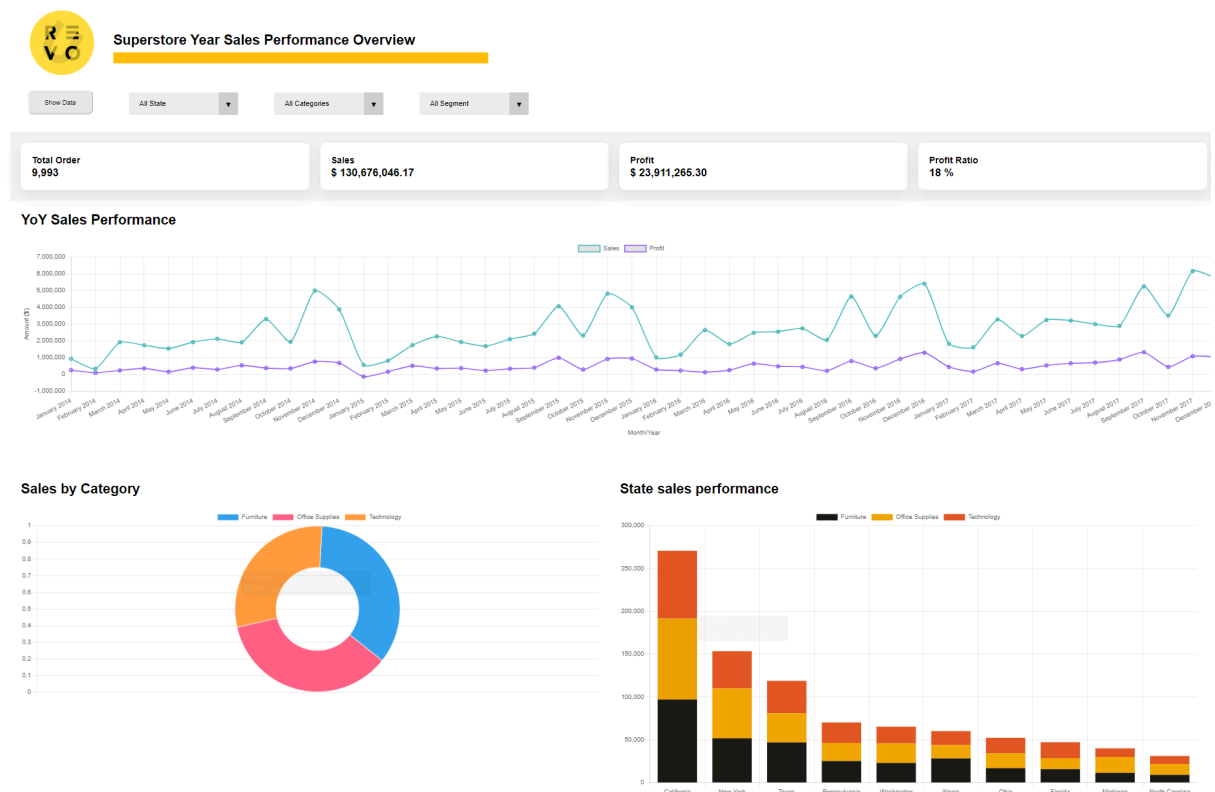
Dashboard Superstore Performance

Project Overview: Project ini dibuat untuk melihat performance dari suatu perusahaan retail. Yang tujuannya mengetahui penjualan dari sebuah superstore. Data yang digunakan adalah data public Superstore. Dashboard dibuat dalam 2 bentuk yaitu menggunakan Looker Studio dan kemudian dibuat ulang menggunakan HTML, CSS, dan Javascript.

Tools: Google Looker Studio, HTML, CSS, Javascript

Skill: Skill yang digunakan disini adalah analisis mengenai bagaimana dashboard dibuat sesuai dengan kebutuhan. selain itu project ini juga memerlukan skill Data Visualization.

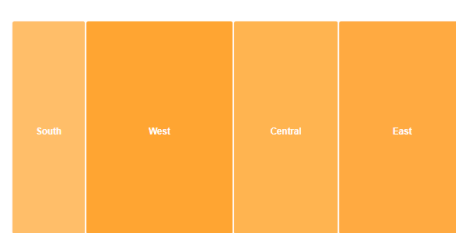
Project:



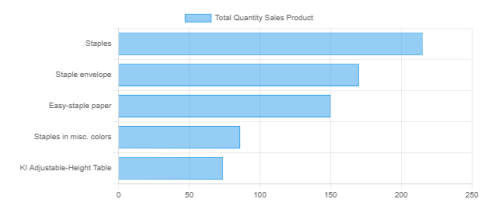
Top Performanced Salesman

Salesman	Total Sales
Thais Sisman	484.00
Lela Donovan	530.00
Carl Jackson	1,652.00
Mitch Gastineau	1,674.00
Roy Skaria	2,233.00

Sales per Region



Top 5 Product Sold based on Quantity Sales



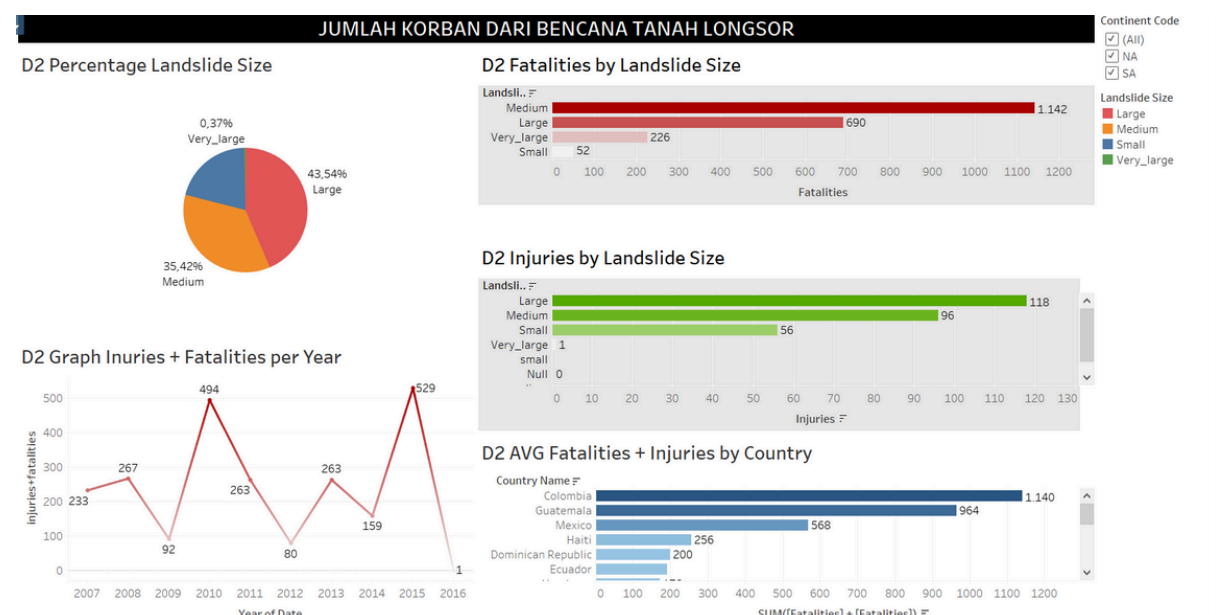
Dashboard Tanah Longsor

Project Overview: Project ini dibuat untuk melihat dampak dari tanah longsor yang terjadi di beberapa tempat di Amerika. Project ini dibuat dengan menggunakan software Visualisasi Tableau dan menggunakan data public dari kaggle.

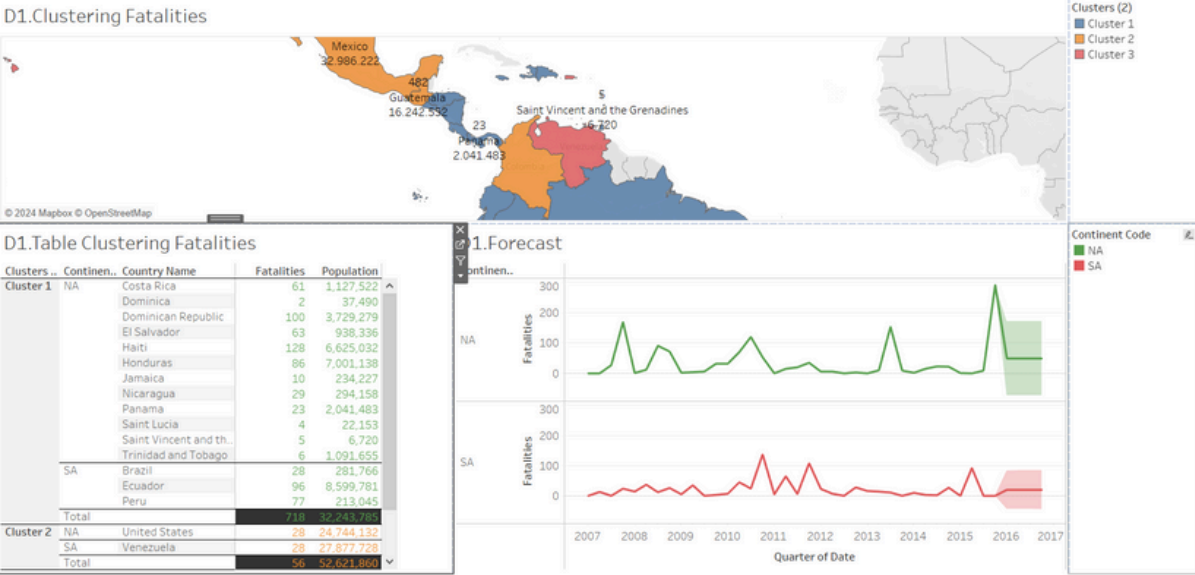
Tools: Tableau

Skill: Data Understanding, Data Visualization

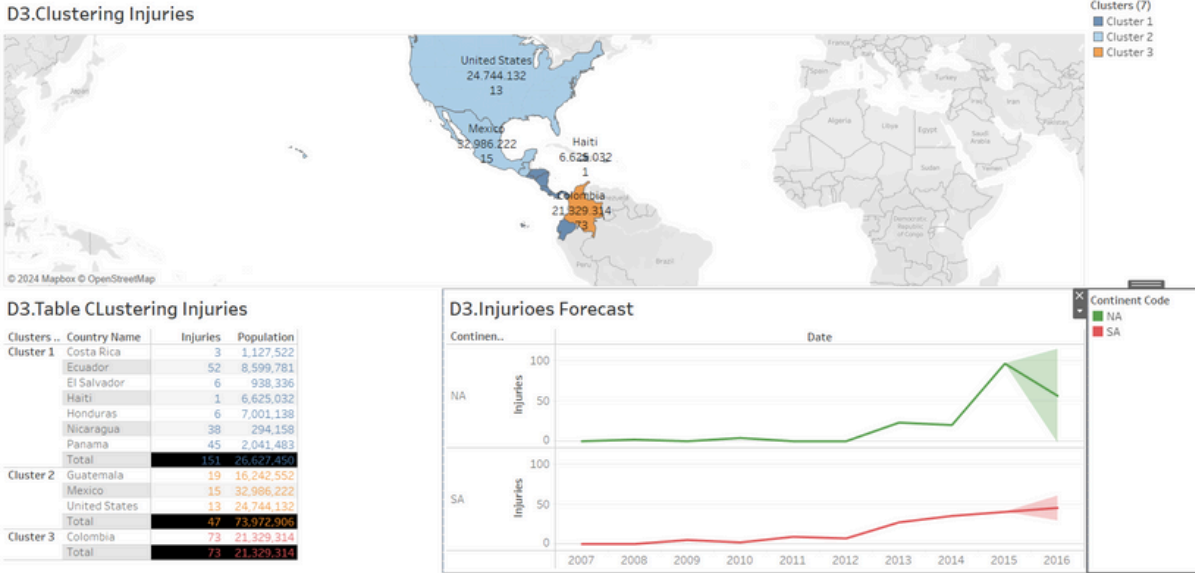
Project:



DASHBOARD CLUSTERING & FORECASTING FATALITIES BY LANDSLIDE



DASHBOARD CLUSTERING & FORECASTING INJURIES BY LANDSLIDE



Dashboard Sales Car (Microsoft Power BI)

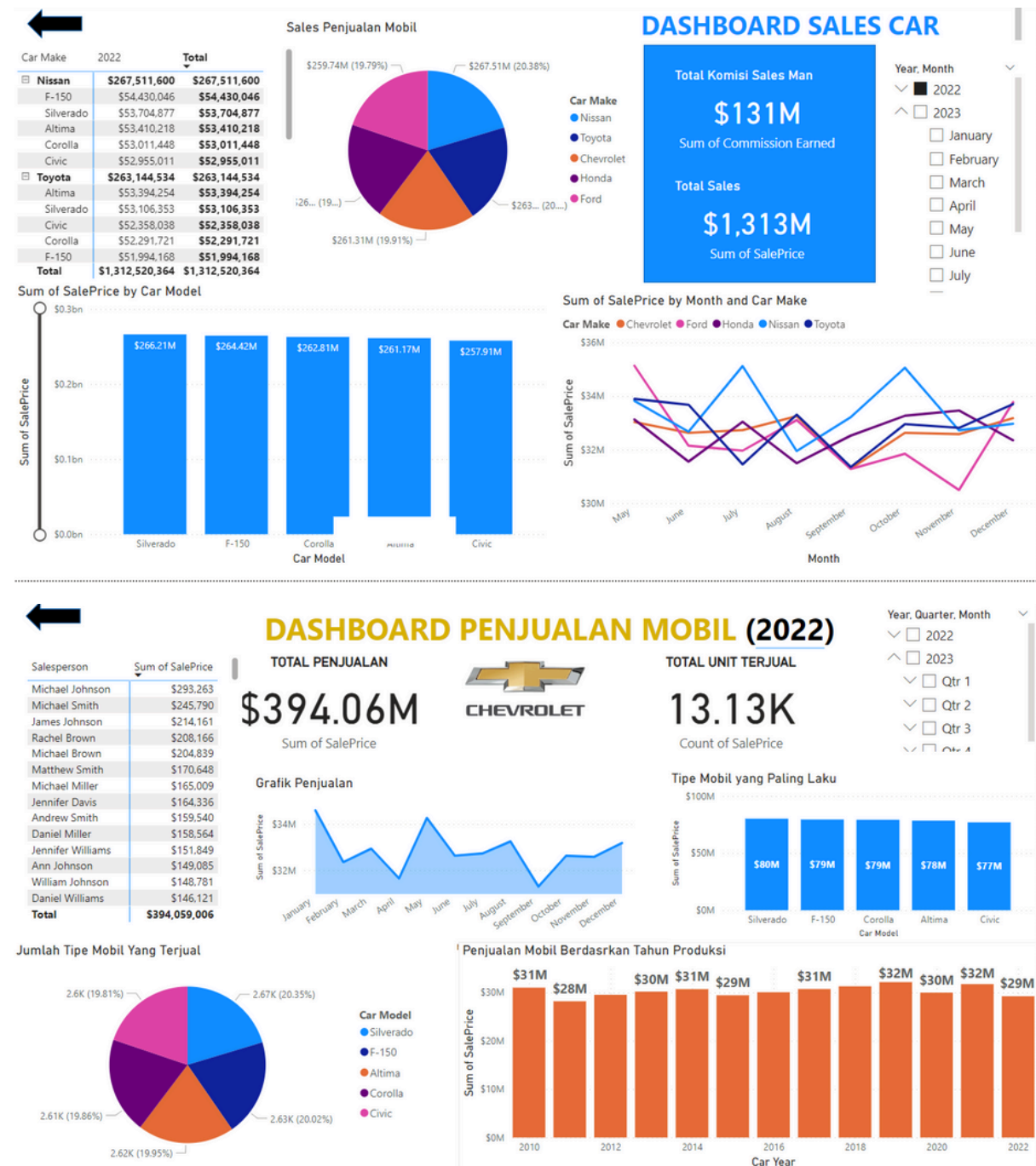
Project Overview: Project ini dibuat untuk melihat performan penjualan Mobil. Dashboard ini menampilkan penjualan, detail mobil dan performa sales. Project

ini dibuat dengan menggunakan software Visualisasi Microsoft Power BI dan menggunakan data public dari kaggle.

Tools: Microsoft Power BI

Skill: Data Understanding, Data Visualization

Project:



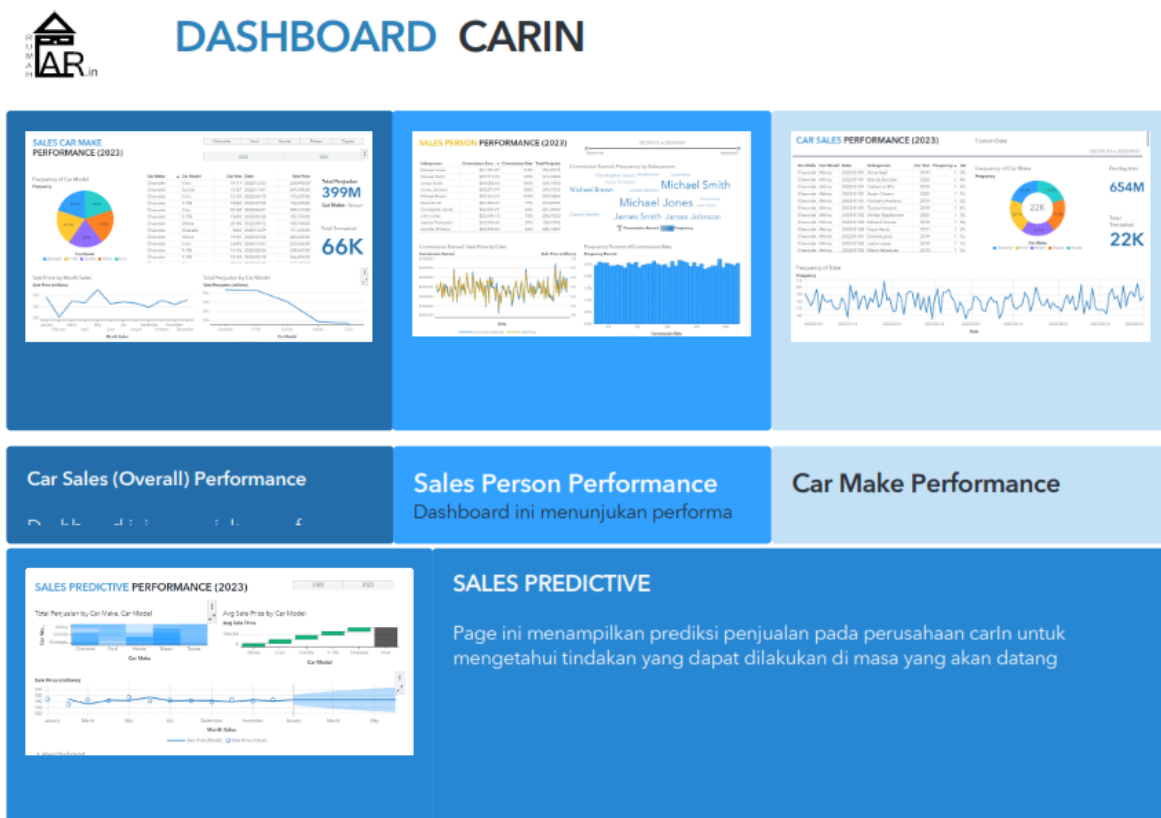
Dashboard Sales Car (SAS Viya)

Project Overview: Project ini dibuat untuk melihat performan penjualan Mobil. Dashboard ini menampilkan penjualan, detail mobil dan performa sales. Project ini dibuat dengan menggunakan software Visualisasi SAS Visual Analytics dan menggunakan data public dari kaggle.

Tools: SAS Visual Analytics

Skill: Data Understanding, Data Visualization

Project:



CAR SALES PERFORMANCE

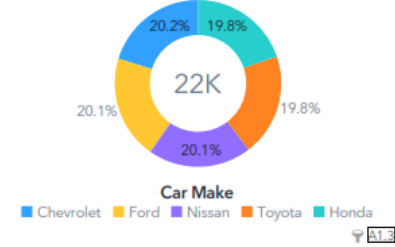
Custom Date

🔍 A1.1

Car Make	Car Model	Date	Salesperson
Chevrolet	Altima	2023/01/01	Anne Neal
Chevrolet	Altima	2023/01/01	Brandy Escobar
Chevrolet	Altima	2023/01/01	Catherine Ellis
Chevrolet	Altima	2023/01/01	Austin Owens
Chevrolet	Altima	2023/01/01	Kimberly Arellano
Chevrolet	Altima	2023/01/01	Tyrone Howard
Chevrolet	Altima	2023/01/02	Amber Stephenson
Chevrolet	Altima	2023/01/02	Edward Graves
Chevrolet	Altima	2023/01/02	Kayla Hardy

🔍 A1.2

Frequency of Car Make



Pendapatan

654M

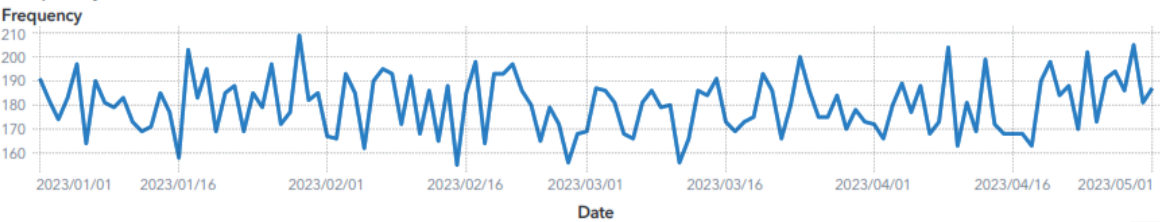
🔍 A1.3

Total Transaksi

22K

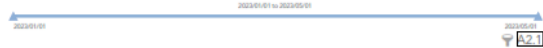
🔍 A1.4

Frequency of Date



🔍 A1.5

SALES PERSON PERFORMANCE



🔍 A2.1

Salesperson	Commission Earned
Michael Jones	\$41,201.09
Michael Smith	\$39,914.94
James Smith	\$34,058.46
James Johnson	\$33,271.97
Michael Brown	\$29,967.63
David Smith	\$27,482.03
Christopher Jones	\$26,039.69

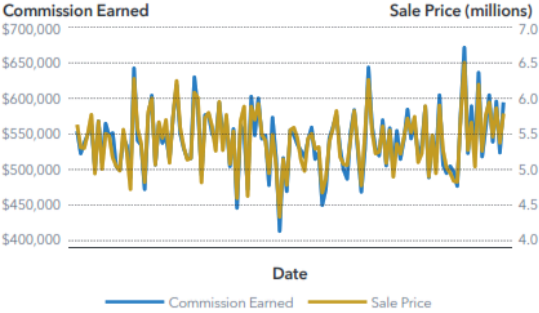
🔍 A2.2

Commission Earned, Frequency by Salesperson



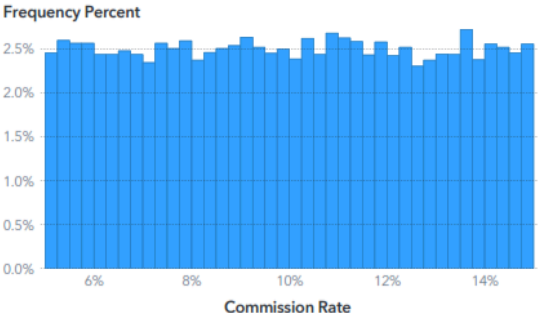
🔍 A2.3

Commission Earned, Sale Price by Date



🔍 A2.4

Frequency Percent of Commission Rate

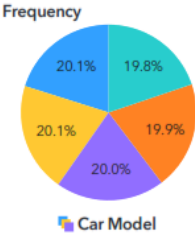


🔍 A2.5

SALES CAR MAKE
PERFORMANCE (2023)

Chevrolet	Ford	Honda	Nissan	Toyota
2022			2023	

Frequency of Car Model



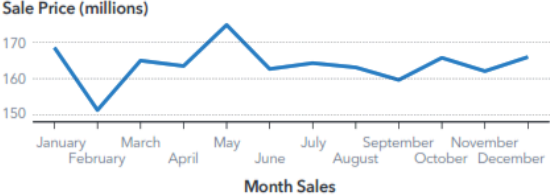
Car Make	Car Model	Car Year	Date	Sale Price
Chevrolet	Civic	14119	2022/12/23	230,490.00
Chevrolet	Corolla	16127	2022/11/01	247,340.00
Chevrolet	Civic	12103	2022/06/10	173,637.00
Chevrolet	F-150	10080	2022/07/05	106,020.00
Chevrolet	Civic	22149	2023/04/01	299,313.00
Chevrolet	F-150	12093	2023/01/30	152,793.00
Chevrolet	Altima	24184	2022/05/12	390,764.00
Chevrolet	Silverado	8062	2022/12/29	117,635.00
Chevrolet	Altima	14107	2022/07/22	223,250.00
Chevrolet	Civic	12092	2022/12/01	212,566.00

Total Transaksi

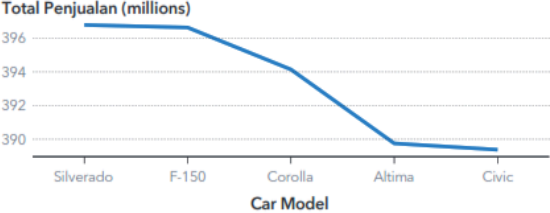
66K

Total Penjualan
399M
Car Make:
Nissan

Sale Price by Month Sales



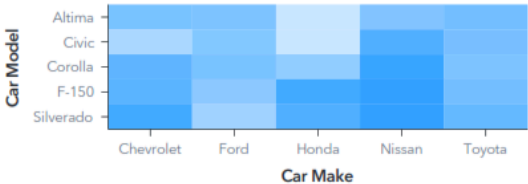
Total Penjualan by Car Model



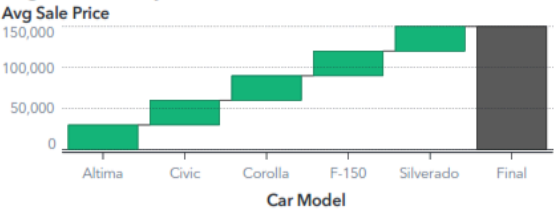
SALES PREDICTIVE PERFORMANCE (2023)

2022	2023
------	------

Total Penjualan by Car Make, Car Model



Avg Sale Price by Car Model



Sale Price (millions)



95% forecast confidence.

Lampiran

