

Superstore Performance January 2017 – December 2020

Project-PS 10-Melani Ayu Azizah



About Superstore

Superstore adalah perusahaan dagang yang menjual berbagai kategori produk seperti peralatan kantor, furniture bahkan teknologi seperti handphone. Superstore perlu melakukan evaluasi penjualan yang nantinya dapat memberikan gambaran bagi Superstore dalam menjalankan usahanya ke depan. Adanya evaluasi ini dapat pula menyelesaikan masalah yang ditemukan pada saat akhir pelaporan.

Data Understanding

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 20 columns):
#   Column              Non-Null Count  Dtype  
---  -
0   Order_ID            9994 non-null   object  
1   Customer_ID         9994 non-null   object  
2   Postal_Code         9994 non-null   int64   
3   Product_ID          9994 non-null   object  
4   Sales               9994 non-null   float64  
5   Quantity            9994 non-null   int64   
6   Discount            9994 non-null   float64  
7   Profit              9994 non-null   float64  
8   Category            9994 non-null   object  
9   Sub-Category        9994 non-null   object  
10  Product_Name         9994 non-null   object  
11  Order_Date           9994 non-null   object  
12  Ship_Date            9994 non-null   object  
13  Ship_Mode            9994 non-null   object  
14  Customer_Name        9994 non-null   object  
15  Segment             9994 non-null   object  
16  Country/Region       9994 non-null   object  
17  City                 9994 non-null   object  
18  State                9994 non-null   object  
19  Region              9994 non-null   object  
dtypes: float64(3), int64(2), object(15)
memory usage: 1.5+ MB
```

Superstore memiliki 20 kolom dengan jumlah baris 9994. Pada 20 kolom tersebut terdapat sales, profit dan discount. Kolom-kolom yang akan dianalisis sales, quantity, discount, profit, order date & ship date serta kolom pendukung seperti category, sub-category, ship mode, segment dan city.

Data Understanding

Berdasarkan kolom yang sudah diketahui, akan dilakukan analisis berupa

1. Apakah profit pada superstore memiliki persentase yang besar dibandingkan sales
2. Apakah discount membuat sales dan profit meningkat
3. Bagaimana performa pengiriman pada superstore

Data Preprocessing

Missing Data

```
✓ [6] df.isnull().sum().sum()
0 d
↔ 0
```

Data superstore tidak memiliki missing data sehingga dilanjutkan ke tahap preprocessing kedua

Duplicate Data

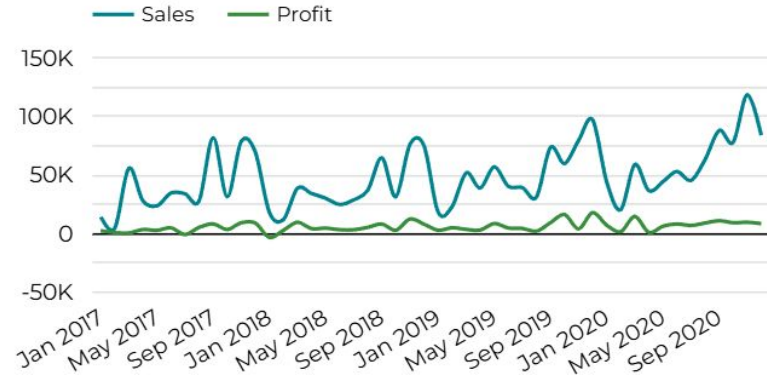
```
✓ [7] df.duplicated().sum()
0 d
↔ 1
```

Data superstore memiliki baris yang duplikat sehingga perlu dilakukan penghapusan

```
✓ [8] df.drop_duplicates(inplace=True)
      df.duplicated().sum()
0 d
↔ 0
```

Data duplikat telah dihapus sehingga data menjadi 9993 baris dan sudah dapat dianalisis lebih lanjut

Data Analytics



Line Chart di atas menunjukkan bahwa penjualan superstore memiliki **pola yang musiman** pada bulan September naik, Oktober turun, November naik serta Desember turun lagi di setiap tahun baik 2017, 2018, 2019 dan 2020.

Pada bulan Februari setiap tahun, penjualannya menurun tajam.

Grafik profit mengikuti sales namun ketika sales naik tajam, profit tidak naik secara drastis, sedangkan saat sales menurun, profit menurun tajam bahkan rugi.

Data Analytics

Sub-Category	Sales	Profit	Discount
Phones	330,007.05	11.92%	0.15
Chairs	328,167.73	4.4%	0.17
Storage	223,843.61	8.91%	0.07
Tables	206,965.53	-14.77%	0.26
Binders	203,412.73	-19.96%	0.37
Machines	189,238.63	-7.2%	0.31
Accessories	167,380.32	21.82%	0.08
Copiers	149,528.03	31.72%	0.16
Bookcases	114,880	-12.66%	0.21
Appliances	107,532.16	-15.69%	0.17
Furnishings	91,705.16	13.71%	0.14
Paper	78,479.21	42.56%	0.07
Supplies	46,673.54	11.2%	0.08
Art	27,118.79	25.16%	0.07
Envelopes	16,476.4	42.31%	0.08
Labels	12,486.31	42.97%	0.07
Fasteners	3,024.28	29.92%	0.08

1 - 17 / 17 < >

Berdasarkan tabel di samping, sales tinggi (berwarna biru tua) salah satu penyebabnya adalah discount yang tinggi (berwarna merah tua) pula kecuali pada storage. Hal ini dapat disimpulkan bahwa discount yang tinggi dapat menyebabkan konsumen terus membeli barang dari superstore sehingga sales atau penjualan di superstore juga tinggi.

Data Analytics

Sub-Category	Sales	Profit	Discount
Labels	12,486.31	42.97%	0.07
Paper	78,479.21	42.56%	0.07
Envelopes	16,476.4	42.31%	0.08
Copiers	149,528.03	31.72%	0.16
Fasteners	3,024.28	29.92%	0.08
Art	27,118.79	25.16%	0.07
Accessories	167,380.32	21.82%	0.08
Furnishings	91,705.16	13.71%	0.14
Phones	330,007.05	11.92%	0.15
Supplies	46,673.54	11.2%	0.08
Storage	223,843.61	8.91%	0.07
Chairs	328,167.73	4.4%	0.17
Machines	189,238.63	-7.2%	0.31
Bookcases	114,880	-12.66%	0.21
Tables	206,965.53	-14.77%	0.26
Appliances	107,532.16	-15.69%	0.17
Binders	203,412.73	-19.96%	0.37

1 - 17 / 17 < >

Persentase profit terhadap sales pada setiap kategori cukup beragam. Pada kategori **Labels** profitnya mencapai 42% namun pada **Binders** rugi sebesar 19,96%. Binders merupakan sub-category paling tidak menguntungkan.

Dapat dilihat pula ternyata kerugian pada binders ini disebabkan oleh sangat besarnya discount yang ditetapkan yaitu 0,37.

Dari 2 analisis terhadap sales, profit, dan discount dapat disimpulkan bahwa adanya **discount memang menaikkan sales namun sangat mengurangi profit**. Sehingga perlu adanya kebijakan penetapan discount pada barang secara lebih selektif.

Data Analytics

Order_ID ^	Shipping Mode	Shipping Performance
CA-2017-100006	Standard Class	83.33%
CA-2017-100090	Standard Class	125%
CA-2017-100293	Standard Class	125%
CA-2017-100328	Standard Class	83.33%
CA-2017-100363	Standard Class	71.43%
CA-2017-100391	Standard Class	125%
CA-2017-100678	Standard Class	125%
CA-2017-100706	Second Class	150%
CA-2017-100762	Standard Class	100%
CA-2017-100860	Second Class	75%
CA-2017-100867	Standard Class	100%
CA-2017-100881	Standard Class	125%

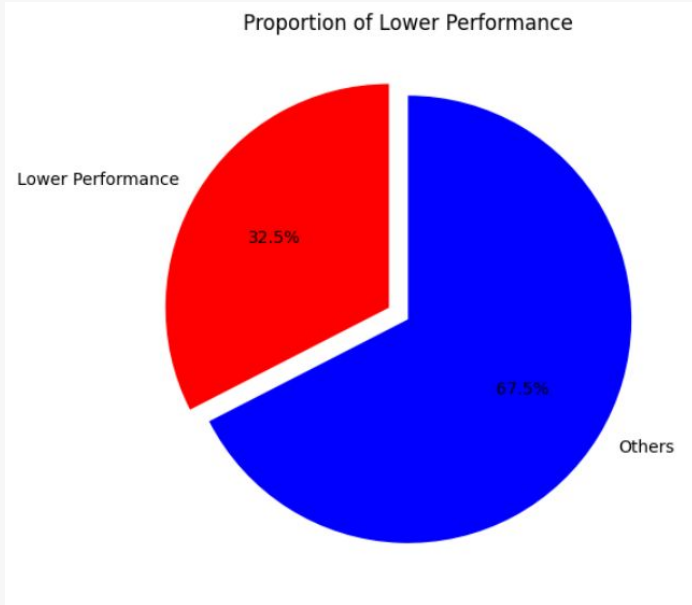
1 - 100 / 5009 < >

[3251 rows x 3 columns]

Selain melihat data mengenai sales dan profit, dapat dilihat pula performa pengiriman produk dari superstore. Ada 3 mode pengiriman yaitu Standard Class, Second Class, First Class dan Same Day. Shipping performance 100% menandakan waktu tempuh produk sama dengan rata rata waktu tempuh pada mode pengiriman produk itu sendiri. Bila <100% maka produk kurang dari rata rata dan performanya kurang baik.

Ada 3251 dari 9993 pesanan yang performanya kurang baik.

Data Analytics



Sebesar 32,5% dari 9993 pesanan, memiliki performa pengiriman yang kurang. Ini perlu dibenahi oleh Superstore agar setiap pesanan yang dikirim memiliki waktu pengiriman yang sama rata pada setiap mode pengiriman.

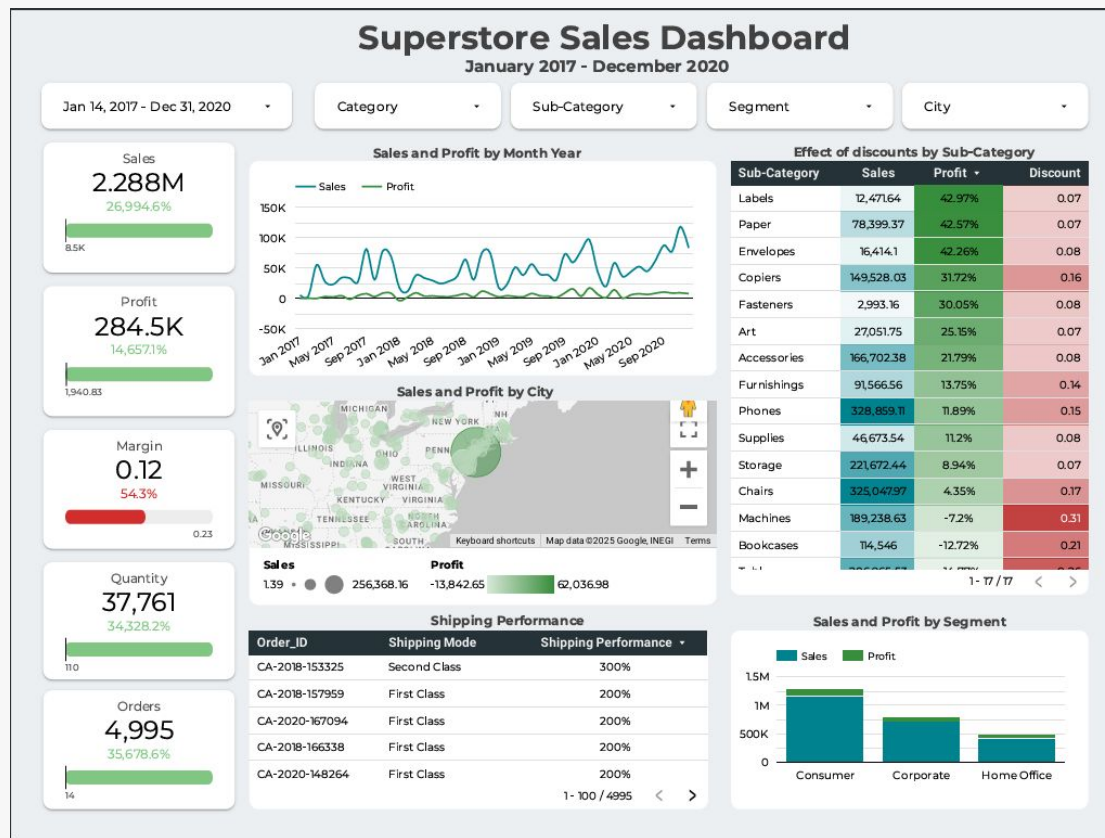
Solutions

Menetapkan diskon yang tepat

Penetapan diskon ini dapat dilakukan dengan menggunakan **analisis regresi**, dengan variabel independen nya adalah sales dan discount serta variabel dependen nya adalah profit. Analisis ini digunakan dengan data historis sehingga bila modelnya sudah ditemukan, dapat dilihat berapa besar discount sebelum profit mencapai nilai 0.

Evaluasi kinerja pengiriman

Melakukan **pengumpulan data** pada sistem kerja pengiriman sehingga dapat melakukan pembaharuan pada sistem kerja dan membenahi performa pengiriman.



Thank you!

Find the Dashboard [here!](#)

Find my Github [here!](#)

Thanks!

