

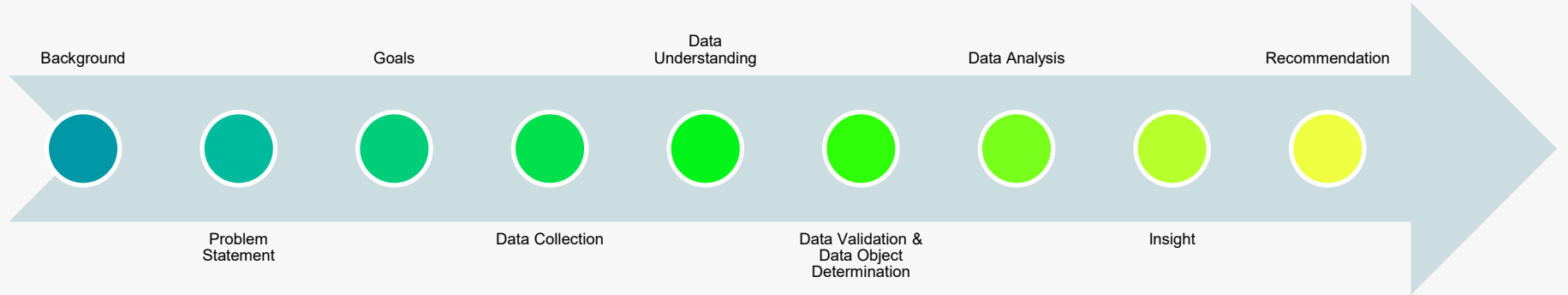
# Analisis Segmentasi Pelanggan dan Diskon SaaS AWS

Final Project Bootcamp Data Analyst  
DQLab Batch 16

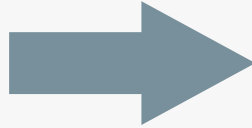
Rahmawati Annisa Salsadilla



# Content

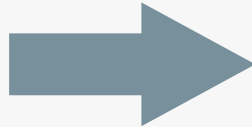


# Background



## Identifikasi & Strategi Aset Data:

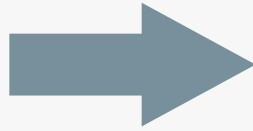
- AWS mengelola data transaksi layanan SaaS B2B sebagai aset utama untuk mendukung strategi penghematan biaya pelanggan melalui model Pay less by using more, Bundled Discounts, dan Savings Plan.
- Strategi ini sejalan dengan tujuan bisnis AWS untuk retensi pelanggan dan pertumbuhan pasar, memastikan kualitas data yang berkelanjutan untuk analisis diskon.



## Tata Kelola, Model Operasional, & Data Stewardship:

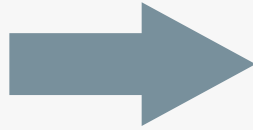
- Tata kelola data di AWS berfokus pada integritas dan ketersediaan aset data transaksi inti yang berbeda dengan tata kelola TI. Model operasionalnya cenderung terpusat untuk kebijakan diskon dan terdistribusi untuk data spesifik pelanggan.
- Aspek data stewardship mencakup pengelolaan metadata rinci, dokumentasi aturan/standar diskon yang jelas, dan penjaminan kualitas data penggunaan untuk akurasi diskon dan efektivitas Savings Plan, didukung oleh aktivitas tata kelola data operasional.

# Background



## Diskon Penggunaan & Bundel:

AWS membantu perusahaan menghemat biaya dengan diskon volume (Pay less by using more) dan Bundled Discounts seiring peningkatan penggunaan.



## Savings Plan:

Pelanggan dapat merencanakan dan menghemat dengan berkomitmen pada pengeluaran minimum per jam selama 1-3 tahun, memberikan diskon signifikan dan prediktabilitas biaya yang tidak dimiliki pesaing.

# Problem Statement

**Perusahaan ingin memahami bagaimana berbagai besaran diskon mempengaruhi sales dan profit.** Oleh karenanya analisis ini akan mengelompokkan pelanggan ke dalam segmen menggunakan rentang interquartile sales dimana:

- kuartil dibawah 25% masuk kedalam segmen low sales,
- kuartil antara 25% - 75% masuk kedalam segmen medium sales,
- kuartil diatas 75% masuk kedalam high sales.

Dari segmen tersebut, akan dianalisis berapa ambang batas diskon yang ideal dapat menaikkan sales tetapi tidak menurunkan profit sampai negatif. Informasi ini akan membantu perusahaan mengembangkan strategi diskon untuk meningkatkan sales dan profit kepada pelanggan di segmen medium sales dan low sales.

Dalam analisis ini, akan mencoba menjawab pertanyaan berikut:

1. Pelanggan mana saja yang dikategorikan sebagai high, medium, dan low sales?
2. Bagaimana berbagai besaran diskon mempengaruhi penjualan dan keuntungan pada setiap kategori pelanggan?

# Goals



Dengan mengidentifikasi segmen pelanggan dan menganalisis besaran diskon yang ideal, dapat membantu perusahaan merumuskan strategi diskon yang lebih tepat sasaran untuk mengoptimalkan sales dan profit, khususnya untuk segmen pelanggan yang berada di kategori medium dan low sales.

# Data Collection

## 1. Perencanaan Data

Data yang digunakan pada analisis ini adalah transaksi AWS SaaS (penjualan, keuntungan, diskon) untuk memahami pengaruh diskon terhadap sales dan profit pelanggan. Data tersedia dalam dataset "SaaS-Sales.csv" pada Kaggle dengan volume yang memadai untuk analisis pengelompokan pelanggan berdasarkan kategori penjualan (low, medium, high).

## 2. Pengambilan Data

Data pada Kaggle diunduh kemudian diunggah ke Python di Google Colab dengan library pandas karena kemampuannya dalam memproses file CSV dan manipulasi data tabular.

## 3. Data Siap untuk Analisis

Data transaksi AWS SaaS berada dalam satu dataframe yang terintegrasi. Ini memastikan bahwa semua informasi yang diperlukan untuk menganalisis dampak diskon terhadap sales dan profit pelanggan telah siap untuk tahap analisis lebih lanjut menggunakan tools python pada google colab.

# Data Understanding

Dataset ini berisi data transaksi dari Perusahaan AWS SaaS yang menjual perangkat lunak ke Perusahaan lain (B2B) dari tahun 2020 sampai 2023. Dalam dataset tersebut, setiap baris mewakili satu transaksi/pesanan (9.994 transaksi), dan kolom-kolomnya meliputi:

Unique Code	Date	Region	Customer	Product	Customer Criteria	Sales
<ul style="list-style-type: none"><li>•Row ID</li><li>•Order ID</li></ul>	<ul style="list-style-type: none"><li>•Order Date</li><li>•Date Key</li></ul>	<ul style="list-style-type: none"><li>•Country</li><li>•City</li><li>•Region</li><li>•Subregion</li></ul>	<ul style="list-style-type: none"><li>•Contact Name</li><li>•Customer ID</li><li>•Customer Name</li></ul>	<ul style="list-style-type: none"><li>•Product Name</li><li>•License</li></ul>	<ul style="list-style-type: none"><li>•Industry</li><li>•Segment</li></ul>	<ul style="list-style-type: none"><li>•Sales</li><li>•Quantity</li><li>•Discount</li><li>•Profit</li></ul>



# Data Understanding

Dilihat dari statistik deskriptif menunjukkan:

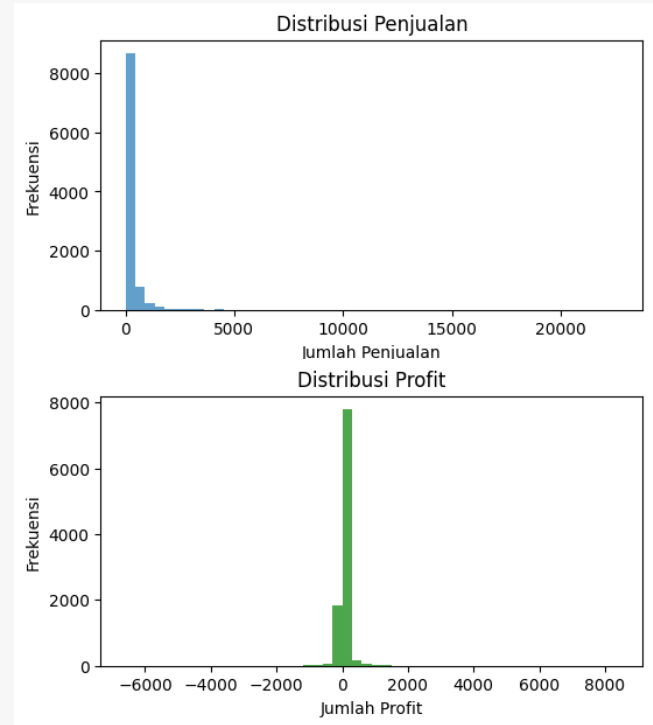
- Sales paling kecil USD 0.444 dan paling besar USD 22638.4, dengan rata-rata USD 229.8
- Kuantitas paling kecil 1 dan paling besar 14, dengan rata-rata 3 pembelian produk
- Pelanggan antara kuartil 50% sampai 75% menggunakan diskon sebesar 0.2, dan diskon paling besar 0.8
- Profit paling kecil USD -6599.9 kemungkinan ini adalah pelanggan yang menggunakan diskon besar. Profit paling besar USD 8399.9 dan rata-rata USD 28.6

	Sales	Quantity	Discount	Profit
count	9994.000000	9994.000000	9994.000000	9994.000000
mean	229.858001	3.789574	0.156203	28.656896
std	623.245101	2.225110	0.206452	234.260108
min	0.444000	1.000000	0.000000	-6599.978000
25%	17.280000	2.000000	0.000000	1.728750
50%	54.490000	3.000000	0.200000	8.666500
75%	209.940000	5.000000	0.200000	29.364000
max	22638.480000	14.000000	0.800000	8399.976000

# Data Understanding

## Distribusi Sales dan Profit

- Berdasarkan diagram sales/penjualan menunjukkan sebagian besar transaksi berada di kisaran nilai sales yang lebih rendah, dengan beberapa transaksi bernilai tinggi.
- Berdasarkan diagram profit juga menunjukkan hal yang serupa seperti penjualan, namun terdapat nilai negatif ini bisa menunjukkan kerugian perusahaan tapi perlu dianalisis lebih lanjut.

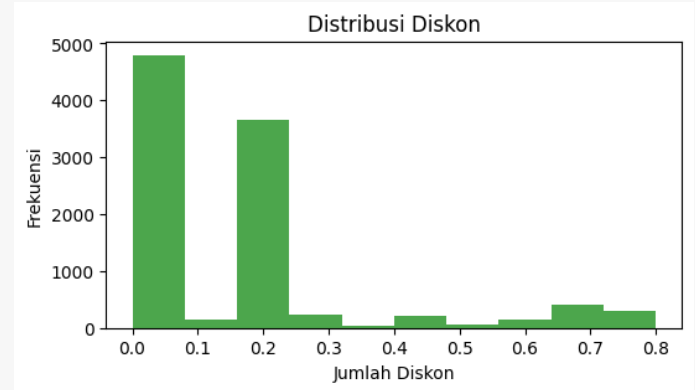


# Data Understanding

## Distribusi Diskon

Berdasarkan diagram pemberian diskon menunjukkan banyak pelanggan tidak menggunakan diskon, sebagian menggunakan diskon sebesar 20%, sisanya terdapat pelanggan yang menggunakan diskon besar.

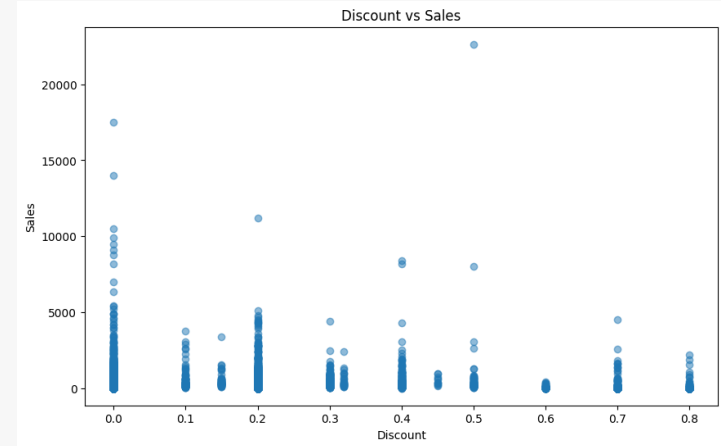
Kita akan mencari tahu pola pemberian diskon yang kecil dan besar ini lalu bagaimana besaran diskon mempengaruhi sales dan profit.



# Data Understanding

## Distribusi Diskon vs Sales

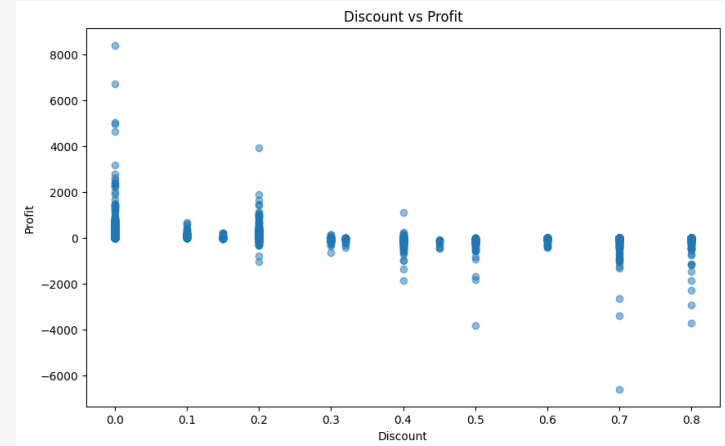
- **Sebaran Data:** Data sales tersebar cukup luas pada berbagai besaran diskon. Sebagian besar sales terjadi pada diskon kecil  $\leq 0.2$ .
- **Konsentrasi Penjualan:** Ada konsentrasi sales yang tinggi pada diskon 0.0. Ini menunjukkan bahwa banyak sales terjadi tanpa memberikan diskon sama sekali.
- **Penjualan Tertinggi:** Sales tertinggi terjadi pada transaksi tanpa diskon atau diskon 0.0. Hal ini menunjukkan bahwa ada transaksi besar yang tidak memerlukan diskon besar.
- **Penjualan dengan Diskon Tinggi:** Sales pada diskon yang besar diatas 0.5 tampaknya lebih sedikit dan tidak ada sales dengan jumlah sangat besar pada besaran diskon yang lebih tinggi ini.
- **Efektivitas Diskon:** Diskon tidak selalu berhubungan langsung dengan peningkatan sales. Ada banyak sales dengan diskon kecil, dan sedikit sales besar dengan diskon besar.



# Data Understanding

## Distribusi Diskon vs Profit

- **Sebaran Data:** Data profit tersebar cukup luas pada berbagai besaran diskon. Sebagian besar profit berada di diskon  $\leq 0.2$ .
- **Profit Tertinggi:** Profit tertinggi terjadi pada tanpa diskon atau diskon 0.0. Ini menunjukkan bahwa sales tanpa diskon atau dengan diskon kecil lebih menguntungkan.
- **Kerugian pada diskon Tinggi:** Pada tingkat diskon yang besar diatas 0.5, terlihat adanya banyak titik yang berada disumbu vertikal negatif menunjukkan kerugian.
- **Profit stabil pada diskon Rendah:** Pada diskon kecil  $< 0.4$ , profit cenderung lebih stabil dan terdapat beberapa titik profit yang besar.
- **Variabilitas profit:** Terdapat variabilitas yang cukup besar pada profit untuk berbagai besaran diskon. Beberapa transaksi dengan diskon  $< 0.4$  tetap menghasilkan kerugian, dan beberapa transaksi dengan diskon tinggi masih menghasilkan keuntungan kecil.



# Data Understanding

## Penjualan Produk Paling Banyak

Total Terjual	
Product	
ContactMatcher	1842
Support	1560
FinanceHub	981
SaaS Connector Pack	957
Site Analytics	889
Marketing Suite - Gold	846
ChatBot Plugin	796
Data Smasher	775
OneView	466
SaaS Connector Pack - Gold	254
Marketing Suite	228
Storage	217
Big OI Database	115
Alchemy	68

Total Industri	
Industry	
Finance	2127
Energy	1316
Tech	1236
Manufacturing	1222
Healthcare	1049
Consumer Products	1021
Retail	972
Communications	593
Transportation	351
Misc	107

Total Segmen	
Segment	
SMB	5191
Strategic	3020
Enterprise	1783

# Data Validation & Data Object Determination

- **Validasi Data:** Memastikan kelengkapan data dengan memeriksa tidak ada nilai kosong.
- **Penyesuaian Tipe Data:** Mengubah tipe data kolom agar sesuai dengan konteks dan menghindari operasi yang tidak relevan ('Customer ID' diubah menjadi tipe object).
- **Penentuan Objek Data:** Menghapus kolom yang tidak relevan ('Row ID', 'Date Key') untuk fokus pada data yang mendukung tujuan analisis. Jadi total kolom yang digunakan hanya 17 kolom.
- **Penanganan Outlier:** Nilai ekstrem atau outlier tidak dihapus karena semua data dianggap penting dalam konteks bisnis ini karena berkaitan dengan transaksi penjualan.

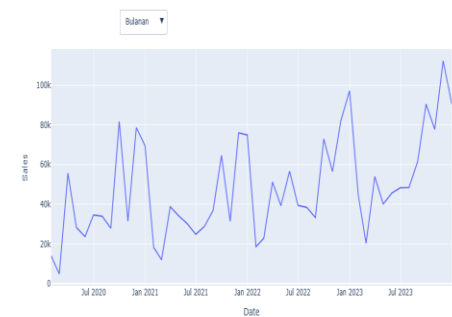
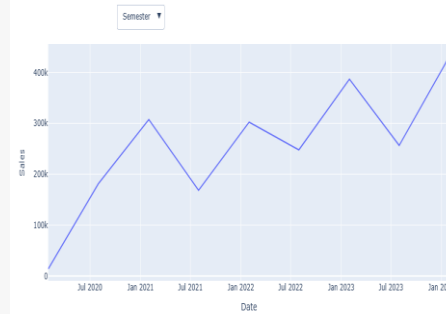
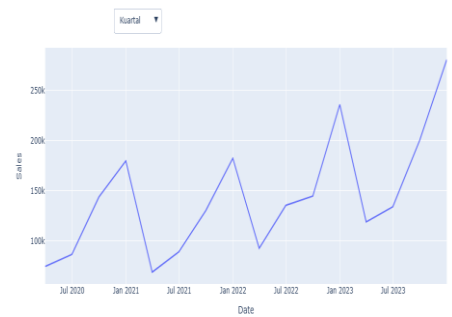
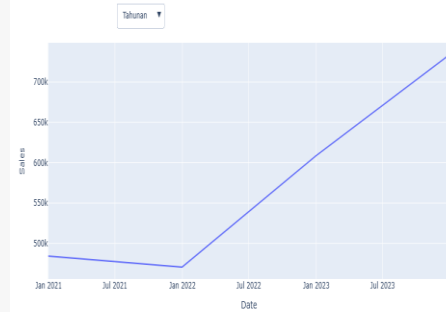
		0			
	Row ID	0.0			
	Order ID	0.0			
	Order Date	0.0			
	Date Key	0.0			
	Contact Name	0.0			
	Country	0.0			
	City	0.0			
	Region	0.0			
	Subregion	0.0			
	Customer	0.0			
	Customer ID	0.0			
	Industry	0.0			
	Segment	0.0			
	Product	0.0			
	License	0.0			
	Sales	0.0			
	Quantity	0.0			
	Discount	0.0			
	Profit	0.0			

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 17 columns):
#   Column              Non-Null Count  Dtype
---  -
0   Order ID            9994 non-null   object
1   Order Date          9994 non-null   datetime64[ns]
2   Contact Name        9994 non-null   object
3   Country             9994 non-null   object
4   City                9994 non-null   object
5   Region              9994 non-null   object
6   Subregion           9994 non-null   object
7   Customer            9994 non-null   object
8   Customer ID         9994 non-null   object
9   Industry            9994 non-null   object
10  Segment             9994 non-null   object
11  Product             9994 non-null   object
12  License             9994 non-null   object
13  Sales               9994 non-null   float64
14  Quantity            9994 non-null   int64
15  Discount            9994 non-null   float64
16  Profit              9994 non-null   float64
```

# Data Analysis

## Kinerja penjualan perusahaan secara umum

- Sales mulai meningkat tahun 2022 sampai 2023.
- Secara kuartal dan semester dari tahun 2020 sampai 2023 tren sales cenderung mengalami kenaikan.
- Secara bulanan dari tahun 2020 sampai 2023 tren sales stagnan.
- Secara keseluruhan kinerja sales baik.

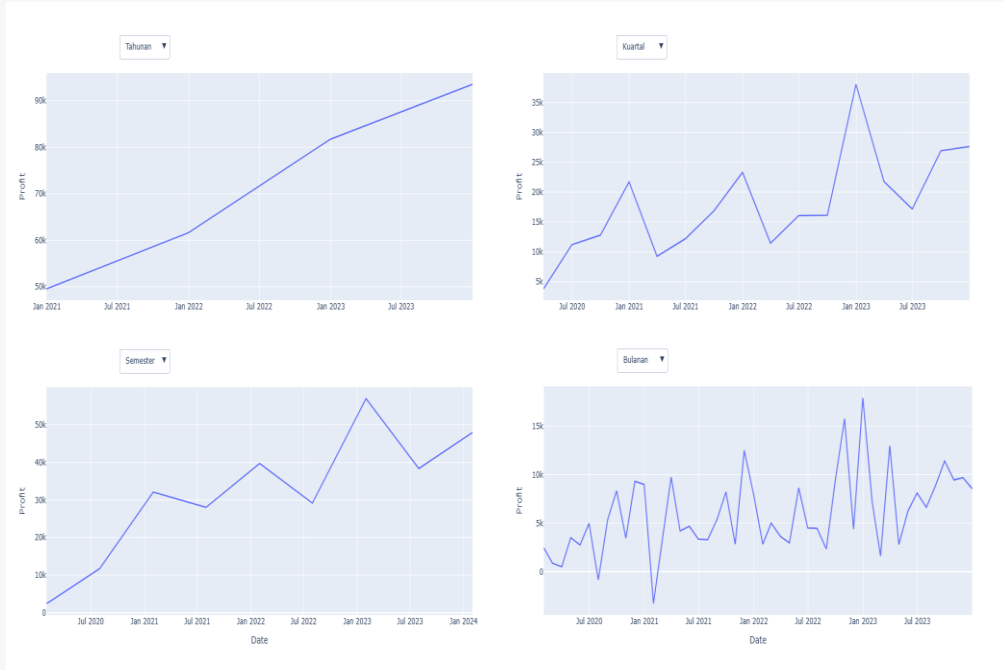




# Data Analysis

## Kinerja penjualan perusahaan secara umum

- Secara tahunan dari tahun 2020 sampai 2023 tren profit konsisten naik.
- Secara kuartal dari tahun 2020 sampai pertengahan 2022 tren profit cenderung stagnan.
- Secara semester dari tahun 2020 sampai 2023 tren profit cenderung mengalami kenaikan.
- Secara bulanan dari tahun 2020 sampai 2023 tren profit juga cenderung stagnan.
- Stagnan kemungkinan disebabkan profit kecil atau negatif.
- Strategi diskon perusahaan berhasil mempertahankan pelanggan dan mendorong penggunaan layanan *cloud*, namun untuk mengoptimalkan profit, penting menemukan ambang batas diskon ideal yang meningkatkan penjualan tanpa menyebabkan profit negatif.



# Data Analysis

## Kinerja penjualan perusahaan secara umum

	Order ID	Order Date	Customer	Industry	Segment	Product	Sales	Quantity	Discount	Profit	Revenue	Total Cost	Cost per Unit
7772	AMER-2022-108196	2022-11-26	Allstate	Finance	SMB	Big OI Database	4499.985	5	0.7	-6599.9780	1349.9955	7949.9735	1589.9947
683	EMEA-2023-168116	2023-11-05	Costco Wholesale	Retail	Strategic	Big OI Database	7999.980	4	0.5	-3839.9904	3999.9900	7839.9804	1959.9951
9774	APJ-2020-169019	2020-07-26	Bosch	Tech	SMB	ContactMatcher	2177.584	8	0.8	-3701.8928	435.5168	4137.4096	517.1762
3011	EMEA-2023-134845	2023-04-18	FedEx	Transportation	Enterprise	Big OI Database	2549.985	5	0.7	-3399.9800	764.9955	4164.9755	832.9951
4991	APJ-2023-122714	2023-12-08	Ford Motor	Manufacturing	Strategic	ContactMatcher	1889.990	5	0.8	-2929.4845	377.9980	3307.4825	661.4965
3151	AMER-2021-147830	2021-12-15	Walgreens	Retail	SMB	Big OI Database	1799.994	2	0.7	-2639.9912	539.9982	3179.9894	1589.9947
5310	APJ-2023-131254	2023-11-20	Itochu	Finance	SMB	ContactMatcher	1525.188	6	0.8	-2287.7820	305.0376	2592.8196	432.1366
9639	EMEA-2021-116638	2021-01-28	Morgan Stanley	Finance	SMB	ContactMatcher	4297.644	13	0.4	-1862.3124	2578.5864	4440.8988	341.6076
1199	APJ-2022-130946	2022-04-09	Volkswagen	Manufacturing	SMB	ContactMatcher	1088.792	4	0.8	-1850.9464	217.7584	2068.7048	517.1762
2697	EMEA-2020-145317	2020-03-18	Anthem	Healthcare	Enterprise	Big OI Database	22638.480	6	0.5	-1811.0784	11319.2400	13130.3184	2188.3864

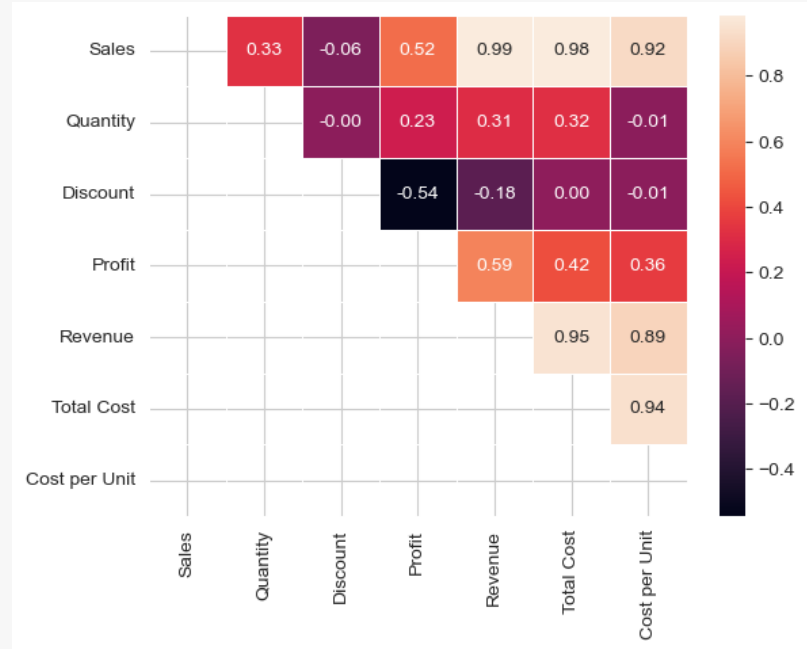
- Tabel diatas hanya menampilkan 10 teratas data.
- Dapat dilihat dari tabel diatas, Ternyata diskon yang besar yang menyebabkan pendapatan negatif.
- Rasio kerugian yang dihitung dari **total profit negatif** (1871) dibagi **jumlah data profit** (9994) sebesar **18.72%**.

# Data Analysis

## Kinerja penjualan perusahaan secara umum

Sekarang coba kita lihat dari sisi korelasi dari Sales sampai Cost per Unit mana saja yang paling saling mempengaruhi. Korelasi disamping menunjukkan memang terdapat asosiasi:

- Korelasi yang sedang ada antara diskon dengan profit dengan arah korelasi negative sebesar  $-0.54$  artinya ketika diskon meningkat maka profit cenderung menurun.
- Korelasi yang lemah antara diskon dengan revenue sebesar  $-0.18$  mengindikasikan juga bahwa ketika diskon meningkat maka revenue cenderung menurun.
- Selanjutnya ada juga korelasi dengan arah positif yang kuat antara sales dengan revenue, total cost, dan cost per unit yang masing-masing  $0.99$ ,  $0.98$ ,  $0.92$ . Artinya kenaikan dari salah satunya akan dipengaruhi kenaikan lainnya.
- Kemudian antara revenue dengan total cost dan cost per unit juga memiliki korelasi positif kuat yang masing-masing  $0.95$  dan  $0.89$ .



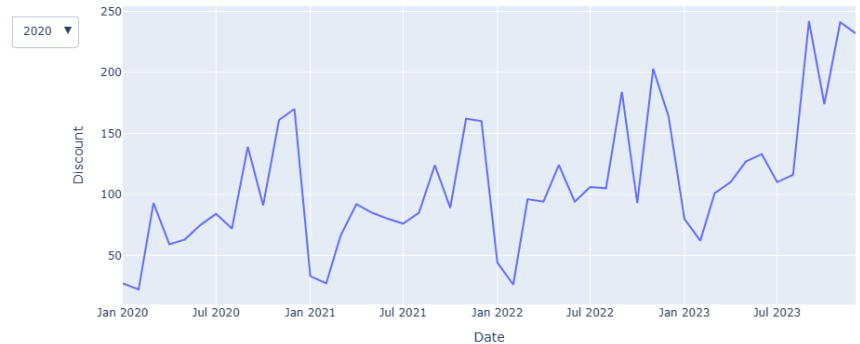
# Data Analysis

## Tren pemberian diskon dari waktu ke waktu

Secara tahunan pemberian diskon paling banyak pada bulan September dan November, di negara Amerika memang ada momen pada kedua bulan tersebut.

- Dibulan September bertepatan dengan musim gugur, perusahaan sering memberikan diskon untuk menarik perhatian konsumen baru dan memancing pelanggan lama berbelanja.
- Sementara dibulan November adalah momen Black Friday yaitu hari Jumat setelah peringatan Thanksgiving di Amerika, dimana toko-toko memberikan diskon besar selama 24 jam dalam upaya untuk menarik konsumen berbelanja hadiah Natal dan tahun baru.

Tren Diskon Perbulan

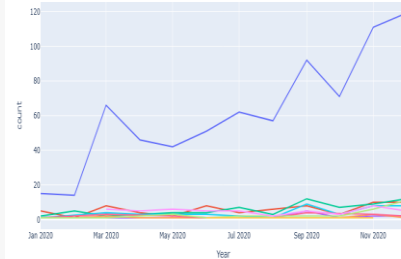


# Data Analysis

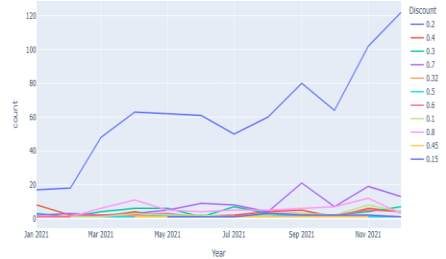
## Tren pemberian diskon dari waktu ke waktu

- Jika dilihat tren pemberian diskon pada setiap besaran diskon sesuai dengan informasi yang kita dapatkan pada visualisasi scatterplot sebelumnya, pemberian diskon paling banyak 20% karena memang angka ini adalah angka yang umum perusahaan gunakan dalam memberikan diskon.
- Walaupun diskon sebesar 20% menunjukkan keberhasilannya baik secara domain knowledge maupun data yang sebelumnya kita lihat grafiknya, tapi kita belum bisa ambil kesimpulan 20% adalah besaran diskon yang ideal, karena perlu mengidentifikasi segmen pelanggan yang menghasilkan sales high, medium, dan low.

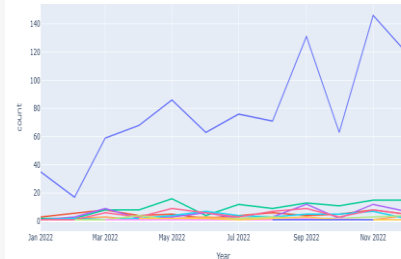
Tren Diskon Perbulan



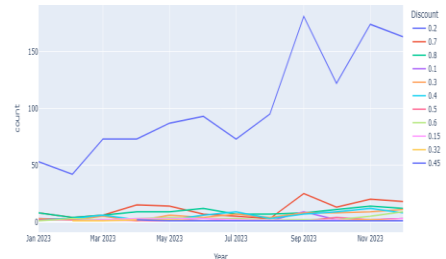
Tren Diskon Perbulan



Tren Diskon Perbulan

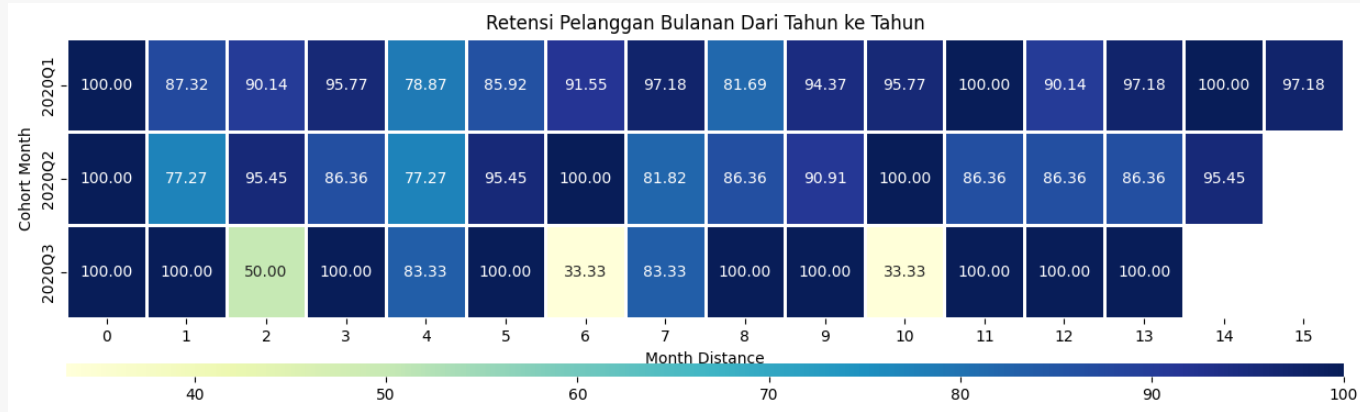


Tren Diskon Perbulan



# Data Analysis

## Analisis retensi pelanggan



- Heatmap tersebut menampilkan persentase retensi pelanggan bulanan dari waktu ke waktu, dengan pelanggan dibagi menjadi tiga kelompok berdasarkan kuartal (2020Q1, 2020Q2, dan 2020Q3). Sumbu horizontal menunjukkan bulan dari bulan 0 (Januari) hingga bulan ke-15 (April tahun berikutnya), sementara sumbu vertikal mewakili kelompok pelanggan.
- Secara keseluruhan, retensi pelanggan menunjukkan tren yang stabil, di mana pelanggan yang bertransaksi sejak tahun 2020 masih aktif bertransaksi hingga akhir tahun 2023, meskipun ada fluktuasi persentase pelanggan yang bertransaksi setiap bulannya dalam setiap kelompok.

# Data Analysis

## Kategori Segmen High Sales

	Customer	first_purchase	last_purchase	total_transactions	total_quantity	total_sales	total_profit	discount_usage	duration	Sales Segment
10	Anthem	2020-01-14	2023-12-24	134	523	55719.2100	5953.2038	0.432836	5 years, 6 months	High
38	Ford Motor	2020-03-03	2023-12-25	176	672	43893.0424	4010.7662	0.482955	5 years, 4 months	High
4	Allianz	2020-03-17	2023-12-30	192	709	42904.1670	5531.8834	0.526042	5 years, 3 months	High
16	Bank of America Corp.	2020-06-15	2023-12-22	132	530	41255.9471	6449.8554	0.439394	5 years, 0 months	High
51	Itochu	2020-02-09	2023-12-29	142	518	29886.8860	1690.4793	0.535211	5 years, 5 months	High
3	Airbus	2020-01-11	2023-12-04	123	458	29798.5930	5593.3657	0.349593	5 years, 6 months	High
54	Kroger	2020-01-28	2023-12-30	135	534	29783.4590	6256.1005	0.496296	5 years, 5 months	High
53	Johnson & Johnson	2020-01-22	2023-12-30	132	457	29489.5720	4632.4593	0.515152	5 years, 5 months	High

- Tabel diatas hanya menampilkan Segmen Pelanggan High Sales dengan 4 total sales tertinggi dan 4 total sales terendah.
- Total Penjualan yang termasuk Segmen High Sales paling USD 55719.2100 dan paling rendah 29489.5720.

# Data Analysis

## Kategori Segmen Medium Sales

	Customer	first_purchase	last_purchase	total_transactions	total_quantity	total_sales	total_profit	discount_usage	duration	Sales Segment
97	Walt Disney	2020-04-02	2023-12-09	103	374	29483.0050	5786.8297	0.563107	5 years, 3 months	Medium
0	AT&T	2020-03-11	2023-12-22	151	536	29462.3240	4520.6641	0.456954	5 years, 4 months	Medium
28	Coca-Cola	2020-02-28	2023-12-22	81	305	28643.8040	9449.0246	0.530864	5 years, 4 months	Medium
63	Morgan Stanley	2020-01-15	2023-12-31	126	484	28631.4500	-361.8099	0.523810	5 years, 6 months	Medium
61	Mitsubishi	2020-01-06	2023-12-14	100	393	16677.8430	1668.9366	0.490000	5 years, 6 months	Medium
6	Alphabet	2020-05-25	2023-12-25	44	181	16116.4610	3980.3457	0.568182	5 years, 1 months	Medium
68	Panasonic	2020-02-03	2023-12-26	83	307	15882.9950	993.8604	0.481928	5 years, 5 months	Medium
85	Target	2020-07-25	2023-12-04	49	180	15736.1290	1207.3113	0.530612	4 years, 11 months	Medium

- Tabel diatas hanya menampilkan Segmen Pelanggan Medium Sales dengan 4 total sales tertinggi dan 4 total sales terendah.
- Total Penjualan yang termasuk Segmen Medium Sales paling USD 29483.0050 dan paling rendah 15736.1290.



# Data Analysis

## Kategori Segmen Low Sales

	Customer	first_purchase	last_purchase	total_transactions	total_quantity	total_sales	total_profit	discount_usage	duration	Sales Segment
42	Glencore	2020-03-28	2023-11-11	56	220	15649.1490	4470.1537	0.410714	5 years, 3 months	Low
52	J.P. Morgan Chase & Co.	2020-05-18	2023-12-25	101	374	15451.0502	1752.5227	0.584158	5 years, 1 months	Low
12	Aviva	2020-01-16	2023-12-31	100	345	14863.5270	1474.3551	0.380000	5 years, 5 months	Low
58	McKesson	2020-05-05	2023-12-25	63	262	14431.0060	1750.2888	0.634921	5 years, 2 months	Low
17	Berkshire Hathaway	2020-03-31	2023-12-27	39	119	7100.9290	812.7311	0.512821	5 years, 3 months	Low
78	Safeway	2020-05-30	2023-10-16	39	158	6633.1370	472.4582	0.461538	5 years, 1 months	Low
27	Citigroup	2020-04-25	2023-05-06	22	97	6132.5210	474.3321	0.545455	5 years, 2 months	Low
39	Gazprom	2020-08-08	2023-11-18	53	193	5402.9020	-13.9695	0.679245	4 years, 11 months	Low

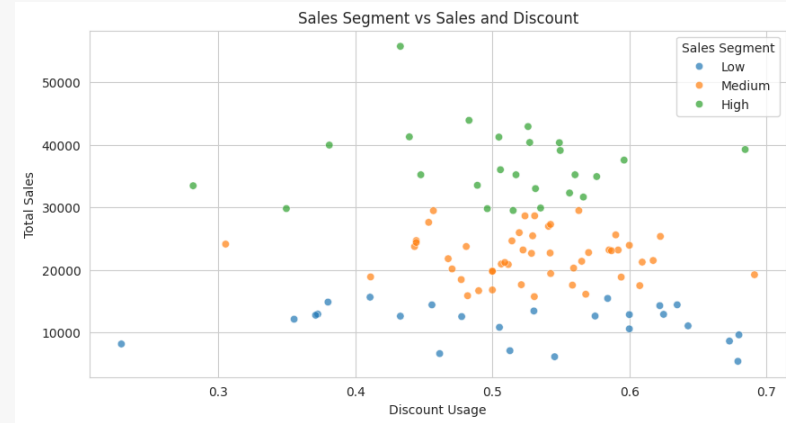
- Tabel diatas hanya menampilkan Segmen Pelanggan Low Sales dengan 4 total sales tertinggi dan 4 total sales terendah.
- Total Penjualan yang termasuk Segmen Low Sales paling USD 15649.1490 dan paling rendah 5402.9020.

# Data Analysis

## Analisis Discount, Sales, dan Profit

Scatterplot disamping kita dapat tarik kesimpulan:

- Pelanggan dalam segmen penjualan high memiliki penggunaan diskon yang bervariasi dari sekitar 0.3 hingga 0.6, dengan tendensi di sekitar 0.5.
- Pelanggan dalam segmen penjualan medium dan low juga memiliki penggunaan diskon yang bervariasi, tetapi cenderung lebih tinggi sekitar 0.4 hingga 0.6.

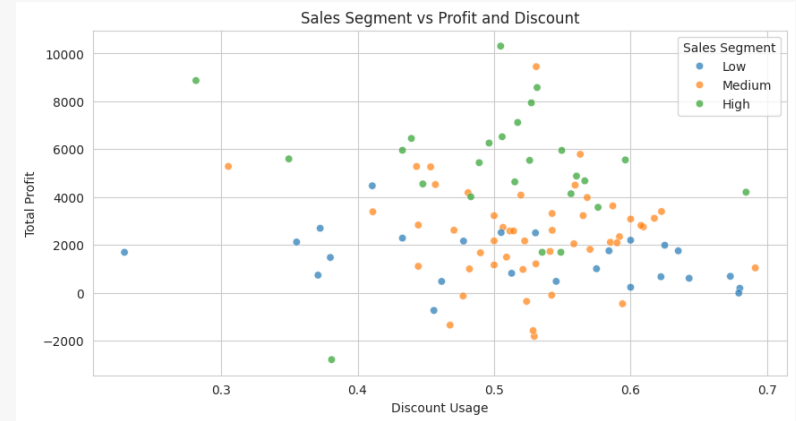


# Data Analysis

## Analisis Discount, Sales, dan Profit

Scatterplot disamping kita dapat tarik kesimpulan:

- Pelanggan dengan profit tertinggi cenderung memiliki penggunaan diskon di sekitar 0.4 hingga 0.5. Diluar interval ini profit cenderung menurun.
- Beberapa pelanggan menyebabkan kerugian profit pada interval diskon yang hampir sama dari 0.4 sampai 0.6.

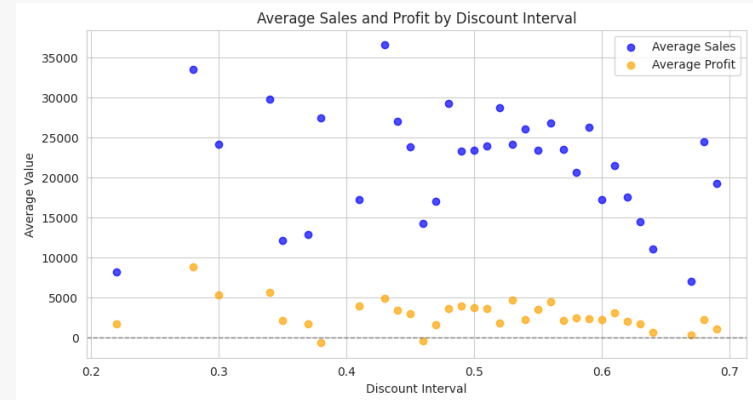


# Data Analysis

## Analisis Discount, Sales, dan Profit

- Dan pada tabel tersebut juga terlihat rata-rata besaran diskon yang digunakan pelanggan disetiap segmen sebesar 5%
- Dapat dilihat pada lineplot disamping diskon yang ideal berada di interval diskon 0.5 hingga 0.6 dimana sales masih cukup tinggi dan profit berada pada nilai positif.

Sales Segment			discount_usage	Sales Segment		total_sales	total_profit
0	Low	0.518270	0	Low	12617.1140	1474.3551	
1	Medium	0.527028	1	Medium	22650.8206	2581.9241	
2	High	0.504075	2	High	35201.6250	5531.8834	



# Insight

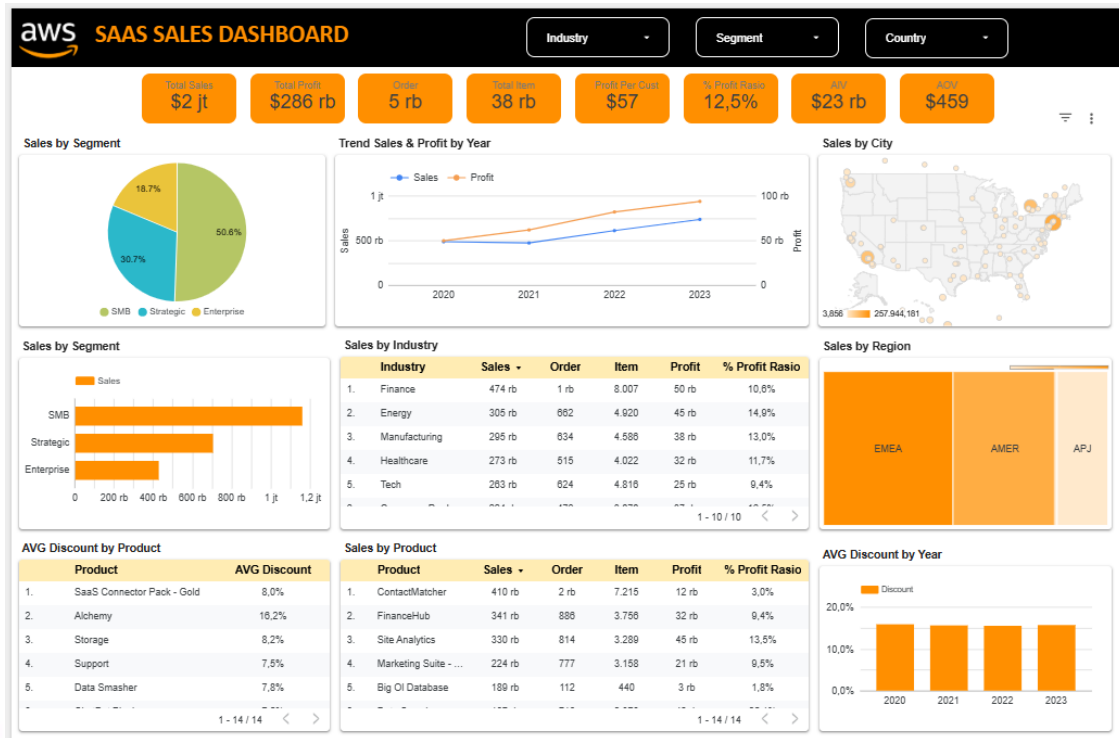
- **Diskon Ideal untuk Sales & Profit**
  - Interval Diskon Optimal: Diskon antara 50% hingga 60% menunjukkan keseimbangan terbaik.
    - Penjualan tetap tinggi dalam rentang ini.
    - Keuntungan tetap positif (tidak melewati batas nol).
  - Rekomendasi Diskon Teraman: Berdasarkan analisis scatterplot dan rata-rata diskon, **diskon 50% adalah pilihan paling aman** untuk menghindari keuntungan negatif.
- **Penargetan Pelanggan Spesifik**
  - Fokus Utama: Mengembangkan strategi diskon untuk segmen pelanggan "Medium Sales" dan "Low Sales".
  - Tujuan:
    - Optimasi Sales: Mendorong pembelian berulang dari kedua segmen ini.
    - Optimasi Profit: Meminimalisir profit negatif yang mungkin terjadi.
  - Data Pelanggan Tersedia: Perusahaan memiliki data detail mengenai pelanggan dalam kategori Medium Sales Segment dan Low Sales Segment untuk penargetan yang tepat.

# Recommendation

- Maksimum pemberian diskon sebesar 50%. Diskon diatas nilai ini cenderung menurunkan keuntungan secara signifikan.
- Berikan diskon secara berjenjang di mana pelanggan mendapatkan diskon yang lebih besar sampai batas maksimal 50% jika mereka memakai service cloud perusahaan sampai tingkat tertentu seperti model Pay less by using more yang sudah diterapkan, atau ketika mereka menambah service tawarkan bundling produk seperti model Bundled Discounts yang sudah diterapkan juga, sehingga mereka mendapatkan diskon jika menyewa beberapa service sekaligus. Kedua cara ini bisa memancing penambahan pembelian pelanggan.
- Strategi dipersonalisasi berdasarkan segmen pelanggan.

# Dashboard

<https://lookerstudio.google.com/reporting/a10a4fdd-42c5-4b0d-834b-57fcfa41ecaf>



# Lampiran 1: Data

[https://github.com/rahmawtnisa/ProjectDQLab\\_BNSP/blob/main/SaaS-Sales.csv](https://github.com/rahmawtnisa/ProjectDQLab_BNSP/blob/main/SaaS-Sales.csv)

Row ID	Order ID	Order Date	Date Key	Contact	Country	City	Region	Subregion	Customer	Customer Industry	Segment	Product	License	Sales	Quantity	Discount	Profit		
1	EMEA-202-#####	20221109	20221109	Nathan Belreland	Dublin	EMEA	UKIR	Chevron	1017	Energy	SMB	Marketing	16GRM07I		261.96	2	0	41.9136	
2	EMEA-202-#####	20221109	20221109	Nathan Belreland	Dublin	EMEA	UKIR	Chevron	1017	Energy	SMB	FinanceHu	QLUW57KZ		731.94	3	0	219.582	
3	AMER-2026/13/2022	20220613	20220613	Deirdre Bi	United Sta	New York	AMER	NAMER	Phillips 66	1056	Energy	Strategic	FinanceHu	JI68VL70H		14.62	2	0	6.8714
4	EMEA-202-#####	20221101	20221101	Zoe Hodg	Germany	Stuttgart	EMEA	EU-WEST	Royal Dut	1031	Energy	SMB	ContactM	DE9GJKGC		957.5775	5	0.45	-383.031
5	EMEA-202-#####	20221101	20221101	Zoe Hodg	Germany	Stuttgart	EMEA	EU-WEST	Royal Dut	1031	Energy	SMB	Marketing	OIF7NY23I		22.368	2	0.2	2.5164
6	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	SaaS Conn	8RM88WF		48.86	7	0	14.1694
7	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	ChatBot	PI4SYTVM0I		7.28	4	0	1.9656
8	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	Site Analy	I07842ES7I		907.152	6	0.2	90.7152
9	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	ContactM	XPBOBXW		18.504	3	0.2	5.7825
10	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	OneView	71J8P9Z1P		114.9	5	0	34.47
11	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	ContactM	43NZ37VE		1706.184	9	0.2	85.3092
12	AMER-202-#####	20200609	20200609	Thomas H	United Sta	New York	AMER	NAMER	BNP Parib	1065	Finance	SMB	Site Analy	JG14MOU		911.424	4	0.2	68.3568
13	EMEA-2024/16/2023	20230416	20230416	Adrian Hu	Sweden	Stockholm	EMEA	NOR	Engie	1062	Energy	SMB	Support	FESQBIUI		15.552	3	0.2	5.4432
14	AMER-202-#####	202211206	202211206	Adrian Ab	Canada	Toronto	AMER	NAMER	Microsoft	1066	Tech	SMB	ContactM	COK66F8I		407.976	3	0.2	132.5922
15	API-2021-11/22/202	20211122	20211122	Emily Mor	Japan	Saitama	API	JAPN	Pfizer	1074	Healthcare	Enterprise	OneView	D8FGC56I		68.81	5	0.8	-123.858
16	API-2021-11/22/202	20211122	20211122	Emily Mor	Japan	Saitama	API	JAPN	Pfizer	1074	Healthcare	Enterprise	ContactM	NRKTSR1A		2.544	3	0.8	-3.816
17	EMEA-202-#####	20201111	20201111	Claire Well	Italy	Naples	EMEA	EU	Toyota M	1087	Manufact	SMB	Marketing	K5XHYWZ		665.88	6	0	13.3176
18	EMEA-2025/13/202C	20200513	20200513	Carol Cam	Turkey	Ankara	EMEA	MEA	Home Def	1047	Retail	SMB	Marketing	LWVKLPY		55.5	2	0	9.99
19	AMER-2028/27/202C	20200827	20200827	Blake Hug	United Sta	Los Angeles	AMER	NAMER	Wells Farg	1099	Finance	SMB	ChatBot	PIZY9KDTYI		8.56	2	0	2.4824
20	AMER-2028/27/202C	20200827	20200827	Blake Hug	United Sta	Los Angeles	AMER	NAMER	Wells Farg	1099	Finance	SMB	Site Analy	I54O29EW		213.48	3	0.2	16.011
21	AMER-2028/27/202C	20200827	20200827	Blake Hug	United Sta	Los Angeles	AMER	NAMER	Wells Farg	1099	Finance	SMB	ContactM	E3TF1V9D		22.72	4	0.2	7.384
22	AMER-202-#####	20221210	20221210	Melanie SI	Costa Rica	San José	AMER	LATAM	Carrefour	1048	Retail	Strategic	ChatBot	PI88G92Q		19.46	7	0	5.0596
23	AMER-202-#####	20221210	20221210	Melanie SI	Costa Rica	San José	AMER	LATAM	Carrefour	1048	Retail	Strategic	OneView	VVDIYAB		60.34	7	0	15.6884



# Lampiran 2: Code Python

[https://github.com/rahmawtnisa/ProjectDQLab\\_BNSP/blob/main/Code\\_AWS\\_SaaS.ipynb](https://github.com/rahmawtnisa/ProjectDQLab_BNSP/blob/main/Code_AWS_SaaS.ipynb)

```
In [9]: # Membuat tabel untuk menampilkan seluruh nilai unik pada kolom Industry, Segment, dan Product
pd.set_option('display.max_colwidth', 1)

nilai_unik = {
    'Daftar Industri': [df['Industry'].unique()],
    'Daftar Segmen Bisnis': [df['Segment'].unique()],
    'Daftar Produk': [df['Product'].unique()],
    'Berapa Saja Jumlah Pembelian': [df['Quantity'].unique()],
    'Berapa Saja Diskon': [df['Discount'].unique()]
} # Membuat dictionary dengan value berupa array dari nilai unik kolom

tabel_nilai_unik = pd.DataFrame(nilai_unik) # Menjadikan dictionary menjadi dataframe dengan kolomnya mengg
tabel_nilai_unik
```

```
Out[9]:
```

	Daftar Industri	Daftar Segmen Bisnis	Daftar Produk	Berapa Saja Jumlah Pembelian	Berapa Saja Diskon
0	[Energy, Finance, Tech, Healthcare, Manufacturing, Retail, Consumer Products, Misc, Transportation, Communications]	[SMB, Strategic, Enterprise]	[Marketing Suite, FinanceHub, ContactMatcher, Marketing Suite - Gold, SaaS Connector Pack, ChatBot Plugin, Site Analytics, OneView, Support, Data Smasher, SaaS Connector Pack - Gold, Storage, Big Ol Database, Alchemy]	[2, 3, 5, 7, 4, 6, 9, 1, 8, 14, 11, 13, 10, 12]	[0.0, 0.45, 0.2, 0.8, 0.3, 0.5, 0.7, 0.6, 0.32, 0.1, 0.4, 0.15]

Do you have any questions?

rahmawati.salsadila@gmail.com  
+62 895 1409 8571



[Rahmawati Annisa Salsadila](#)

*Thanks!*