

Capstone Project 2

Ni Putu Dittarani Pramesti

JCDS-OL016

Mendorong Inovasi Produk: Latar Belakang Perusahaan

Latar Belakang Perusahaan



Dalam dunia bisnis yang semakin kompetitif, terutama dalam industri teknologi informasi, penting bagi perusahaan untuk terus berinovasi dan mengembangkan produk yang sesuai dengan kebutuhan pasar. Microsoft Azure, sebagai salah satu layanan cloud computing terkemuka, berperan krusial dalam membantu perusahaan menghadapi tantangan ini. Tim pengembangan produk di Azure bertanggung jawab untuk menciptakan dan memperbarui layanan dan solusi yang relevan agar tetap menarik bagi pelanggan. Untuk mencapai tujuan ini, pemahaman yang mendalam tentang preferensi pelanggan dan pola penggunaan layanan sangatlah penting.

Namun, dalam proses pengembangan produk, seringkali ada berbagai kendala yang harus dihadapi. **Salah satu kendala utama adalah bagaimana memastikan bahwa solusi yang ditawarkan oleh Azure sesuai dengan tuntutan dan harapan pelanggan di berbagai industri. Setiap industri memiliki karakteristik dan kebutuhan yang unik, yang mempengaruhi cara pelanggan berinteraksi dengan layanan cloud. Oleh karena itu, tim pengembangan Azure harus mampu menyesuaikan produk mereka agar bisa memenuhi ekspektasi dan tuntutan tersebut.**

Dengan memanfaatkan wawasan yang diperoleh dari analisis data industri dan segmen pelanggan, tim Azure dapat menciptakan solusi yang lebih tepat sasaran dan efisien. Misalnya, mereka dapat menyelaraskan pengembangan layanan dengan permintaan pasar dan tren teknologi terkini. Dengan cara ini, tim pengembangan tidak hanya dapat menciptakan fitur yang lebih bernilai bagi pelanggan, tetapi juga meningkatkan kepuasan pengguna serta mempertahankan daya saing Azure di pasar.

Mendorong Inovasi Produk: Mengatasi Kendala dan Menetapkan Target

Stakeholder, Kendala, dan Target

Dalam proyek capstone ini, pihak yang menjadi fokus adalah tim pengembangan produk.

Kendala:

1. Bagaimana preferensi pelanggan dan pemahaman pola penjualan bisa membantu mengoptimalkan pengembangan produk?
2. Bagaimana kebutuhan pelanggan dan keputusan pembelian bisa membantu tim pengembangan meningkatkan produk mereka?
3. Bagaimana cara memenuhi tuntutan khusus dari industri secara efektif?

Target:

1. Menyelaraskan pengembangan produk dengan lonjakan permintaan dengan memahami preferensi pelanggan dan pola penjualan.
2. Membantu tim pengembangan menciptakan fitur yang lebih bernilai dengan mengetahui kebutuhan pelanggan dan pola pembeliannya.
3. Memanfaatkan wawasan industri dan data segmen pelanggan untuk menyesuaikan solusi, sehingga tim bisa memenuhi tuntutan spesifik industri dengan lebih efisien.

Preparasi Package : Eksplanasi Lengkap Data Saas

Persiapan Data

Data Understanding

Analisis ini menyajikan ringkasan dari dataset yang terdiri dari 9,994 baris dan 19 kolom.

Data Checking

1. Struktur Data dan Tipe Kolom

- **Tipe Data:** Dataset ini memiliki kolom dengan tipe int64, float64, dan object. Kolom Order Date sebaiknya diubah menjadi datetime agar lebih mudah dianalisis.
- **Kolom Numerik:** Kolom seperti Sales, Quantity, Discount, dan Profit memberikan metrik penting untuk analisis penjualan dan profitabilitas.
- **Kolom Kategori:** Terdapat kolom kategori seperti Country, City, Region, dan Industry, yang membantu memberikan konteks demografis dan geografis untuk analisis pasar.

2. Distribusi dan Keunikan Data

- **Cakupan Wilayah:** Dataset ini mencakup 228 kota unik, 49 negara unik, dan tiga wilayah (EMEA, AMER, APJ), menunjukkan sifat internasional dari data ini.
- **Profit:** Data profit bervariasi, penting untuk mengidentifikasi produk atau wilayah yang paling menguntungkan.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 19 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Row ID          9994 non-null   int64
1   Order ID        9994 non-null   object
2   Order Date      9994 non-null   object
3   Date Key        9994 non-null   int64
4   Contact Name    9994 non-null   object
5   Country         9994 non-null   object
6   City            9994 non-null   object
7   Region          9994 non-null   object
8   Subregion       9994 non-null   object
9   Customer        9994 non-null   object
10  Customer ID     9994 non-null   int64
11  Industry        9994 non-null   object
12  Segment         9994 non-null   object
13  Product         9994 non-null   object
14  License         9994 non-null   object
15  Sales           9994 non-null   float64
16  Quantity        9994 non-null   int64
17  Discount        9994 non-null   float64
18  Profit          9994 non-null   float64
dtypes: float64(3), int64(4), object(12)
memory usage: 1.4+ MB
```

```
'Milwaukee' 'Frankfurt' 'Kilger' 'Frankfurt' 'Porto' 'Antofagasta'
'Genoa' 'Toluca' 'San Francisco' 'Hamilton' 'Tampere' 'Miami' 'Yonkers'
'Manaus' 'Rosario' 'Ottawa' 'Hyogo' 'Tijuana' 'Gothenburg' 'Lyon'
'Edinburgh' 'Leeds' 'Brisbane' 'Belo Horizonte' 'Ostrava' 'Amsterdam'
'Roswell' 'Murcia' 'Baltimore' 'Brno' 'Montreal' 'Atlanta' 'Milan' 'Cork'
'Jackson' 'Nice' 'Athens' 'Shanghai' 'Alicante' 'Cancun' 'Auckland'
'Springfield' 'Sapporo' 'Richmond' 'Seville' 'Huntsville' 'Calgary'
'Vienna' 'Woking' 'San Diego' 'Shenzhen' 'Louisville' 'Dubai' 'Tampa'
'Jerusalem' 'Zaragoza' 'Virginia Beach' 'Guangzhou' 'Manchester'
'Kharkiv' 'Philadelphia' 'Vantaa' 'Oslo' 'Novosibirsk' 'Columbia'
'Belfast' 'Oakland' 'Izmir' 'Cabo San Lucas' 'Daegu' 'Dallas' 'Phoenix'
'Bristol' 'Venice' 'Quebec' 'Shizuoka' 'Kyoto' 'Antwerp' 'Ibaraki'
'Barcelona' 'Delhi' 'Beijing' 'Turin' 'Jakarta' 'Jeddah' 'Valencia'
'Anaheim' 'Gwangju' 'Lawrence' 'Austin' 'Burlington' 'Sacramento'
'Nantes' 'Fort Worth' 'Brasília' 'Cairo' 'Copenhagen' 'Odessa' 'Zagreb'
'Nizhny Novgorod' 'Houston' 'Nashville' 'Hobart' 'Medellín' 'Granada'
'Glasgow' 'Bangalore' 'Gold Coast' 'Wilmington' 'Bologna' 'Indianapolis'
'Nuremberg' 'Hiroshima' 'Lakewood' 'Chennai' 'Florence' 'Ghent' 'Taipei'
'Arlington' 'San Jose' 'Fresno' 'Busan' 'Lakeland' 'Rotterdam' 'Mannheim'
'Calif' 'Brescia' 'Bari' 'Regina' 'Surabaya' 'Newcastle' 'Hyderabad'
'Doha' 'Palermo' 'Donetsk' 'Lancaster' 'Espoo' 'Catania' 'Tianjin'
'Bursa' 'Buffalo' 'Limerick' 'Quezon City' 'Perth' 'Marseille'
'San Antonio' 'Cleveland' 'Charlotte' 'Tucson' 'Ljubljana' 'Tulsa'
'Winnipeg' 'Thessaloniki' 'Denver' 'Victoria' 'Alexandria' 'Las Vegas'
'Minneapolis' 'Omaha' 'Bandung' 'Franklin' 'Reykjavík' 'Tallahassee'
'Suzhou' 'Chengdu' 'Everett' 'Clinton' 'Memphis']
Negara Unik: ['Ireland' 'United States' 'Germany' 'Sweden' 'Canada' 'Japan' 'Italy'
'Turkey' 'Costa Rica' 'France' 'Australia' 'Luxembourg' 'Brazil'
'Philippines' 'India' 'United Kingdom' 'Chile' 'Spain' 'South Africa'
'Portugal' 'Colombia' 'Belgium' 'Russia' 'Netherlands' 'Mexico' 'Poland'
'Singapore' 'New Zealand' 'Czech Republic' 'Saudi Arabia' 'Argentina'
'South Korea' 'Finland' 'Israel' 'Norway' 'Ukraine'
'United Arab Emirates' 'Greece' 'China' 'Austria' 'Indonesia' 'Egypt'
'Denmark' 'Croatia' 'Taiwan' 'Qatar' 'Slovenia' 'Iceland']
Region Unik: ['EMEA' 'AMER' 'APJ']
Jumlah Baris Duplikat: 0
Profit Unik: [ 41.9136 219.582   6.8714 ... 16.124   4.1028 72.948 ]
```


Preparasi Package : Analisa Lanjutan Data Saas

1. Identifikasi Outlier

Tujuan

Mengidentifikasi nilai ekstrem atau outlier dapat mengindikasikan transaksi yang tidak biasa, kesalahan entri data, atau kasus khusus yang mungkin perlu diteliti lebih lanjut.

Kesimpulan

- **Outlier Positif Dominan:** Produk seperti Alchemy dan Big Ol Database sering kali melampaui ekspektasi, menandakan kinerja tinggi di situasi tertentu.
- **Outlier Negatif Dominan:** Produk seperti Big Ol Database dan Marketing Suite perlu diperhatikan karena lebih banyak outlier negatif.
- **Konsistensi di Rentang Normal:** Produk seperti ChatBot Plugin dan Storage lebih stabil dan mungkin cocok untuk kebutuhan yang mengutamakan hasil konsisten.

Persentase dari setiap 'Tipe Outlier' untuk setiap 'Produk':		
Tipe_Outlier	Outlier Negatif	Outlier Positif \
Product		
Alchemy	0.000000	95.588235
Big Ol Database	29.565217	54.782609
ChatBot Plugin	0.000000	2.135678
ContactMatcher	13.192182	39.250814
Data Smasher	0.387097	22.064516
FinanceHub	7.441386	24.362895
Marketing Suite	25.877193	24.561404
Marketing Suite - Gold	6.146572	19.148936
OneView	9.012876	27.467811
SaaS Connector Pack	3.866249	19.017764
SaaS Connector Pack - Gold	0.000000	11.023622
Site Analytics	5.849269	29.696288
Storage	0.000000	2.764977
Support	0.576923	9.166667
Tidak Ada Outlier		
Tipe_Outlier		
Product		
Alchemy	4.411765	
Big Ol Database	15.652174	
ChatBot Plugin	97.864322	
ContactMatcher	47.557003	
Data Smasher	77.548387	
FinanceHub	68.195719	
Marketing Suite	49.561404	
Marketing Suite - Gold	74.704492	
OneView	63.519313	
SaaS Connector Pack	77.115987	
SaaS Connector Pack - Gold	88.976378	
Site Analytics	64.454443	
Storage	97.235023	
Support	90.256410	

Preparasi Package : Analisa Lanjutan Data Saas

2. Pemeriksaan Distribusi Data

Analisis Uji Normalitas Menggunakan Shapiro-Wilk

Hasil uji Shapiro-Wilk menunjukkan bahwa distribusi data untuk variabel **Sales**, **Quantity**, **Discount**, dan **Profit** kemungkinan besar tidak normal, terlihat dari *p-value* yang sangat rendah (di bawah 0.05).

1. Distribusi Sales

- *p-value*: 4.07×10^{-105}
- Interpretasi: Distribusi Sales tidak normal, mengindikasikan adanya outliers atau skewness, dengan beberapa nilai penjualan yang jauh lebih tinggi atau lebih rendah dari yang lain.

2. Distribusi Quantity

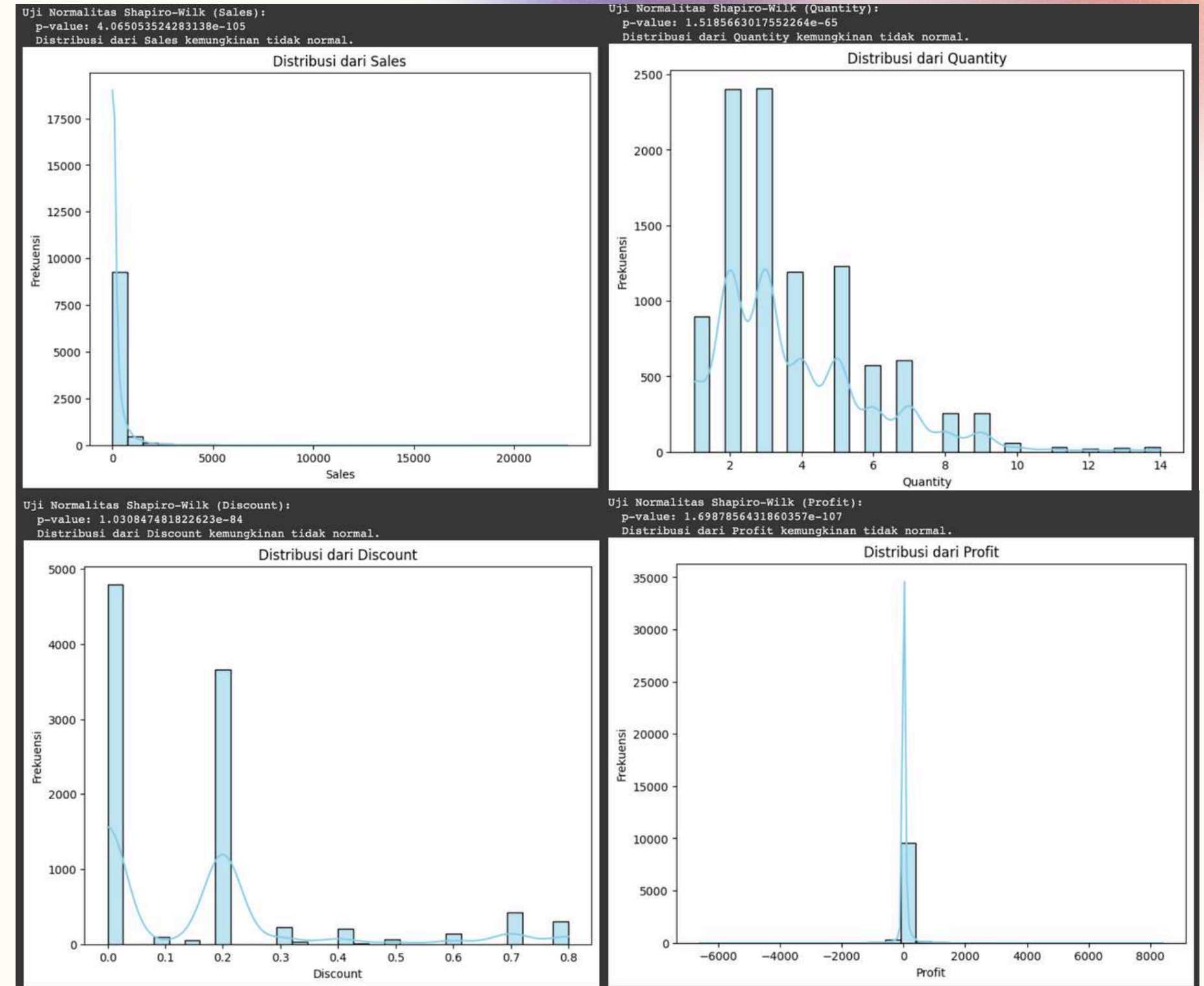
- *p-value*: 1.52×10^{-65}
- Interpretasi: Distribusi Quantity juga tidak normal, kemungkinan disebabkan oleh banyak transaksi dengan kuantitas rendah dan beberapa dengan kuantitas tinggi.

3. Distribusi Discount

- *p-value*: 1.03×10^{-84}
- Interpretasi: Distribusi Discount tidak normal, mungkin karena variasi dalam strategi diskon yang diterapkan, menghasilkan skewness.

4. Distribusi Profit

- *p-value*: 1.70×10^{-107}
- Interpretasi: Distribusi Profit tidak normal, disebabkan oleh transaksi dengan profit yang sangat tinggi atau rendah, berkaitan dengan variasi pada sales, quantity, dan discount.



Preparasi Package : Analisa Lanjutan Data Saas

3. Analisis Efektivitas Penjualan per Produk

Analisis Penjualan Median per Produk

Grafik ini memberikan gambaran mengenai performa penjualan masing-masing produk.

1. Produk dengan Penjualan Median Tertinggi

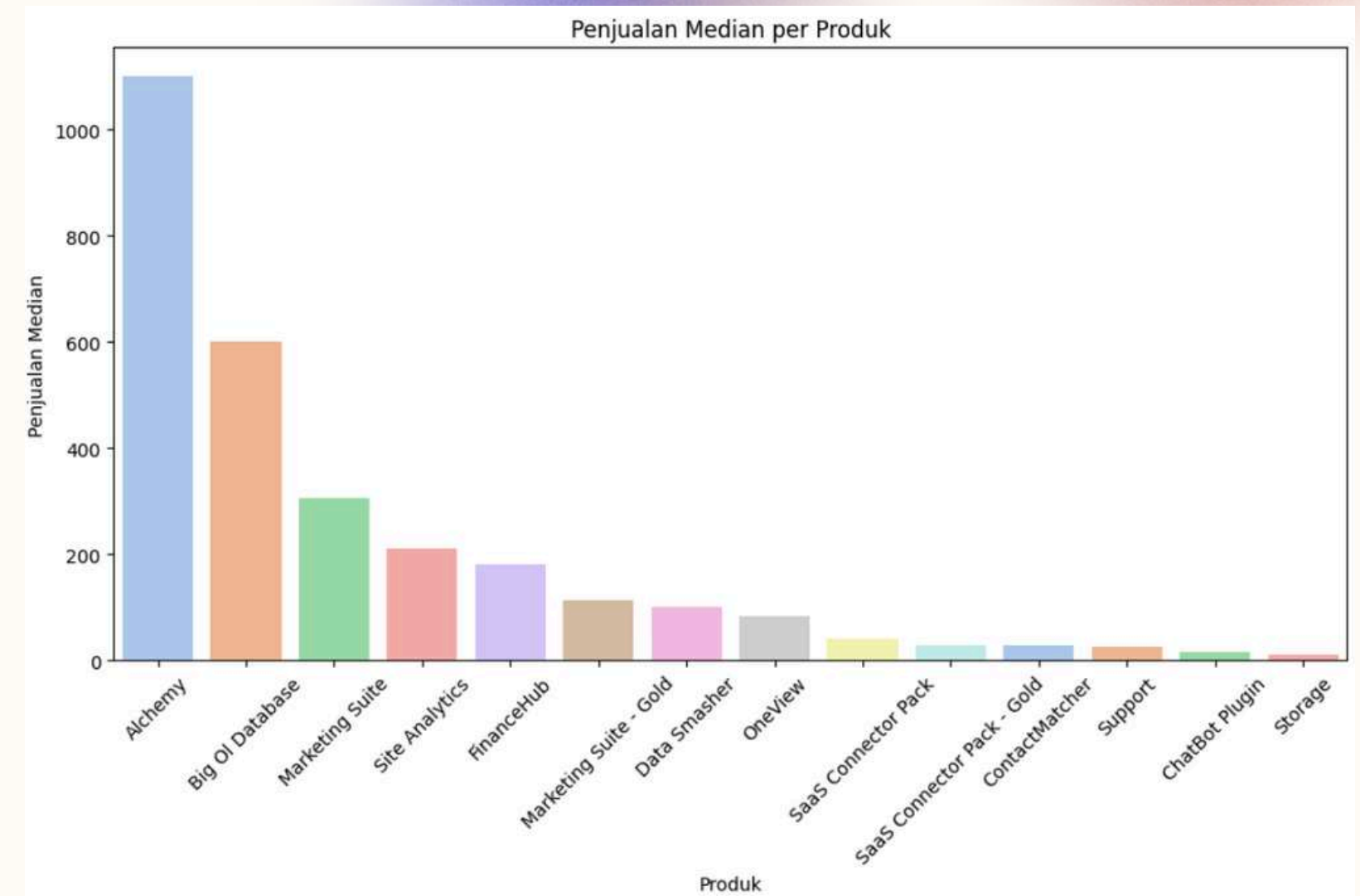
- Alchemy menempati posisi teratas dengan penjualan median yang jauh lebih tinggi dari produk lain, menunjukkan popularitas atau harga jual yang tinggi.
- Big Ol Database berada di posisi kedua, juga menunjukkan performa penjualan yang baik meski tidak setinggi Alchemy.

2. Produk dengan Penjualan Median Sedang

- Produk seperti FinanceHub, Marketing Suite – Gold, dan Data Smasher memiliki penjualan median yang lebih rendah, tetapi tetap menarik bagi sebagian pelanggan.

3. Produk dengan Penjualan Median Terendah

- SaaS Connector Pack, Support, ChatBot Plugin, dan Storage memiliki penjualan median yang rendah, menandakan permintaan yang lebih sedikit.
- Storage memiliki penjualan median terendah, menunjukkan performa penjualan yang paling lemah.



Preparasi Package : Analisa Lanjutan Data Saas

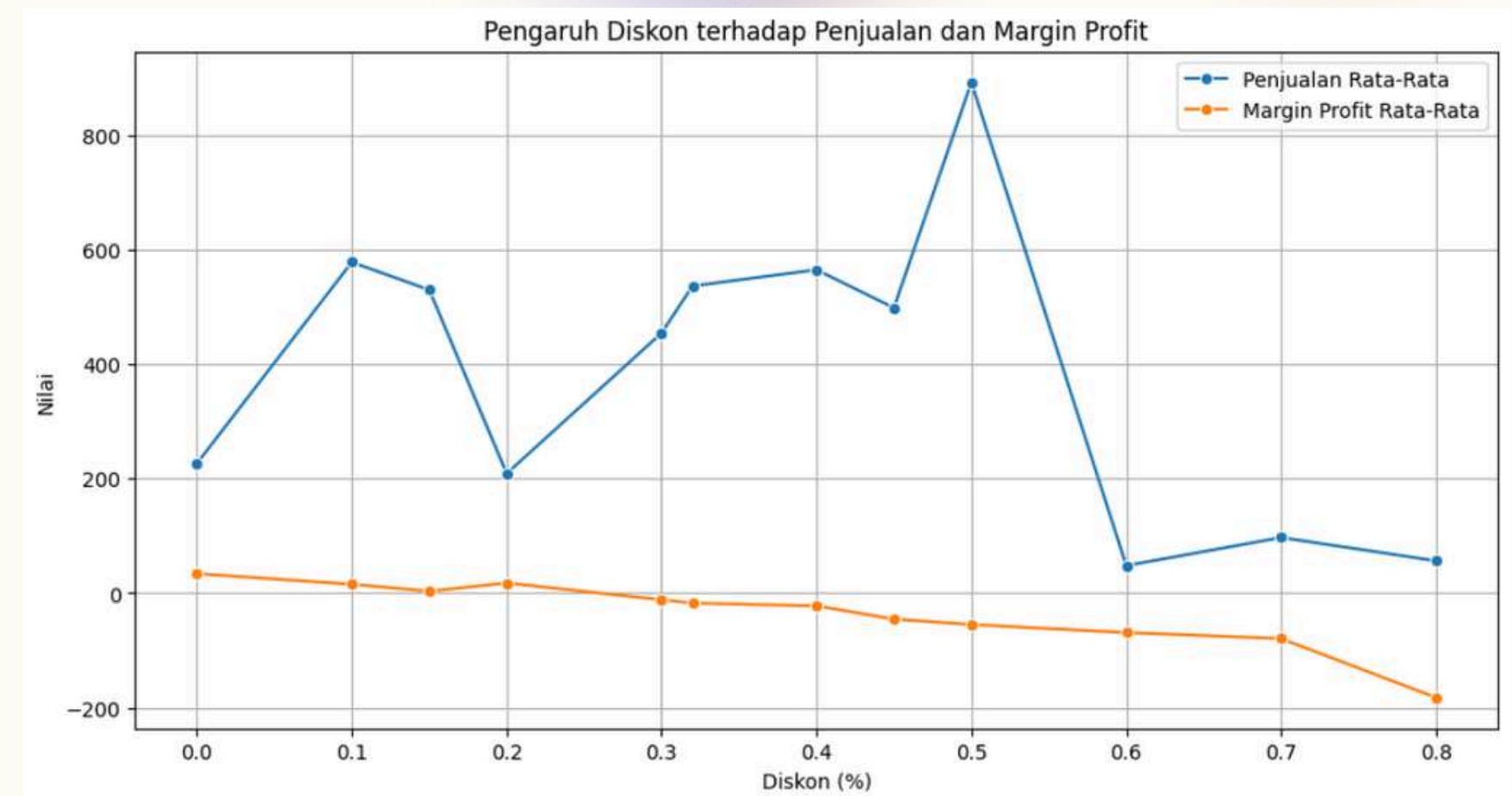
4. Analisis Pengaruh Diskon terhadap Penjualan dan Margin Profit

Hasil Analisis

- **Penjualan Rata-Rata:**
 - Penjualan rata-rata meningkat seiring dengan peningkatan diskon, namun ada penurunan signifikan pada diskon sekitar 0,6.
 - Diskon 50% menghasilkan penjualan tertinggi, menandakan daya tarik yang kuat pada level ini.
- **Margin Profit Rata-Rata:**
 - Margin profit rata-rata cenderung menurun dengan meningkatnya diskon. Diskon tinggi dapat meningkatkan penjualan, tetapi margin profit dapat menjadi negatif.
 - Diskon di kisaran 20%–30% menunjukkan margin yang masih positif, menjadi batas optimal dalam memberikan diskon.

Kesimpulan

- **Optimalisasi Diskon:** Diskon hingga 20%–30% bisa meningkatkan penjualan tanpa mengurangi margin profit secara signifikan.
- **Diskon Tinggi vs Profitabilitas:** Diskon yang terlalu besar berisiko merugikan profitabilitas, sehingga harus dihindari kecuali untuk alasan khusus.



Preparasi Package : Analisa Lanjutan Data Saas

5. Analisis Customer Lifetime Value (CLV) Rata- Rata per Segmen

Tujuan Analisis

Untuk mengetahui Customer Lifetime Value (CLV) rata-rata pada setiap segmen pelanggan, yang diestimasi selama 12 bulan.

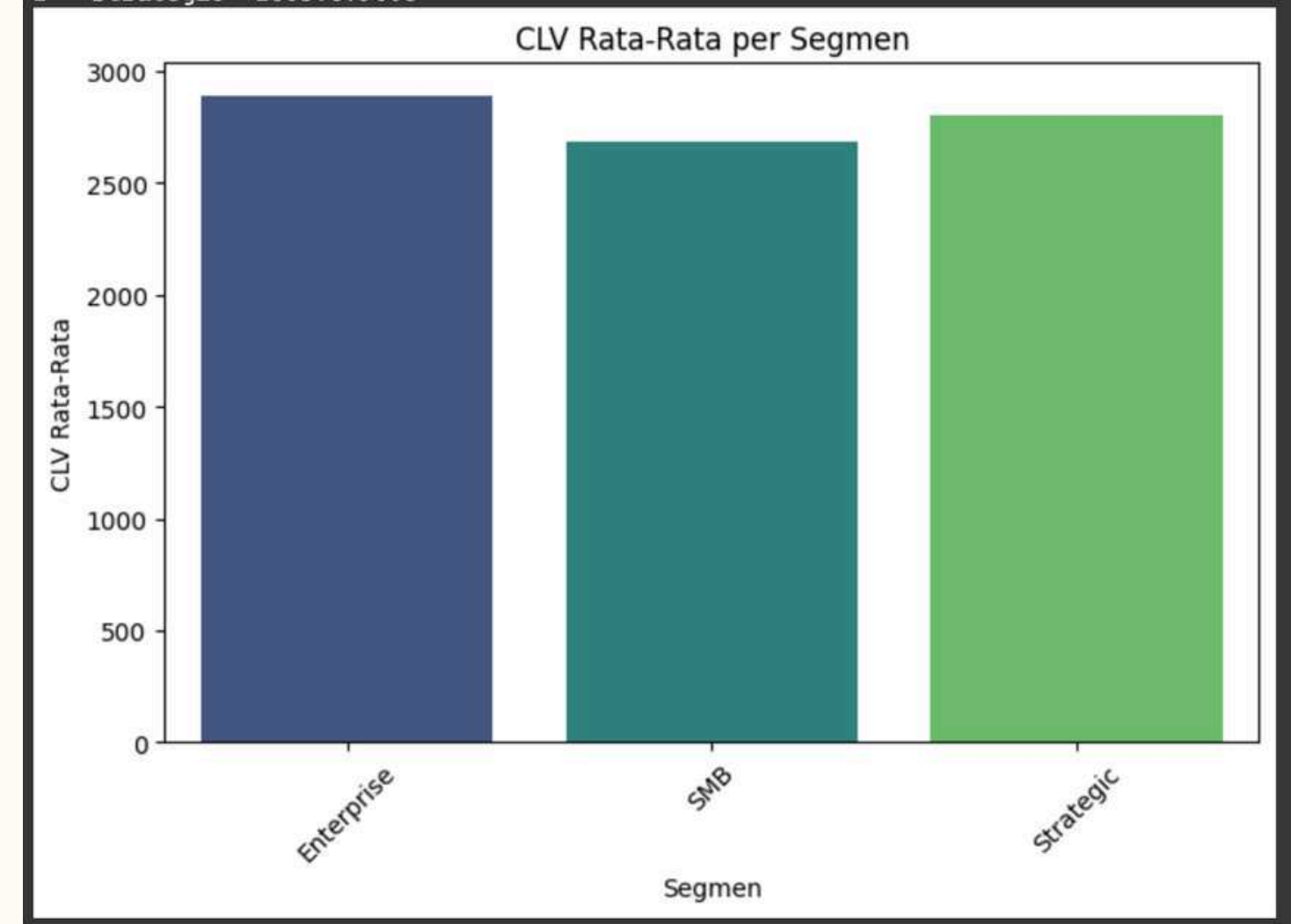
Hasil Analisis CLV

- 1.Enterprise: Segmen ini memiliki CLV tertinggi, yaitu sekitar 2891.66, menunjukkan kontribusi terbesar terhadap pendapatan jangka panjang.
- 2.Strategic: CLV rata-rata untuk segmen ini adalah 2805.88, sedikit di bawah Enterprise, tetapi tetap menunjukkan nilai yang signifikan.
- 3.SMB: Segmen Small and Medium Business memiliki CLV rata-rata 2684.80, dengan perbedaan yang tidak terlalu jauh dari segmen lainnya, sehingga tetap memberikan kontribusi berarti.

Kesimpulan

- Segmen Enterprise: Dengan CLV tertinggi, perusahaan disarankan untuk menerapkan strategi khusus untuk menjaga pelanggan di segmen ini.
- Segmen Strategic: Karena CLV yang tinggi, strategi retention yang lebih fokus bisa diterapkan untuk pelanggan di segmen ini.
- Segmen SMB: Walaupun CLV-nya lebih rendah, segmen ini tetap memiliki potensi untuk ditingkatkan melalui strategi promosi atau upselling.

CLV Rata-Rata per Segmen:		
	Segment	CLV
0	Enterprise	2891.664488
1	SMB	2684.803726
2	Strategic	2805.879603



Preparasi Package : Analisa Lanjutan Data Saas

6. Analisis Total Penjualan dan Profit Berdasarkan Rentang Diskon

1. Rentang Diskon 0–10%

- Total Penjualan: 54.369
- Total Profit: 9.029
- Pada rentang diskon ini, perusahaan tetap dapat meraih profit meski total penjualan relatif kecil. Diskon rendah ini cukup baik untuk mempertahankan profitabilitas.

2. Rentang Diskon 10–20%

- Total Penjualan: 792.153 (tertinggi)
- Total Profit: 91.756
- Diskon ini sangat efektif karena menghasilkan total penjualan dan profit tertinggi. Rentang diskon ini menarik bagi konsumen tanpa mengorbankan margin profit.

3. Rentang Diskon 20–30%

- Total Penjualan: 103.227
- Total Profit: -10.369
- Meskipun ada penjualan, perusahaan mulai mengalami kerugian. Diskon yang lebih besar menggerus margin profit.

4. Rentang Diskon 30–40%

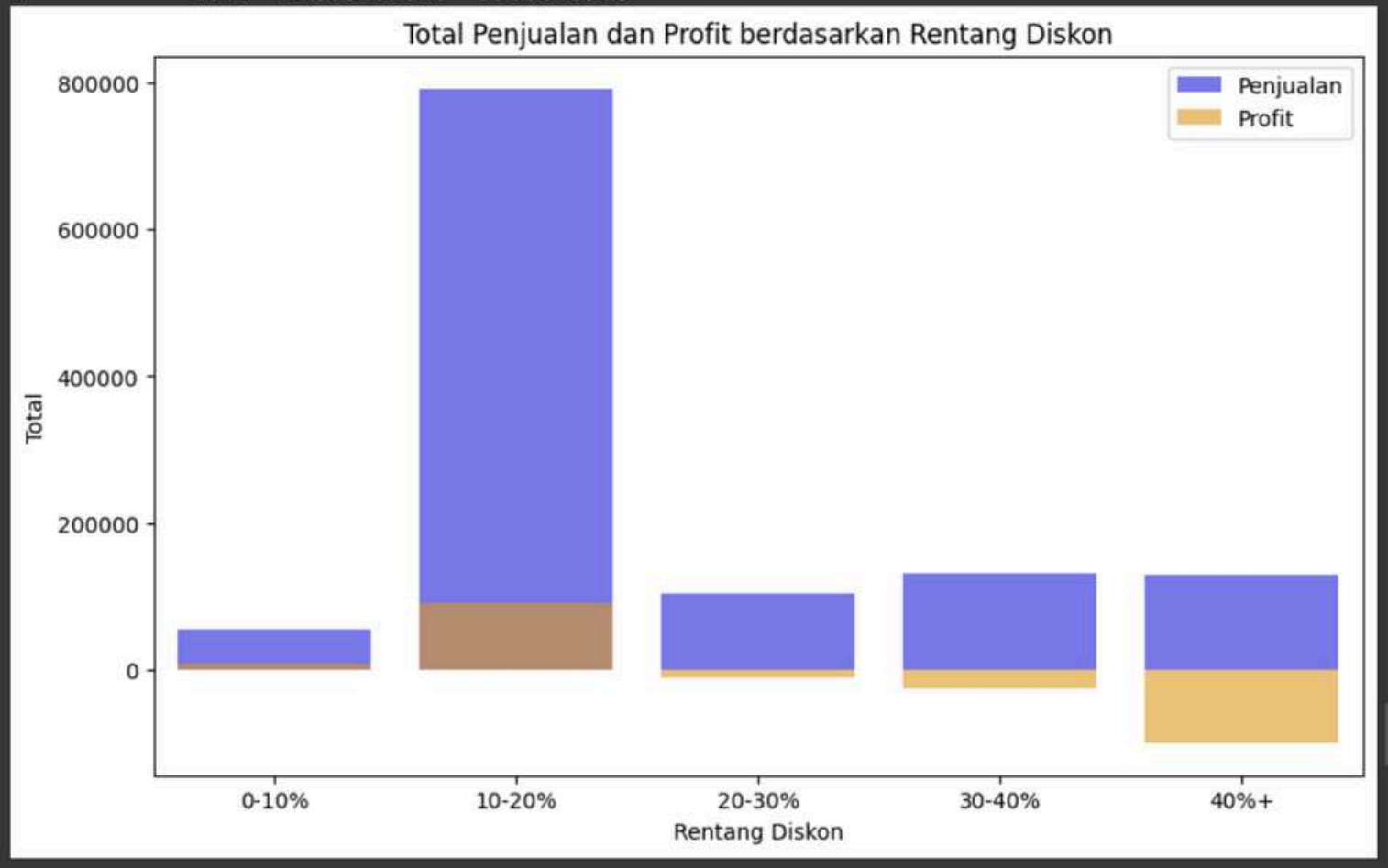
- Total Penjualan: 130.911
- Total Profit: -25.448
- Penjualan sedikit meningkat, tetapi kerugian semakin besar. Diskon tinggi mulai mengurangi profitabilitas.

5. Rentang Diskon 40%+

- Total Penjualan: 128.632
- Total Profit: -99.558 (kerugian terbesar)
- Meskipun penjualan mirip dengan rentang sebelumnya, kerugian sangat besar. Diskon ini terlalu membebani profit perusahaan.

Total Penjualan dan Profit berdasarkan Rentang Diskon:

Rentang Diskon	Sales	Profit
0	0-10%	54369.3510 9029.1770
1	10-20%	792152.8895 91756.2975
2	20-30%	103226.6550 -10369.2774
3	30-40%	130911.2428 -25448.1881
4	40%+	128632.2520 -99558.5905



Preparasi Package : Analisa Lanjutan Data Saas

7. Analisis Hasil Clustering Berdasarkan Penjualan dan Profit

Klaster 0 (Warna Ungu)

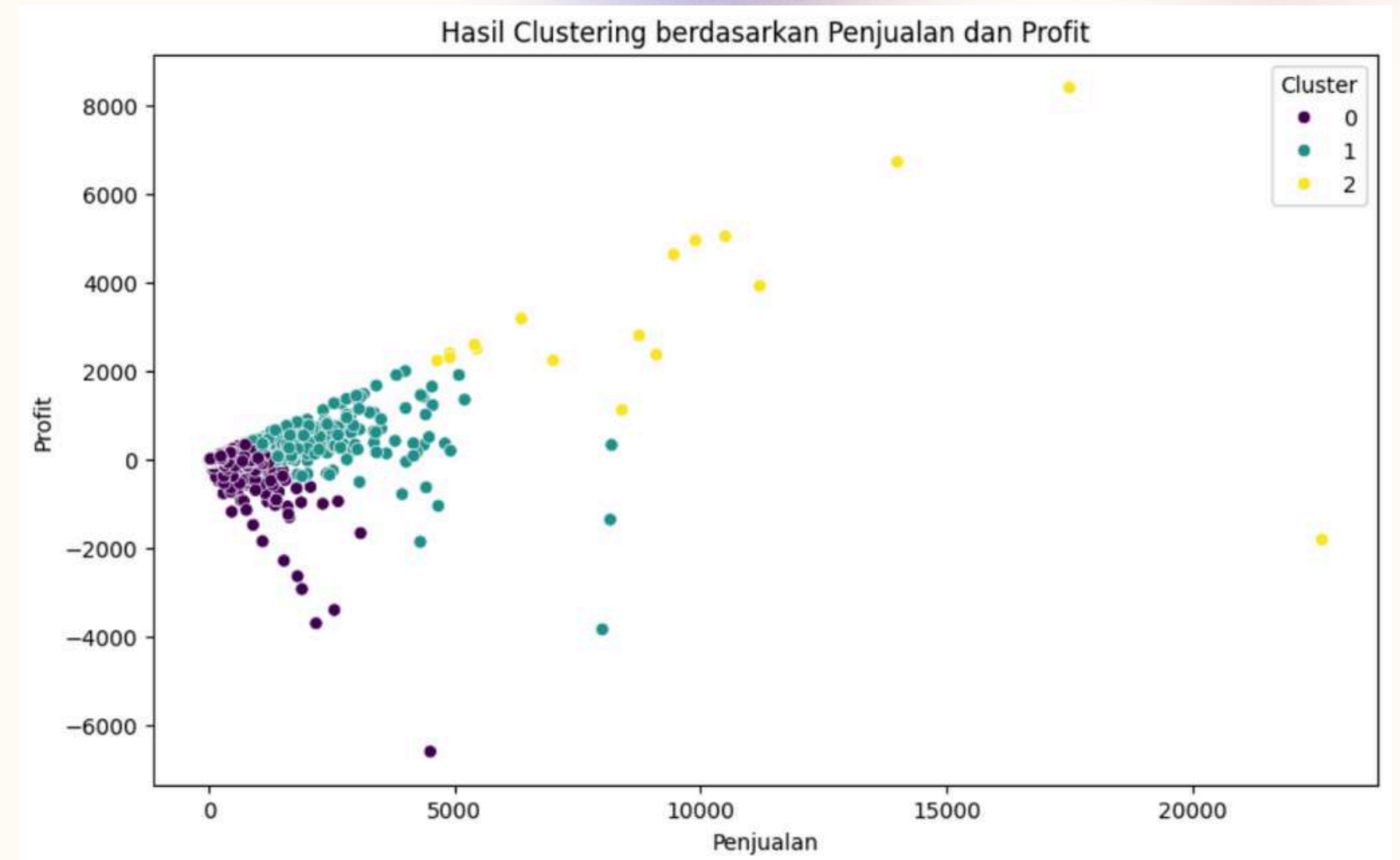
- Ciri-ciri: Klaster ini berada di area dengan nilai penjualan dan profit yang rendah, bahkan ada yang negatif.
- Interpretasi: Klaster ini mungkin terdiri dari produk atau pelanggan yang memiliki tingkat penjualan yang rendah dan menghasilkan profit yang kecil, atau bahkan mengalami kerugian.

Klaster 1 (Warna Hijau)

- Ciri-ciri: Klaster ini terletak di area dengan nilai penjualan dan profit yang sedang.
- Interpretasi: Klaster ini mungkin mewakili produk atau pelanggan yang memberikan penjualan yang cukup baik dengan profit positif, meskipun tidak terlalu tinggi.

Klaster 2 (Warna Kuning)

- Ciri-ciri: Klaster ini berada di area dengan nilai penjualan dan profit yang tinggi.
- Interpretasi: Klaster ini merupakan yang paling menguntungkan bagi perusahaan, di mana produk atau pelanggan di dalamnya memberikan kontribusi penjualan dan profit yang signifikan.



Preparasi Package : Analisa Lanjutan Data Saas

8. Analisis Hasil Uji ANOVA Profit Berdasarkan Segmen

1. Hasil Uji ANOVA

- p-value yang diperoleh dari analisis ini adalah 0.4074. Dalam konteks statistik, nilai p-value ini akan dibandingkan dengan tingkat signifikansi yang umum digunakan, yaitu 0.05, untuk menentukan apakah hasilnya signifikan atau tidak.
- Karena p-value (0.4074) > 0.05, kita tidak memiliki cukup bukti untuk mengatakan bahwa ada perbedaan signifikan dalam profit rata-rata di antara segmen-segmen pelanggan yang berbeda.

2. Kesimpulan

- Dari hasil analisis ini, dapat disimpulkan bahwa tidak ada perbedaan signifikan dalam profit rata-rata antara segmen-segmen pelanggan. Ini menunjukkan bahwa profit rata-rata yang dihasilkan dari segmen pelanggan Enterprise, SMB, dan Strategic cenderung cukup mirip.

```
[ ] # Uji ANOVA untuk profit rata-rata per segmen
segments = [df['Profit'][df['Segment'] == seg] for seg in df['Segment'].unique()]
stat, p_value = f_oneway(*segments)

print("\nHasil Uji ANOVA Profit Berdasarkan Segmen:")
print(f"p-value: {p_value}")
if p_value < 0.05:
    print("Terdapat perbedaan signifikan dalam profit antar segmen pelanggan.")
else:
    print("Tidak terdapat perbedaan signifikan dalam profit antar segmen pelanggan.")
```



```
Hasil Uji ANOVA Profit Berdasarkan Segmen:
p-value: 0.4073892685202365
Tidak terdapat perbedaan signifikan dalam profit antar segmen pelanggan.
```




Kesimpulan

1. Bagaimana Preferensi Pelanggan dan Pemahaman Pola Penjualan Berkontribusi pada Optimalisasi Pengembangan Produk?

Hambatan utama yang terlihat dalam analisis ini adalah pola data penjualan (Sales) dan jumlah pembelian (Quantity) yang menunjukkan distribusi tidak normal. Ini mengindikasikan bahwa preferensi pelanggan bervariasi. Misalnya, segmen pelanggan tertentu mungkin melakukan pembelian dalam jumlah besar, menunjukkan minat pada fitur atau manfaat produk spesifik. Sebaliknya, pelanggan yang melakukan pembelian kecil mungkin lebih memilih harga yang lebih rendah atau diskon yang lebih besar. Dengan melakukan segmentasi yang lebih mendalam berdasarkan pola pembelian ini, tim pengembangan dapat menciptakan produk atau fitur yang lebih sesuai dengan keinginan masing-masing kelompok pelanggan.

2. Dengan Cara Apa Memenuhi Kebutuhan Pelanggan dan Keputusan Pembelian Dapat Memandu Tim Pengembangan dalam Meningkatkan Produk?

Analisis terhadap variabel Discount dan Profit menunjukkan bahwa pemahaman mengenai kebutuhan pelanggan dapat membantu tim pengembangan. Distribusi diskon yang tidak merata mencerminkan adanya pola promosi yang berbeda-beda untuk segmen pelanggan yang berbeda. Dengan menganalisis variabel ini, tim pengembangan dapat menyesuaikan penawaran harga dan promosi untuk pelanggan yang lebih sensitif terhadap diskon. Misalnya, untuk segmen yang lebih menghargai diskon, tim bisa merancang penawaran khusus yang dapat meningkatkan kepuasan pelanggan dan memperkuat loyalitas mereka.



Kesimpulan

3. Dengan Cara Apa Memenuhi Kebutuhan Pelanggan dan Keputusan Pembelian Dapat Memandu Tim Pengembangan dalam Meningkatkan Produk?

Variabel Profit menunjukkan adanya distribusi yang tidak normal, yang berarti ada produk tertentu yang memberikan margin keuntungan yang lebih tinggi pada segmen tertentu, sedangkan produk lain mungkin tidak menguntungkan. Tim pengembangan harus fokus pada analisis margin profit ini untuk memahami produk mana yang perlu diperkuat atau dimodifikasi.

Misalnya, jika sebuah produk memberikan profit tinggi pada segmen Enterprise, tim dapat memperkuat fitur yang paling menarik bagi segmen tersebut atau mempertimbangkan untuk menyesuaikan produk lain agar lebih sesuai dengan kebutuhan mereka. Dengan cara ini, perusahaan tidak hanya memenuhi kebutuhan pelanggan, tetapi juga meningkatkan potensi keuntungan secara keseluruhan.

terima kasih.