

Assignment Day 28 Marketing Channel Analisis

Nama : Fuad Hasyim DS B35

Link PBI :

<https://drive.google.com/file/d/1m9bMnHyILkpZVg5IFW7dMGNhwWtT5e4e/view?usp=sharing>

Link Google Collab :

https://colab.research.google.com/drive/18hN1OgVKrmIyYmG_OHeVKY_E1K0hz7GA?usp=sharing

1. Memahami Dataset

- Unduh dataset yang telah disediakan.
- Identifikasi variabel utama seperti **Customer ID**, transaksi, frekuensi pembelian, total belanja, dan atribut geografis pelanggan.
- Lakukan **pembersihan data (data cleaning)** jika diperlukan.

Goals: Memastikan data **bersih dan siap digunakan** dalam analisis profitabilitas, perilaku pelanggan, dan segmentasi RFM.

Info data yang diperlukan

```
df_customers.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 793 entries, 0 to 792
Data columns (total 8 columns):
 #   Column                Non-Null Count  Dtype  
---  --
 0   customer_id           793 non-null   object 
 1   purchase_frequency    793 non-null   int64  
 2   total_spent           793 non-null   float64
 3   avg_spent_per_order   793 non-null   float64
 4   total_profit          793 non-null   float64
 5   city                  793 non-null   object 
 6   state                 793 non-null   object 
 7   region                793 non-null   object 
dtypes: float64(3), int64(1), object(4)
memory usage: 49.7+ KB
```

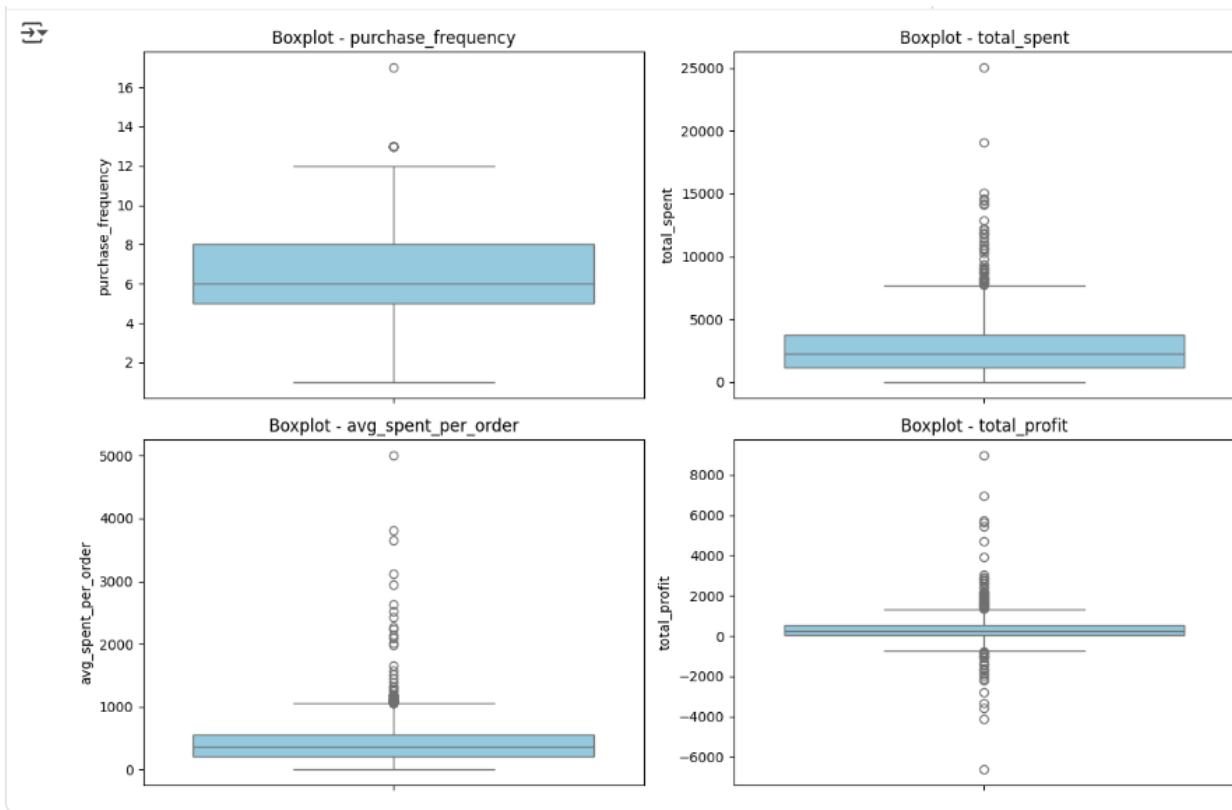
Cek missing value

```
df_customers.isnull().sum()
```

```
customer_id    0
purchase_frequency    0
total_spent      0
avg_spent_per_order  0
total_profit      0
city             0
state            0
region           0
dtype: int64
```

Tidak ada missing value

Cek Outlier

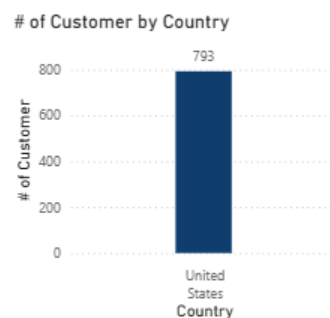


Ditemukan beberapa outliers pada variabel numerik.

Namun outliers tersebut dibiarkan apa adanya karena kemungkinan merepresentasikan perilaku asli customer (high-value customers).

2. Analisis Profitabilitas dan Jumlah Pelanggan

- Gunakan dataset **Superstore** untuk menghitung **profit tertinggi berdasarkan negara**.
- Hitung jumlah pelanggan yang berasal dari negara tersebut.
- Visualisasikan data dengan tabel atau grafik.



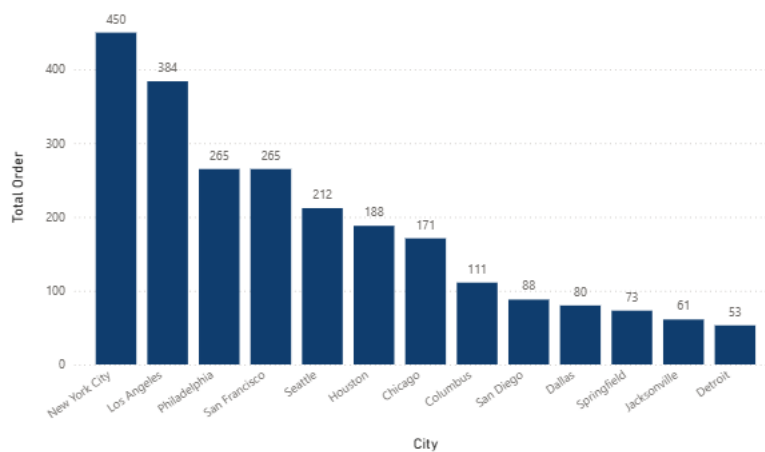
Total profit yang dihasilkan di seluruh transaksi Superstore di AS adalah **sekitar 290 ribu** dengan jumlah customer sebanyak (**793**).

3. Analisis Perilaku Pembelian Pelanggan

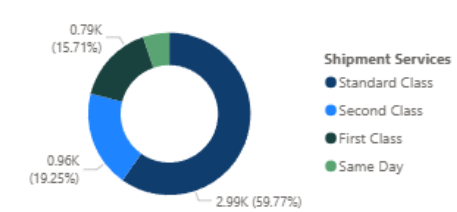
- Lakukan eksplorasi data terkait **kota dengan jumlah transaksi terbanyak**.
- Identifikasi **kategori produk paling laris**.

- Temukan **jenis pengiriman yang paling sering digunakan** oleh pelanggan.
- Gunakan visualisasi yang sesuai untuk memperjelas hasil analisis.

Total Order by City

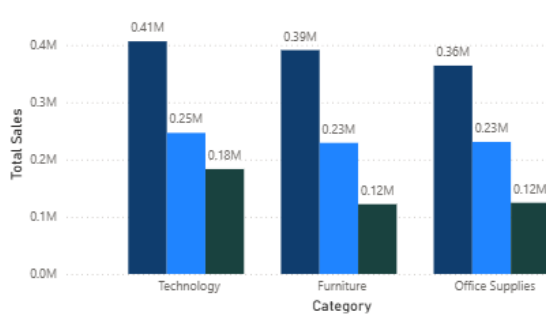


Total Order by Shipment Services

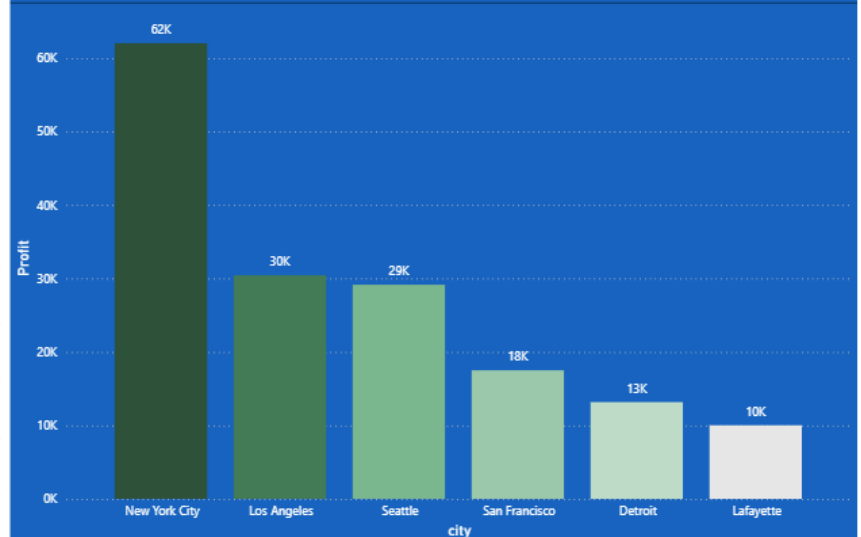


Total Sales by Category and Segment

Segment ● Consumer ● Corporate ● Home Office



Profit by city



- **New York City** menghasilkan profit tertinggi (62K), jauh di atas Los Angeles (30K) dan Seattle (29K).
- Segmen **Champions** mendominasi (202 pelanggan), diikuti oleh **Lost Customer** (146) dan **Cannot Lose Them** (131).
- **Technology** menyumbang penjualan tertinggi (410K), disusul **Furniture** (390K) dan **Office Supplies** (360K).
- **Standard Class** paling dominan (59.8% dari total order), diikuti oleh Second Class (19.2%).
- Transaksi paling banyak berasal dari **New York City** (450) dan **Los Angeles** (384).

4. Segmentasi Pelanggan dengan RFM Analysis

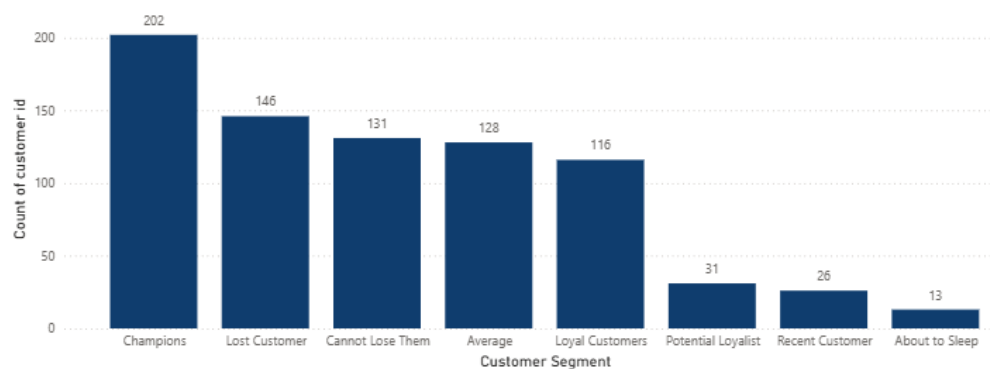
- Lakukan **perhitungan Recency, Frequency, dan Monetary (RFM)** berdasarkan data transaksi.
- Kategorikan pelanggan ke dalam **9 segmen berdasarkan skor RFM**.
- Buat visualisasi distribusi pelanggan dalam tiap segmen.

```

Customer Segment =
SWITCH(
  TRUE(),
  -- 1. Champions
  'RFMTable'[RFM Score] >= "451" && 'RFMTable'[RFM Score] <= "555", "Champions",
  -- 2. Loyal Customers
  'RFMTable'[RFM Score] >= "401" && 'RFMTable'[RFM Score] <= "450", "Loyal Customers",
  -- 3. Potential Loyalist
  'RFMTable'[RFM Score] >= "351" && 'RFMTable'[RFM Score] <= "400", "Potential Loyalist",
  -- 4. Recent Customers
  'RFMTable'[RFM Score] >= "301" && 'RFMTable'[RFM Score] <= "350", "Average",
  -- 5. Need Attention
  'RFMTable'[RFM Score] >= "251" && 'RFMTable'[RFM Score] <= "300", "Recent Customer",
  -- 6. About To Sleep
  'RFMTable'[RFM Score] >= "201" && 'RFMTable'[RFM Score] <= "250", "Cannot Lose Them",
  -- 7. At Risk
  'RFMTable'[RFM Score] >= "151" && 'RFMTable'[RFM Score] <= "200", "About to Sleep",
  -- 8. Lost
  'RFMTable'[RFM Score] >= "101" && 'RFMTable'[RFM Score] <= "150", "Lost Customer",
  -- 9. Hibernating
  'RFMTable'[RFM Score] < "101", "Bye Bye",
  -- Default fallback
  "Uncategorized"
)

```

Count of customer id by Customer Segment



1. **Champions (202 pelanggan)**
Segmen terbesar. Ini adalah pelanggan paling bernilai — mereka sering bertransaksi dan memiliki engagement tinggi. Mereka perlu dijaga dengan program loyalitas atau penawaran eksklusif agar tetap aktif.
2. **Lost Customer (146 pelanggan)**
Segmen ini cukup besar. Artinya banyak pelanggan yang dulunya aktif tapi sekarang sudah berhenti bertransaksi. Penting untuk dilakukan **reaktivasi** misalnya lewat email marketing atau promo comeback.
3. **Cannot Lose Them (131 pelanggan)**
Pelanggan penting yang kontribusinya besar. Perlu perhatian khusus agar tidak berpindah ke kompetitor.
4. **Average (128 pelanggan)**
Pelanggan dengan tingkat aktivitas menengah. Bisa diarahkan untuk naik ke segmen yang lebih tinggi melalui program upselling atau loyalty program.
5. **Loyal Customers (116 pelanggan)**
Mereka rutin bertransaksi meski tidak sebanyak Champions. Potensial untuk dijadikan brand advocate.

6. **Potential Loyalist (31 pelanggan)**

Baru mulai menunjukkan pola loyalitas. Diperlukan dorongan agar menjadi pelanggan tetap, misalnya dengan diskon khusus atau poin reward.

7. **Recent Customer (26 pelanggan)**

Pelanggan baru. Masih tahap awal hubungan, jadi penting untuk memberi pengalaman positif agar mereka bertahan.

8. **About to Sleep (13 pelanggan)**

Hampir tidak aktif. Segmen ini perlu **intervensi cepat** agar tidak menjadi Lost Customer.

9. **Hibernating (0 Pelanggan)**

Conclusion & Rekomendasi

1. Product Strategy

- Fokuskan promosi pada **Technology**, karena kategori ini menjadi penggerak penjualan tertinggi.
- Lakukan **cross-selling** untuk Furniture & Office Supplies, khususnya ke pelanggan korporat.

2. Geographical Expansion

- Tingkatkan kampanye pemasaran di kota-kota besar (New York, LA, Seattle) karena kontribusinya signifikan.
- Buat strategi **targeted marketing** di kota kecil untuk meningkatkan awareness dan penetrasi pasar.

3. Logistics Optimization

- Perkuat **Standard Class shipping** agar lebih cepat dan efisien, mengingat mayoritas pelanggan menggunakan layanan ini.
- Evaluasi harga & kecepatan layanan **Same Day**, karena masih sedikit diminati — mungkin karena biaya terlalu tinggi.

Mayoritas pelanggan berada pada segmen **Champions dan Lost Customer**, menunjukkan adanya **kesenjangan antara pelanggan aktif dan pelanggan yang hilang**.

Strategi bisnis sebaiknya fokus pada dua hal:

- **Retensi pelanggan terbaik (Champions, Loyal, Cannot Lose Them).**
- **Reaktivasi pelanggan yang hilang atau hampir tidak aktif (Lost & About to Sleep).**

Channel Marketing Analysis Superstore Dataset

793

Customers

6.32

Average of Frequency

2.90K

Average of Monetary

531

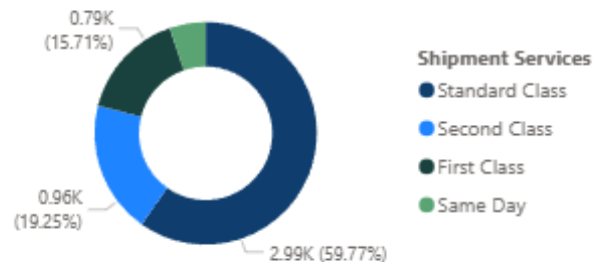
City

146.80

Average of Recency

Customer Segment (0/8)

Total Order by Shipment Services



Insight

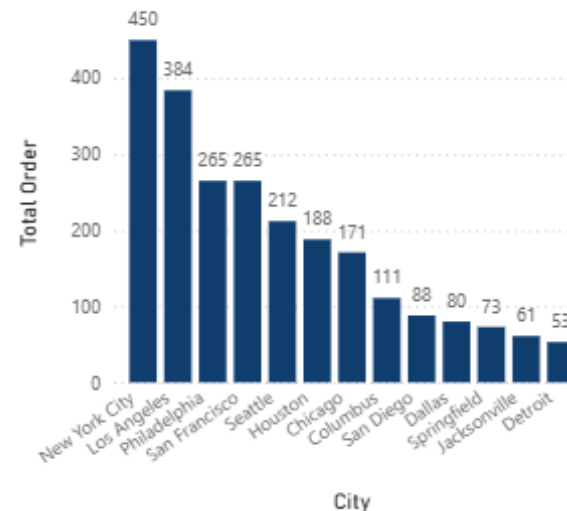
New York City menghasilkan profit tertinggi (62K), jauh di atas Los Angeles (30K) dan Seattle (29K). Segmen **Champions** mendominasi (202 pelanggan), diikuti oleh **Lost Customer** (146) dan **Cannot Lose Them** (131). **Technology** menyumbang penjualan tertinggi (410K), disusul **Furniture** (390K) dan **Office Supplies** (360K). **Standard Class** paling dominan (59.8% dari total order), diikuti oleh **Second Class** (19.2%). Transaksi paling banyak berasal dari **New York City** (450) dan **Los Angeles** (384).

Business Recommendations:

1. Customer Engagement

- Berikan program **reward & eksklusif** untuk **Champions** agar tetap loyal.
- Lakukan **re-engagement campaign** untuk **Lost Customer** dan **About to Sleep** dengan promosi khusus atau reminder email

Total Order by City



1. Product Strategy

- Fokuskan promosi pada **Technology**, karena kategori ini menjadi penggerak penjualan tertinggi.
- Lakukan **cross-selling** untuk Furniture & Office Supplies, khususnya ke pelanggan korporat.

2. Geographical Expansion

- Tingkatkan kampanye pemasaran di kota-kota besar (New York, LA, Seattle) karena kontribusinya signifikan.
- Buat strategi **targeted marketing** di kota kecil untuk meningkatkan awareness dan penetrasi pasar.

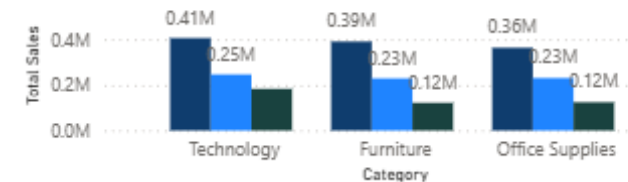
3. Logistics Optimization

- Perkuat **Standard Class shipping** agar lebih cepat dan efisien, mengingat mayoritas pelanggan menggunakannya.
- Evaluasi harga & kecepatan layanan **Same Day**, karena masih sedikit diminati — mungkin karena biaya terlalu tinggi.

[Link Google Collabs](#)

Total Sales by Category and Segment

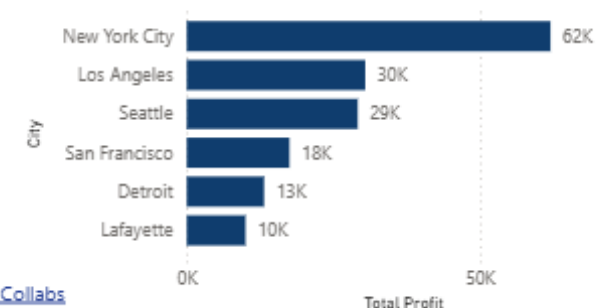
Segment ● Consumer ● Corporate ● Home Office



Total Customer by Customer Segment



Total Profit by City



531

City

793

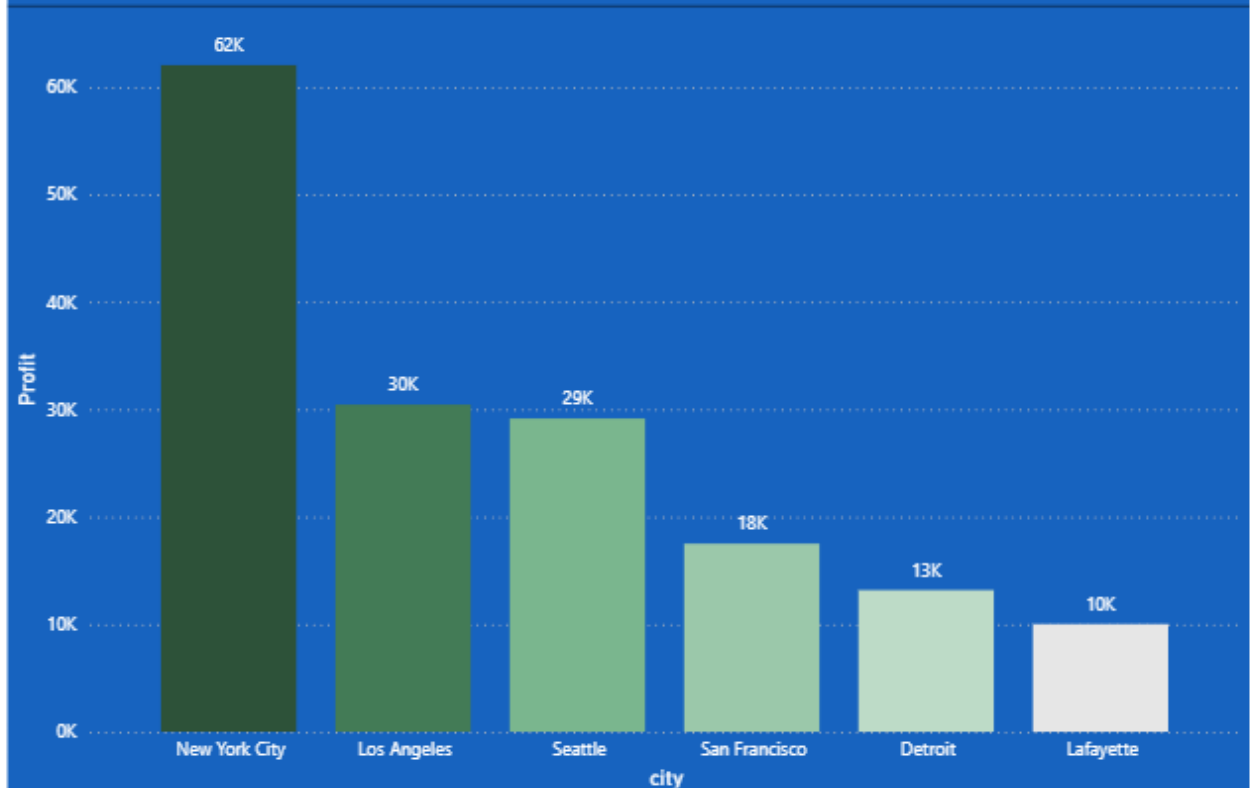
Customers

Profitability by City

Kota mana yang paling profit & berapa jumlah user-nya?

Count of customer id	city	Sum of profit
355	New York City	62,036.98
304	Los Angeles	30,440.76
188	Seattle	29,156.10
232	San Francisco	17,507.39
50	Detroit	13,181.79
17	Lafayette	10,018.39
35	Jackson	7,581.68
17	Atlanta	6,993.66
13	Minneapolis	6,824.58
86	San Diego	6,377.20
66	Springfield	6,200.70
97	Columbus	5,897.10
44	Newark	5,793.76
33	Columbia	5,606.12
40	Richmond	5,567.25
13	Providence	5,505.66
19	Henderson	4,911.86
31	Arlington	4,169.70
28	Lakewood	3,234.03
23	Milwaukee	2,793.09
6	Yonkers	2,767.76
12	Indianapolis	2,741.65
9	Roswell	2,427.44
21	Lawrence	2,382.15
793		286,397.02

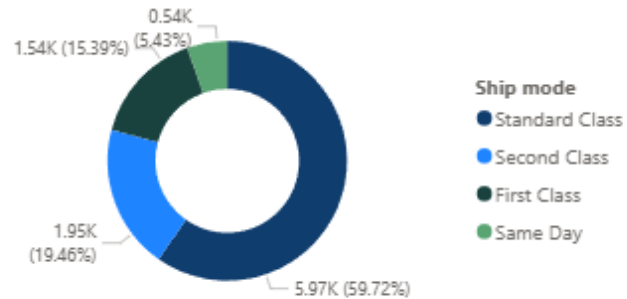
Profit by city



Customer Behavior

Jenis pengiriman paling sering digunakan

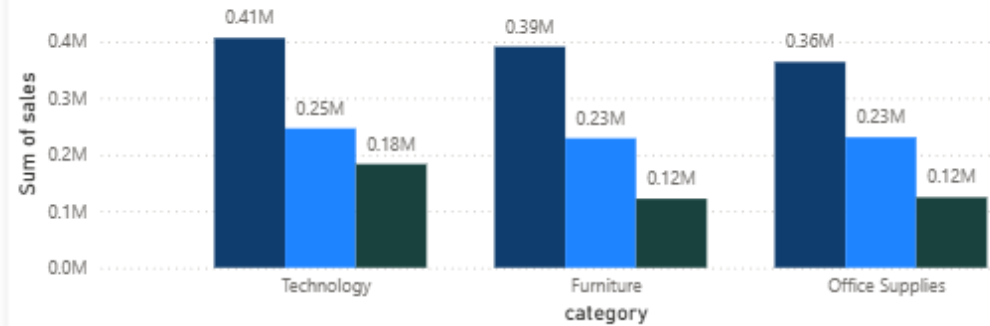
Count of order id by Ship mode



Kategori produk paling banyak dibeli

Sum of sales by category and segment

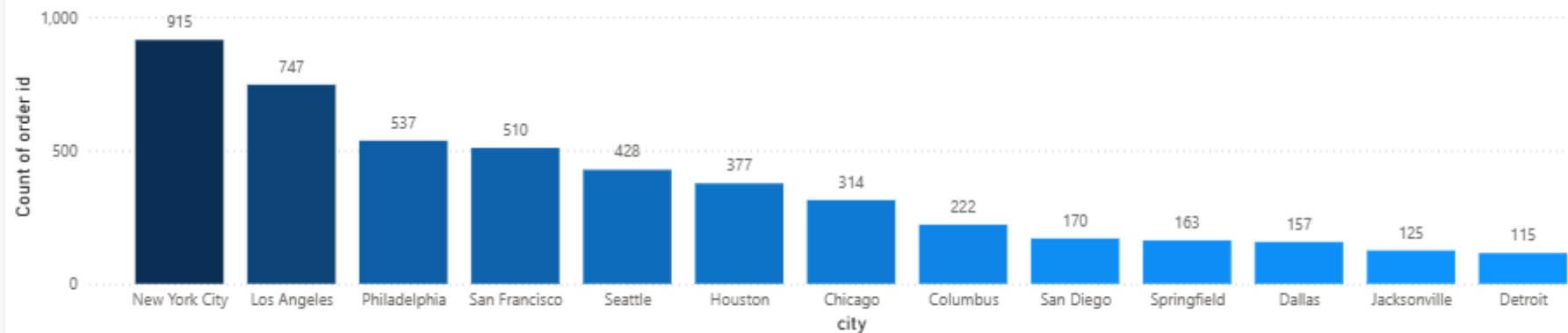
segment ● Consumer ● Corporate ● Home Office



Kota dengan pembelian terbanyak

Count of order id by city

Count of city 115 915



146.80

Average of Recency

2.90K

Average of Monetary

6.32

Average of Frequency

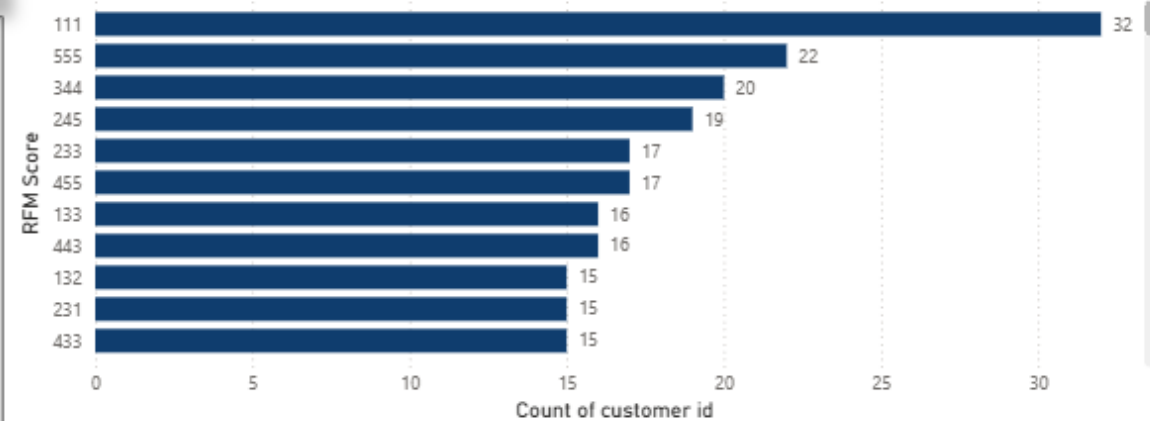
customer_id	Recency	Frequency	Monetary	RFM Score	Customer Segment
AA-10315	184	5	5,563.56	235	Cannot Lose Them
AA-10375	19	9	1,056.39	552	Champions
AA-10480	259	4	1,790.51	123	Lost Customer
AA-10645	55	6	5,086.94	335	Average
AB-10015	415	3	886.16	111	Lost Customer
AB-10060	54	8	7,755.62	345	Average
AB-10105	41	10	14,473.57	455	Champions
AB-10150	41	5	966.71	432	Loyal Customers
AB-10165	25	8	1,113.84	542	Champions
AB-10255	166	9	914.53	251	Recent Customer
AB-10600	50	4	1,515.86	422	Loyal Customers
AC-10420	232	3	925.80	111	Lost Customer
AC-10450	11	7	5,527.85	545	Champions
AC-10615	6	5	2,537.69	533	Champions
AC-10660	282	5	657.32	131	Lost Customer
AD-10180	29	8	6,106.88	445	Loyal Customers
AF-10870	41	7	2,052.91	443	Loyal Customers
AF-10885	415	4	861.57	121	Lost Customer
AG-10270	278	5	2,582.90	133	Lost Customer
AG-10300	118	4	367.55	221	Cannot Lose Them
AG-10330	8	5	660.97	531	Champions
AG-10390	114	5	200.95	231	Cannot Lose Them
AG-10495	78	8	2,356.86	343	Average
AG-10525	845	4	1,455.04	122	Lost Customer
AG-10675	97	7	3,489.04	344	Average

RFM Summary & Segmentation

Customer Segment (0/8)



Count of customer id by RFM Score



Count of customer id by Customer Segment

