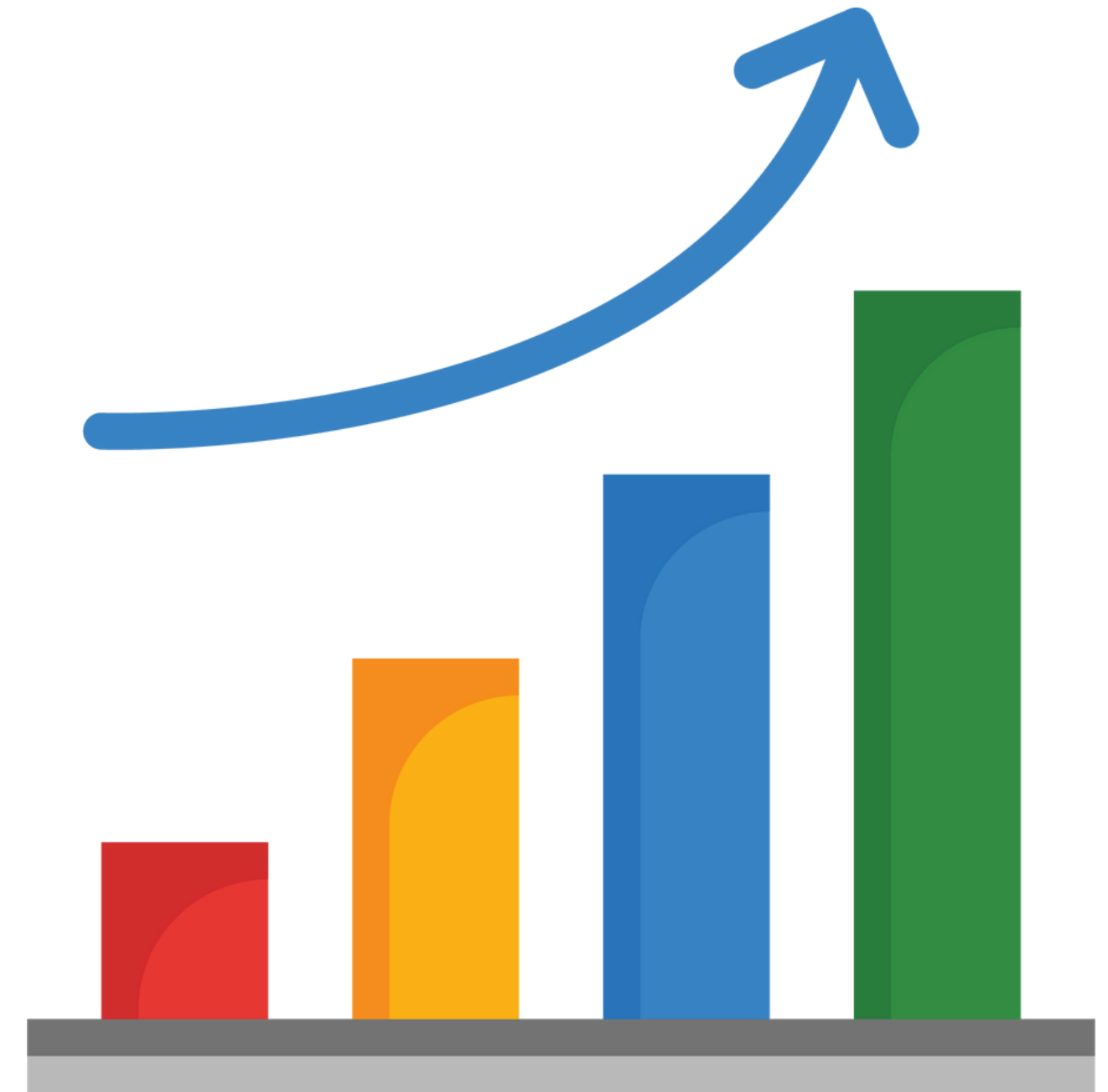




Analisis Penjualan dan Profitabilitas Terhadap Dataset Toko

Analisis saya dalam hal ini menghitung total penjualan setiap kategori dan sub-kategori produk guna mengidentifikasi kontribusi masing-masing terhadap total penjualan. Selain itu, dataset ini memungkinkan visualisasi tren penjualan dari waktu ke waktu untuk mendeteksi pola musiman dan fluktuasi permintaan. Terakhir, menghitung profit per produk bertujuan untuk menentukan produk yang paling menguntungkan, sehingga perusahaan dapat memfokuskan strategi bisnis pada produk yang memberikan margin laba tertinggi.



Dalam hal ini saya, mendapatkan dataset dari dengan judul dataset
“stores_sales_forecasting.csv”

kaggle

Dataset

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Postal Code	Region	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit
0	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-BO-10001798	Furniture	Bookcases	Bush Somerset Collection Bookcase	26196	2	0	419136
1	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	73194	3	0	219582
2	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	33311	South	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular Table	9575775	5	45	-383031
3	6	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	...	90032	West	FUR-FU-10001487	Furniture	Furnishings	Eldon Expressions Wood and Plastic Desk Access...	4886	7	0	141694
4	11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	...	90032	West	FUR-TA-10001539	Furniture	Tables	Chromcraft Rectangular Conference Tables	1706184	9	2	853092
5 rows × 21 columns																					

Dalam dataset terdapat 2121 baris dan 21 kolom

Data Info

Dari gambar disamping, memberikan informasi bahwa bahwa disetiap kolom yang ada memiliki tipe data yang berbeda - beda

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 2121 entries, 0 to 2120
```

```
Data columns (total 21 columns):
```

#	Column	Non-Null Count	Dtype
0	Row ID	2121 non-null	int64
1	Order ID	2121 non-null	object
2	Order Date	2121 non-null	object
3	Ship Date	2121 non-null	object
4	Ship Mode	2121 non-null	object
5	Customer ID	2121 non-null	object
6	Customer Name	2121 non-null	object
7	Segment	2121 non-null	object
8	Country	2121 non-null	object
9	City	2121 non-null	object
10	State	2121 non-null	object
11	Postal Code	2121 non-null	int64
12	Region	2121 non-null	object
13	Product ID	2121 non-null	object
14	Category	2121 non-null	object
15	Sub-Category	2121 non-null	object
16	Product Name	2121 non-null	object
17	Sales	2121 non-null	int64
18	Quantity	2121 non-null	int64
19	Discount	2121 non-null	int64
20	Profit	2121 non-null	int64

```
dtypes: int64(6), object(15)
```

```
memory usage: 348.1+ KB
```

- Menghapus kolom 'Row Id'

```
data = data.drop(columns='Row ID')
```

Menghapus kolom Row Id, kolom tersebut tidak terlalu dibutuhkan dalam analisis ini.

Missing Values

Dalam dataset tersebut tidak ditemukan data yang kosong, semua kolom terisi

```
data.isnull().sum()

Order ID      0
Order Date    0
Ship Date     0
Ship Mode     0
Customer ID   0
Customer Name 0
Segment       0
Country       0
City          0
State         0
Postal Code   0
Region        0
Product ID    0
Category      0
Sub-Category  0
Product Name  0
Sales         0
Quantity      0
Discount      0
Profit        0
dtype: int64
```

Data Duplicate

Dalam dataset ini, terdapat data yang duplikat berjumlah 1. Saya mengambil langkah dengan menghapus data tersebut dengan syantak disamping. Menurut saya apabila terdapat data yang duplikat pada dataset dapat mengganggu keakuratan data itu sendiri

```
data.duplicated().sum()
```

```
np.int64(1)
```

```
data = data.drop_duplicates()
```

```
data.duplicated().sum()
```

```
np.int64(0)
```


Mengubah Tipe Data

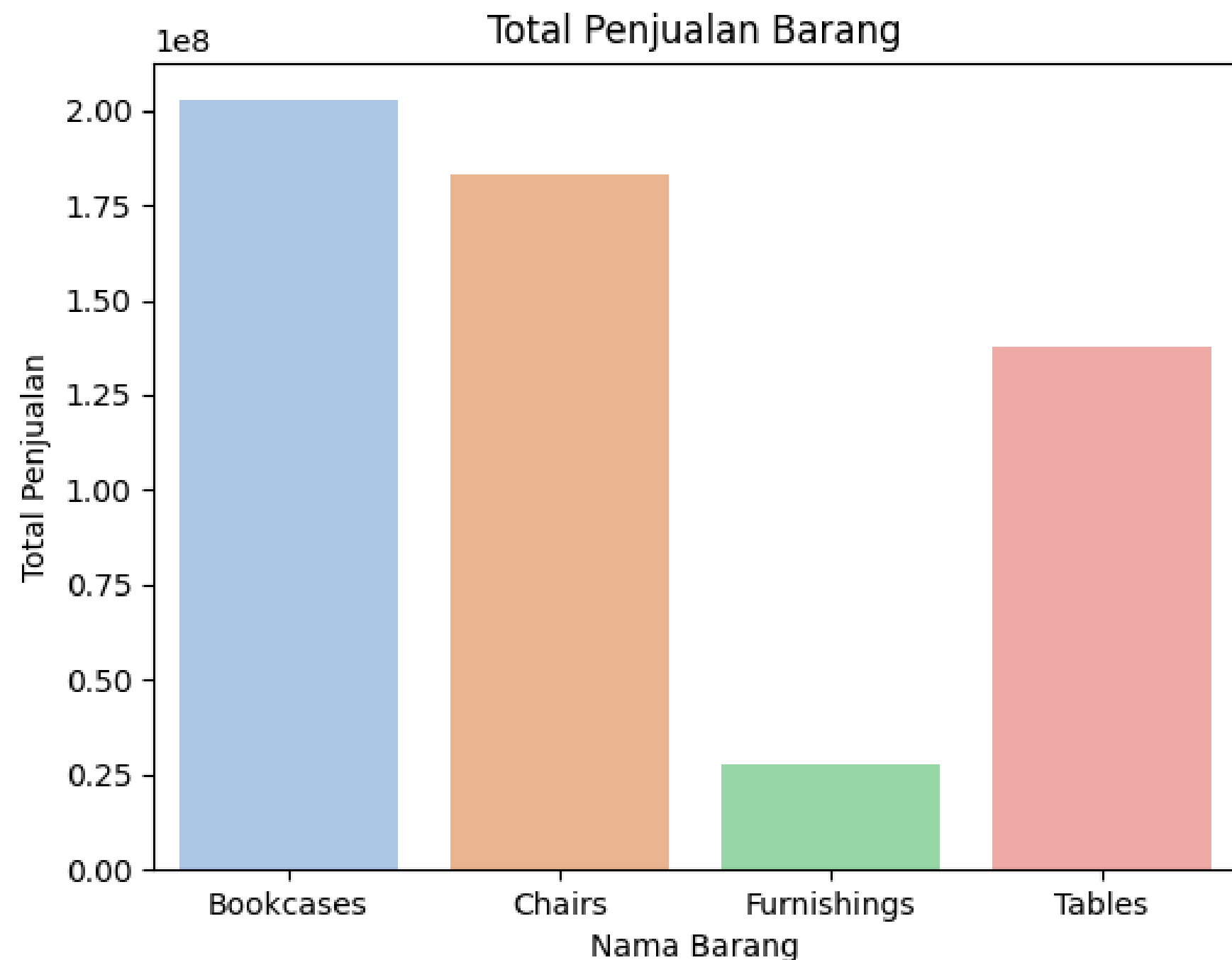
```
data['Order Date'] = pd.to_datetime(data['Order Date'])  
data['Ship Date'] = pd.to_datetime(data['Ship Date'])  
data['Sales'] = pd.to_numeric(data['Sales'])
```

Kolom 'Order Date' dan 'Ship Date' sebelumnya bertipe data object yang berisi tanggal pemesanan, kemudian saya mengubah tipe data tersebut menjadi datetime untuk memaksimalkan pemrosesan data

Kolom 'Sales' sebelumnya bertipe data object yang dimana isi kolom tersebut berupa angka-angka penjualan produk, kemudian saya mengubah tipe data kolom tersebut ke numerik supaya maksimal pemrosesan data

Data Visualization

Menghitung total penjualan untuk setiap kategori dan sub-kategori produk



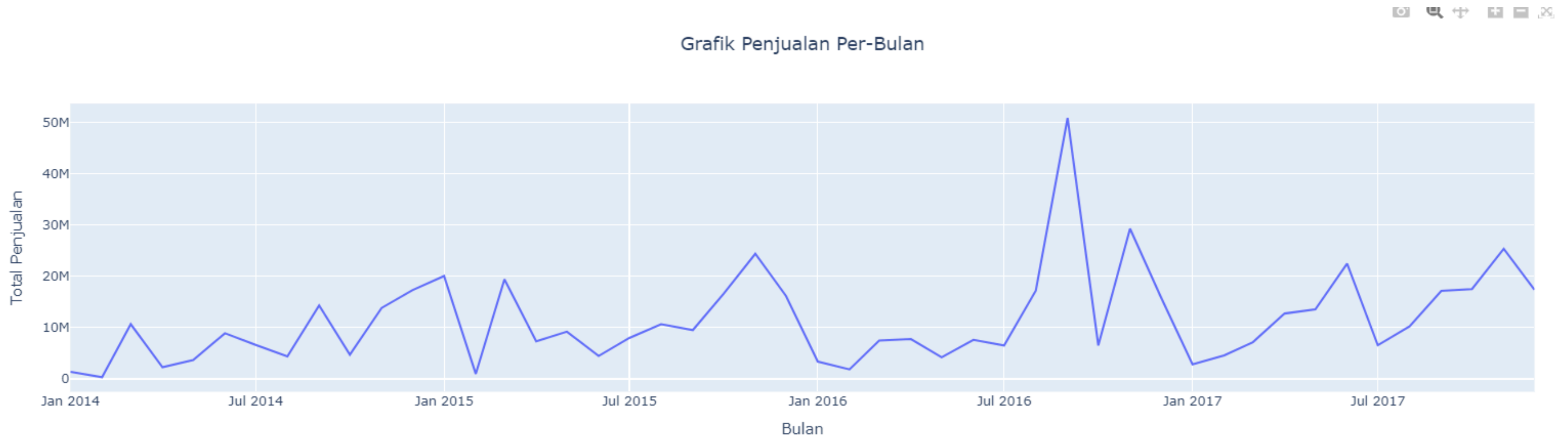
Pada grafik ini menunjukkan bahwa penjualan yang paling tinggi yaitu Bookcases dengan total penjualan yang mencapai sekitar 2 juta lebih, ini menjadikan bookcases adalah kategori barang yang paling laku.

Disusul dengan chairs dan tables yang memiliki penjualan yang lebih sedikit dibandingkan bookcases, tetapi lebih tinggi daripada furnishings.

Diantara semua kategori barang, ada barang yang memiliki penjualan barang yang paling terendah yaitu Furnishings. Dimana Furnishings ini memiliki penjualan sekitar 250.000 lebih dan ini membuktikan bahwa barang ini kurang diminati.

Data Visualization

Visualisasi tren penjualan dari waktu ke waktu-perbulan



Secara keseluruhan, grafik menunjukkan total penjualan dari Januari 2014 sampai Juli 2017, terjadi beberapa periode mengalami peningkatan maupun penurunan secara signifikan. Lonjakan penjualan yang paling mencolok terjadi pada tahun 2016 lebih spesifik pada September 2016, mungkin disebabkan dengan beberapa faktor tertentu.

Lonjakan penurunan yang paling mencolok setelah terjadi lonjakan penjualan yang paling mencolok terjadi pada bulan Oktober 2016. Mungkin disebabkan beberapa faktor tertentu. Yang menjadi catatan meningkatkan strategi penjualan untuk mengatasi penurunan pasca-lonjakan dan memastikan pertumbuhan penjualan konsisten

Hitung profit per produk untuk menentukan produk yang paling menguntungkan

```
profit_produk = data.groupby('Product Name')['Profit'].sum().reset_index().sort_values(by = 'Profit', ascending=False)  
profit_produk
```

✓ 0.0s

	Product Name	Profit
175	Global Deluxe High-Back Manager's Chair	15585910
217	Hon 4070 Series Pagoda Armless Upholstered Sta...	13886348
313	Office Star - Professional Matrix Back Chair w...	11792928
218	Hon 4070 Series Pagoda Round Back Stacking Chairs	10528144
263	Lifetime Advantage Folding Chairs, 4/Carton	9922640
...
211	Hon 2090 Pillow Soft Series Mid Back Swivel/...	-8246763
21	Balt Solid Wood Round Tables	-12010581
321	Riverside Palais Royal Lawyers Bookcase, Royal...	-15813591
42	Bush Advantage Collection Racetrack Conference...	-19343976
63	Chromcraft Bull-Nose Wood Oval Conference Tabl...	-28761156

380 rows × 2 columns

Produk yang paling menguntungkan adalah 'Global Deluxe High-Back Manager's Chair' memiliki profit tertinggi sebesar 15.585.910 ini menunjukkan produk yang paling menguntungkan di antara produk lainnya.

Produk 'Global Deluxe High-Back Manager's Chair' produk yang memiliki daya tarik pasar yang lebih besar atau biaya yang lebih rendah dibandingkan dengan yang lain. Produk 'Global Deluxe High-Back Manager's Chair' yang paling menguntungkan dapat membantu dalam merencanakan strategi penjualan dan pemasaran mungkin dengan cara meningkatkan kuantitas atau kualitas produk tersebut. Fokus pada produk dengan profit tinggi dapat meningkatkan keseluruhan profitabilitas bisnis.



Terima Kasih