

Analysis Penjualan dari tahun 2003 - 2005

Proyek ini berisi tentang Analisis data mengenai penjualan Mobil Klasik, Sepeda Motor, Pesawat, Kapal, Kereta Api, Truk, Bus dan Mobil Antik.

Data set didapat dari Kaagle: <https://www.kaggle.com/datasets/kyanyoga/sample-sales-data>

IMPORT MODUL

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

IMPORT DATASET

```
data = pd.read_csv("sales_data_sample.csv", encoding='latin-1')
print(data.head())
```

```
➡
```

	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	\
0	10107	30	95.70	2	2871.00	
1	10121	34	81.35	5	2765.90	
2	10134	41	94.74	2	3884.34	
3	10145	45	83.26	6	3746.70	
4	10159	49	100.00	14	5205.27	

	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	...	\
0	2/24/2003 0:00	Shipped	1	2	2003	...	
1	5/7/2003 0:00	Shipped	2	5	2003	...	
2	7/1/2003 0:00	Shipped	3	7	2003	...	
3	8/25/2003 0:00	Shipped	3	8	2003	...	
4	10/10/2003 0:00	Shipped	4	10	2003	...	

	ADDRESSLINE1	ADDRESSLINE2	CITY	STATE	\
0	897 Long Airport Avenue	NaN	NYC	NY	

2818	False	False	False	False	False	...	False
2819	False	False	False	False	False	...	False
2820	False	False	False	False	False	...	False
2821	False	False	False	False	False	...	False
2822	False	False	False	False	False	...	False


	ADDRESSLINE2	CITY	STATE	POSTALCODE	COUNTRY	TERRITORY	\
0	True	False	False	False	False	True	
1	True	False	True	False	False	False	
2	True	False	True	False	False	False	
3	True	False	False	False	False	True	
4	True	False	False	True	False	True	
...	
2818	True	False	True	False	False	False	
2819	True	False	True	False	False	False	
2820	True	False	True	False	False	False	
2821	True	False	True	False	False	False	
2822	True	False	False	False	False	True	

	CONTACTLASTNAME	CONTACTFIRSTNAME	DEALSIZE
0	False	False	False
1	False	False	False
2	False	False	False
3	False	False	False
4	False	False	False
...
2818	False	False	False
2819	False	False	False
2820	False	False	False
2821	False	False	False
2822	False	False	False

[2823 rows x 25 columns]

MENGHITUNG JUMLAH KOSONG DISETIAP KOLOM DATA

```
missing_values_per_column = data.isnull().sum()
print(missing_values_per_column)
```



```
ORDERNUMBER      0
QUANTITYORDERED  0
PRICEEACH         0
ORDERLINENUMBER  0
SALES             0
ORDERDATE        0
STATUS           0
QTR_ID           0
MONTH_ID         0
YEAR_ID          0
PRODUCTLINE      0
MSRP             0
PRODUCTCODE      0
CUSTOMERNAME     0
PHONE            0
ADDRESSLINE1     0
ADDRESSLINE2     2521
CITY             0
STATE            1486
POSTALCODE       76
COUNTRY          0
TERRITORY        1074
CONTACTLASTNAME  0
CONTACTFIRSTNAME 0
DEALSIZE         0
dtype: int64
```

MENGHAPUS BARIS DATA KOSONG DAN MENAMPILKAN DATA YANG SUDAH BERSIH

```
data_cleaned_columns = data.dropna(axis=1)
print(data_cleaned_columns)
```



	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	...	MSRP	\
0	2/24/2003 0:00	Shipped	1	2	2003	...	95	
1	5/7/2003 0:00	Shipped	2	5	2003	...	95	
2	7/1/2003 0:00	Shipped	3	7	2003	...	95	
3	8/25/2003 0:00	Shipped	3	8	2003	...	95	
4	10/10/2003 0:00	Shipped	4	10	2003	...	95	
...	
2818	12/2/2004 0:00	Shipped	4	12	2004	...	54	
2819	1/31/2005 0:00	Shipped	1	1	2005	...	54	
2820	3/1/2005 0:00	Resolved	1	3	2005	...	54	
2821	3/28/2005 0:00	Shipped	1	3	2005	...	54	
2822	5/6/2005 0:00	On Hold	2	5	2005	...	54	

	PRODUCTCODE	CUSTOMERNAME	PHONE	\
0	S10_1678	Land of Toys Inc.	2125557818	
1	S10_1678	Reims Collectables	26.47.1555	
2	S10_1678	Lyon Souveniers	+33 1 46 62 7555	
3	S10_1678	Toys4GrownUps.com	6265557265	
4	S10_1678	Corporate Gift Ideas Co.	6505551386	
...	
2818	S72_3212	Euro Shopping Channel	(91) 555 94 44	
2819	S72_3212	Oulu Toy Supplies, Inc.	981-443655	
2820	S72_3212	Euro Shopping Channel	(91) 555 94 44	
2821	S72_3212	Alpha Cognac	61.77.6555	
2822	S72_3212	Gifts4AllAges.com	6175559555	

	ADDRESSLINE1	CITY	COUNTRY	CONTACTLASTNAME	\
0	897 Long Airport Avenue	NYC	USA	Yu	
1	59 rue de l'Abbaye	Reims	France	Henriot	
2	27 rue du Colonel Pierre Avia	Paris	France	Da Cunha	
3	78934 Hillside Dr.	Pasadena	USA	Young	
4	7734 Strong St.	San Francisco	USA	Brown	
...	

```

1      Paul      Small
2      Daniel    Medium
3      Julie     Medium
4      Julie     Medium
...
2818    Diego     Small
2819    Pirkko    Medium
2820    Diego     Medium
2821    Annette   Small
2822     Juri     Medium

```

```
[2822 rows x 31 columns]
```

RATA RATA HARGA PER PRODUK

```

avg_price_per_product = data.groupby('PRODUCTLINE')['PRICEEACH'].mean()
print(avg_price_per_product)

```

```

➡ PRODUCTLINE
Classic Cars      87.335781
Motorcycles       82.997553
Planes            81.740915
Ships             83.855470
Trains            75.654675
Trucks and Buses  87.527940
Vintage Cars      78.148204
Name: PRICEEACH, dtype: float64

```

TOTAL PENJUALAN PER PRODUK

```

total_sales_per_category = data.groupby('PRODUCTLINE')['PRICEEACH'].sum()
print(total_sales_per_category)

```

```

➡ PRODUCTLINE
Classic Cars      84453.70
Motorcycles       27472.19
Planes            25012.72

```

```
Ships          19622.18
Trains         5825.41
Trucks and Buses 26345.91
Vintage Cars   47435.96
Name: PRICEEACH, dtype: float64
```

TOTAL PENJUALAN PER BULAN TAHUN 2003

```
data['Month'] = data['ORDERDATE'].dt.month
data_2003 = data[data['ORDERDATE'].dt.year == 2003]
total_sales_per_month_2003 = data_2003.groupby('Month')['PRICEEACH'].sum()

print(total_sales_per_month_2003)
```

```
➡ Month
1      3203.31
2      3441.55
3      4086.67
4      4925.11
5      4753.17
6      3879.33
7      4310.41
8      4826.26
9      6407.33
10     13367.06
11     24717.89
12      5875.05
Name: PRICEEACH, dtype: float64
```

TOTAL PENJUALAN PER BULAN TAHUN 2004

```
data['Month'] = data['ORDERDATE'].dt.month
data_2004 = data[data['ORDERDATE'].dt.year == 2004]
total_sales_per_month_2004 = data_2004.groupby('Month')['PRICEEACH'].sum()

print(total_sales_per_month_2004)
```

```
➡ Month
1      7576.40
2      7222.72
3      4686.99
4      5281.88
5      6282.94
6      6994.18
7      7795.36
8     11147.85
9      7870.55
10     13246.16
11     25272.54
12      9110.53
Name: PRICEEACH, dtype: float64
```

TOTAL PENJUALAN PER BULAN TAHUN 2005

```
data['Month'] = data['ORDERDATE'].dt.month
data_2005 = data[data['ORDERDATE'].dt.year == 2005]
total_sales_per_month_2005 = data_2005.groupby('Month')['PRICEEACH'].sum()

print(total_sales_per_month_2005)
```

```
➡ Month
1      8332.07
2      8238.22
3      8658.95
4      4556.29
5     10101.30
Name: PRICEEACH, dtype: float64
```


TOTAL PENJUALAN PER BULAN DARI TAHUN 2003 - 2005

```
data['Month'] = data['ORDERDATE'].dt.month
total_sales_per_month = data.groupby('Month')['PRICEEACH'].sum()
print(total_sales_per_month)
```

```
Month
1      19111.78
2      18902.49
3      17432.61
4      14763.28
5      21137.41
6      10873.51
7      12105.77
8      15974.11
9      14277.88
10     26613.22
11     49990.43
12     14985.58
Name: PRICEEACH, dtype: float64
```

TOTAL PENJUALAN PER TAHUN

```
data['Year'] = data['ORDERDATE'].dt.year
total_sales_per_year = data.groupby('Year')['PRICEEACH'].sum()
print(total_sales_per_year)
```

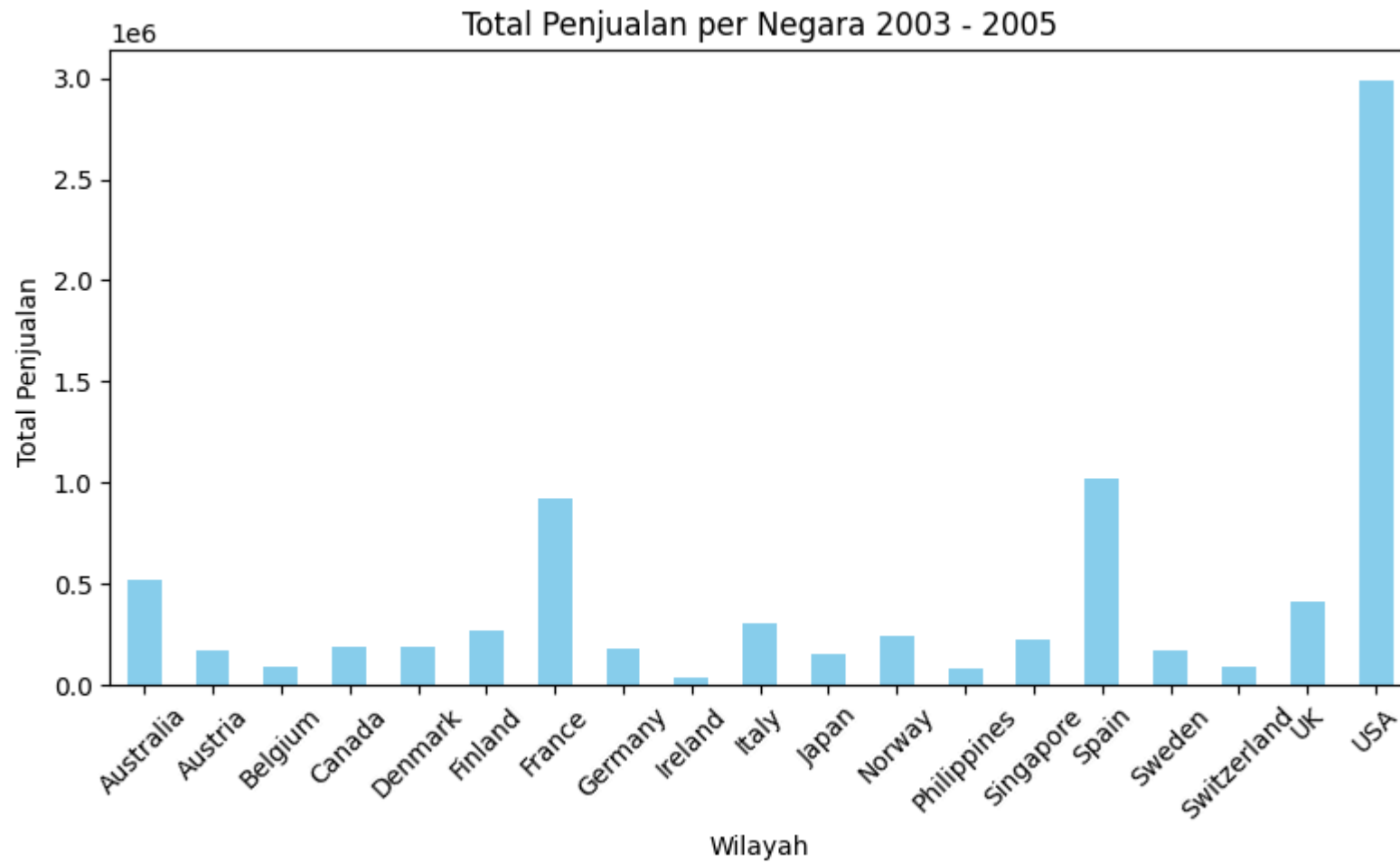
```
Year
2003     83793.14
2004    112488.10
2005     39886.83
Name: PRICEEACH, dtype: float64
```

TOTAL PENJUALAN PER NEGARA

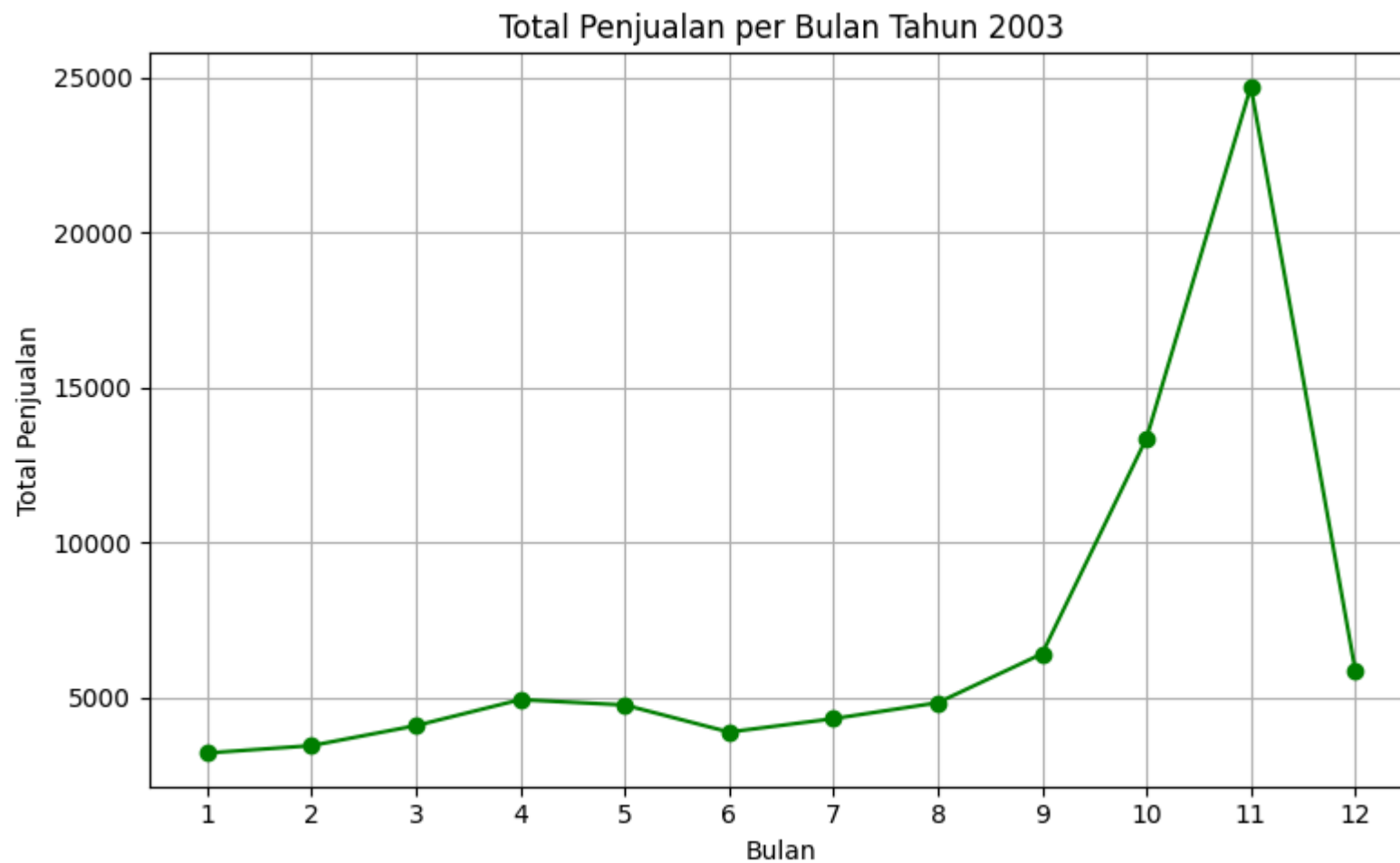
```
total_sales_per_country = data.groupby('COUNTRY')['SALES'].sum()
print(total_sales_per_country)
```

```
⇒ COUNTRY
Australia      521598.46
Austria         172793.05
Belgium         94528.88
Canada         193504.34
Denmark        192747.63
Finland        268714.70
France         919257.85
Germany        178689.08
Ireland         43237.24
Italy          309402.87
Japan          153076.69
Norway         246115.80
Philippines     80291.17
Singapore      227985.50
Spain          1021705.97
Sweden         174264.10
Switzerland     93344.91
UK             413203.34
USA            2986425.21
Name: SALES, dtype: float64
```

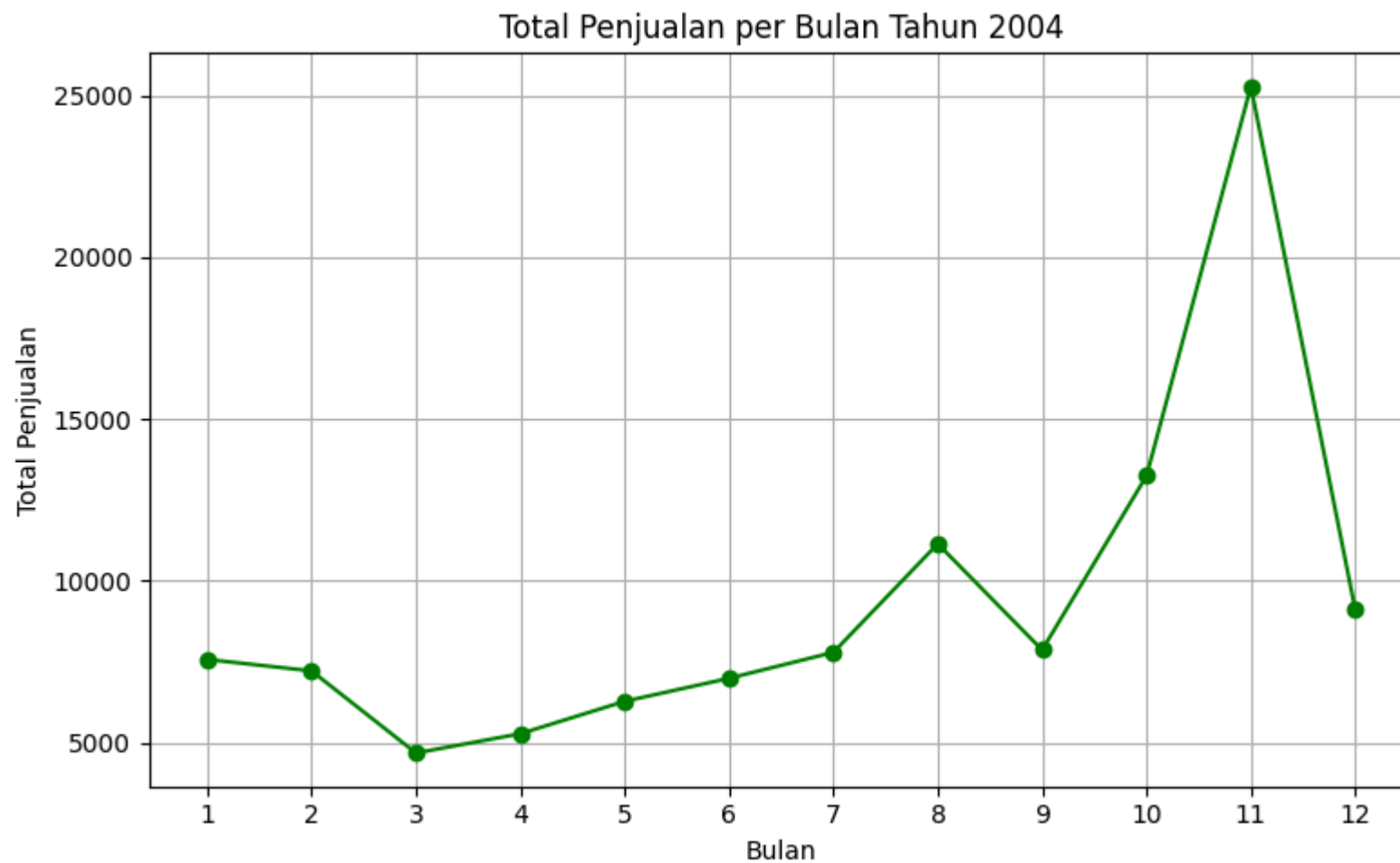
```
plt.figure(figsize=(8, 5))
total_sales_per_country.plot(kind='bar', color='skyblue')
plt.title('Total Penjualan per Negara 2003 - 2005')
plt.xlabel('Wilayah')
plt.ylabel('Total Penjualan ')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



```
plt.figure(figsize=(8, 5))
total_sales_per_month_2003.plot(kind='line', marker='o', color='green')
plt.title('Total Penjualan per Bulan Tahun 2003')
plt.xlabel('Bulan')
plt.ylabel('Total Penjualan')
plt.xticks(np.arange(1, 13, step=1))
plt.grid(True)
plt.tight_layout()
plt.show()
```



```
plt.figure(figsize=(8, 5))
total_sales_per_month_2004.plot(kind='line', marker='o', color='green')
plt.title('Total Penjualan per Bulan Tahun 2004')
plt.xlabel('Bulan')
plt.ylabel('Total Penjualan')
plt.xticks(np.arange(1, 13, step=1))
plt.grid(True)
plt.tight_layout()
plt.show()
```



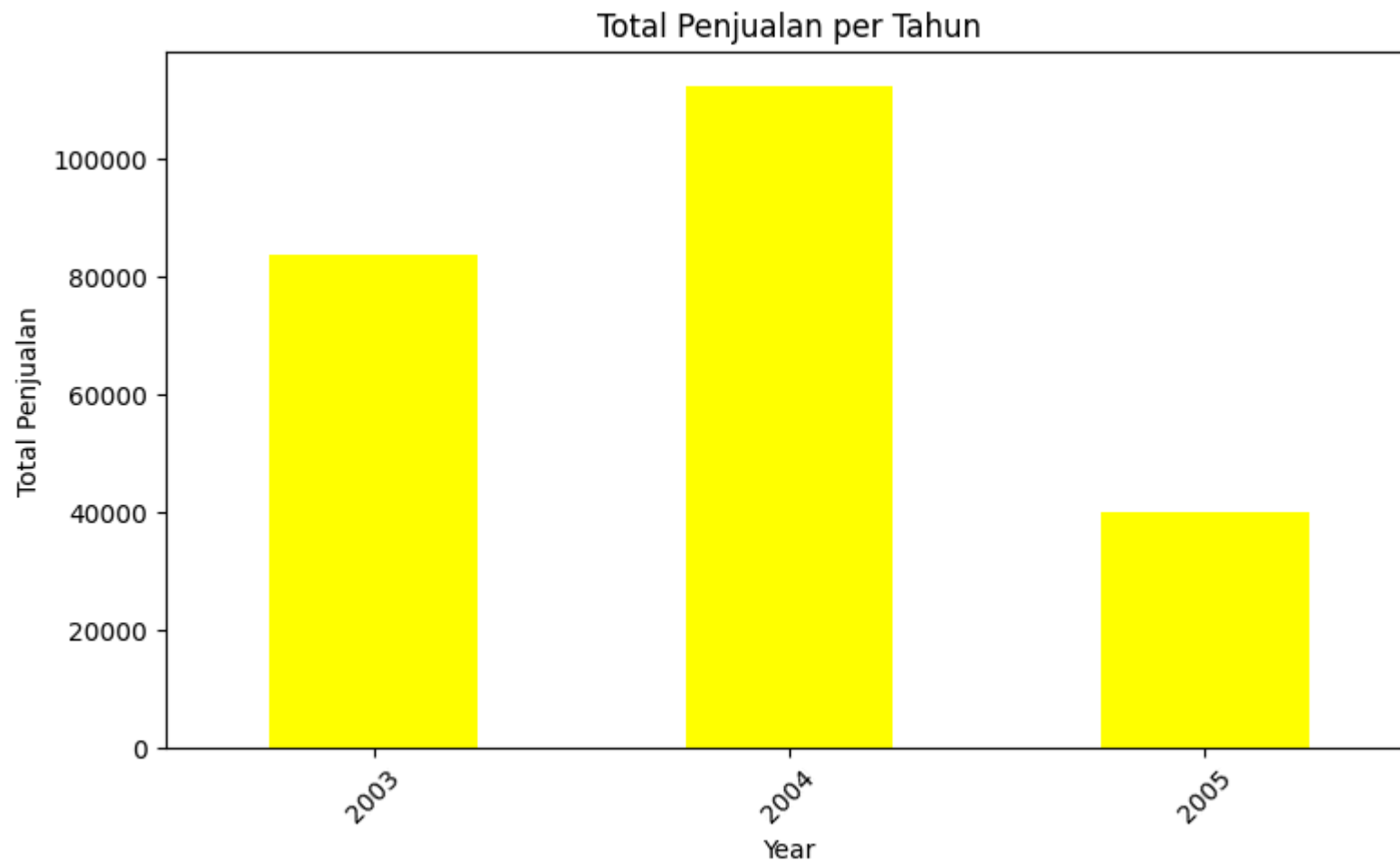
```
plt.figure(figsize=(8, 5))
total_sales_per_month_2005.plot(kind='line', marker='o', color='green')
plt.title('Total Penjualan per Bulan Tahun 2005')
plt.xlabel('Bulan')
plt.ylabel('Total Penjualan')
plt.xticks(np.arange(1, 13, step=1))
plt.grid(True)
plt.tight_layout()
plt.show()
```



```
plt.figure(figsize=(8, 5))
total_sales_per_month.plot(kind='line', marker='o', color='green')
plt.title('Total Penjualan per Bulan dari Tahun 2003 - 2005')
plt.xlabel('Bulan')
plt.ylabel('Total Penjualan')
plt.xticks(np.arange(1, 13, step=1))
plt.grid(True)
plt.tight_layout()
plt.show()
```



```
plt.figure(figsize=(8, 5))
total_sales_per_year.plot(kind='bar', color='yellow')
plt.title('Total Penjualan per Tahun')
plt.ylabel('Total Penjualan')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



KESIMPULAN

Penurunan di Awal Tahun Penjualan cenderung lebih rendah di awal tahun, terutama di bulan Januari dan Februari.

Musim Puncak Penjualan Bulan November menunjukkan penjualan tertinggi selama dua tahun berturut turut 2003 dan 2004. Ini mungkin menandakan adanya musim puncak penjualan seperti peningkatan permintaan dari konsumen

Penjualan awal tahun 2005 Januari hingga Mei menunjukkan peningkatan yang signifikan dibandingkan dengan tahun-tahun sebelumnya. Karena data hanya tersedia hingga Mei 2005, sulit untuk membuat kesimpulan yang lengkap mengenai tren penjualan di seluruh tahun 2005. Namun tren kenaikan di awal tahun menunjukkan bahwa penjualan tahun 2005 mungkin lebih tinggi secara keseluruhan jika dibandingkan dengan 2003 dan 2004.