churn

July 15, 2025

0.1 Import Library + Load Data

[11]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[12]:
```
df = pd.read_csv('~/Untitled Folder/Portfolio-Magang/dataset.csv')
```

[13]:
```
df.head()
```

[13]:

	Nama	Kota	Paket	Tgl_beli	Durasi_Bulan	Frekuensi_Topup	Kuota_Bulan_GB
0	John	Surabaya	Freedom Combo	2024-10-08	9	5	16
1	Willie	Jakarta	Freedom Combo	2025-03-15	4	1	7
2	Samuel	Medan	Freedom Combo	2024-07-22	12	2	20
3	Chris	Semarang	Unlimited 2GB	2025-01-11	2	3	5
4	Andrew	Surabaya	Yellow	2025-03-30	12	1	1

0.2 Hitung Churn Rate Customer

[14]:
```
# mengetahui churn rate
# Frekuensi_Topup  1 dalam 3+ bulan → dianggap churn
df['Churn'] = ((df['Frekuensi_Topup'] <= 1) & (df['Durasi_Bulan'] >=
3))
# hitung churn
churn_rate = df['Churn'].sum() / len(df) * 100
print(f'Hasil Persentase Churn Rate Customer: {churn_rate:.2f}%')
```

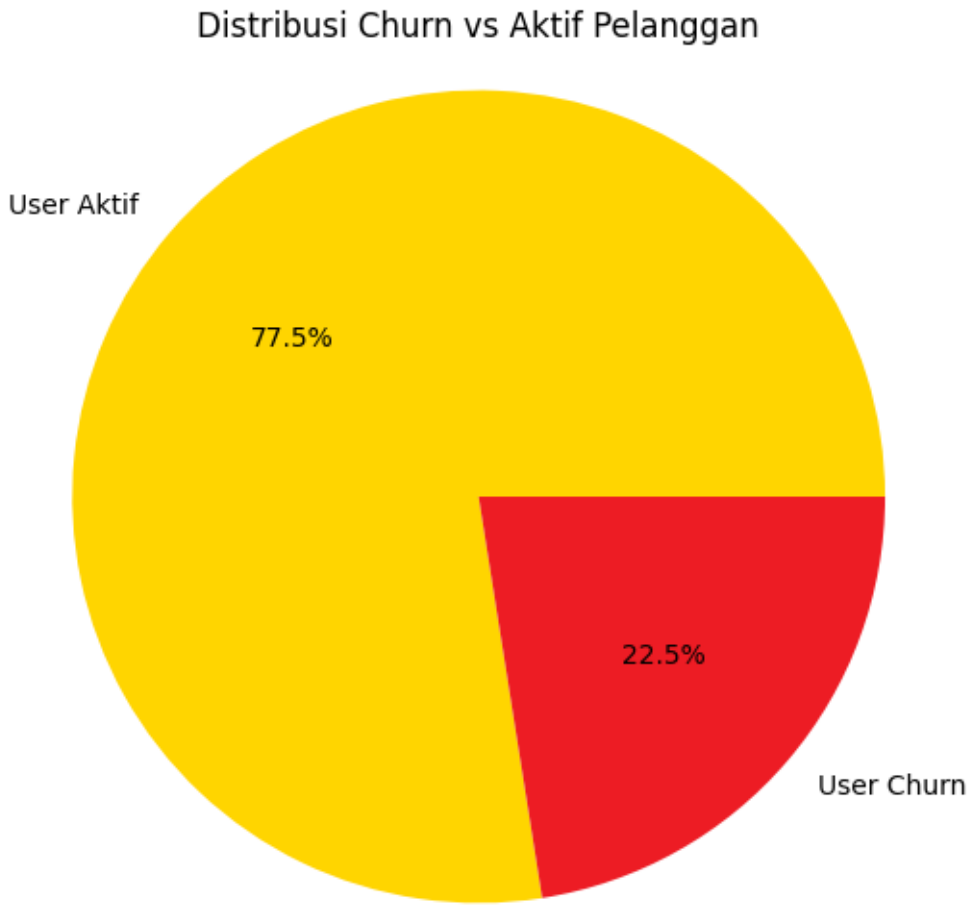
```
Hasil Persentase Churn Rate Customer: 22.50%
```

0.3 Tampilkan Dalam Grafik

[15]:
```
# Buat pie chart churn vs aktif
churn_counts = df['Churn'].value_counts()
labels = ['User Aktif', 'User Churn']
sizes = [churn_counts[0], churn_counts[1]]
colors = ['#FFD500', '#ED1C24']
plt.figure(figsize=(6, 6))
plt.pie(sizes, labels=labels, autopct='%1.1f%%',
colors=colors)
plt.title('Distribusi Churn vs Aktif Pelanggan')
plt.axis('equal')
```

```
/tmp/ipykernel_75737/1900604499.py:4: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  sizes = [churn_counts[0], churn_counts[1]]
```

[15]:
```
(np.float64(-1.0999999318335567),
 np.float64(1.0999999967539789),
 np.float64(-1.0999998974046947),
 np.float64(1.0999999790110455))
```


Distribusi Churn vs Aktif Pelanggan
User Aktif
77.5%
22.5%
User Churn

[ ]: