Untitled-1

```
# Telco Churn Analysis Whole Code %%
!pip install pandas


# %%
!pip install matplotlib
!pip install seaborn
!pip install numpy

# %%
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
df = pd.read_csv("customer_churn.csv")
df.head()

# %%
df.info()

# %%
df['TotalCharges'] = df['TotalCharges'].replace(" ","0")
df['TotalCharges'] = df['TotalCharges'].astype("float")

# %%
df.info()

# %%
df.isnull().sum().sum()

# %%
df.describe()

# %%
df.duplicated().sum()

# %%
df["customerID"].duplicated().sum()

# %%
def conv(value):
    if value == 1:
        return "yes"
    else:
        return "no"
df["SeniorCitizen"]=df["SeniorCitizen"].apply(conv)

# %%
```

```
df.tail()

# %%


ax =sns.countplot(x="Churn", data=df)
ax.bar_label(ax.containers[0])
plt.show()

# %%
gb=df.groupby("Churn").agg({"Churn":"count"})
plt.figure(figsize=(3,3))
plt.pie(gb["Churn"], labels=gb.index, autopct ="%1.2f%%")
plt.title("Percentage of Churn Customers", fontsize=10)

plt.show()


# %%
plt.figure(figsize=(4,4))
sns.countplot(x="gender", data=df , hue="Churn")
plt.title("churn By Gender")
plt.show()

# %%
plt.figure(figsize=(4,4))
sns.countplot(x="SeniorCitizen", data=df , hue="Churn")
plt.title("churn By SeniorCitizen")
plt.show()

# %% [markdown]
# #on the basis of senior citizen there is more churned out
#

# %%
sns.histplot(x="tenure", data=df, hue="Churn", bins=72)
plt.show()

# %% [markdown]
# #people who have used our services fro long time stayed while people who have used our
services for one-two months have churned out

# %%
plt.figure(figsize=(4,4))
ax=sns.countplot(x="Contract",data=df,hue="Churn")
ax.bar_label(ax.containers[0])
plt.title("Count of customers by Contract")
plt.show()

# %% [markdown]
```

```
# #people who have month to month contract are likely to churn then who have one or two month of contract

# %%
df.columns.values

# %%
# import pandas as pd
# import seaborn as sns
# import matplotlib.pyplot as plt

cols = ['PhoneService', 'MultipleLines', 'InternetService',
        'OnlineSecurity', 'OnlineBackup', 'DeviceProtection',
        'TechSupport', 'StreamingTV', 'StreamingMovies'
        ]

n_cols = 3
n_rows = int(len(cols) / n_cols) + (len(cols) % n_cols > 0)

fig, axs = plt.subplots(n_rows, n_cols, figsize=(5 * n_cols, 4 * n_rows))
axs = axs.flatten()

for i, col in enumerate(cols):
    sns.countplot(data=df, x=col, ax=axs[i], hue="Churn")
    axs[i].set_title(col)
    axs[i].set_xlabel('')
    axs[i].set_ylabel('Count')
    axs[i].tick_params(axis='x')

# Hide any unused subplots
for j in range(i + 1, len(axs)):
    axs[j].axis('off')

plt.tight_layout()
plt.show()


# %% [markdown]
# #People using fiber-optic internet were more likely to leave than those using DSL or no internet.
#
# People who had extra services — like online security, backup, device protection, or tech support — stayed more often.
#
# Having fewer or no of those extra services was linked with higher chance of leaving.
#
# Simple takeaway: More services (especially protection/support extras) → less churn. Fiber-optic internet alone (maybe because it’s more expensive) → more churn.

# %%
plt.figure(figsize=(6,6))
```

```
ax=sns.countplot(x="PaymentMethod",data=df,hue="Churn")
ax.bar_label(ax.containers[0])
plt.xticks(rotation=45)
plt.title("Churn By Payment Method")
plt.show()

# %% [markdown]
# #customer is likely to churn when they are using Electronic check as a Payment method

# %% [markdown]
#



```