```
from google.colab import files
uploaded = files.upload()
```

```
Choose Files titanic.csv.csv
titanic.csv.csv(text/csv) - 61194 bytes, last modified: 11/1/2025 - 100% done
Saving titanic.csv.csv to titanic.csv (2).csv
```

```
import os
os.listdir()
```

```
['.config',
 'titanic.csv.csv',
 'titanic.csv (1).csv',
 'Titanic (1).csv',
 'Titanic.csv',
 'titanic.csv (2).csv',
 'sample_data']
```

```
import pandas as pd

titanic = pd.read_csv("Titanic.csv")
titanic.head()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticke
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 2117
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs	female	38.0	1	0	PC 1759

Next steps: Generate code with titanic | New interactive sheet

```
# 📘 Step 1: Import libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Display settings
pd.set_option('display.max_columns', None)
sns.set(style="whitegrid")

# 📊 Step 2: Load data (already loaded)
# titanic = pd.read_csv("Titanic.csv")   # Skip this line if already loaded

# 📈 Step 3: Basic Info
print("🔹 Dataset Information:")
print(titanic.info())
```

```
print("\n 🔹 Summary Statistics:")
print(titanic.describe())

print("\n 🔹 Missing Values:")
print(titanic.isnull().sum())

# 🧮 Step 4: Value counts
print("\n 🔹 Survived Count:")
print(titanic['Survived'].value_counts())

print("\n 🔹 Passenger Class Count:")
print(titanic['Pclass'].value_counts())

# 🎨 Step 5: Visualizations
plt.figure(figsize=(6,4))
sns.countplot(x='Survived', data=titanic, palette='pastel')
plt.title("Survival Count")
plt.show()

plt.figure(figsize=(6,4))
sns.countplot(x='Pclass', hue='Survived', data=titanic, palette='Set2')
plt.title("Survival by Passenger Class")
plt.show()

plt.figure(figsize=(6,4))
sns.histplot(titanic['Age'], bins=30, kde=True, color='skyblue')
plt.title("Age Distribution")
plt.show()

plt.figure(figsize=(6,4))
sns.boxplot(x='Survived', y='Age', data=titanic, palette='muted')
plt.title("Age vs Survival")
plt.show()

plt.figure(figsize=(6,4))
sns.barplot(x='Sex', y='Survived', data=titanic, palette='cool')
plt.title("Survival Rate by Gender")
plt.show()

# 💡 Step 6: Correlation Heatmap
plt.figure(figsize=(8,6))
sns.heatmap(titanic.corr(numeric_only=True), annot=True, cmap='Blues')
plt.title("Correlation Heatmap")
plt.show()

# 📝 Step 7: Observations
print("\n✅ Observations:")
print("""
1. More passengers did not survive (0) than those who did (1).
2. Passengers in 1st class had a higher survival rate than those in 3rd class
3. Females had a higher chance of survival compared to males.
4. Younger passengers were more likely t
```

```
  File "/tmp/ipython-input-732729534.py", line 64
    print("""
          ^
SyntaxError: incomplete input
```

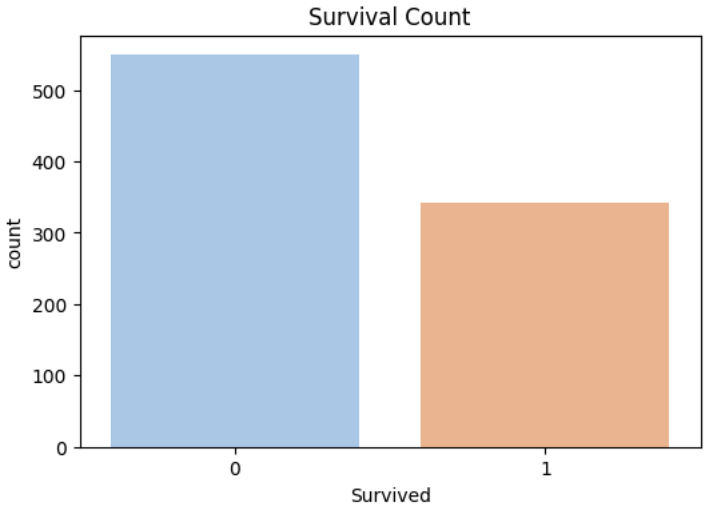
Next steps: Explain error

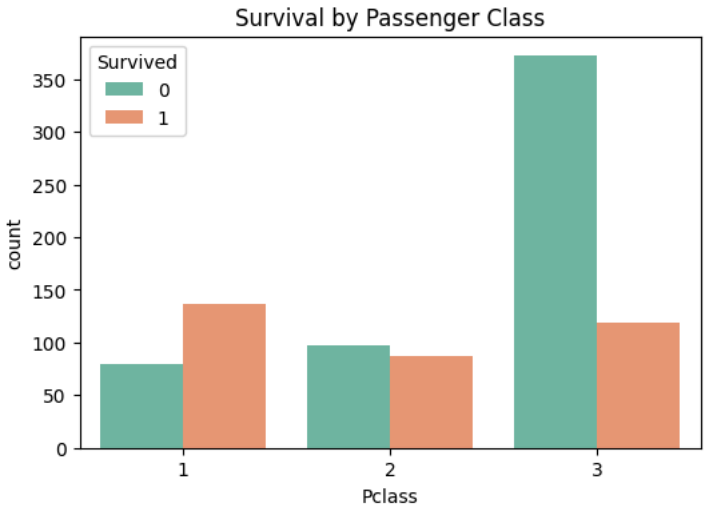
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# STEP 2 – Load the Dataset
titanic = pd.read_csv("Titanic.csv")

# STEP 3 – Basic Info
print("🔹 Dataset Information:")
print(titanic.info())

print("\n🔹 Summary Statistics:")
print(titanic.describe())

print("\n🔹 Missing Values:")
print(titanic.isnull().sum())

# STEP 4 – Value Counts
print("\n🔹 Survived Count:")
print(titanic['Survived'].value_counts())

print("\n🔹 Passenger Class Count:")
print(titanic['Pclass'].value_counts())

# STEP 5 – Visualizations
plt.figure(figsize=(6,4))
sns.countplot(x='Survived', data=titanic, palette='pastel')
plt.title("Survival Count")
plt.show()

plt.figure(figsize=(6,4))
sns.countplot(x='Pclass', hue='Survived', data=titanic, palette='Set2')
plt.title("Survival by Passenger Class")
plt.show()

plt.figure(figsize=(6,4))
sns.histplot(titanic['Age'].dropna(), bins=30, kde=True, color='skyblue')
plt.title("Age Distribution")
plt.show()

plt.figure(figsize=(6,4))
sns.boxplot(x='Survived', y='Age', data=titanic, palette='muted')
plt.title("Age vs Survival")
plt.show()

plt.figure(figsize=(6,4))
sns.barplot(x='Sex', y='Survived', data=titanic, palette='cool')
plt.title("Survival Rate by Gender")
```

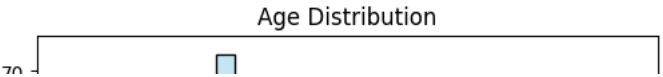
```
plt.title("Survival Rate by Gender")
plt.show()

# STEP 6 — Correlation Heatmap
plt.figure(figsize=(8,6))
sns.heatmap(titanic.corr(numeric_only=True), annot=True, cmap='Blues')
plt.title("Correlation Heatmap")
plt.show()

# STEP 7 — Observations
print("\n✅ Observations:")
print("""
1. More passengers did not survive (0) than those who did (1).
2. Passengers in 1st class had higher survival than 3rd class.
3. Females had a higher survival chance than males.
4. Younger passengers tended to survive more.
5. 'Cabin' and 'Age' columns have missing values.
""")

# STEP 8 — Summary
print("\n📋 Summary:")
print("""
The Titanic dataset shows that passenger class, gender, and age
played major roles in determining survival chances.
First-class passengers and females were prioritized during rescue,
while many third-class male passengers did not survive.
""")
```

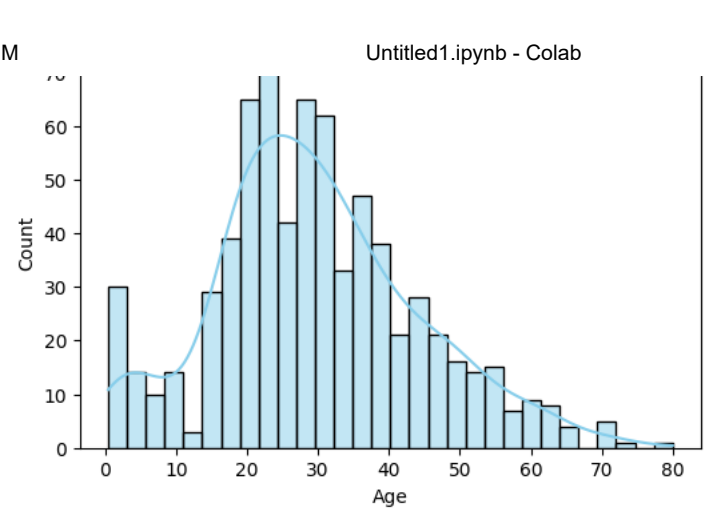
```
🔹 Dataset Information:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  891 non-null    int64
 1   Survived     891 non-null    int64
 2   Pclass       891 non-null    int64
 3   Name         891 non-null    object
 4   Sex          891 non-null    object
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64
 7   Parch        891 non-null    int64
 8   Ticket       891 non-null    object
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object
 11  Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
None

🔹 Summary Statistics:
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000
mean    446.000000    0.383838    2.308642   29.699118    0.523008
std     257.353842    0.486592    0.836071   14.526497    1.102743
min       1.000000    0.000000    1.000000    0.420000    0.000000
25%     223.500000    0.000000    2.000000   20.125000    0.000000
50%     446.000000    0.000000    3.000000   28.000000    0.000000
75%     668.500000    1.000000    3.000000   38.000000    1.000000
max     891.000000    1.000000    3.000000   80.000000    8.000000

            Parch        Fare
count  891.000000  891.000000
mean     0.381594   32.204208
std      0.806057   49.693429
min      0.000000    0.000000
25%      0.000000    7.910400
50%      0.000000   14.454200
75%      0.000000   31.000000
max      6.000000  512.329200

🔹 Missing Values:
PassengerId      0
Survived         0
Pclass           0
Name             0
Sex              0
Age            177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          687
Embarked         2
dtype: int64

🔹 Survived Count:
Survived
0    549
1    342
```

```
Name: count, dtype: int64

🔹 Passenger Class Count:
Pclass
3    491
1    216
2    184
Name: count, dtype: int64
/tmp/ipython-input-1097384752.py:27: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in

  sns.countplot(x='Survived', data=titanic, palette='pastel')
```

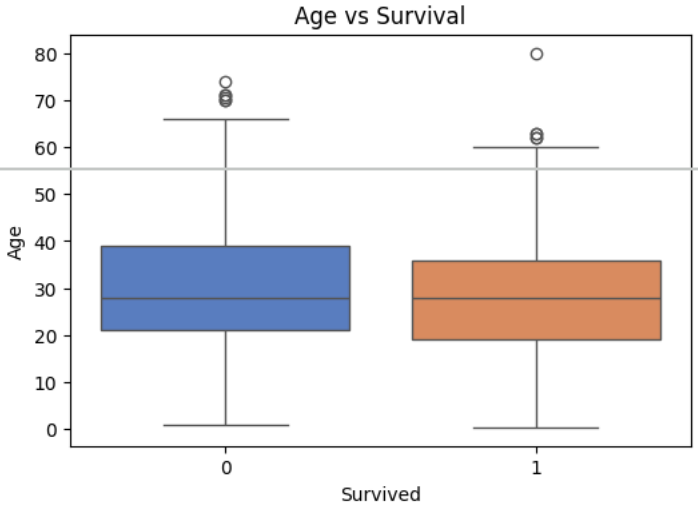









```
/tmp/ipython-input-1097384752.py:42: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in

  sns.boxplot(x='Survived', y='Age', data=titanic, palette='muted')
```




```
/tmp/ipython-input-1097384752.py:47: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in

  sns.barplot(x='Sex', y='Survived', data=titanic, palette='cool')
```