

Cell 1: Import and Load Dataset

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# Load the dataset
df = pd.read_csv('train.csv') # Make sure train.csv is in the same folder
df.head()
```

Cell 2: Dataset Info

```
# Dataset Information
df.info()
df.describe(include='all')
```

Cell 3: Missing Values

```
# Check for missing values
df.isnull().sum()
```

Cell 4: Survival Count Plot

```
# Plot survival count
sns.countplot(data=df, x='Survived')
plt.title("Survival Count")
plt.show()
```

Cell 5: Survival by Sex

```
# Survival by Sex
sns.countplot(data=df, x='Sex', hue='Survived')
plt.title("Survival by Sex")
plt.show()
```

Cell 6: Age Distribution

```
# Age Distribution
sns.histplot(df['Age'], bins=10, kde=True)
plt.title("Age Distribution")
plt.show()
```

Cell 7: Correlation Heatmap

```
# Correlation Heatmap
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title("Correlation Heatmap")
plt.show()
```

Cell 8: Summary (Markdown)

Summary

- Females had a higher survival rate.
- Most passengers were young adults.
- Some correlation observed between Fare, Pclass, and Survived.