

Exploratory Data Analysis (EDA)

Code:

```
# 🚢 Titanic Dataset - Premium EDA Visualization  
# -----  
  
# 1 Import Libraries  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns  
  
# 2 Set Visual Theme  
sns.set_theme(style="whitegrid", context="talk", palette="deep")  
plt.rcParams['figure.facecolor'] = 'whitesmoke'  
plt.rcParams['axes.facecolor'] = 'white'  
plt.rcParams['axes.titlesize'] = 16  
plt.rcParams['axes.labelsize'] = 13  
  
# 3 Load Dataset  
path = r"C:\Users\bewna\Downloads\titanic\train.csv"  
df = pd.read_csv(path)  
print("✅ Data Loaded Successfully!\n")  
  
# 4 Quick Overview  
print("📘 Dataset Info:")  
print(df.info(), "\n")  
print("📊 Summary Statistics:")  
print(df.describe(), "\n")  
print("✖ Missing Values:")  
print(df.isnull().sum(), "\n")
```

```
# 5 Missing Data Heatmap
```

```
plt.figure(figsize=(10, 4))

sns.heatmap(df.isnull(), cbar=False, cmap='Reds', yticklabels=False)

plt.title("🔍 Missing Value Map")

plt.show()
```

```
# 6 Dashboard: Survival Overview
```

```
fig, ax = plt.subplots(1, 3, figsize=(18, 5))

sns.countplot(data=df, x='Survived', palette='crest', ax=ax[0])

ax[0].set_title("Survival Count")

ax[0].bar_label(ax[0].containers[0])

sns.countplot(data=df, x='Pclass', hue='Survived', palette='viridis', ax=ax[1])

ax[1].set_title("Passenger Class vs Survival")

sns.countplot(data=df, x='Sex', hue='Survived', palette='coolwarm', ax=ax[2])

ax[2].set_title("Gender vs Survival")

fig.suptitle("🌟 Titanic Survival Overview", fontsize=20)

plt.tight_layout()

plt.show()
```

```
# 7 Age, Fare, and Embarked Analysis
```

```
fig, ax = plt.subplots(2, 2, figsize=(16, 10))

sns.histplot(data=df, x='Age', kde=True, bins=30, color='teal', ax=ax[0,0])

ax[0,0].set_title("Age Distribution")

sns.histplot(data=df, x='Fare', kde=True, bins=30, color='orange', ax=ax[0,1])

ax[0,1].set_title("Fare Distribution")

sns.countplot(data=df, x='Embarked', hue='Survived', palette='Set2', ax=ax[1,0])

ax[1,0].set_title("Embarked vs Survival")
```

```
sns.scatterplot(data=df, x='Age', y='Fare', hue='Survived', palette='cool', s=80,
alpha=0.8, ax=ax[1,1])

ax[1,1].set_title("Age vs Fare (Colored by Survival)")

fig.suptitle("📊 Demographics & Embarkation Insights", fontsize=20)

plt.tight_layout()

plt.show()

# 8 Boxplots: Age/Fare vs Survival

fig, ax = plt.subplots(1, 2, figsize=(14, 6))

sns.boxplot(data=df, x='Survived', y='Age', palette='mako', ax=ax[0])

sns.boxplot(data=df, x='Survived', y='Fare', palette='flare', ax=ax[1])

ax[0].set_title("Age by Survival")

ax[1].set_title("Fare by Survival")

fig.suptitle("🎯 Survival by Age and Fare", fontsize=20)

plt.tight_layout()

plt.show()

# 9 Correlation Heatmap

corr = df.select_dtypes(include='number').corr()

plt.figure(figsize=(10,6))

sns.heatmap(corr, annot=True, cmap='coolwarm', fmt=".2f", linewidths=0.5)

plt.title("🔥 Feature Correlation Matrix")

plt.show()

# 10 Pairplot for Relationships

sns.pairplot(df[['Survived', 'Pclass', 'Age', 'Fare', 'SibSp', 'Parch']], hue='Survived',
palette='husl', corner=True)

plt.suptitle("🔍 Pairwise Relationships Among Key Features", y=1.03, fontsize=18)

plt.show()
```

1 1 Correlation Ranking

```
corr_sorted = corr['Survived'].sort_values(ascending=False)[1:6]

plt.figure(figsize=(8,5))

corr_sorted.plot(kind='bar', color='seagreen')

plt.title("Top Correlated Features with Survival")

plt.ylabel("Correlation Strength")

plt.xticks(rotation=45)

plt.show()
```

1 2 Smart Insight Summary

```
print("\n💡 Key Insights Summary:")

insights = [
    " 1 Women had a much higher survival rate than men.",
    " 2 1st-class passengers had significantly better survival chances.",
    " 3 Younger and wealthier passengers survived more often.",
    " 4 Fare and Pclass show a strong negative correlation.",
    " 5 Embarked = C passengers had higher survival probability."
]

for tip in insights:
    print(tip)

print("\n✅ Titanic EDA Completed Successfully!")
```





