# 📊 Mini Visualization Dashboard - Titanic Dataset

# Task 4: Data Science with Python Internship

# Created by: Sirigi Reddy Kyathi Durga

# Import libraries

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

# Set visualization style

sns.set(style="whitegrid", palette="pastel")

plt.rcParams['figure.figsize'] = [8,5]

# Load Titanic dataset (built-in from Seaborn)

titanic = sns.load\_dataset('titanic')

# Display first few rows

print("Dataset Preview:")

print(titanic.head())

# Basic info

print("\nDataset Info:")

print(titanic.info())

# Check for missing values

print("\nMissing Values Count:")

print(titanic.isnull().sum())

# ============================

# 🔹 Data Cleaning

# ============================

# Fill missing age with median

titanic['age'].fillna(titanic['age'].median(), inplace=True)

# Fill missing embarked with mode

titanic['embarked'].fillna(titanic['embarked'].mode()[0], inplace=True)

# Drop rows with missing 'deck' for simplicity

titanic.drop(columns=['deck'], inplace=True)

# ============================

# 🔹 Feature Engineering

# ============================

# Create new column 'family\_size'

titanic['family\_size'] = titanic['sibsp'] + titanic['parch'] + 1

# Create categorical age group

titanic['age\_group'] = pd.cut(titanic['age'], bins=[0,12,18,35,60,100],

labels=['Child','Teen','Young Adult','Adult','Senior'])

# ============================

# 🔹 Visualization Dashboard

# ============================

# 1️⃣ Survival Rate by Gender

sns.countplot(x='sex', hue='survived', data=titanic)

plt.title("Survival Count by Gender")

plt.xlabel("Gender")

plt.ylabel("Count")

plt.show()

print("🟢 Insight: More females survived compared to males, indicating women had higher rescue priority.")

# 2️⃣ Survival Rate by Passenger Class

sns.countplot(x='pclass', hue='survived', data=titanic)

plt.title("Survival Count by Passenger Class")

plt.xlabel("Passenger Class")

plt.ylabel("Count")

plt.show()

print("🟢 Insight: Higher-class passengers (1st class) had better survival rates than lower classes.")

# 3️⃣ Age Distribution by Survival

sns.kdeplot(data=titanic, x='age', hue='survived', fill=True)

plt.title("Age Distribution by Survival Status")

plt.show()

print("🟢 Insight: Younger passengers had slightly higher survival chances compared to older ones.")

# 4️⃣ Correlation Heatmap

sns.heatmap(titanic.corr(numeric\_only=True), annot=True, cmap="coolwarm")

plt.title("Correlation Heatmap of Titanic Features")

plt.show()

print("🟢 Insight: 'fare' and 'pclass' show a moderate negative correlation — higher fare means higher class.")

# 5️⃣ Boxplot of Fare by Class

sns.boxplot(x='pclass', y='fare', data=titanic)

plt.title("Fare Distribution by Passenger Class")

plt.show()

print("🟢 Insight: First-class passengers paid much higher fares than others.")

# 6️⃣ Survival Rate by Embarkation Port

sns.countplot(x='embarked', hue='survived', data=titanic)

plt.title("Survival Count by Embarked Port")

plt.show()

print("🟢 Insight: Passengers from port 'C' (Cherbourg) had the highest survival rate.")

# 7️⃣ Family Size vs Survival

sns.barplot(x='family\_size', y='survived', data=titanic, ci=None)

plt.title("Survival Rate by Family Size")

plt.show()

print("🟢 Insight: Small families (2–4 members) had better survival odds than solo travelers or large families.")

# 8️⃣ Pie Chart - Overall Survival Distribution

survival\_counts = titanic['survived'].value\_counts()

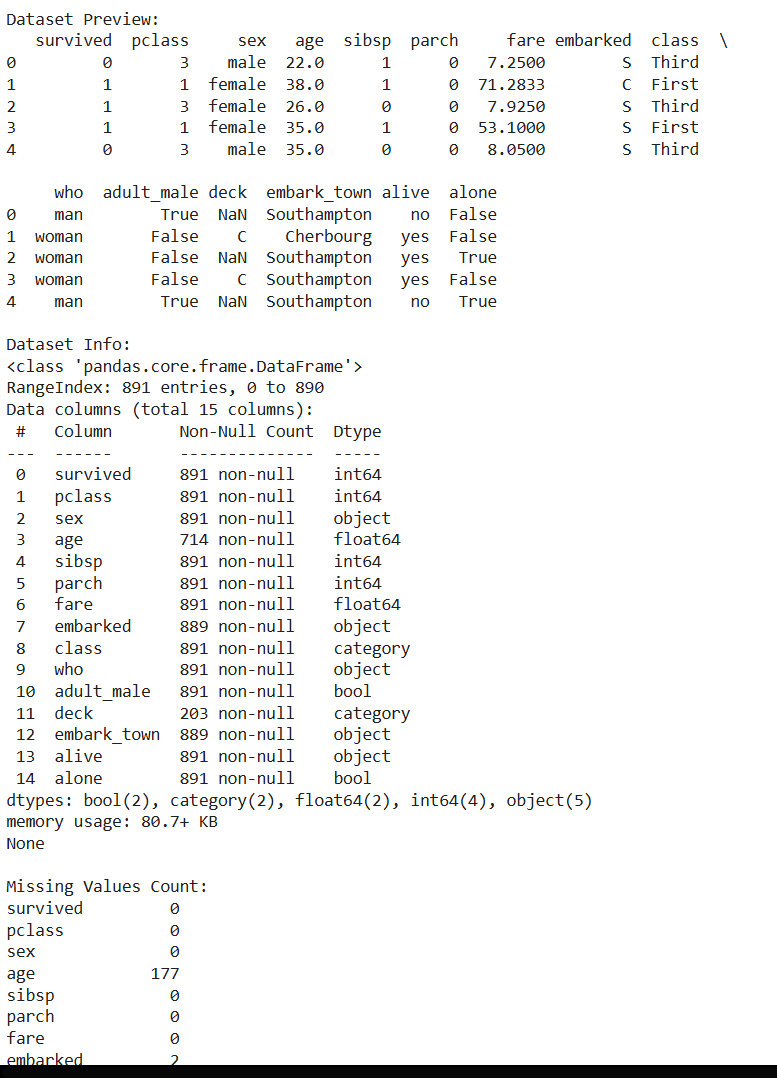
plt.pie(survival\_counts, labels=['Not Survived', 'Survived'], autopct='%1.1f%%', startangle=90, colors=['#ff9999','#66b3ff'])

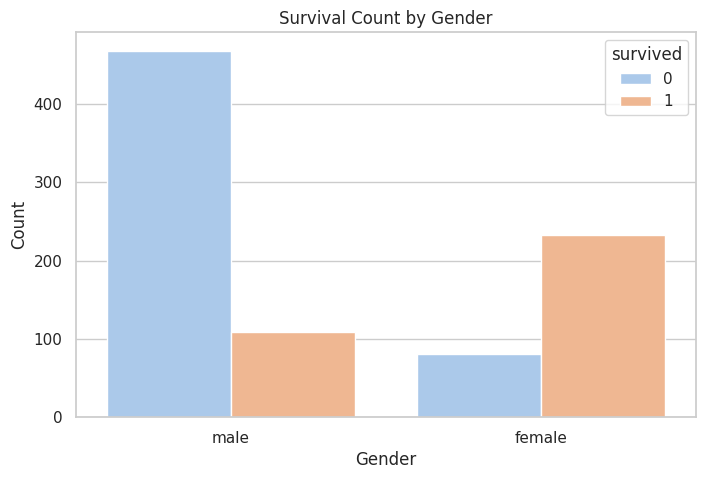
plt.title("Overall Survival Percentage")

plt.show()

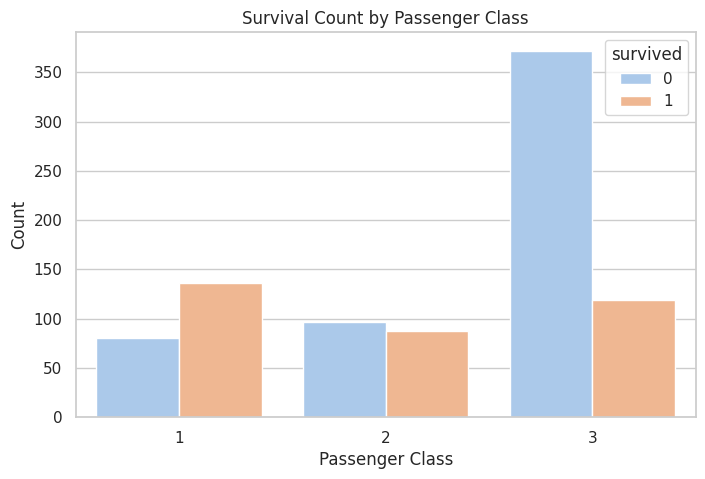
print("🟢 Insight: Around 38% of total passengers survived the Titanic disaster.")

OUTPUT:

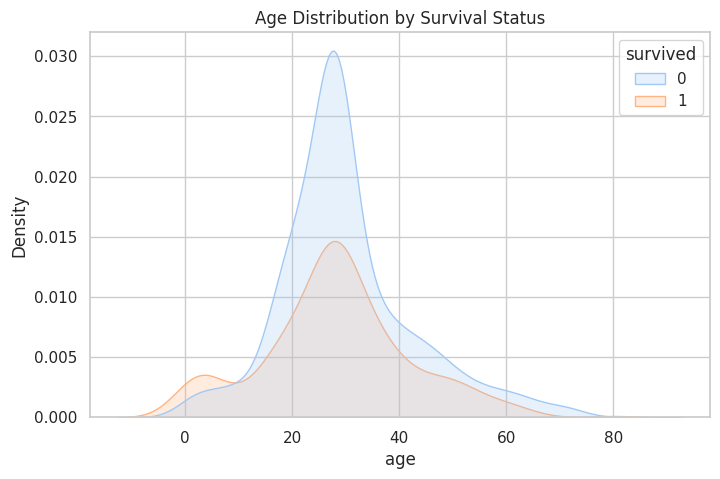




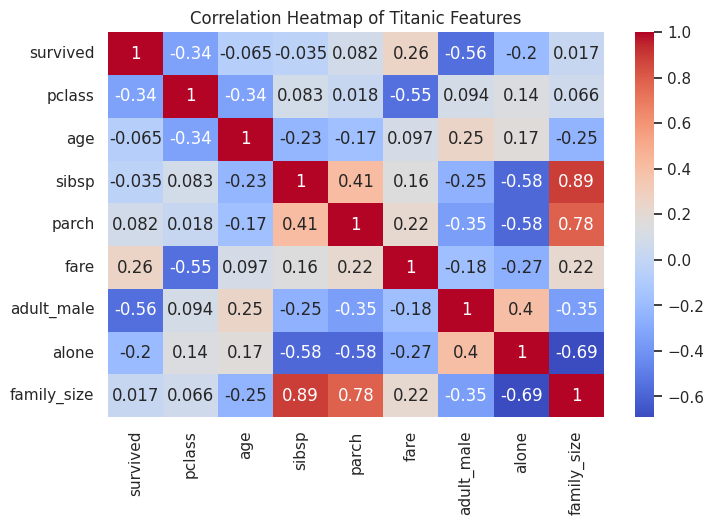
🟢 Insight: More females survived compared to males, indicating women had higher rescue priority.



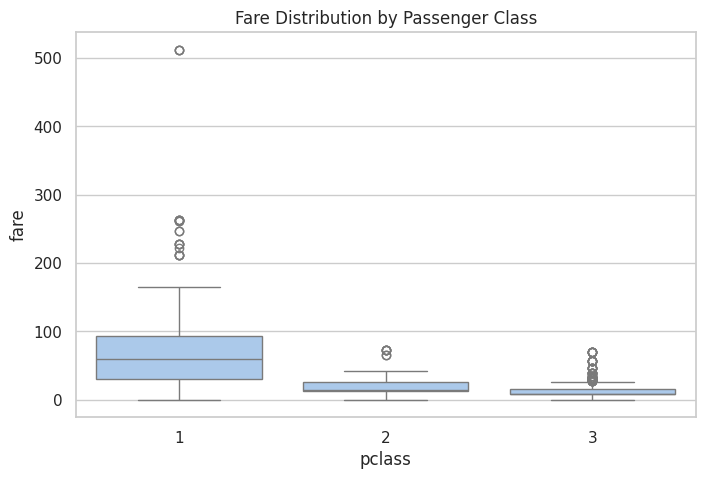
🟢 Insight: Higher-class passengers (1st class) had better survival rates than lower classes.



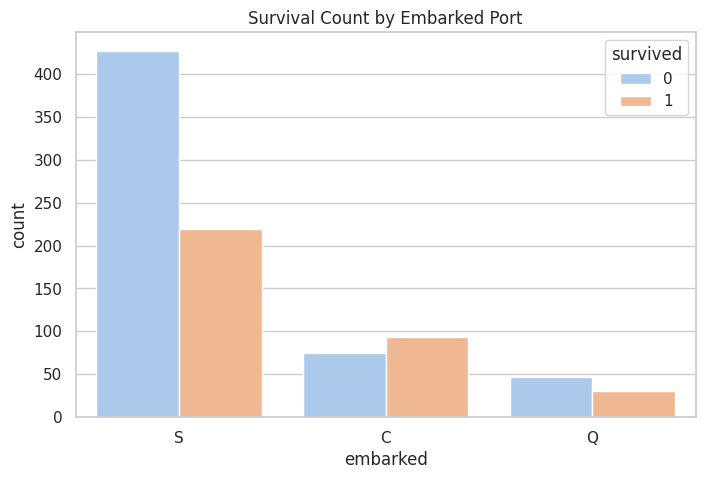
🟢 Insight: Younger passengers had slightly higher survival chances compared to older ones.



🟢 Insight: 'fare' and 'pclass' show a moderate negative correlation — higher fare means higher class.



🟢 Insight: First-class passengers paid much higher fares than others.

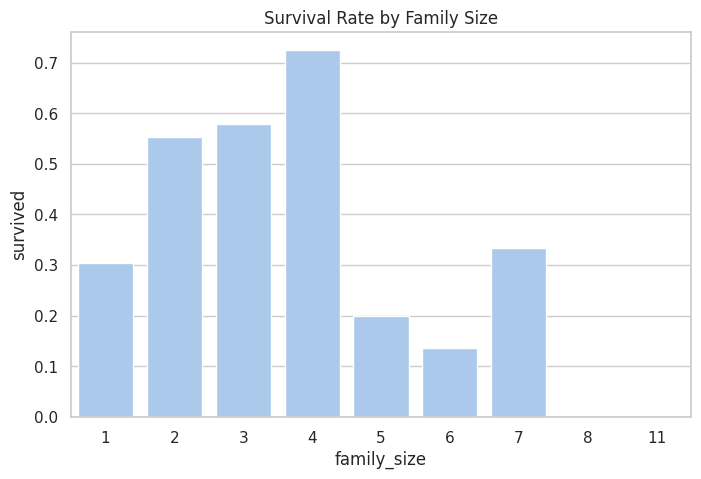


🟢 Insight: Passengers from port 'C' (Cherbourg) had the highest survival rate.

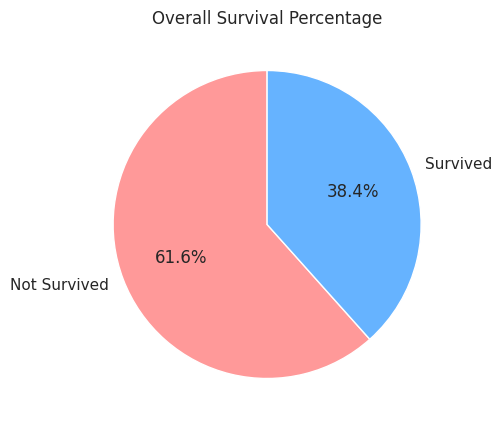
/tmp/ipython-input-490769266.py:99: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

sns.barplot(x='family\_size', y='survived', data=titanic, ci=None)



🟢 Insight: Small families (2–4 members) had better survival odds than solo travelers or large families.



🟢 Insight: Around 38% of total passengers survived the Titanic disaster.