

**AI/ML Programming**

**MCA-475**

**CIA – 01**

***BY***

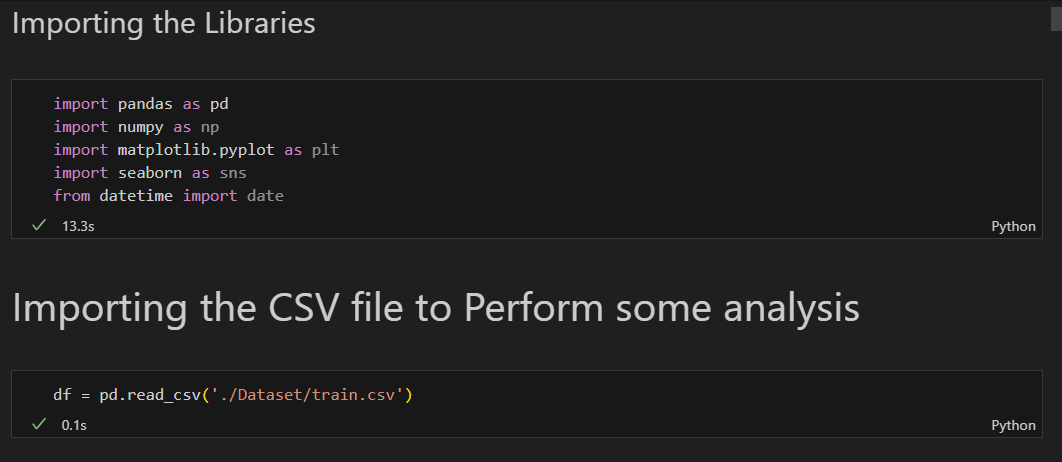
**Himanshu Heda (24225013)**

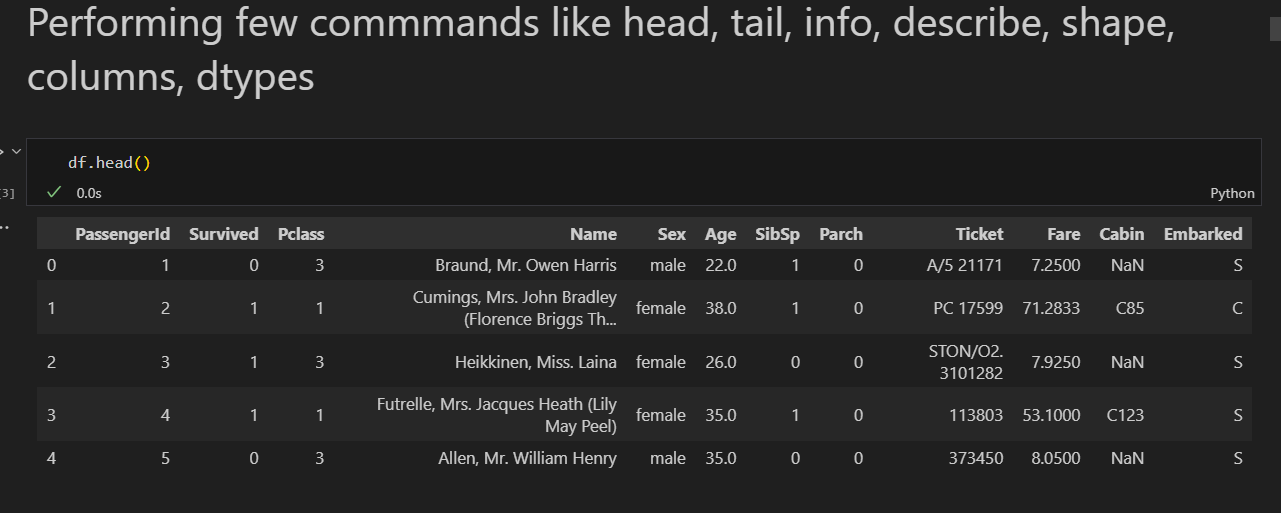
**SUBMITTED TO**

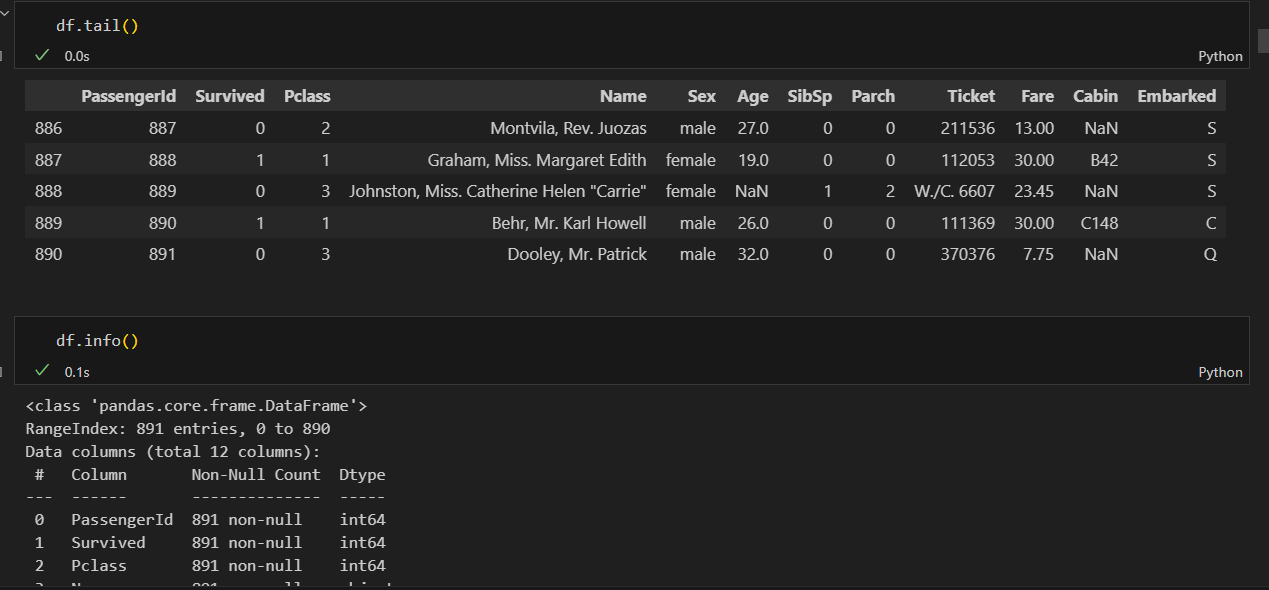
**Dr. Manjula Shannhog**

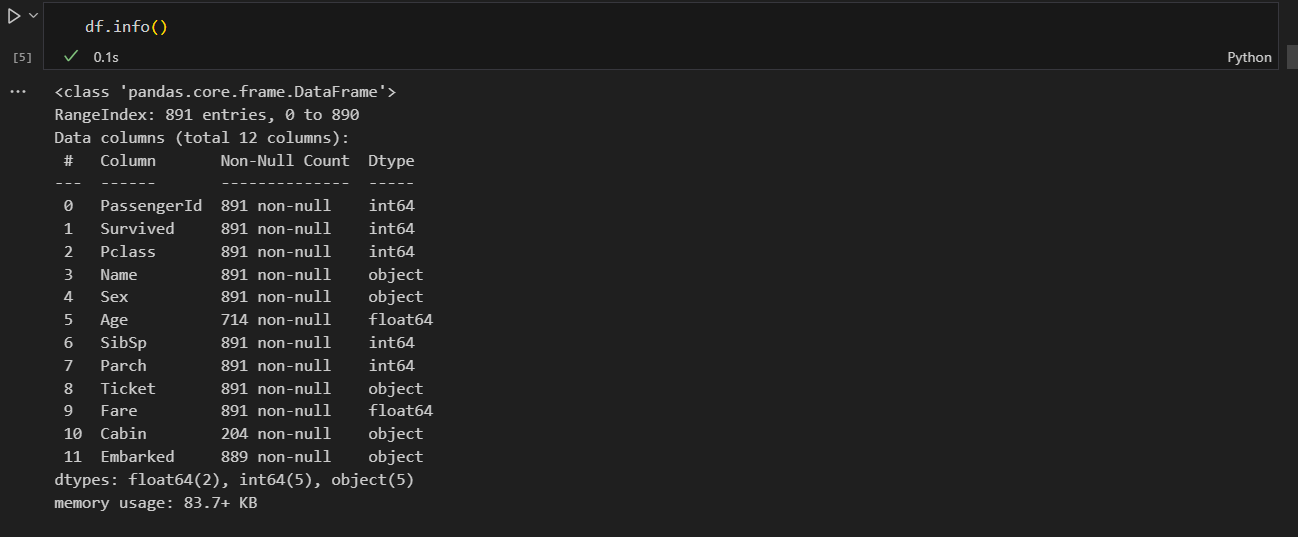
**SCHOOL OF SCIENCES**

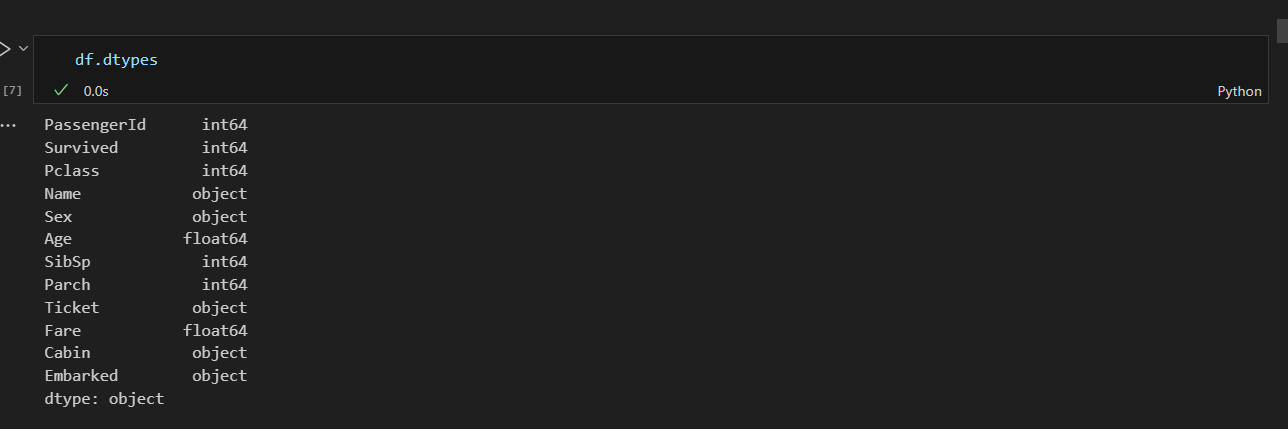
**2024-25**

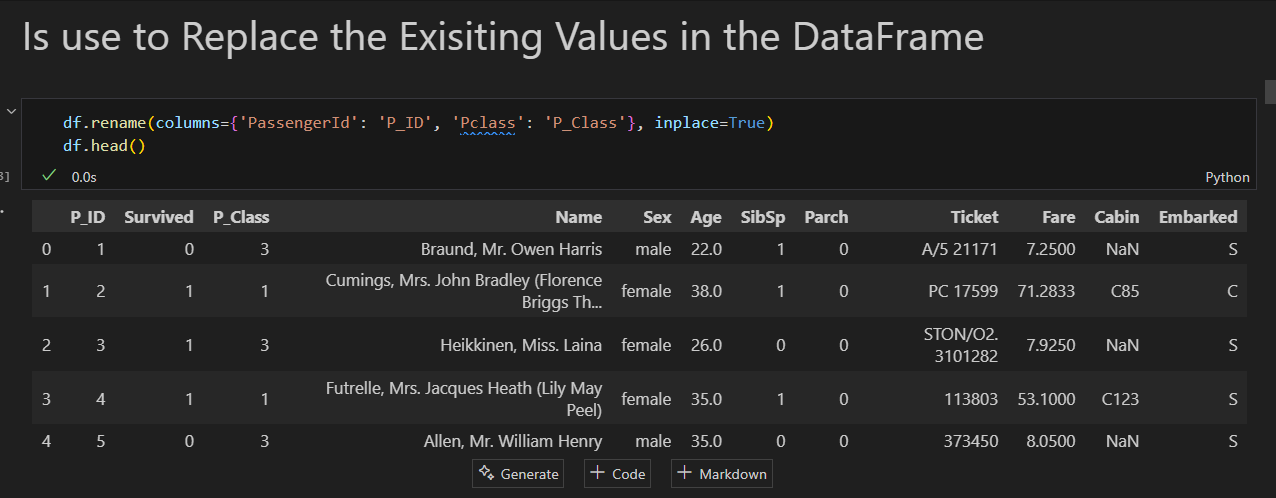
****

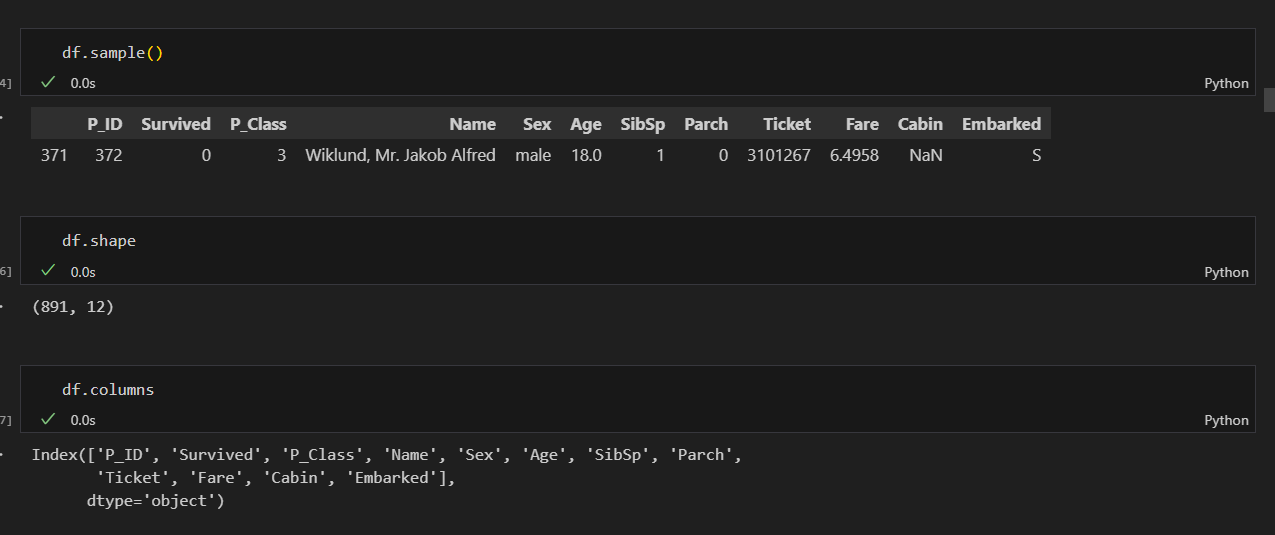
****

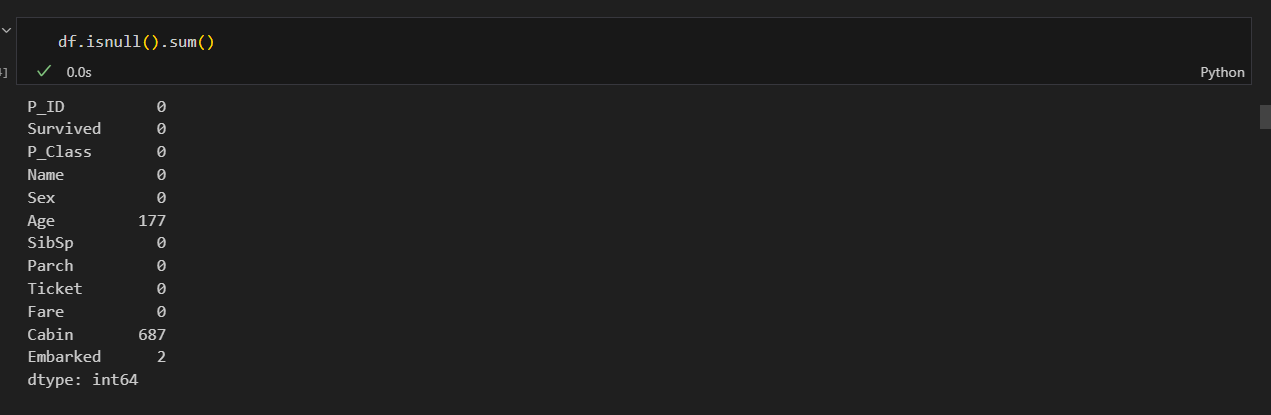
****

****

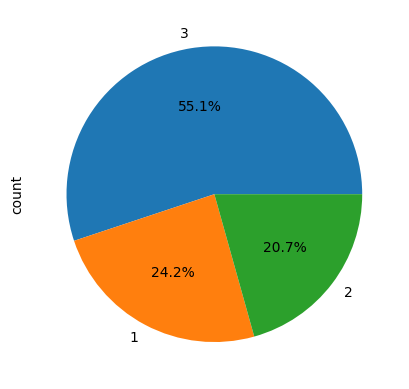
****

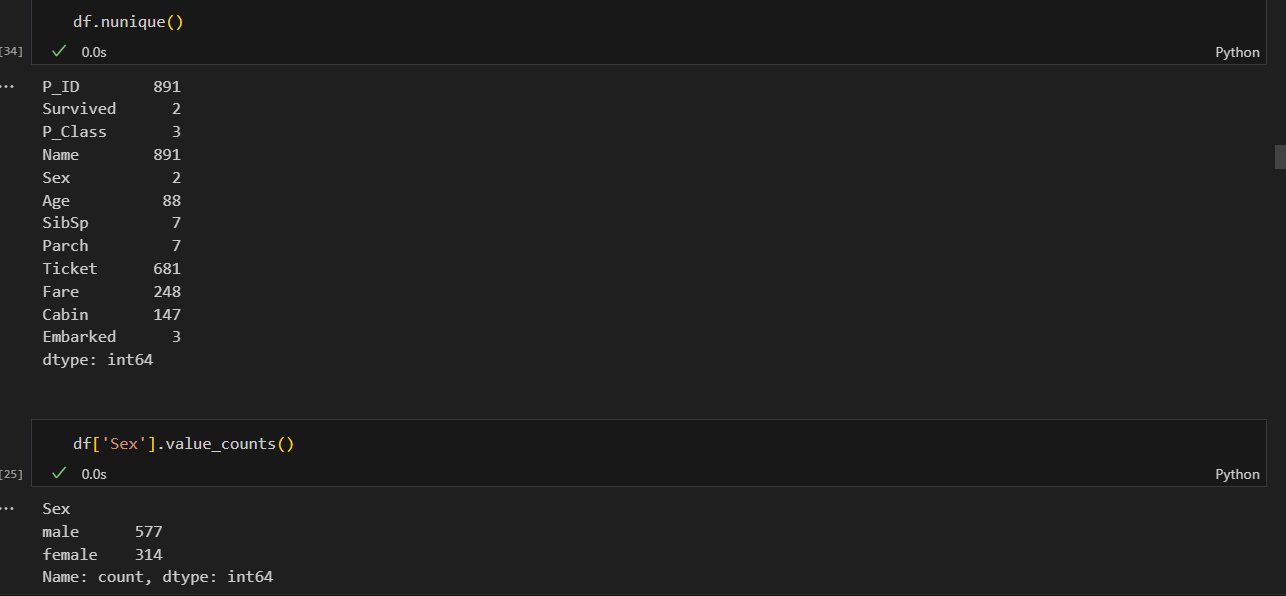
****

****

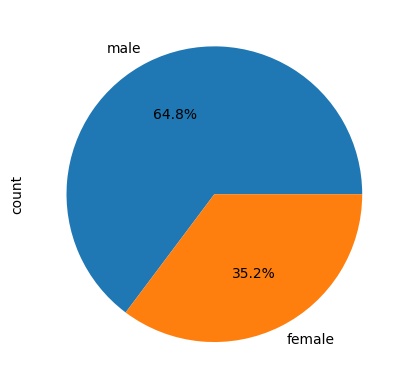
****

df['P\_Class'].value\_counts().plot(kind='pie', autopct='%1.1f%%')

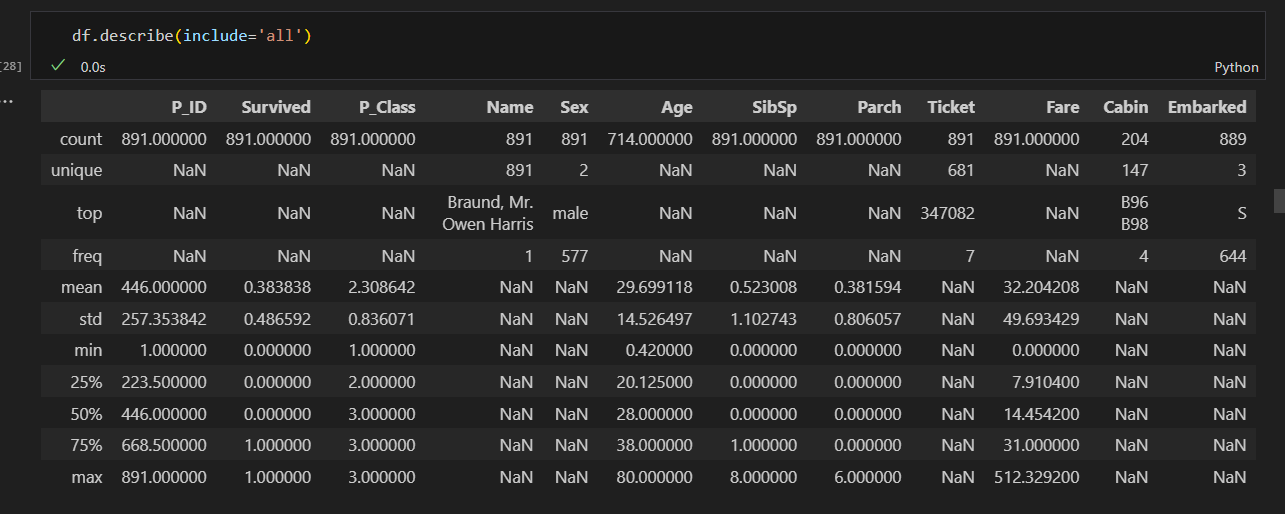
****

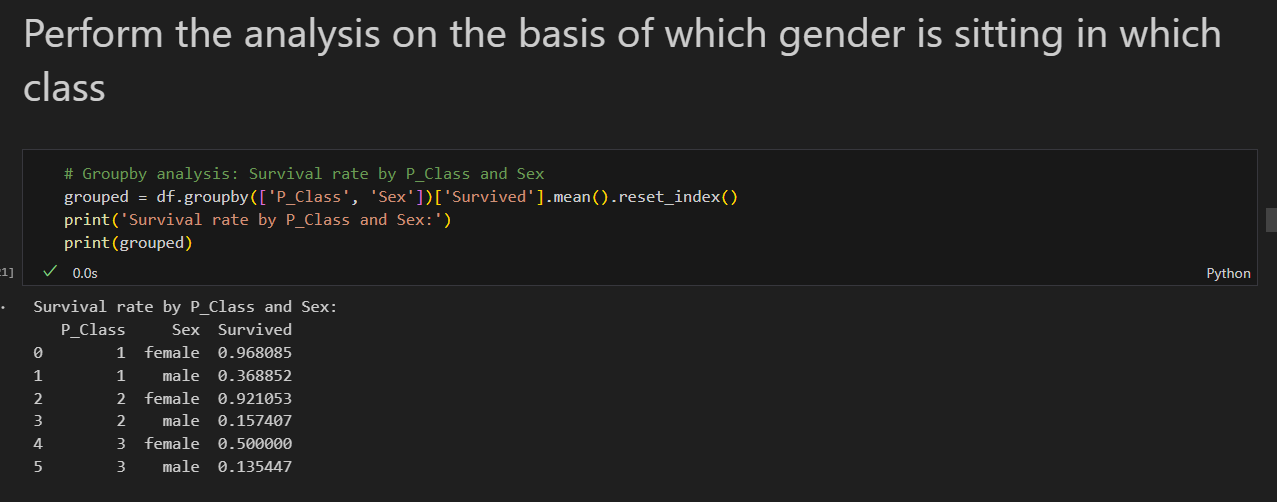
****

df['Sex'].value\_counts().plot(kind='pie', autopct='%1.1f%%')

****

****

****

****

# Plotting the survival rate by P\_Class and Sex

plt.figure(figsize=(8,6))

sns.barplot(data=grouped, x='P\_Class', y='Survived', hue='Sex', palette='Set2')

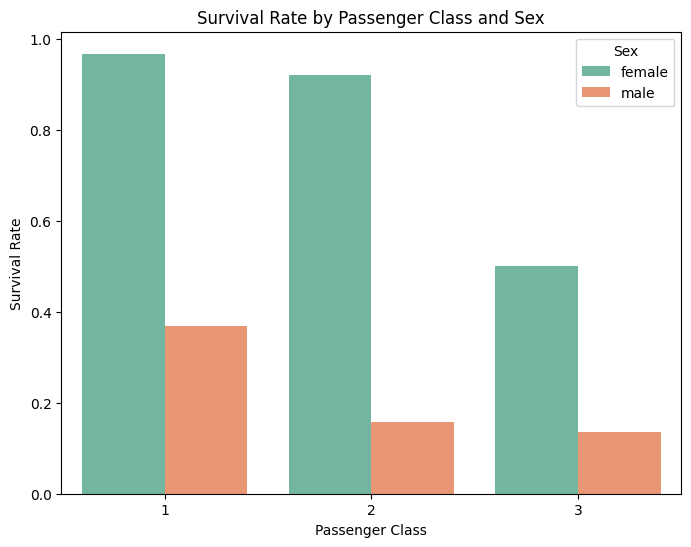
plt.title('Survival Rate by Passenger Class and Sex')

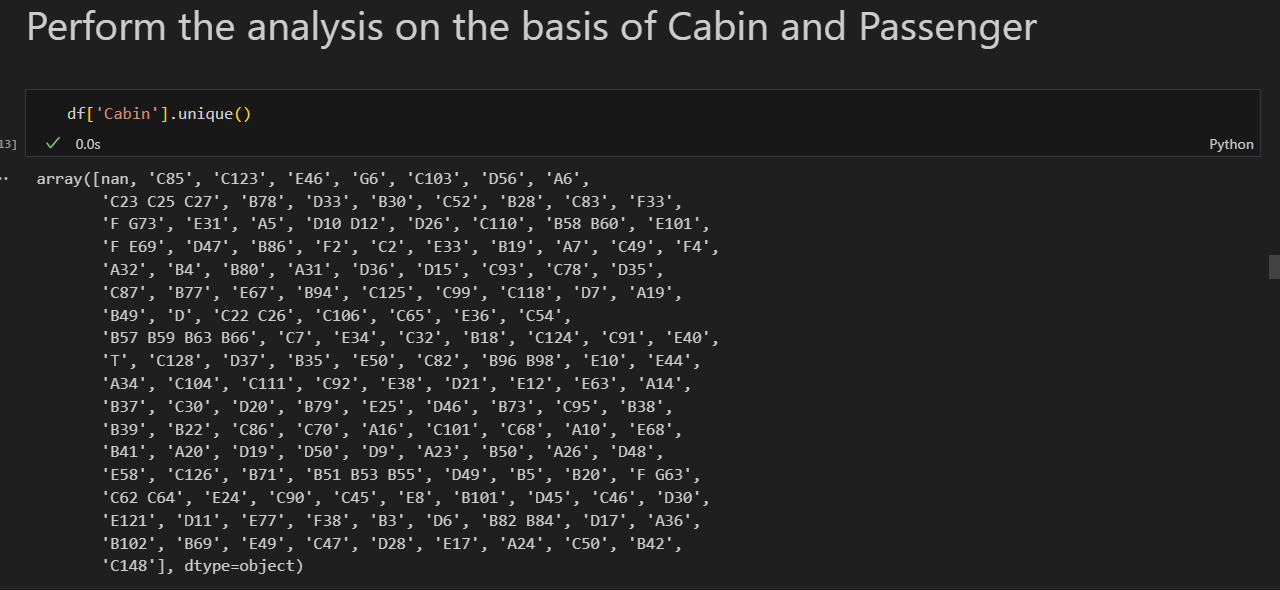
plt.xlabel('Passenger Class')

plt.ylabel('Survival Rate')

plt.legend(title='Sex')

plt.show()

****

****

# Graphical presentation of Cabin data

plt.figure(figsize=(10,5))

cabin\_missing = df['Cabin'].isnull().sum()

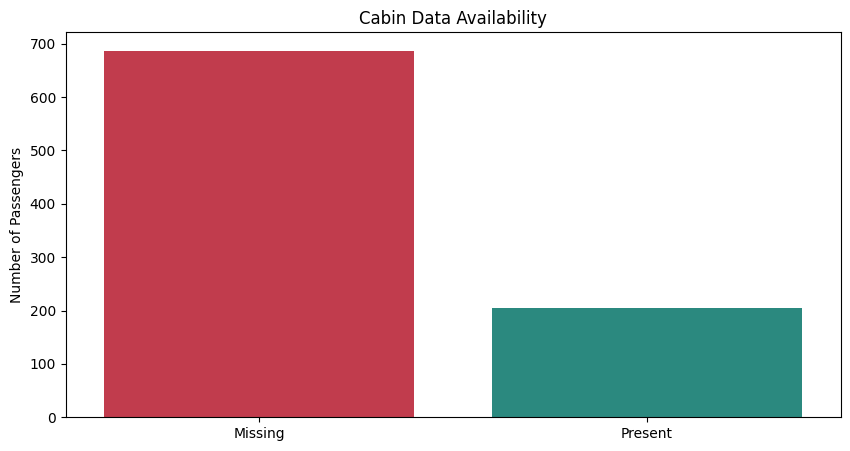
cabin\_present = df['Cabin'].notnull().sum()

sns.barplot(x=['Missing', 'Present'], y=[cabin\_missing, cabin\_present], palette=['#d7263d', '#1b998b'])

plt.title('Cabin Data Availability')

plt.ylabel('Number of Passengers')

plt.show()

****

# Most common cabins (excluding missing)

cabin\_counts = df['Cabin'].value\_counts().head(10)

plt.figure(figsize=(12,6))

sns.barplot(x=cabin\_counts.index, y=cabin\_counts.values, palette='viridis')

plt.title('Top 10 Most Common Cabins')

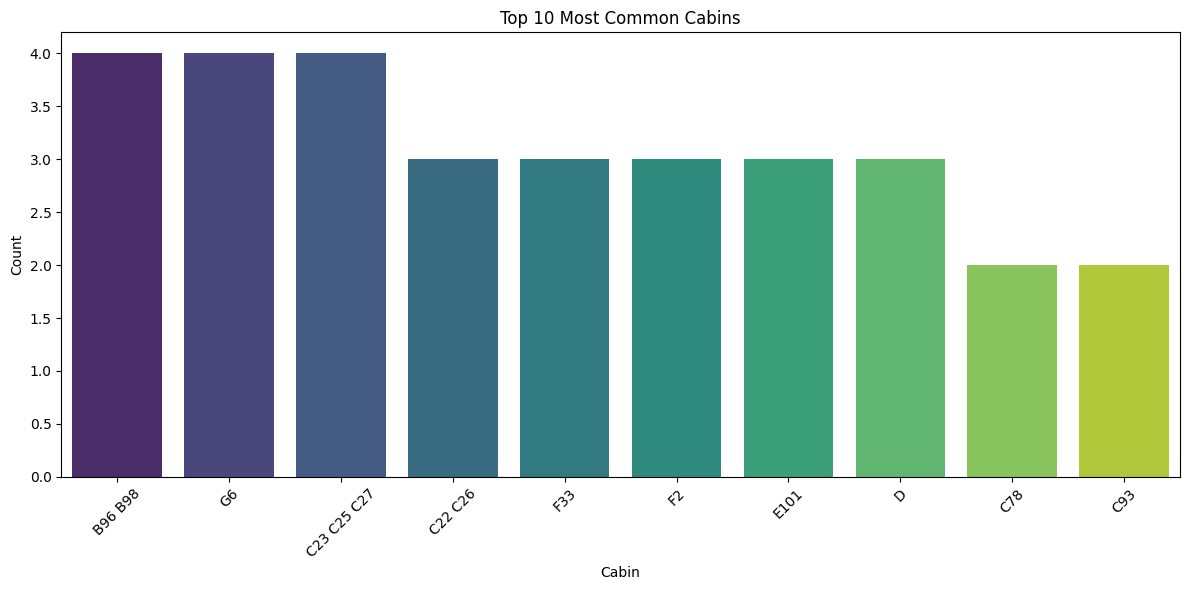
plt.xlabel('Cabin')

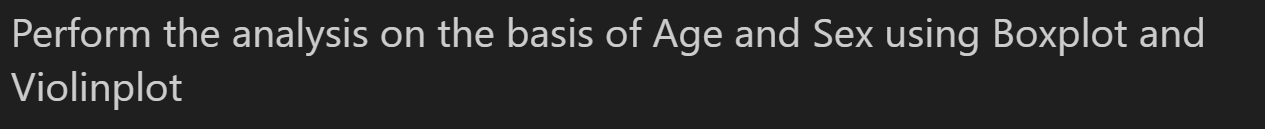
plt.ylabel('Count')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

****

****

# Age distribution by Sex: Boxplot and Violinplot

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

sns.boxplot(data=df, x='Sex', y='Age', palette='Set2')

plt.title('Age Distribution by Sex (Boxplot)')

plt.xlabel('Sex')

plt.ylabel('Age')

plt.show()

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

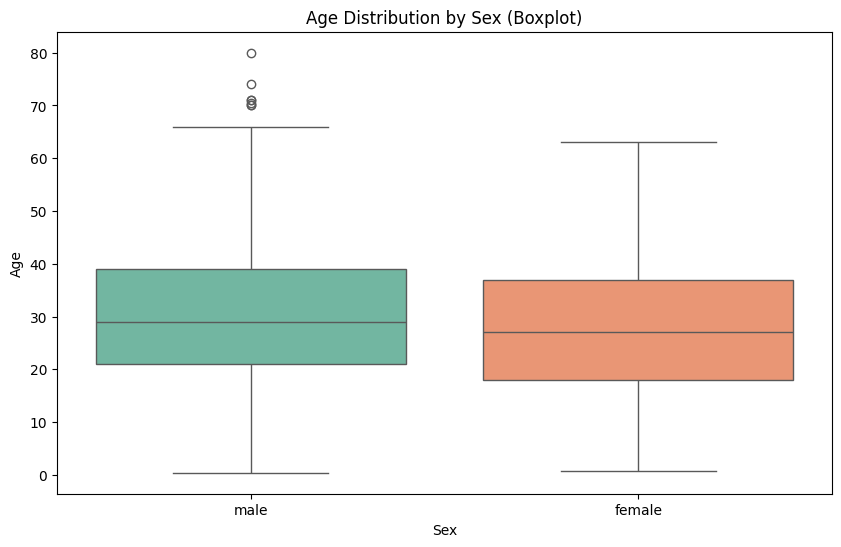
sns.violinplot(data=df, x='Sex', y='Age', palette='Set2')

plt.title('Age Distribution by Sex (Violinplot)')

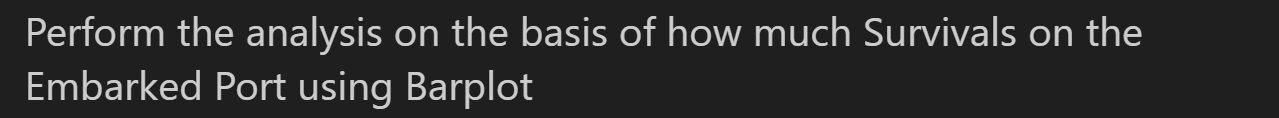
plt.xlabel('Sex')

plt.ylabel('Age')

plt.show()

****

****

****

# Survival rate by Embarked port

embarked\_survival = df.groupby('Embarked')['Survived'].mean().reset\_index()

print('Survival rate by Embarked port:')

print(embarked\_survival)

plt.figure(figsize=(6,4))

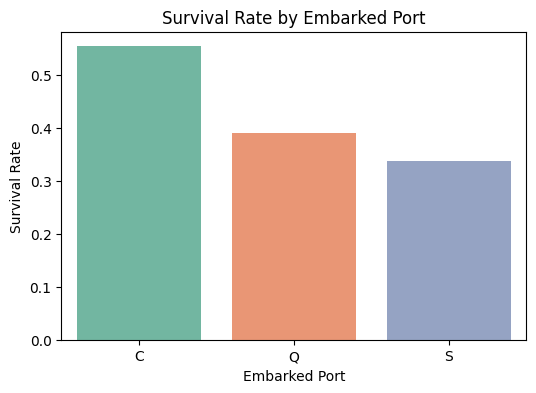
sns.barplot(data=embarked\_survival, x='Embarked', y='Survived', palette='Set2')

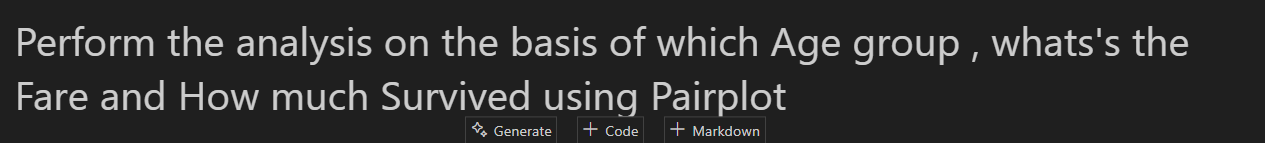
plt.title('Survival Rate by Embarked Port')

plt.ylabel('Survival Rate')

plt.xlabel('Embarked Port')

plt.show()

****

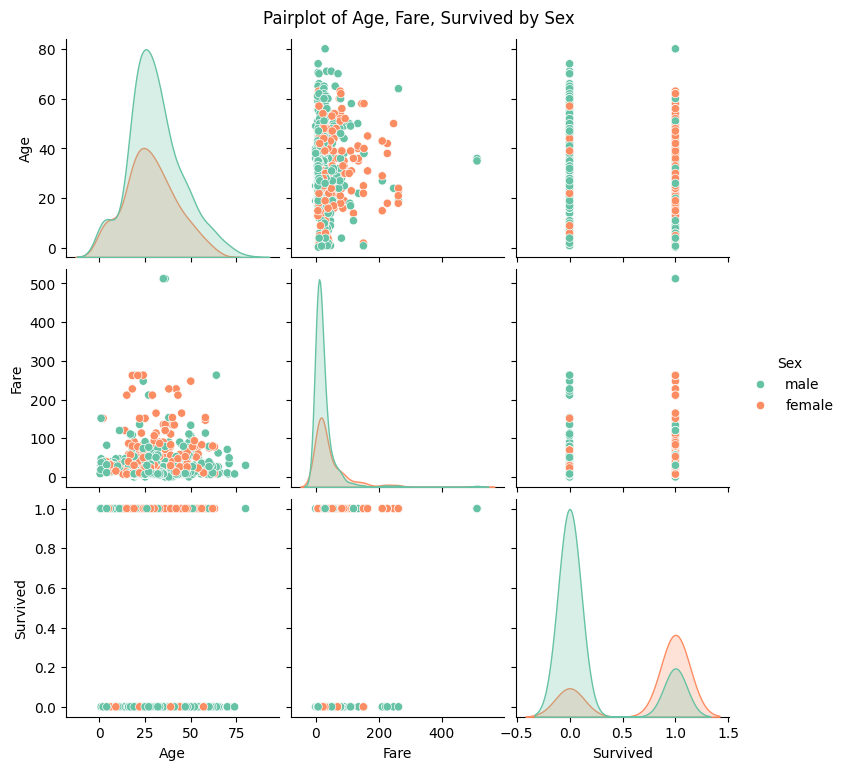
****

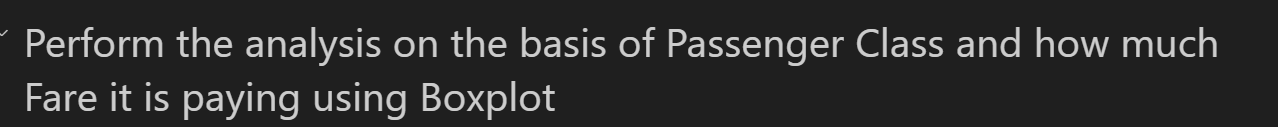
# Pairplot for selected features

sns.pairplot(df, vars=['Age', 'Fare', 'Survived'], hue='Sex', palette='Set2')

plt.suptitle('Pairplot of Age, Fare, Survived by Sex', y=1.02)

plt.show()

****

****

# Distribution of Fare by Passenger Class

plt.figure(figsize=(8,6))

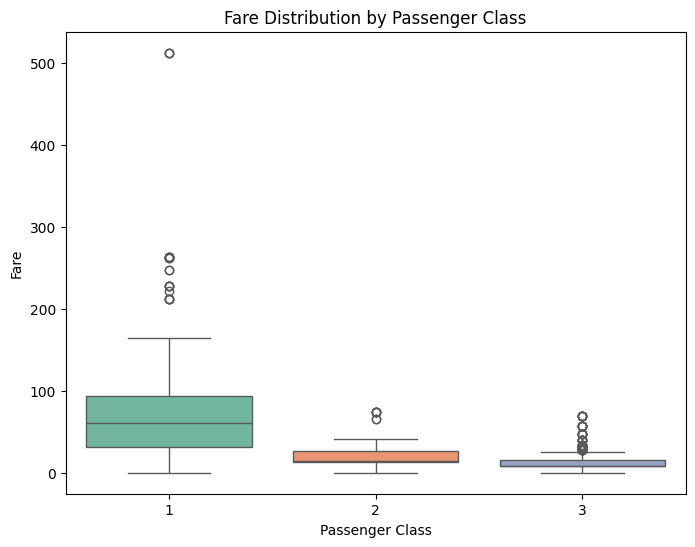
sns.boxplot(data=df, x='P\_Class', y='Fare', palette='Set2')

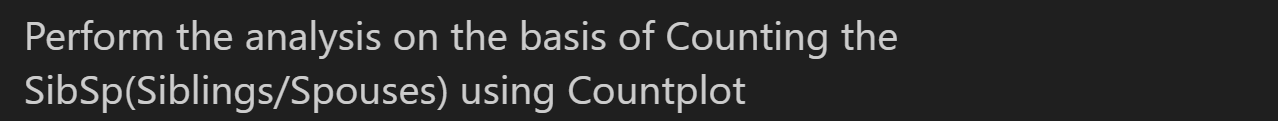
plt.title('Fare Distribution by Passenger Class')

plt.xlabel('Passenger Class')

plt.ylabel('Fare')

plt.show()

****

****

# Countplot for number of siblings/spouses aboard (SibSp)

plt.figure(figsize=(7,5))

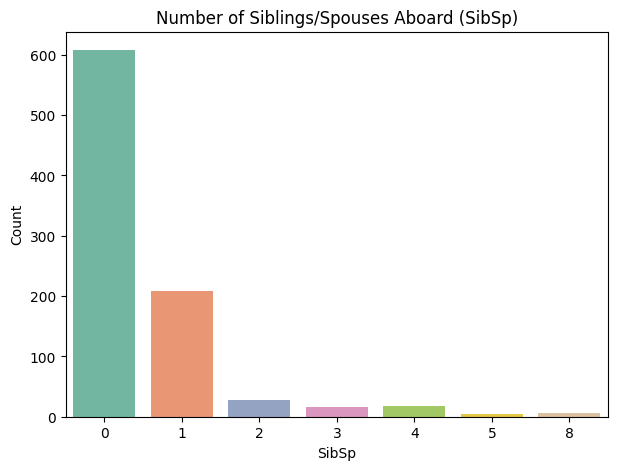
sns.countplot(data=df, x='SibSp', palette='Set2')

plt.title('Number of Siblings/Spouses Aboard (SibSp)')

plt.xlabel('SibSp')

plt.ylabel('Count')

plt.show()

****