**Future Intern as Data Analysis**

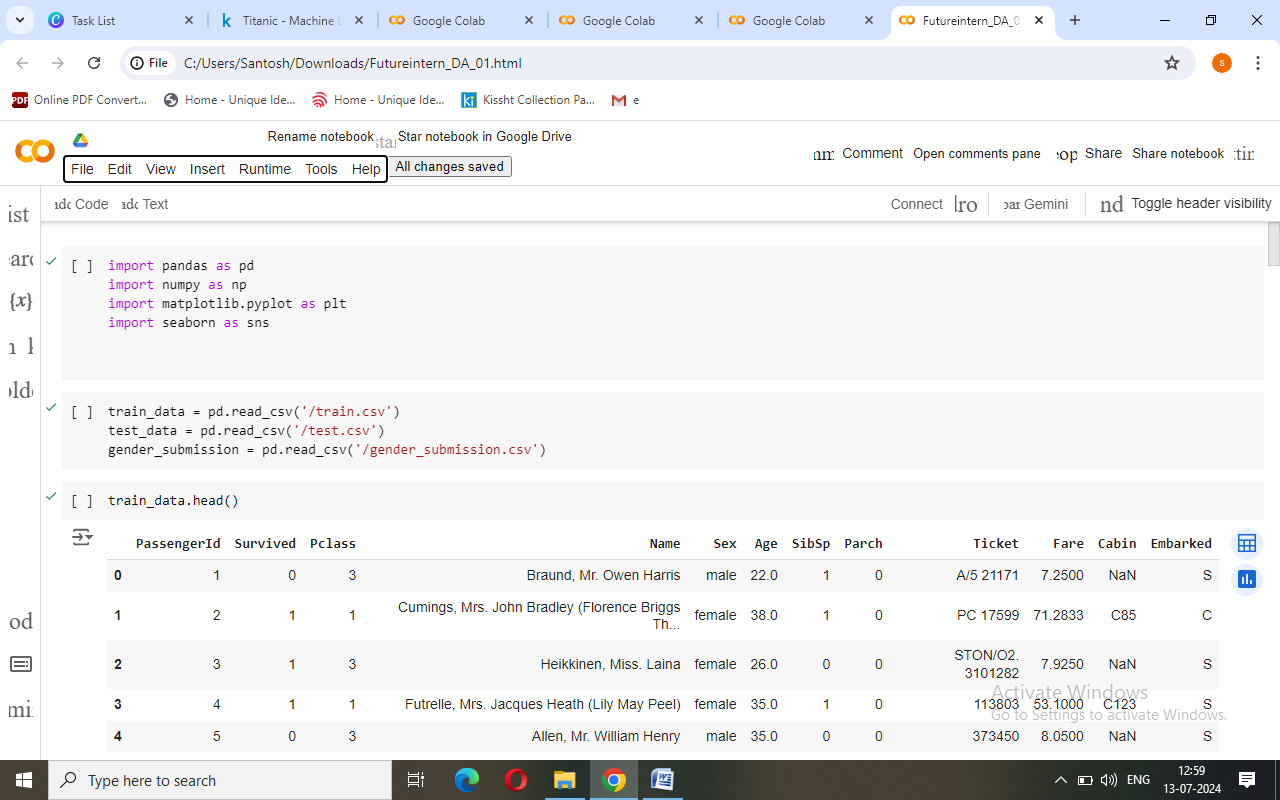
**Task 1**

**Task: Clean a dataset by removing missing values and outliers**

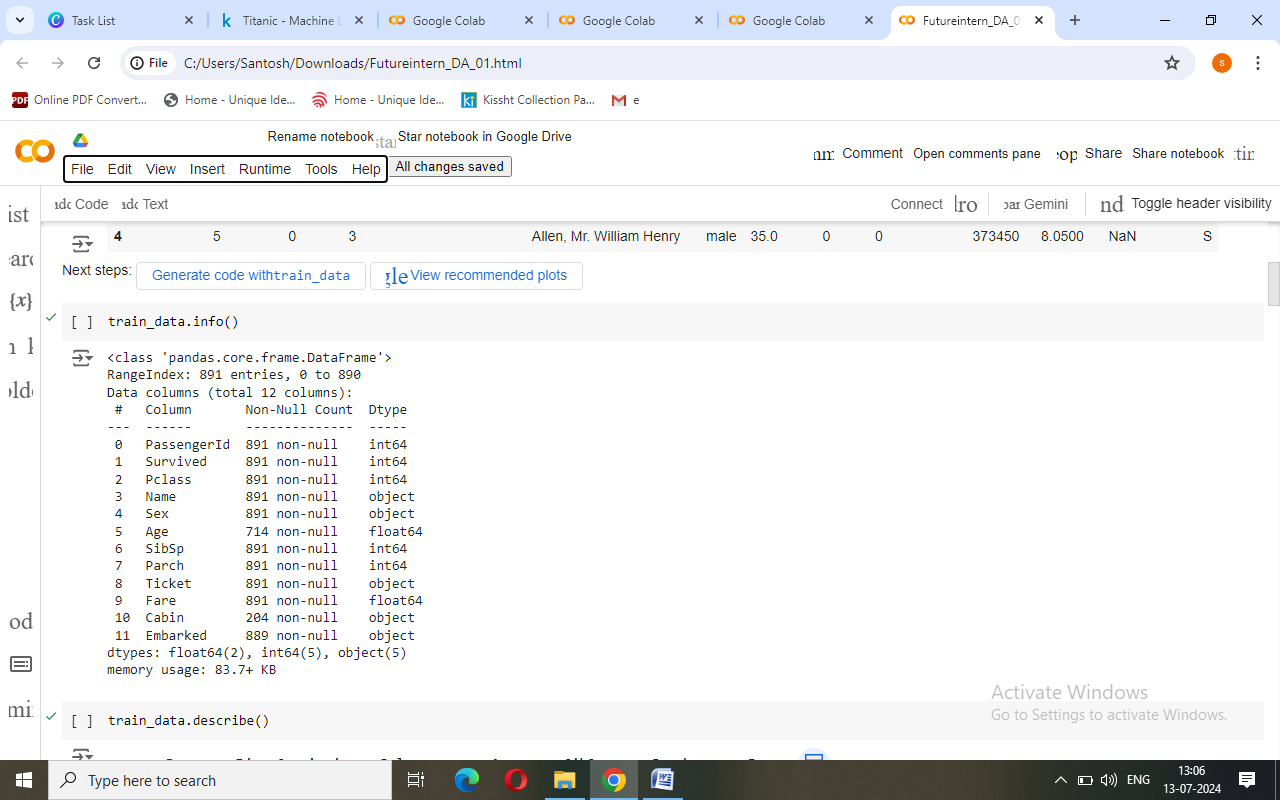
**Solution:**

**Step:**

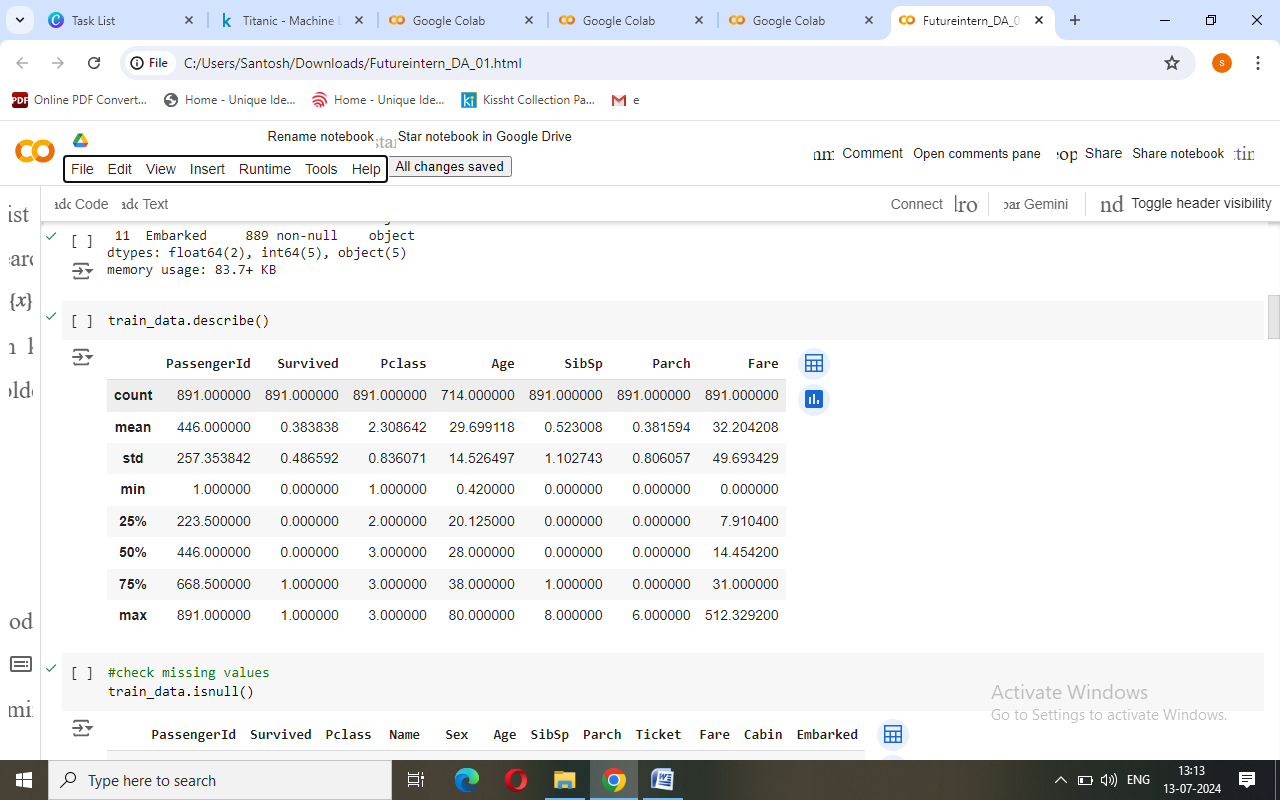
**1.Import packages and display train dataset and test dataset**

****

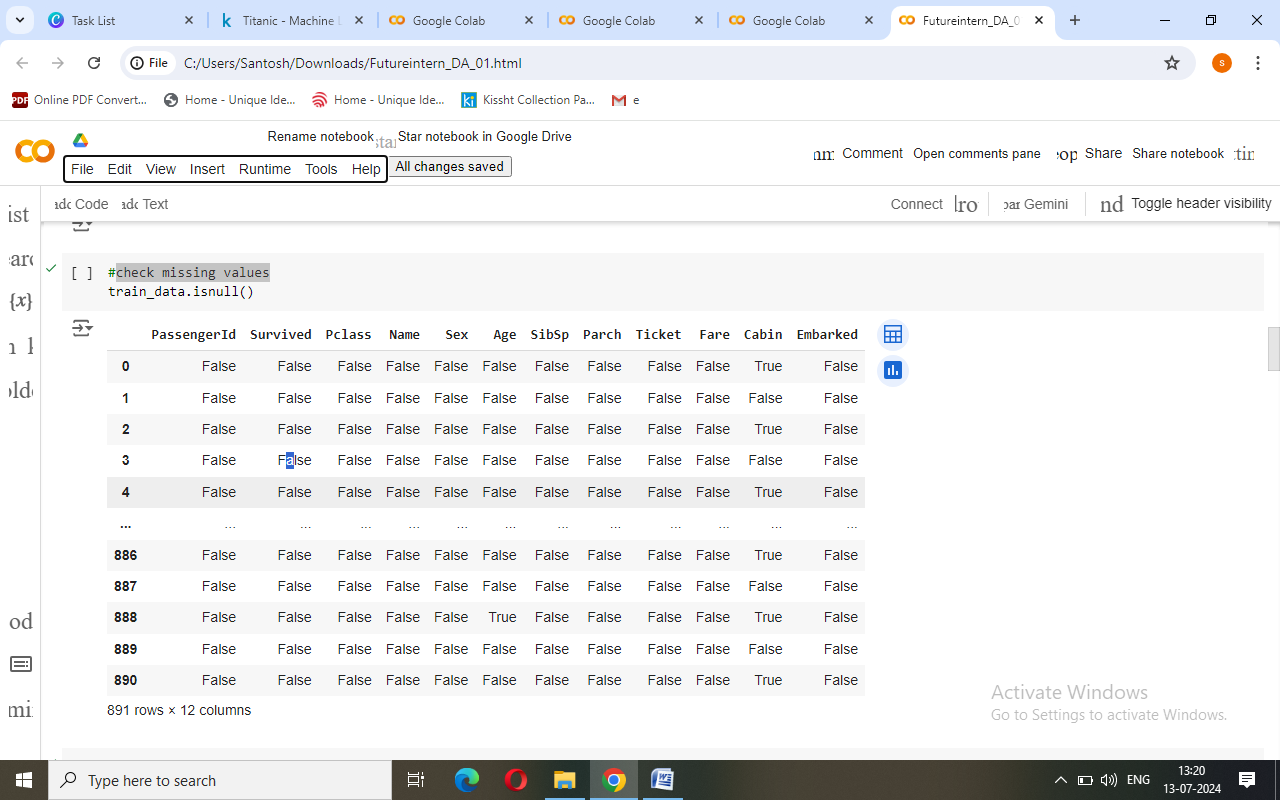
**2. Display data types**



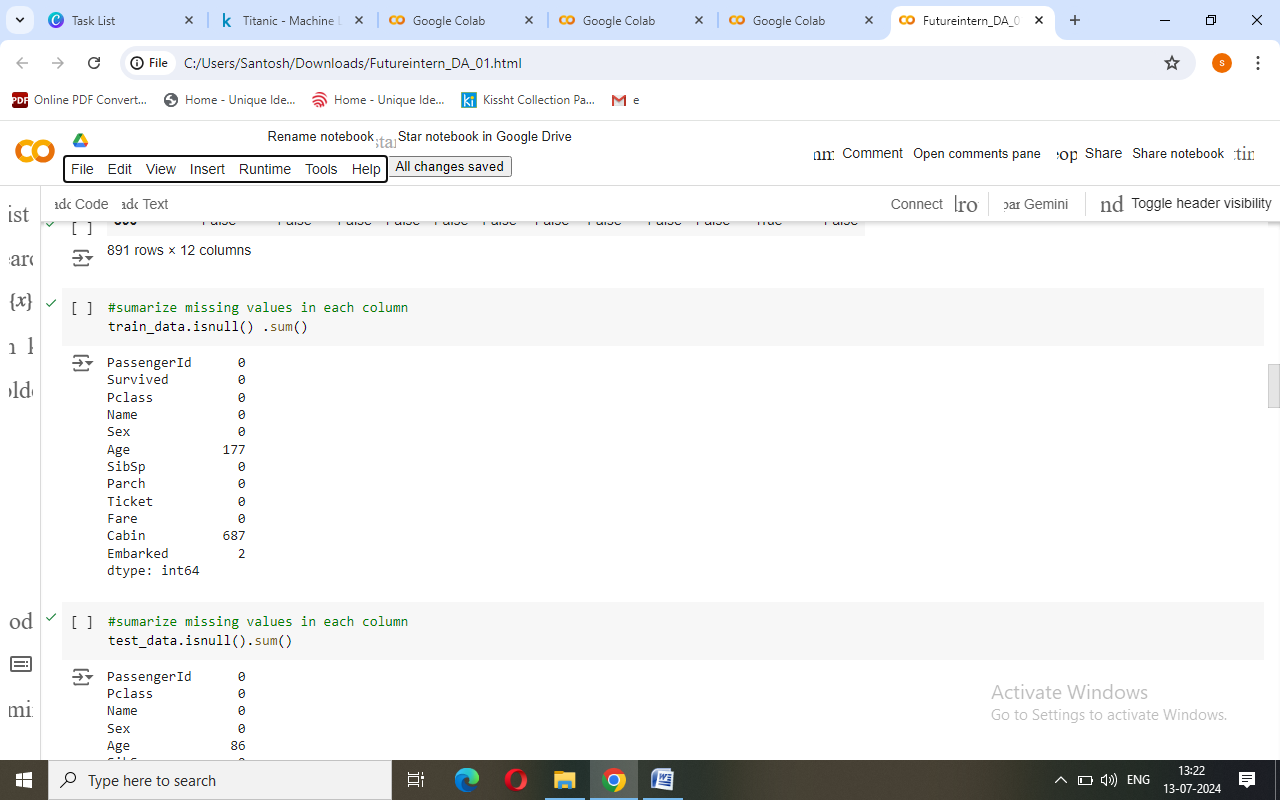
**3. Get Summary Statistics for numerical column**

****

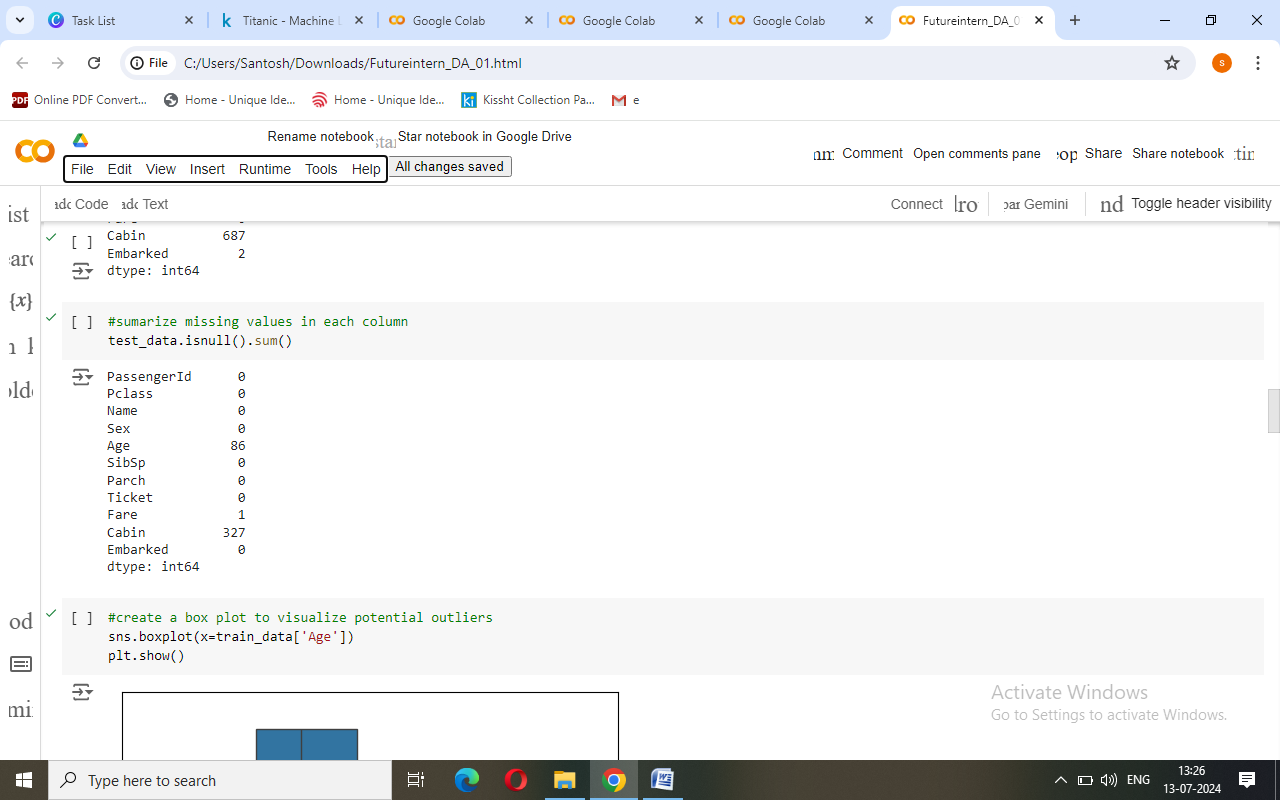
**4. Get the missing values in train dataset**

****

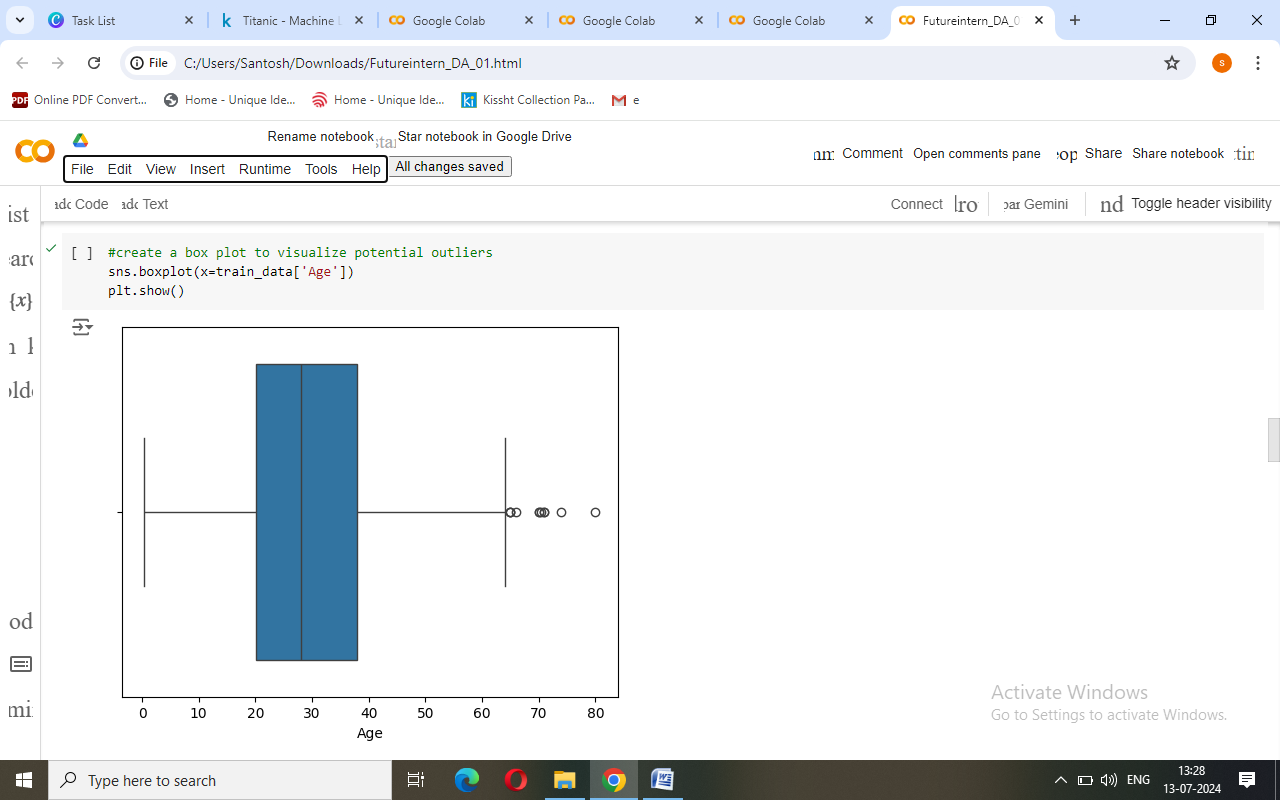
**5. Get missing values from train dataset in each column**

****

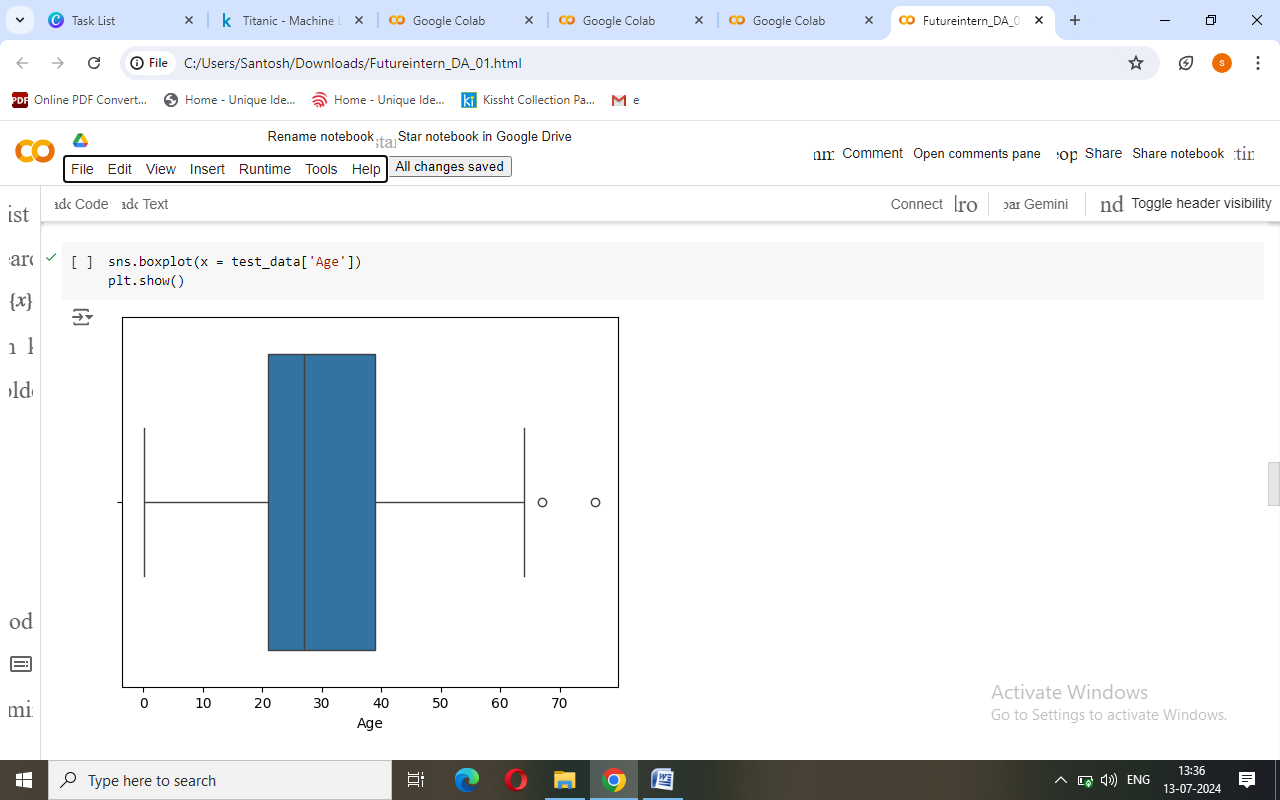
**6. Get missing values from test dataset in each column**

****

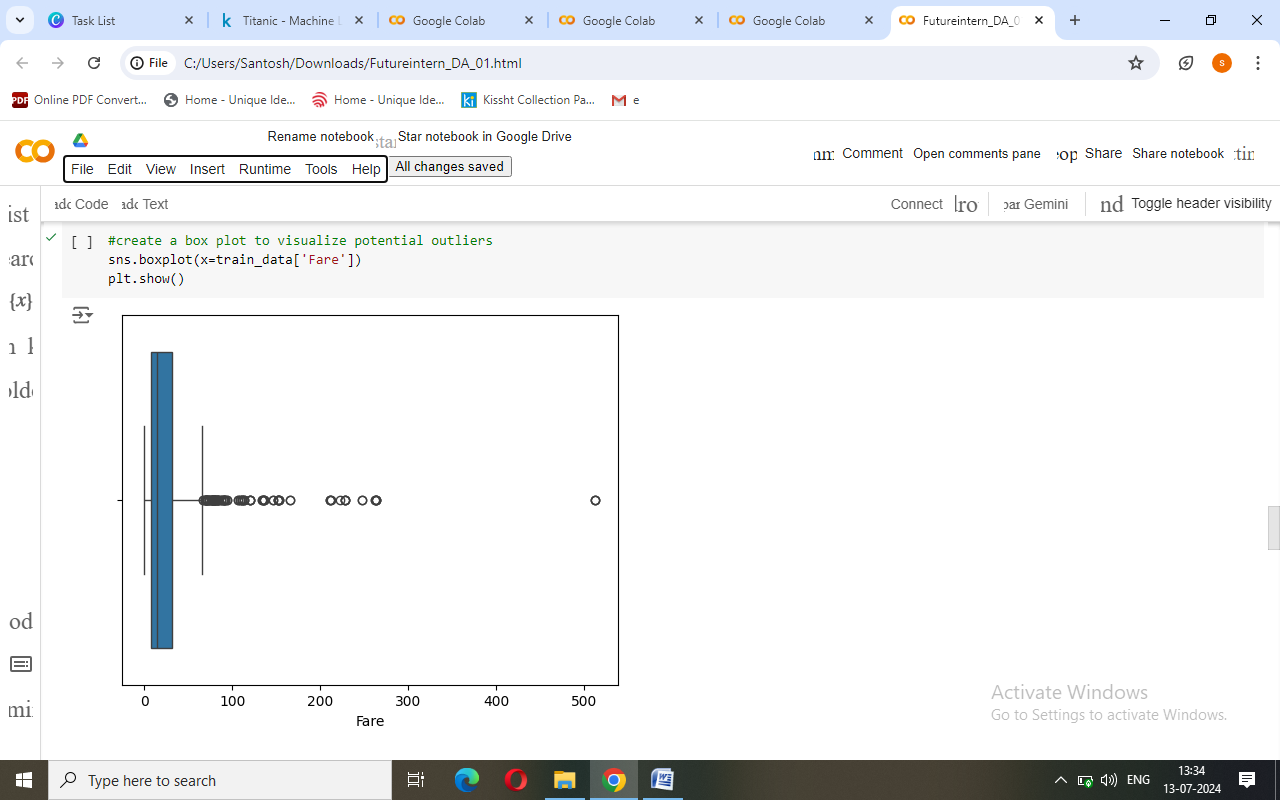
**7. Create a box plot to visualize potential outliers for train dataset(Age)**

****

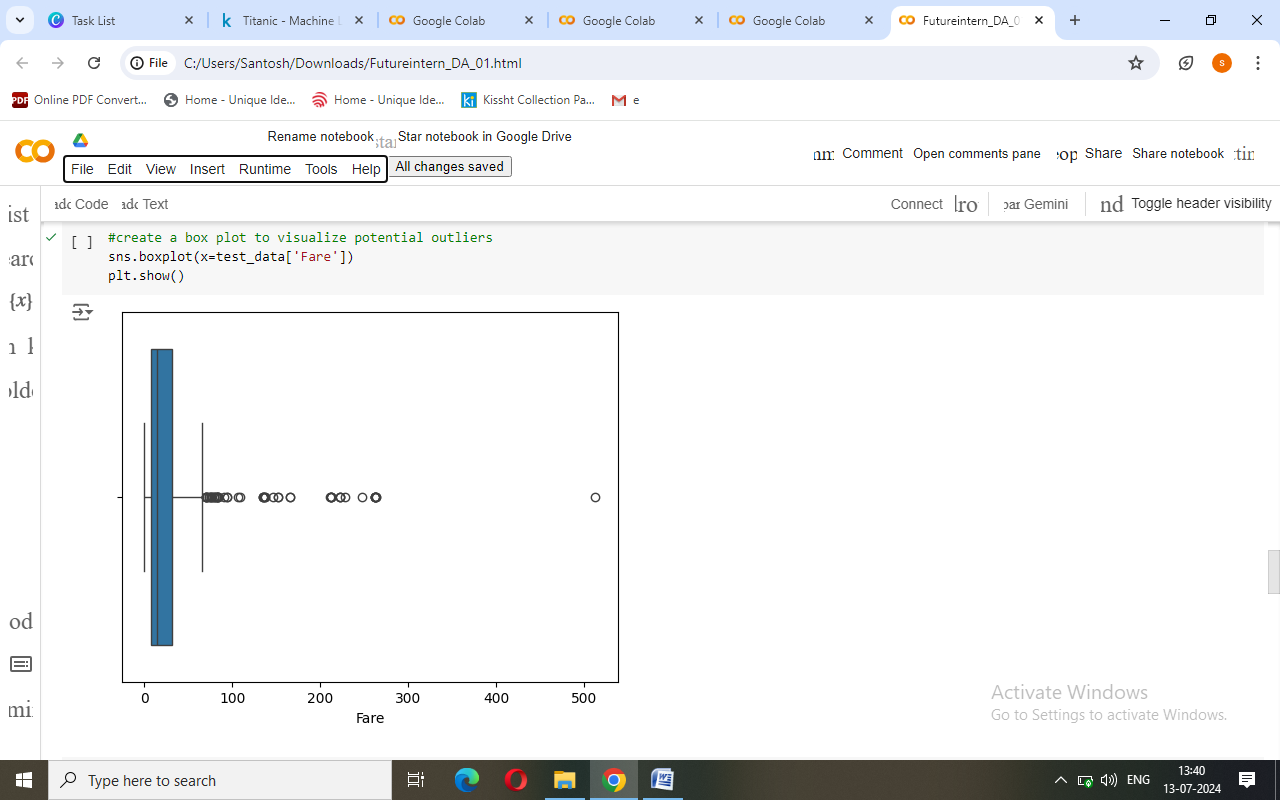
**8. Create a box plot to visualize potential outliers for test dataset (Age)**

****

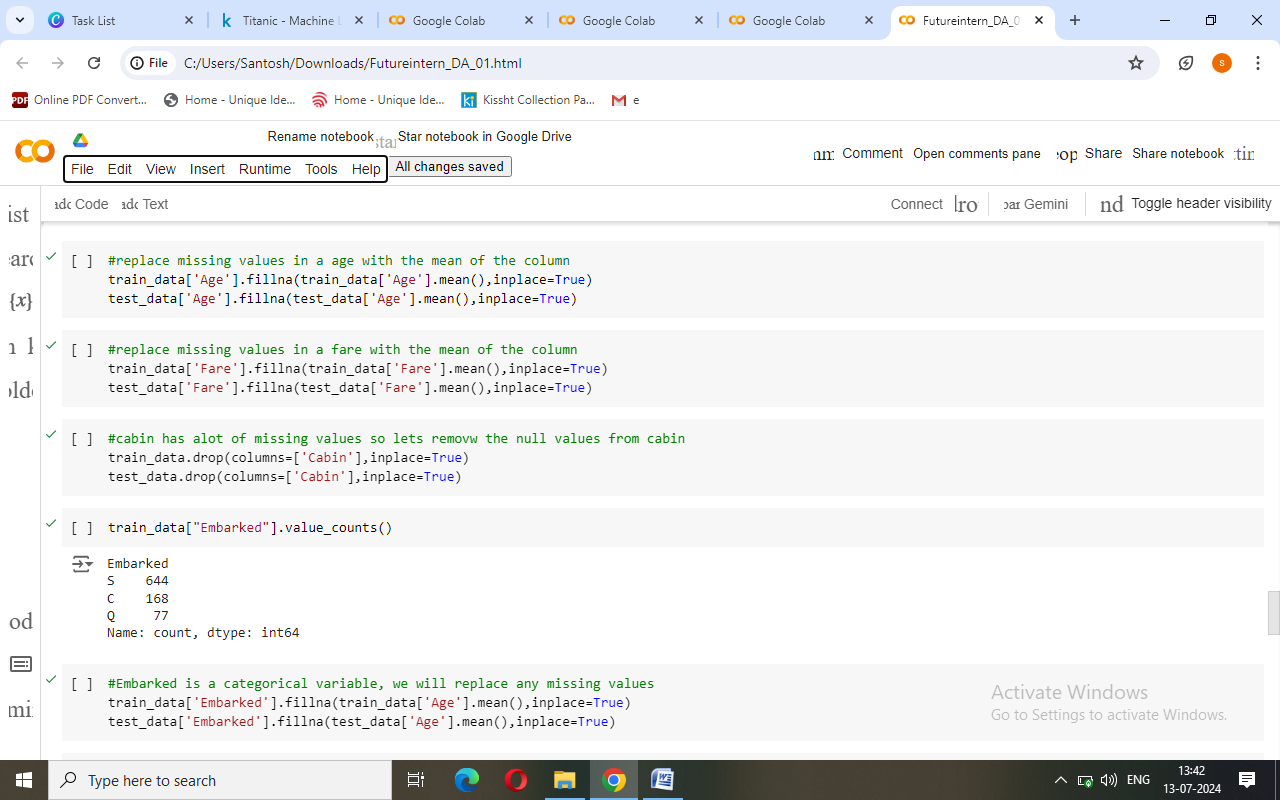
**9. Create a box plot to visualize potential outliers for train dataset (Fare)**

****

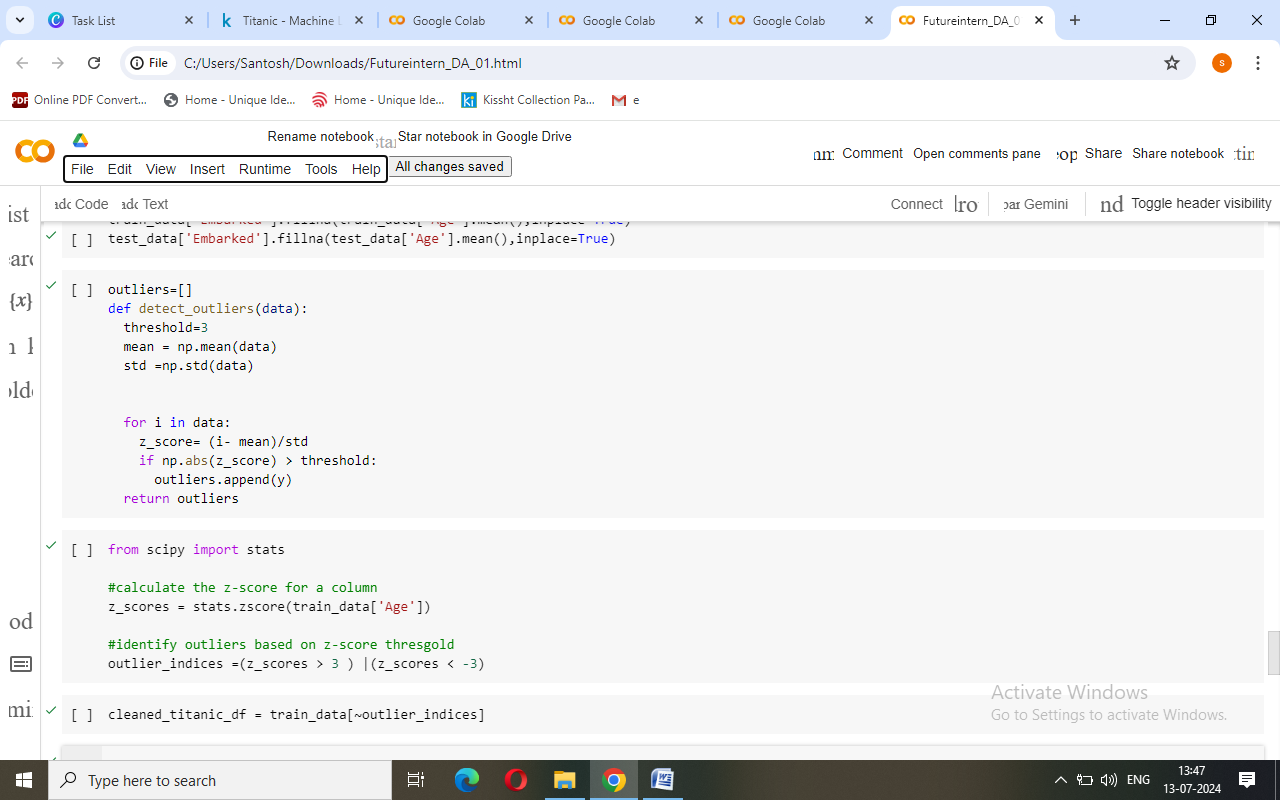
**10. Create a box plot to visualize potential outliers for test dataset(Fare)**

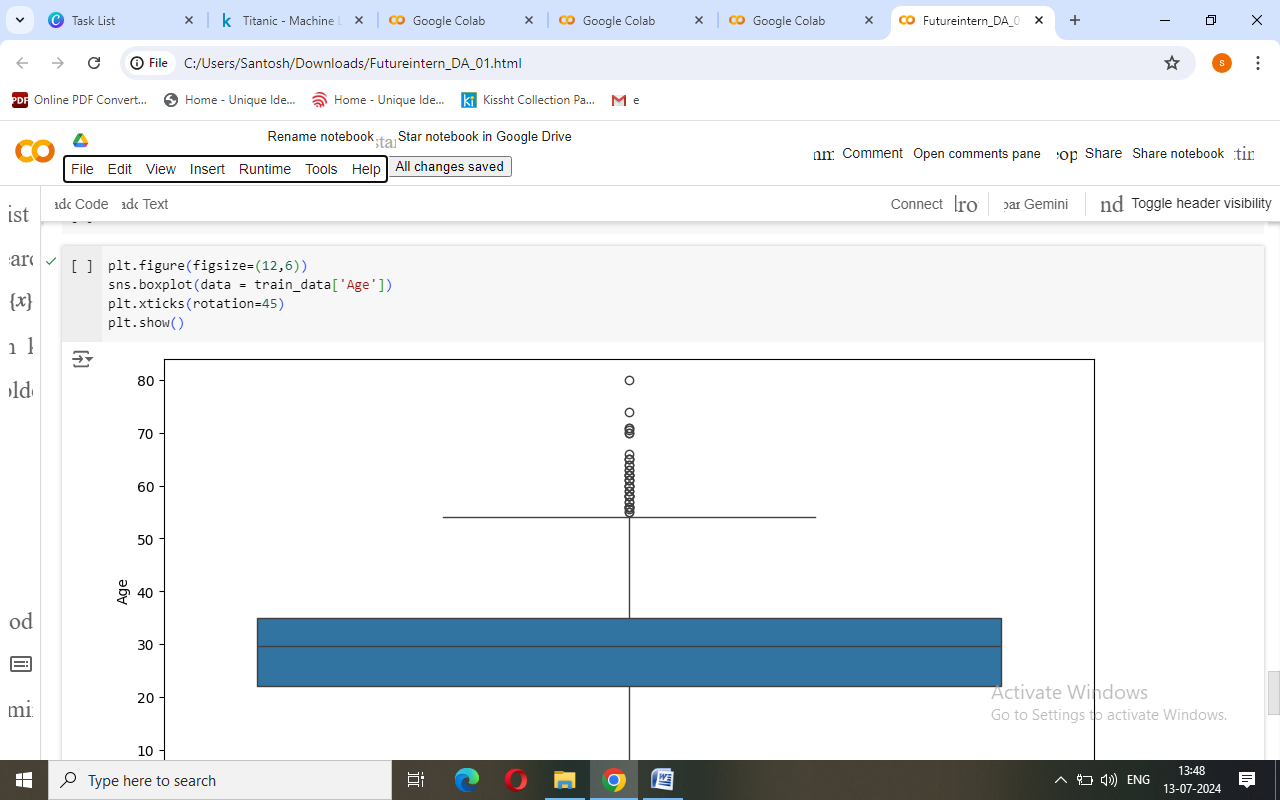
****

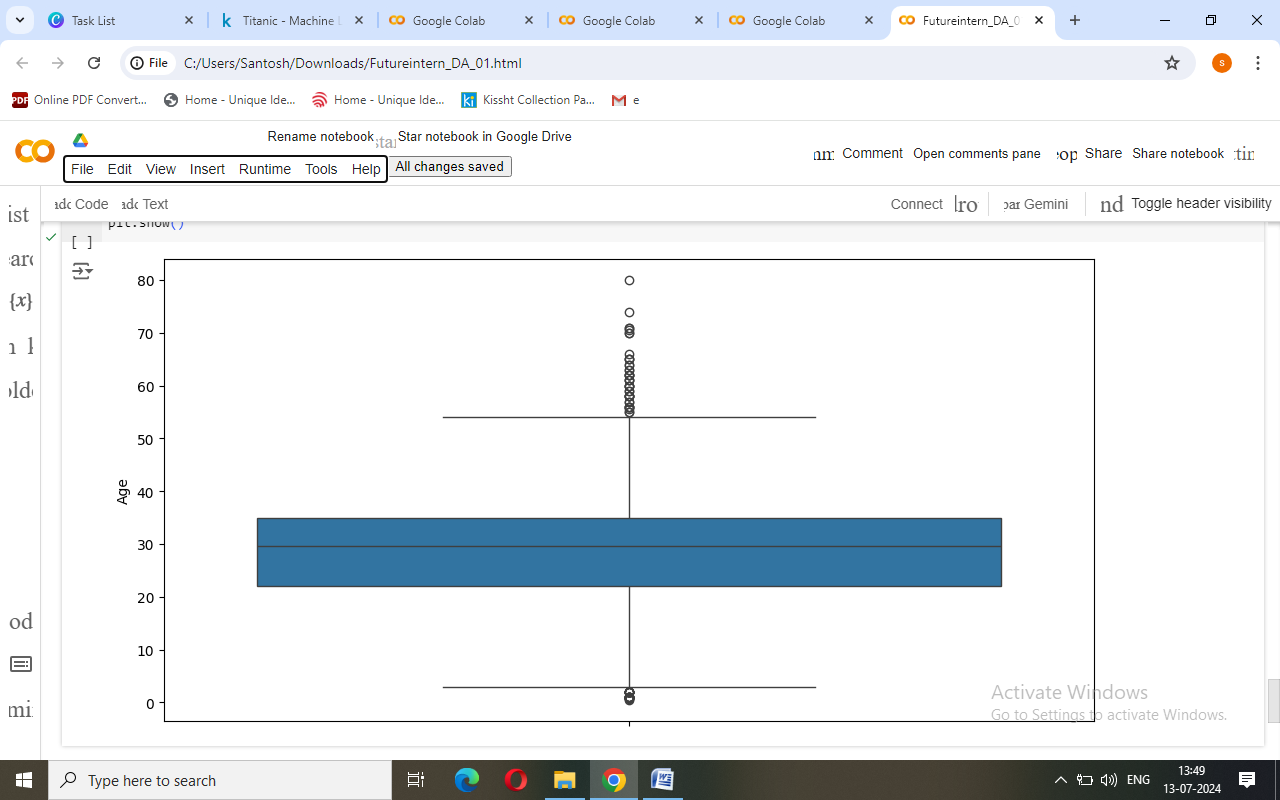
**11. Replace the missing values with mean of the column**

****

**12. outlier**

****

****

****