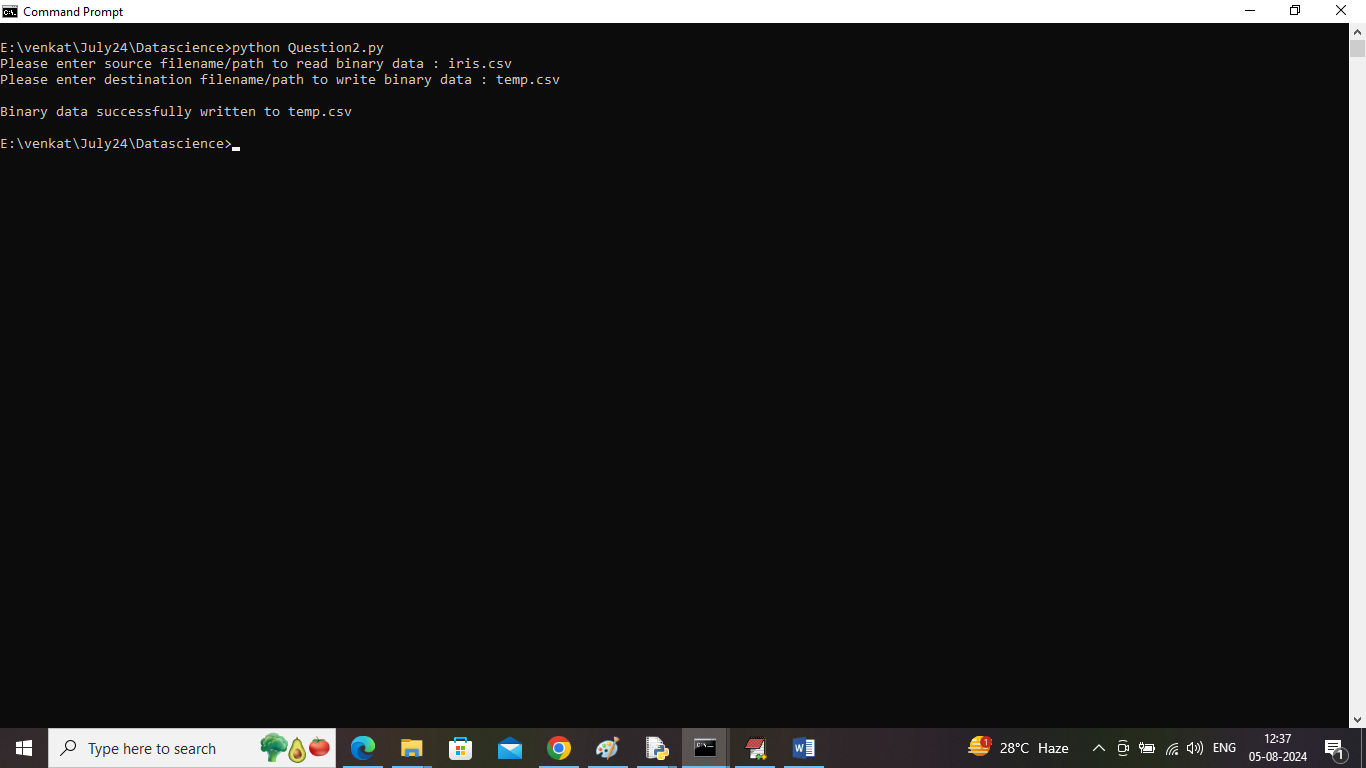
Dataset Science Assignment

In this assignment we need to implement total 7 tasks but I manages to complete 2, 3, 4, 6 and 7. In below screen I am showing output of each task.

Question 2) Program to read and write binary file

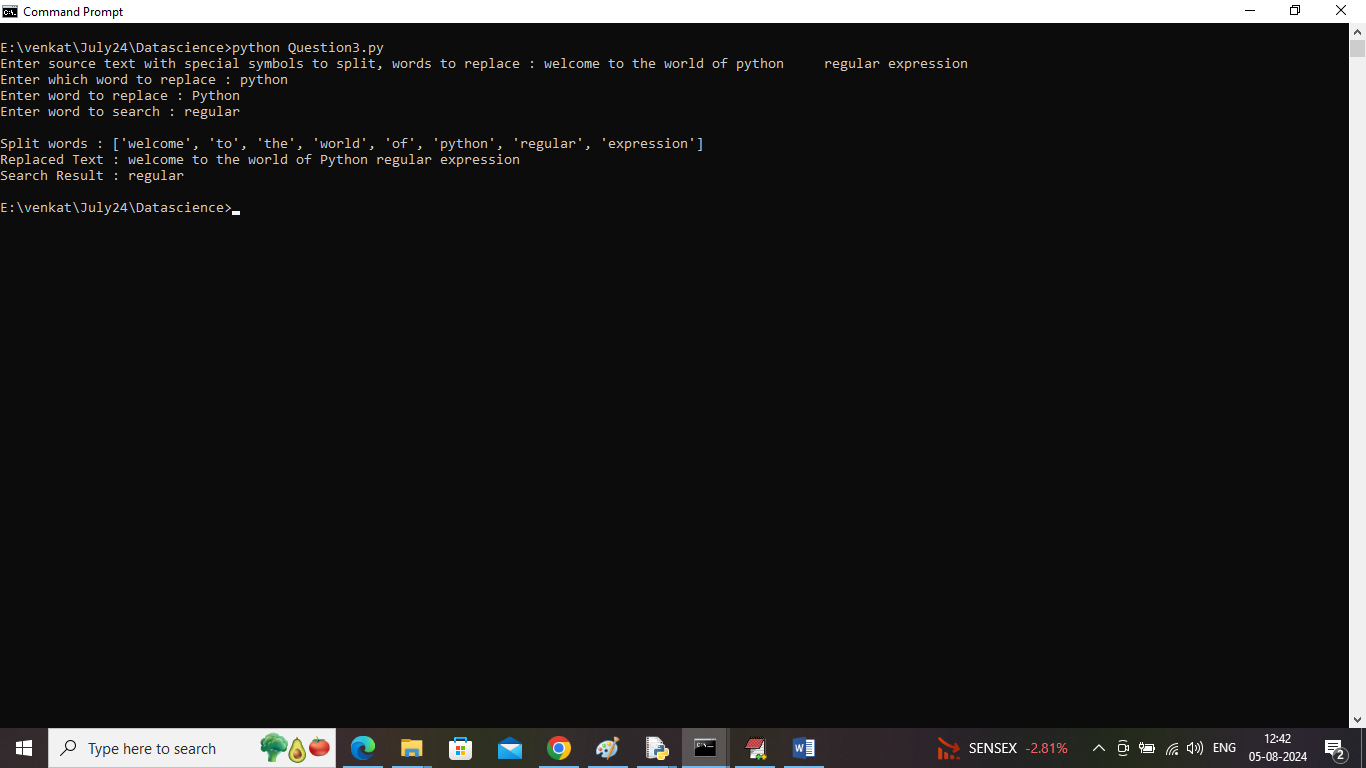
Answer) this program written in a file called Question2.py and below is the execution of this program



In above screen using command ‘python Question2.py’ and then press enter key to input source file path and destination file path and then application will read data in binary format and save to destination file in binary format.

Question 3) write a program to search, split and replace string using pattern matching and regular expression

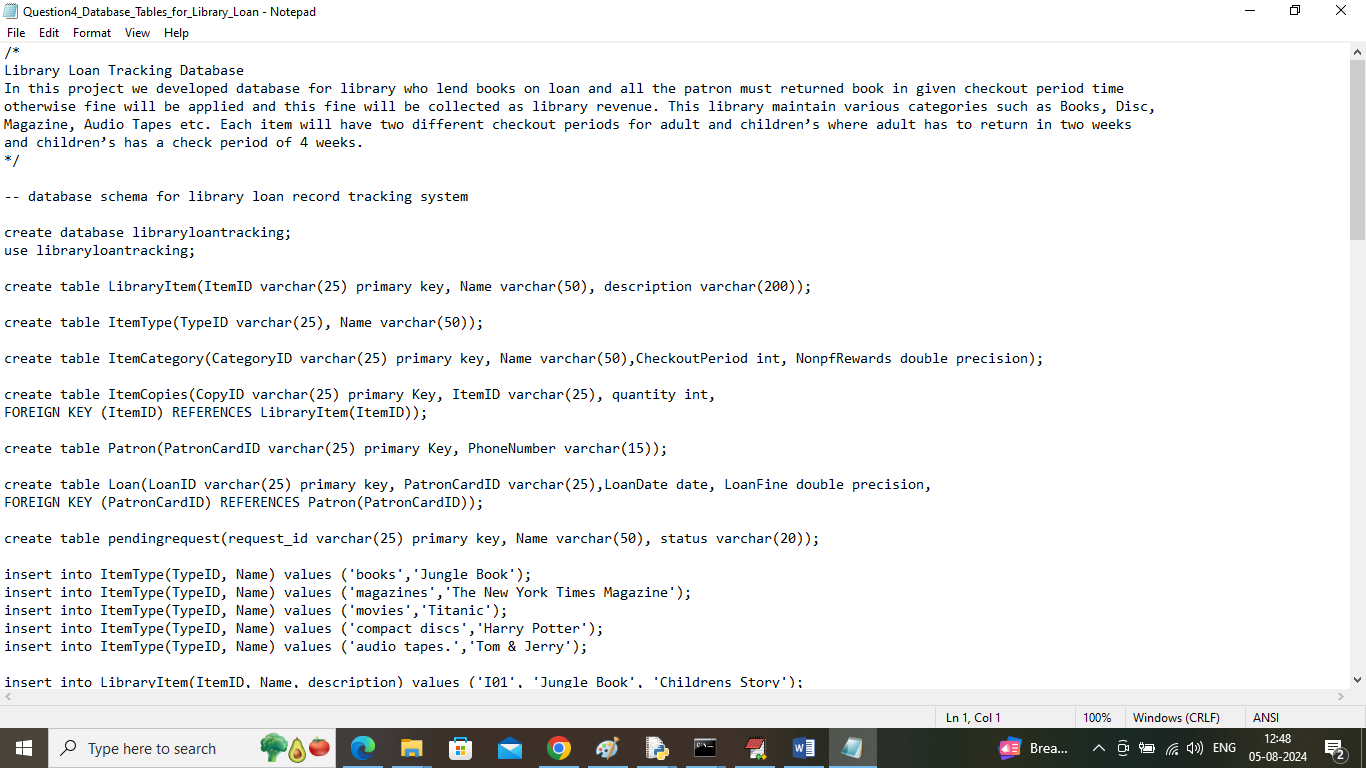
Answer) this program written in a file called Question3.py which will split data using single and multiple spaces and then replace word with another given word and in below screen showing output for this question



In above screen by using command ‘python Question3.py’ and then press enter key and then input some sentence with single or multiple spaces for split and then enter source and destination word to replace and then enter word to search and after giving all input will get above output with split array and then replaced and search work using pattern matching and regular expression

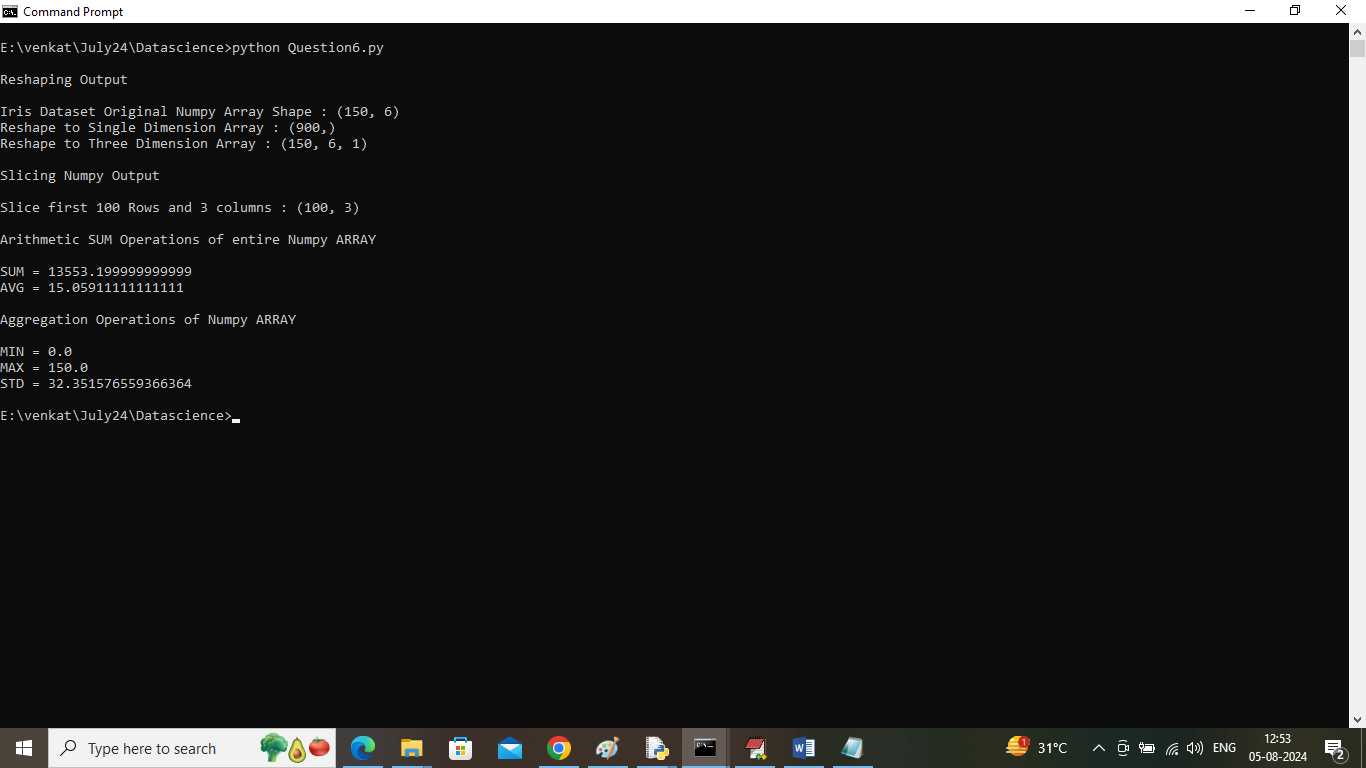
Question 4) designed a relational database for small application using SQL and perform CRUD operations

Answer) for this question I designed database for ‘Library Loan’ project which contains Table creation, insert, search, delete and update queries and in below screen showing screen of all database queries. Just copy entire queries from ‘Question4\_Database\_Tables\_for\_Library\_Loan.txt’ file and paste in MYSQL. In below screen showing database query details



Question 6) create NUMPY array with different shapes, perform arithmetic operation and slicing etc.?

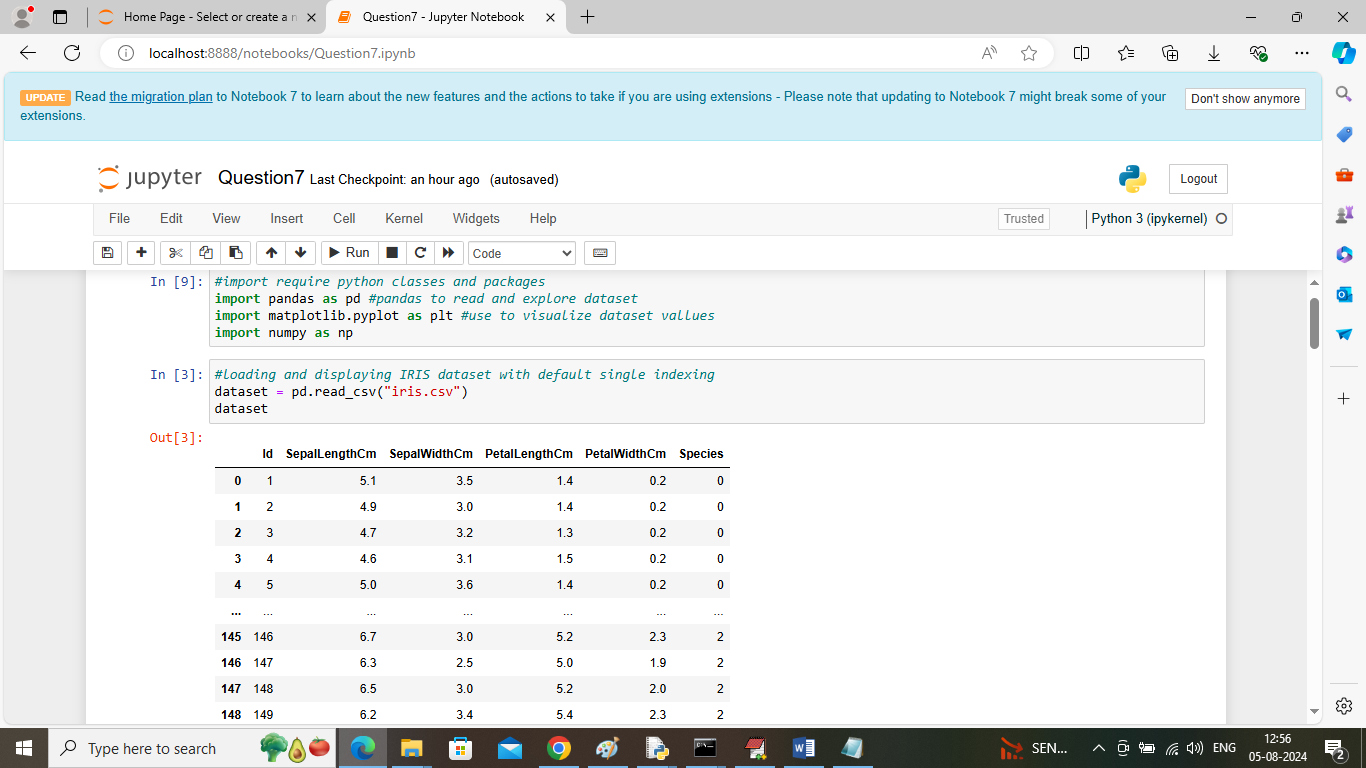
Answer) This program written in a file called ‘Question6.py’ and can execution this program using command ‘python Question6.py’ and then press enter key and then NUMPY will create array from ‘iris.csv’ file and then reshape to different shapes such as single and three dimension. Perform various operations such as SLICING, arithmetic and aggregation operations. In below screen showing output of this question



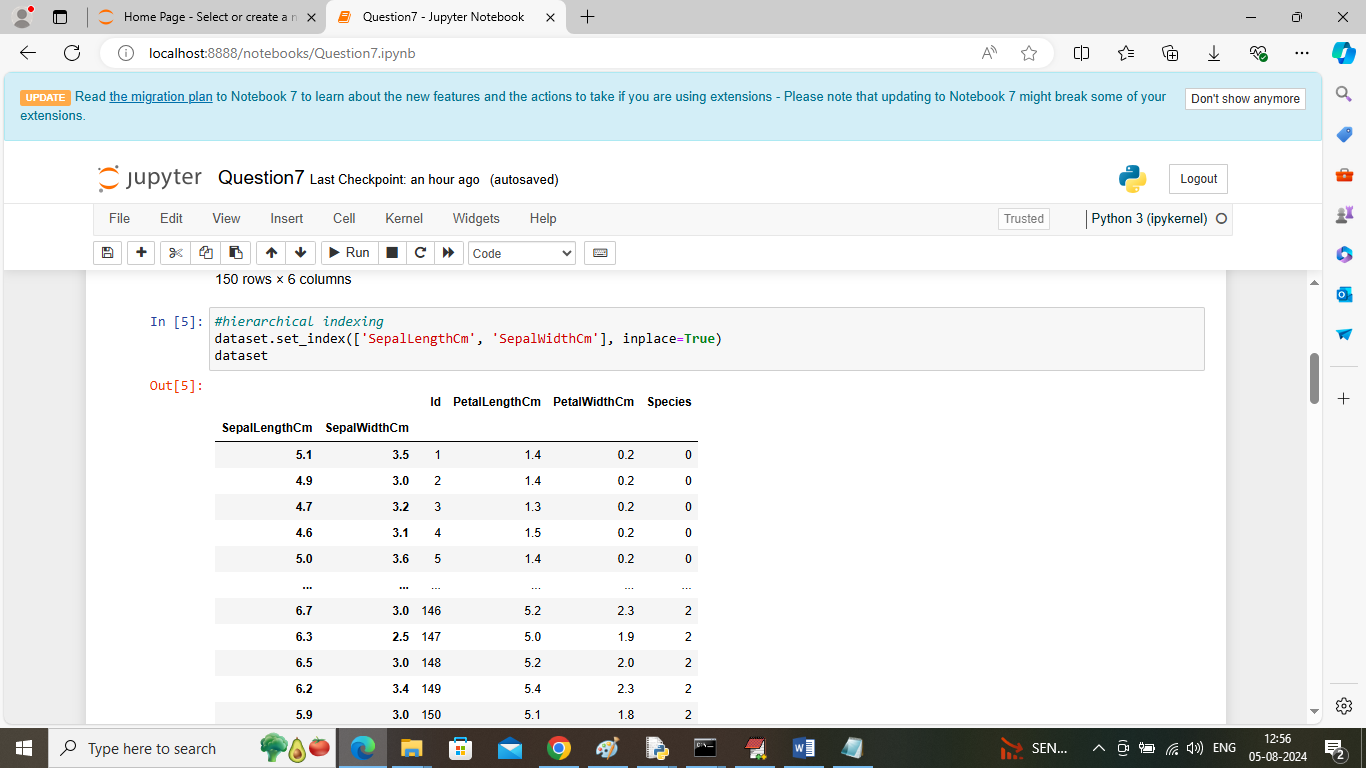
In above screen can see NUMPY array creation, reshaping, slicing, arithmetic and aggregation output.

Question 7) program for Pandas data structure with single and hierarchical indexing, handling missing values, arithmetic and plotting operations?

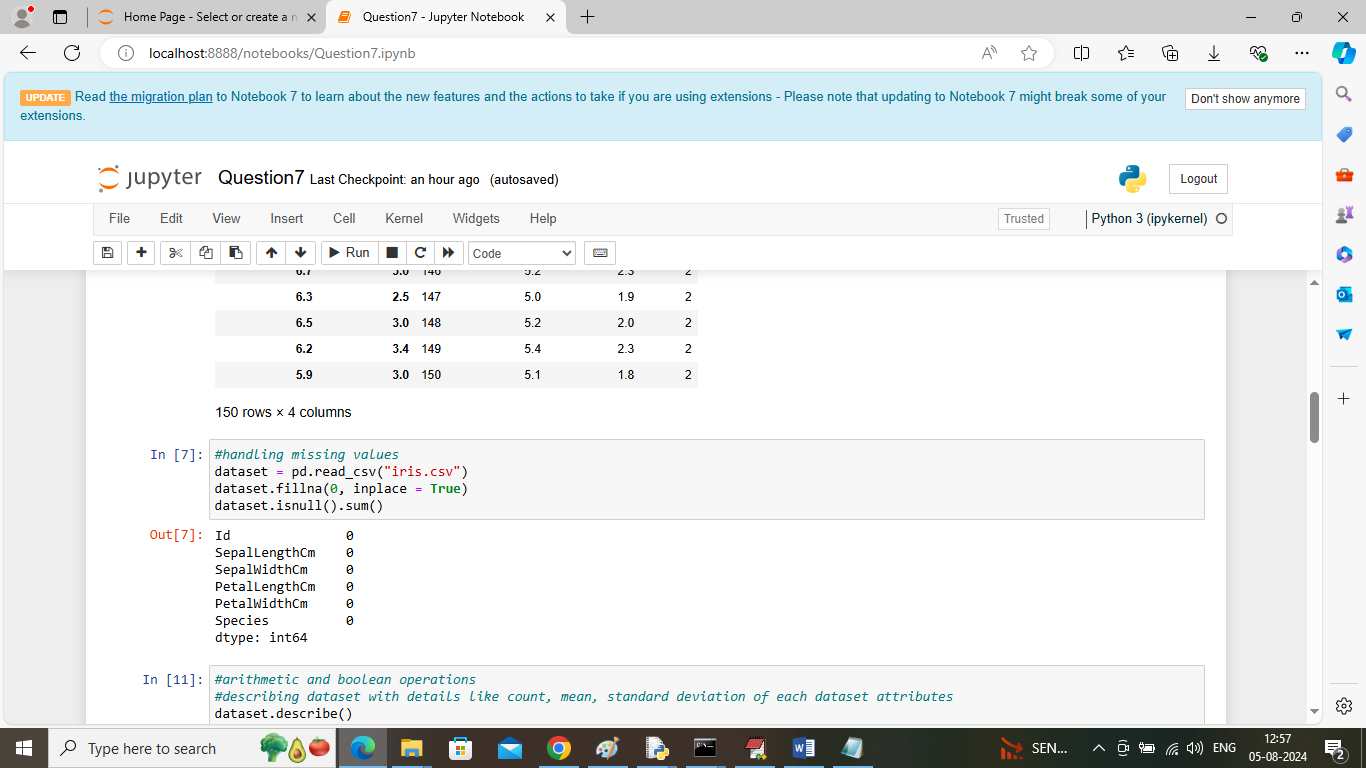
Answers) this program implementation we did using JUPYTER notebook and then implement all points. In below screen showing output for each points



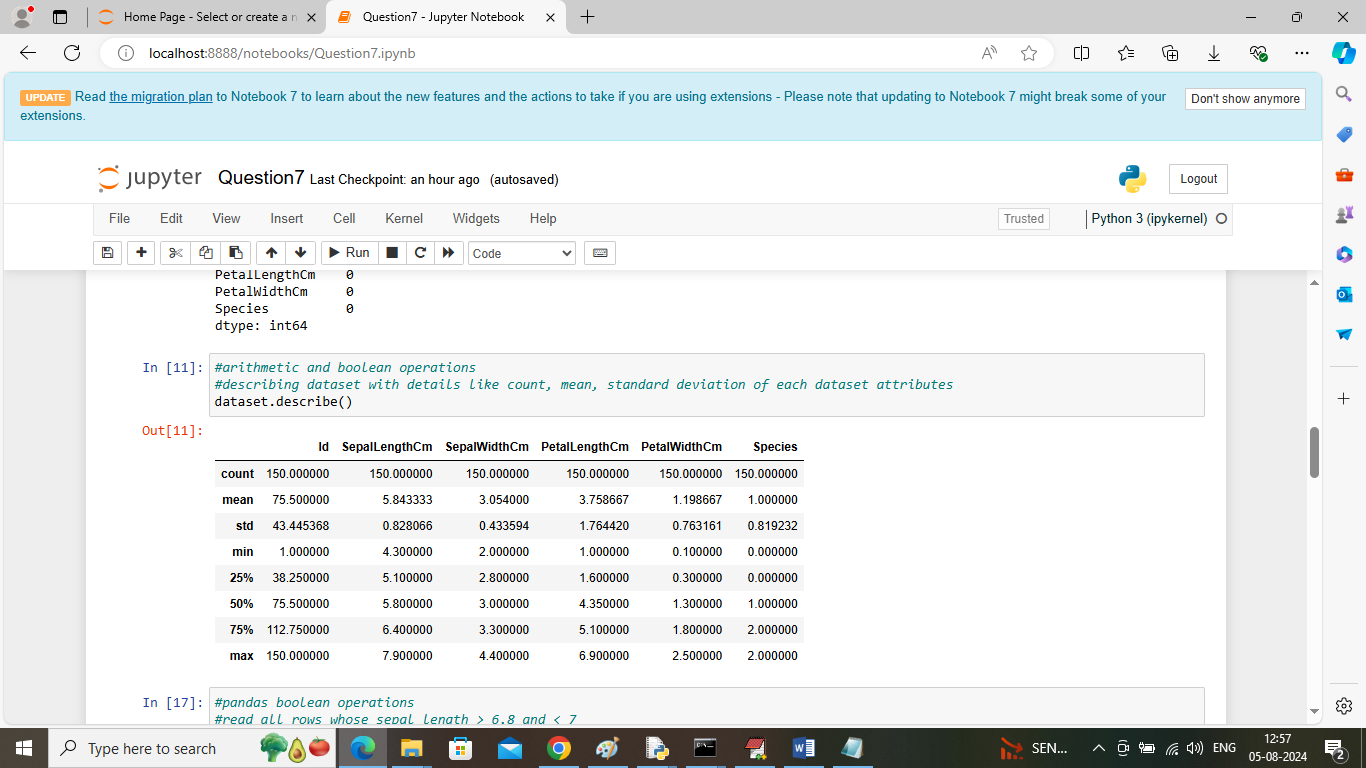
In above screen loading and displaying dataset with default single indexing



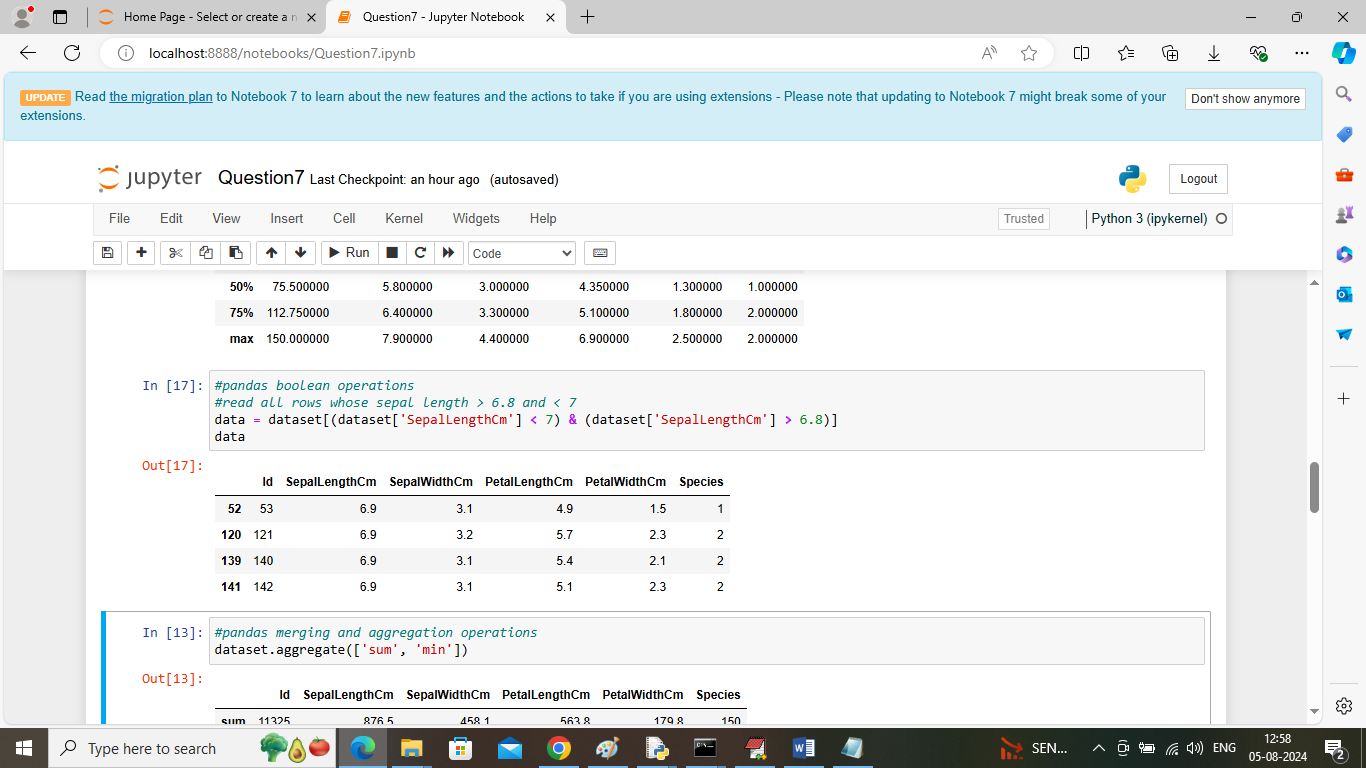
In above screen performing ‘Hierarchical’ indexing with multiple columns and then printing new index table



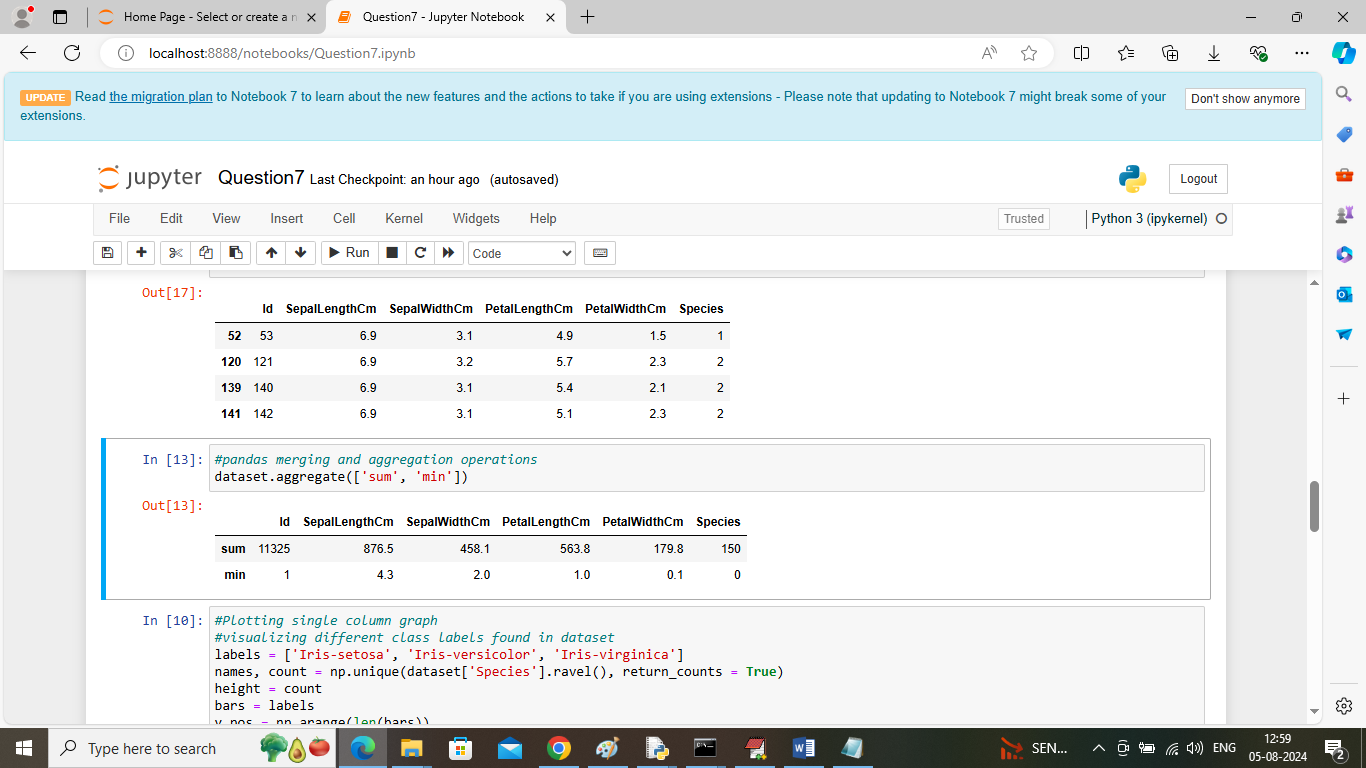
In above screen handling and replacing missing values and then printing sum of all missing values which is 0 as we already handle missed values



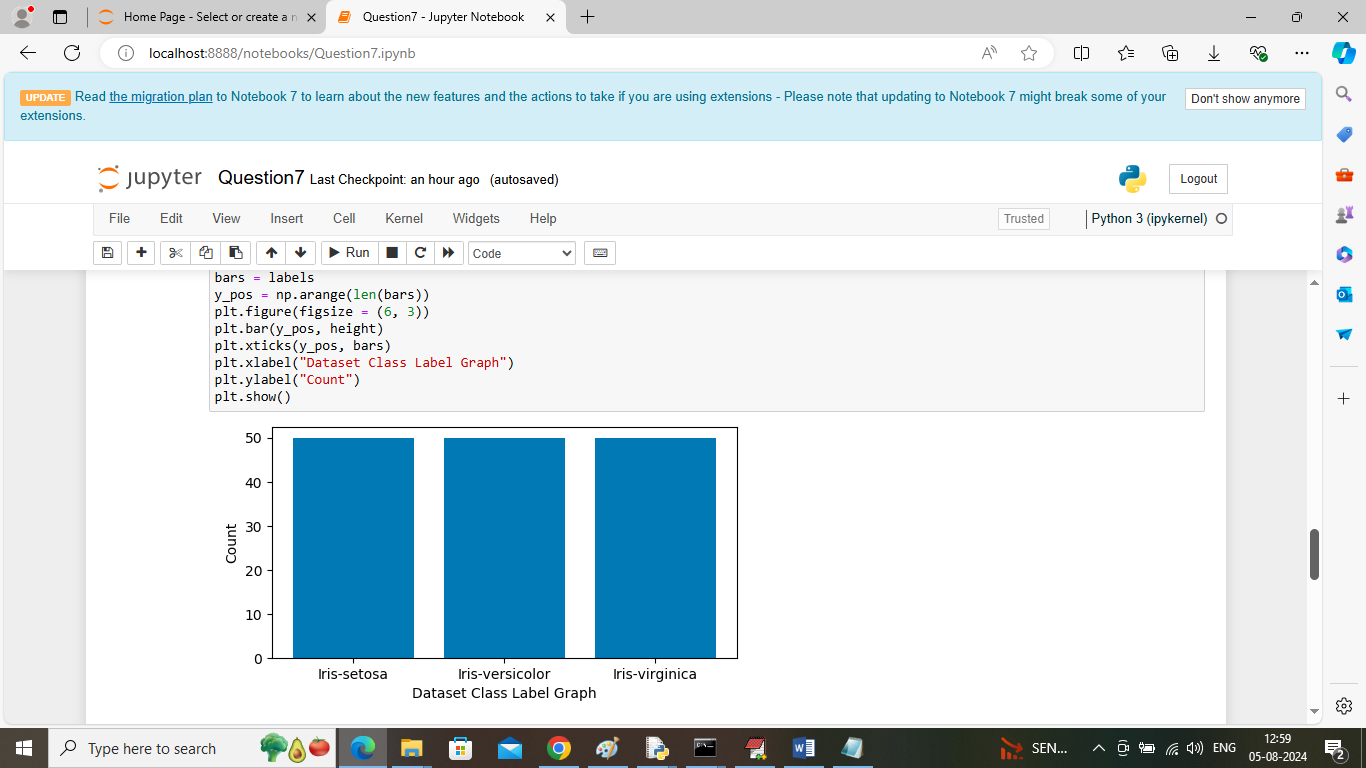
In above screen describing dataset to print all arithmetic operations such as Min, Max, count, average and many more



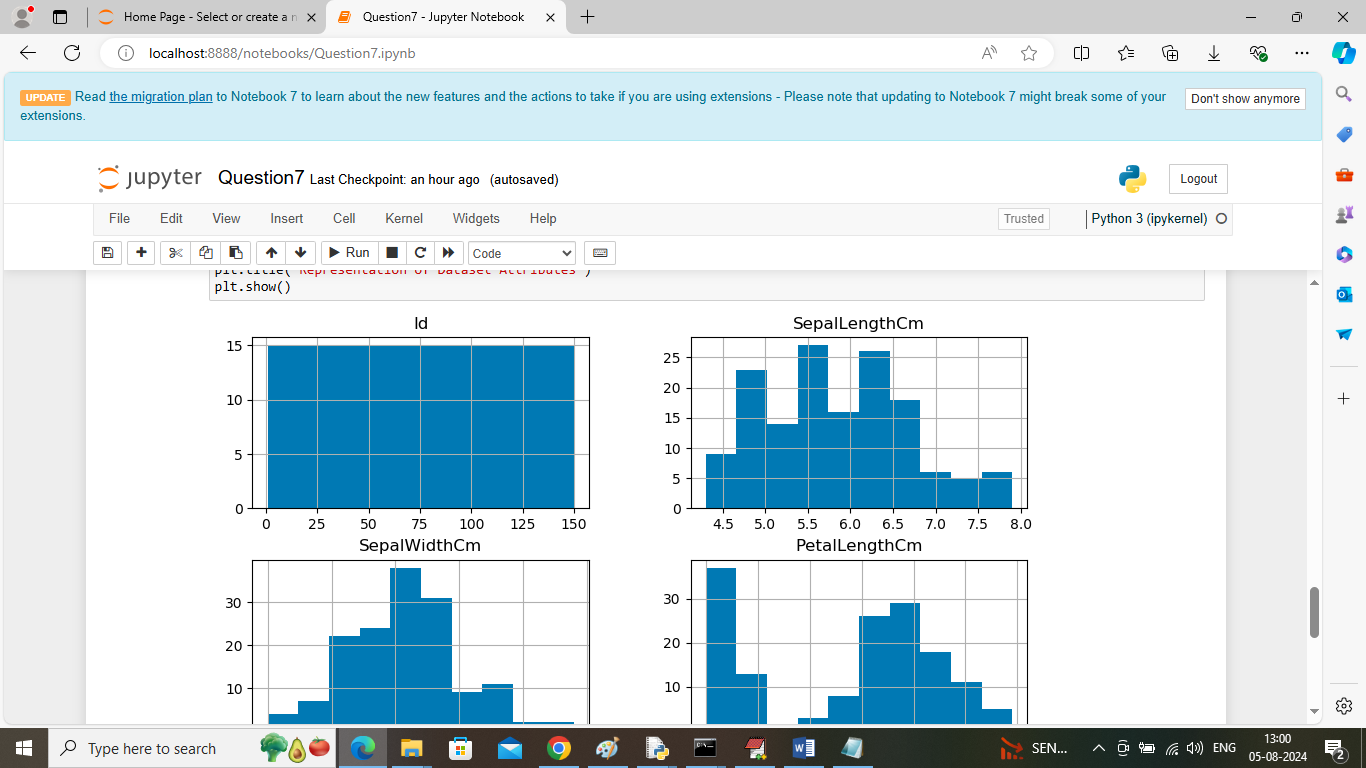
In above screen performing logical operation by selecting all rows whose sepal length fall between given two values

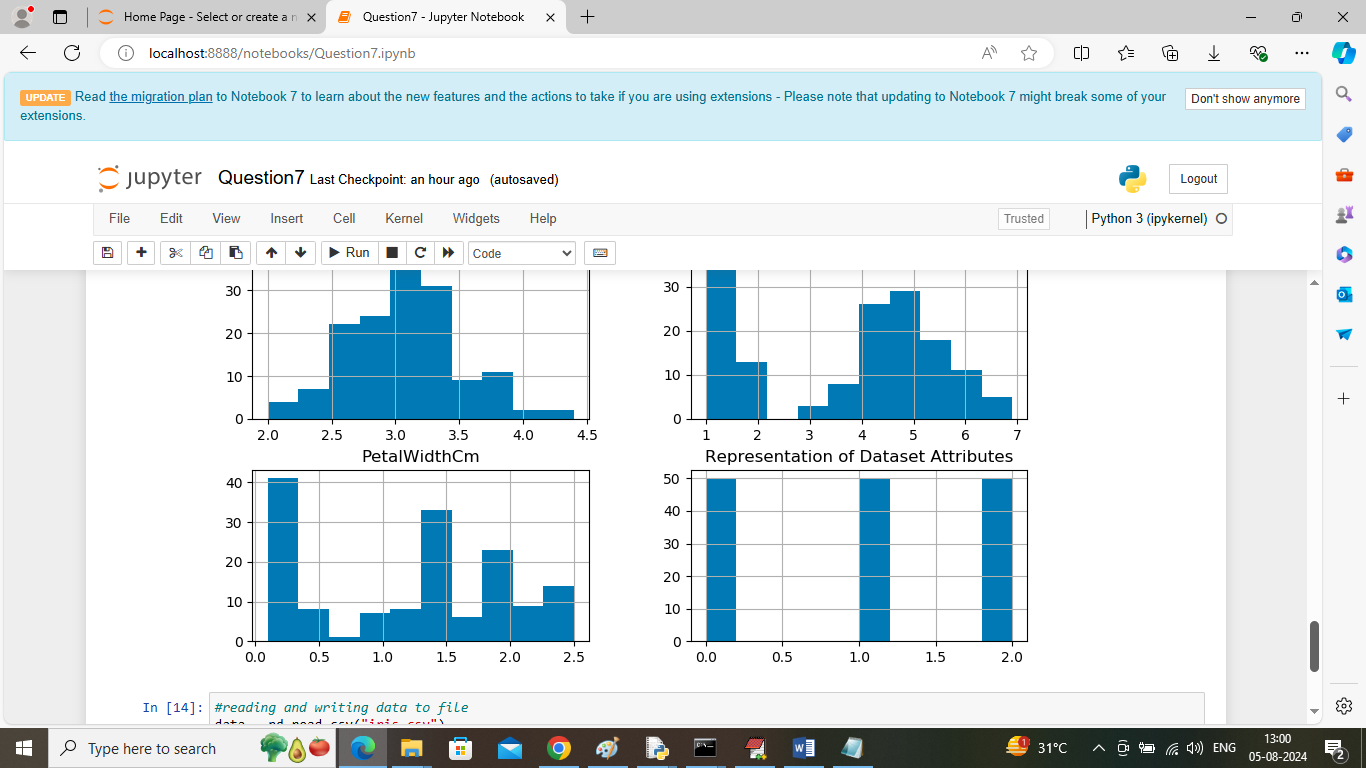


In above screen performing aggregation operations

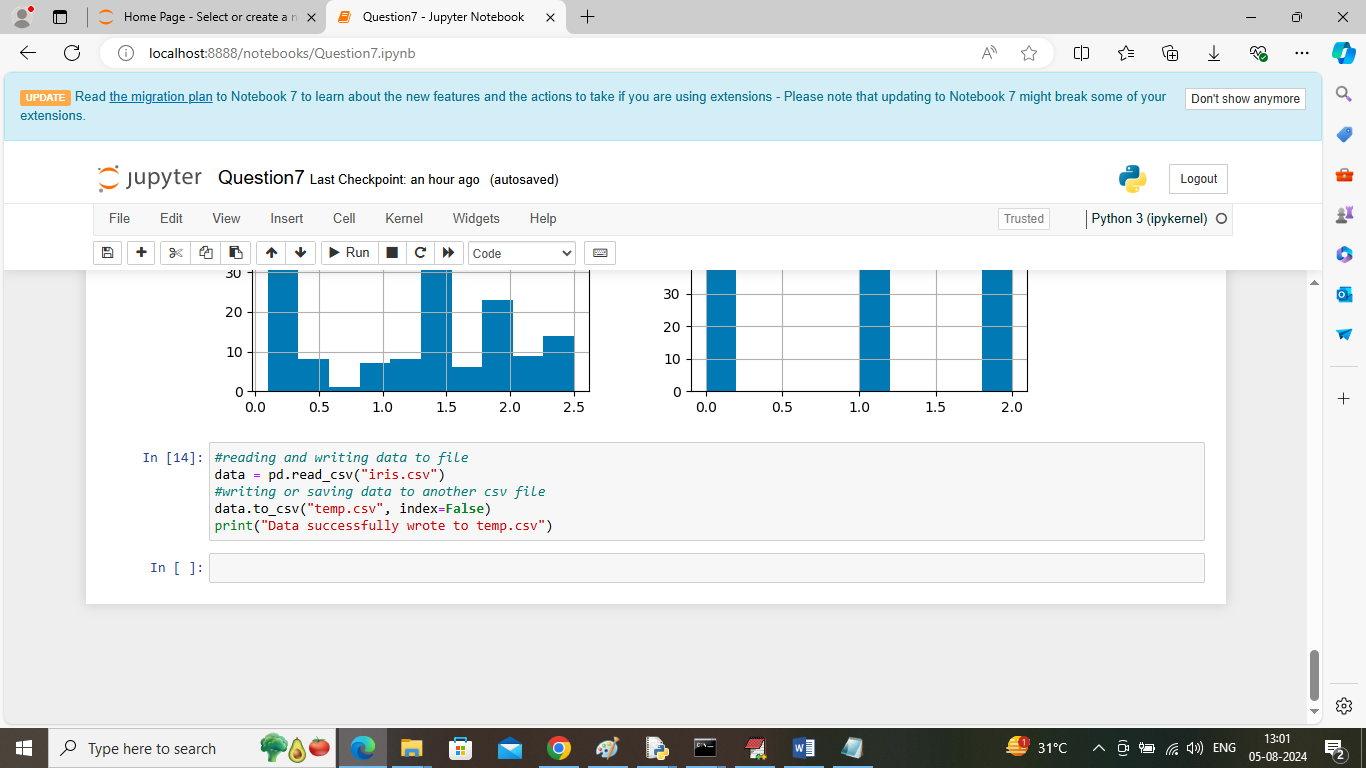


In above screen visualizing single column from Pandas table where x-axis represents class labels from IRIS dataset and y-axis represents counts of each class labels





In above two screens visualizing all columns from pandas table where x-axis represents unique values from dataset and y-axis represents ranges of those values



In last block reading and writing data to another file.

So in above screens we shown output of 5 tasks.