**Spring 2024: CS5720 Neural Networks & Deep Learning - ICP-1**

**Assignment-4**

**Praneeth Kumar** (700747594)

GITHUB LINK: https://github.com/praneethk0910/Neural-Networks-Deep-Learning---ICP-4.git

**1. Data Manipulation**

a. Read the provided CSV file ‘data.csv’.

b.https://drive.google.com/drive/folders/1h8C3mLssosIOLsvoYwPLzy2fJ4IOF?usp=sharing

c. Show the basic statistical description about the data.

d. Check if the data has null values.

i. Replace the null values with the mean

e. Select at least two columns and aggregate the data using: min, max, count, mean.

f. Filter the data frame to select the rows with calories values between 500 and 1000.

g. Filter the data frame to select the rows with calories values > 500 and pulse < 100.

h. Create a new “df\_modified” data frame that contains all the columns from df except for

“Maxpulse”.

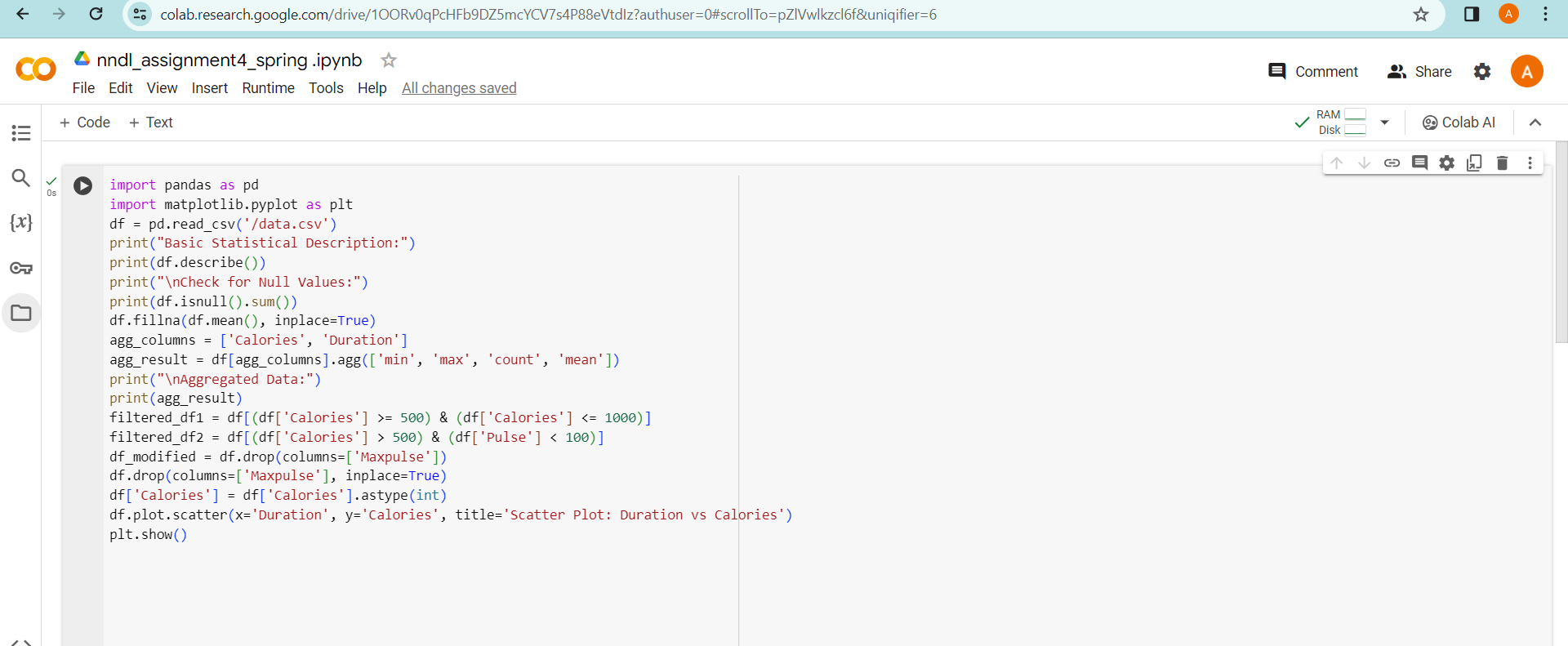
i. Delete the “Maxpulse” column from the main df data frame

j. Convert the datatype of Calories column to int datatype.

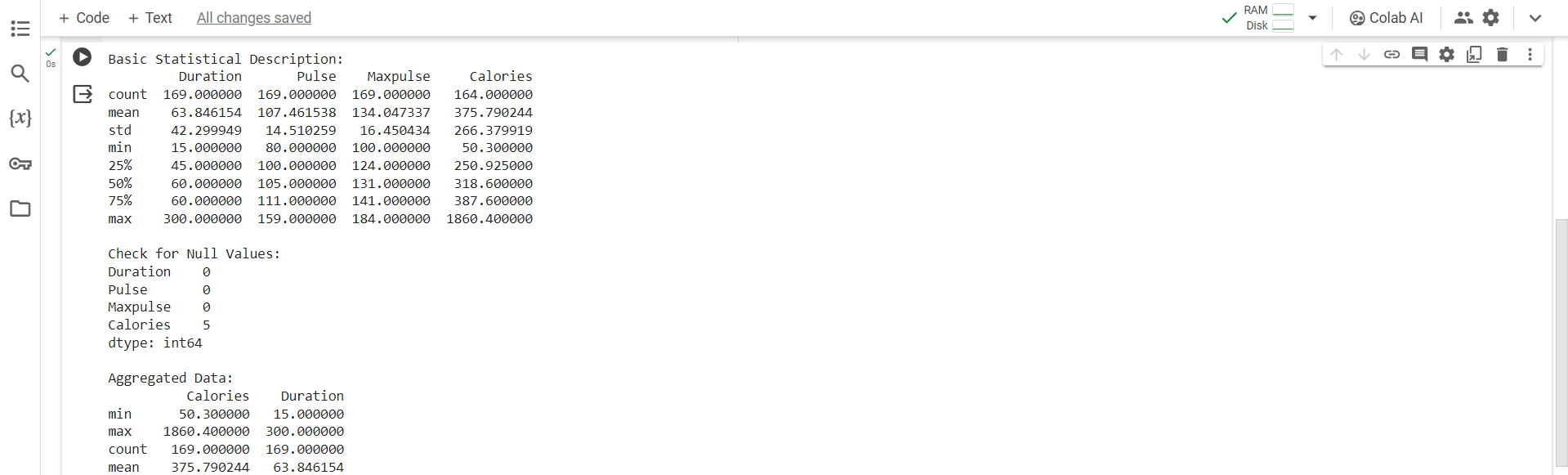
k. Using pandas create a scatter plot for the two columns (Duration and Calories).

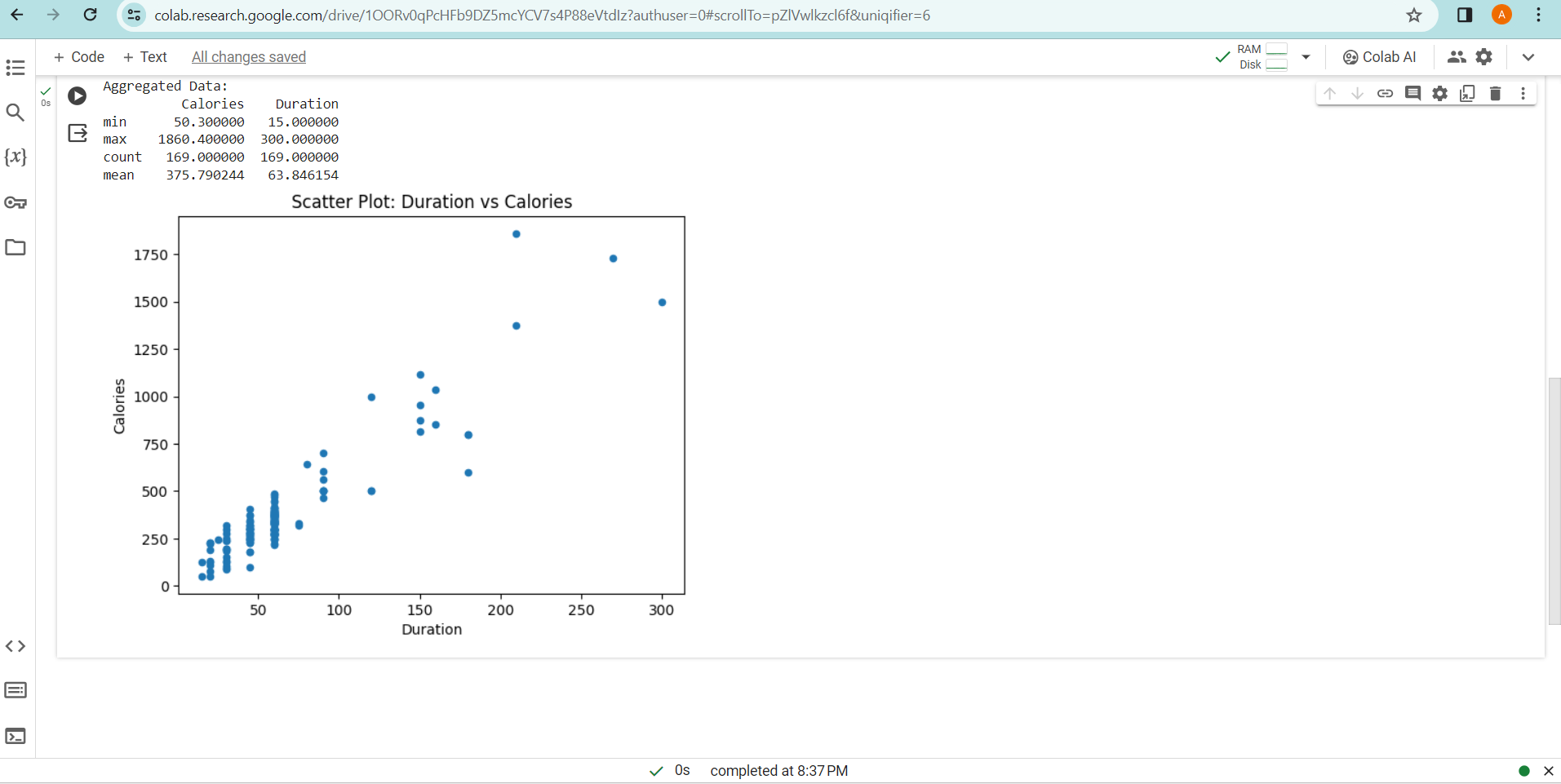
Example.

**Source code:**



**Output:**





**2. Linear Regression**

a) Import the given “Salary\_Data.csv”

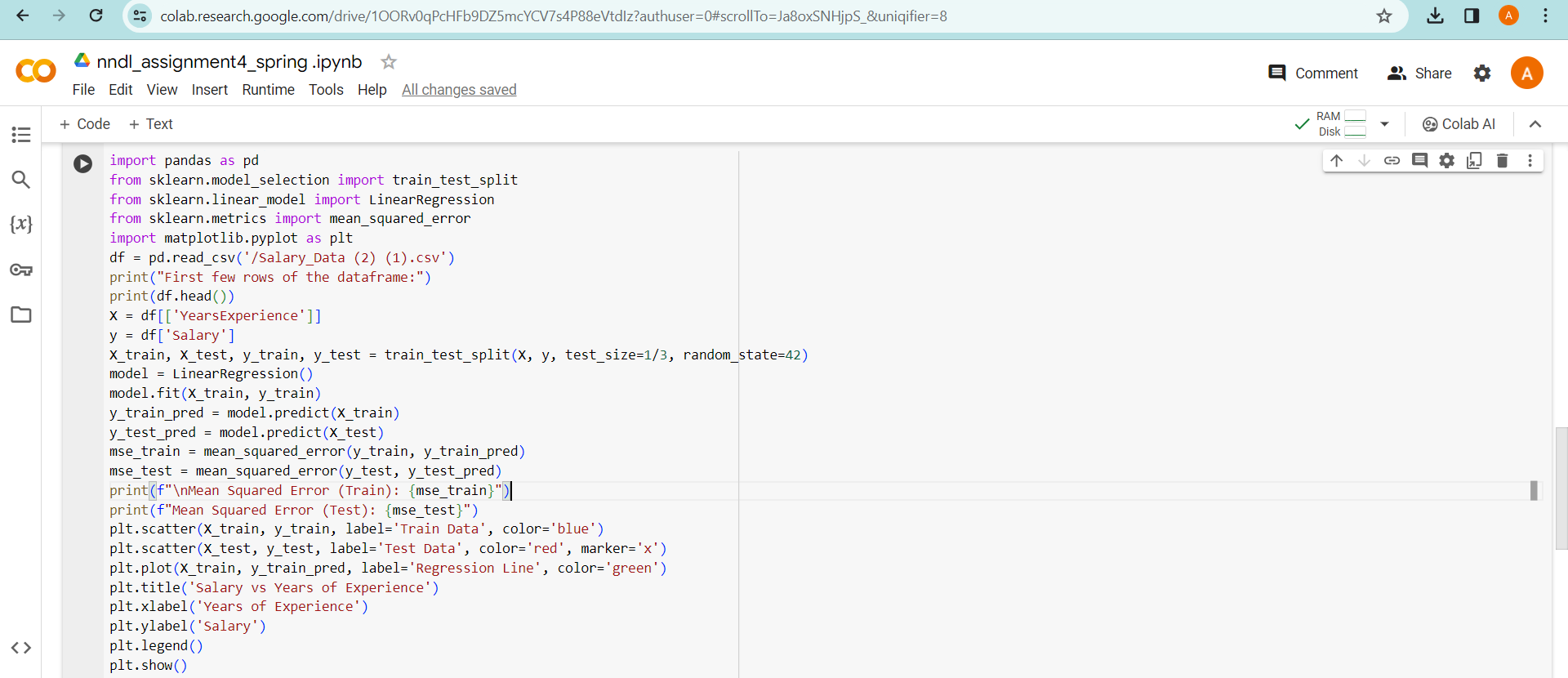
b) Split the data in train test partitions, such that 1/3 of the data is reserved as test subset.

c) Train and predict the model.

d) Calculate the mean squared error

e) Visualize both train and test data using scatter plot.

**Source code:**



**Output:**

