# Fall 2023: CS5720 Neural Networks & Deep Learning - ICP-4 Assignment-4

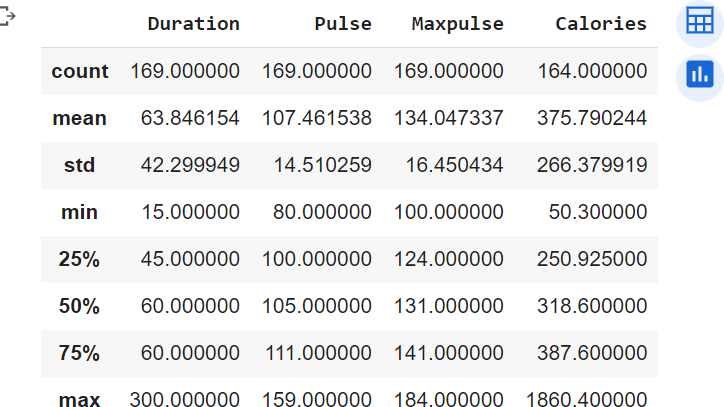
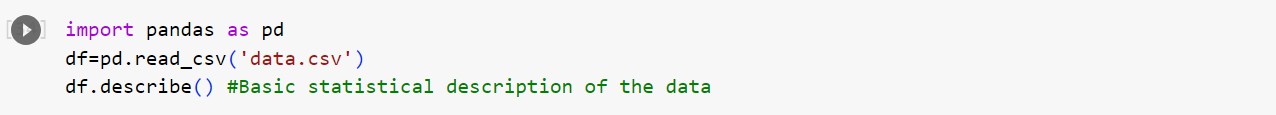
**NAME: Akhila Katepalli**

**STUDENT ID:700745186**

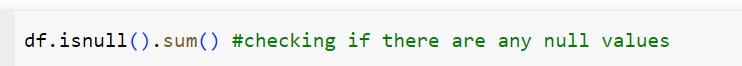
Github Link:https://github.com/Akhilakatepalli/ICP4/upload/main

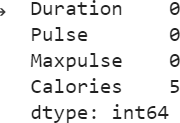
# Data Manipulation

* 1. Read the provided CSV file ‘data.csv’.
  2. https://drive.google.com/drive/folders/1h8C3mLsso-R-sIOLsvoYwPLzy2fJ4IOF?usp=sharing
  3. Show the basic statistical description about the data.

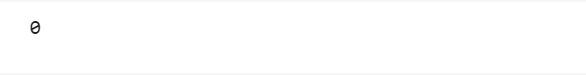
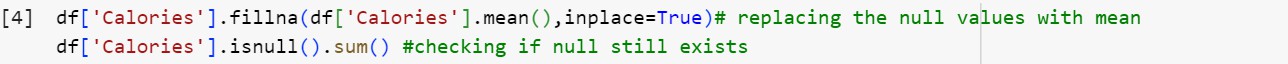


* 1. Check if the data has null values.

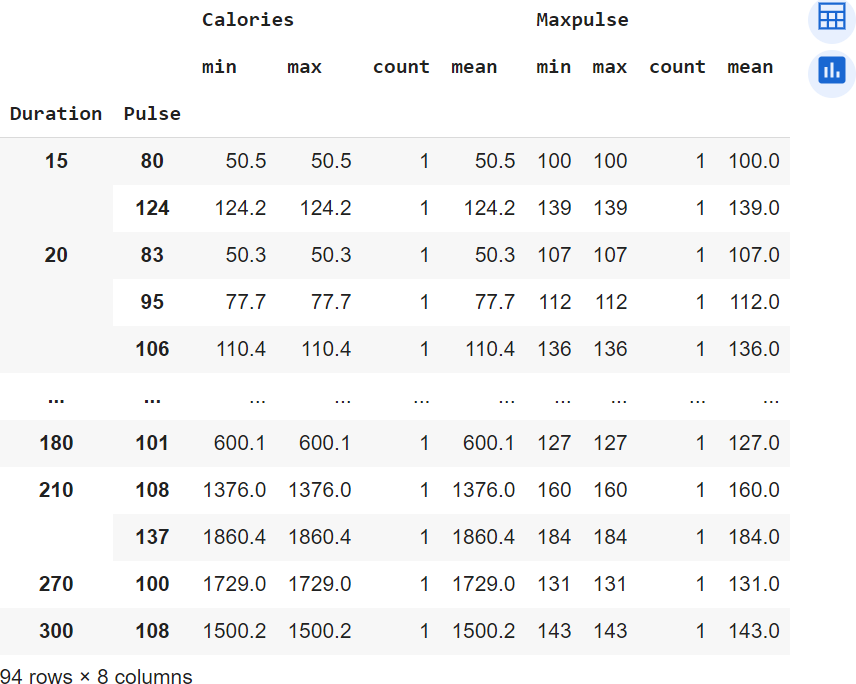
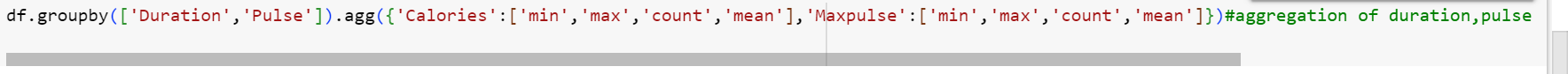




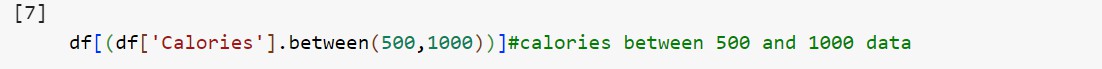
i. Replace the null values with the mean

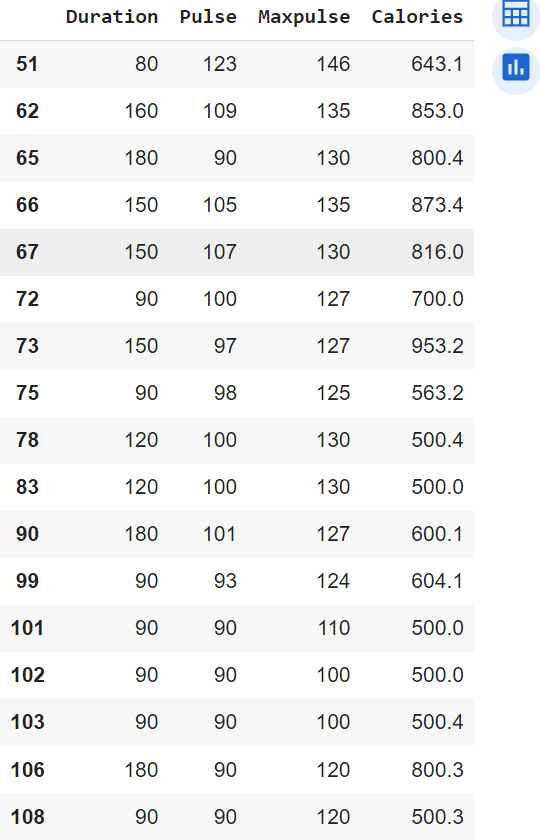


* 1. Select at least two columns and aggregate the data using: min, max, count,mean.



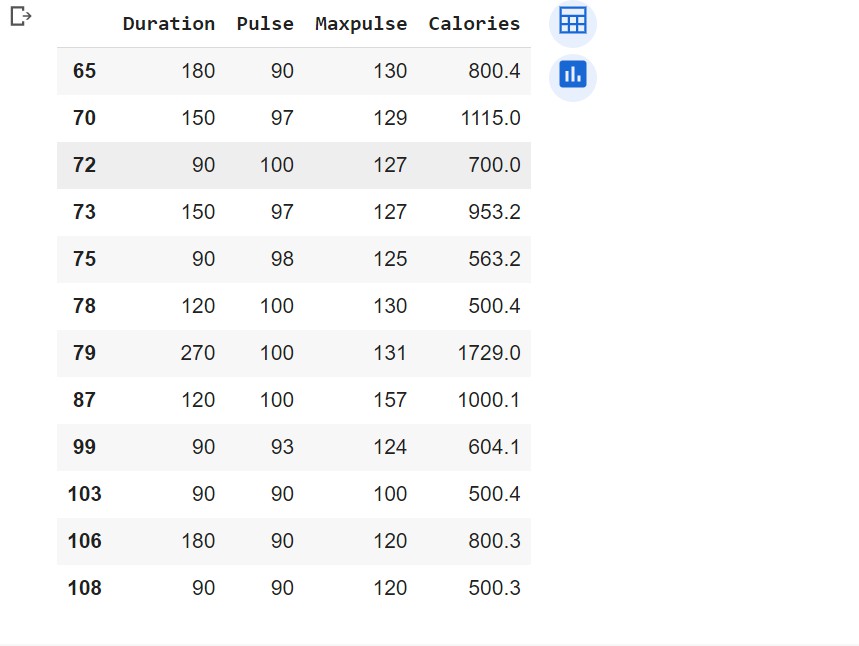
* 1. Filter the dataframe to select the rows with calories values between 500 and1000.

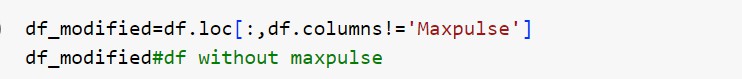


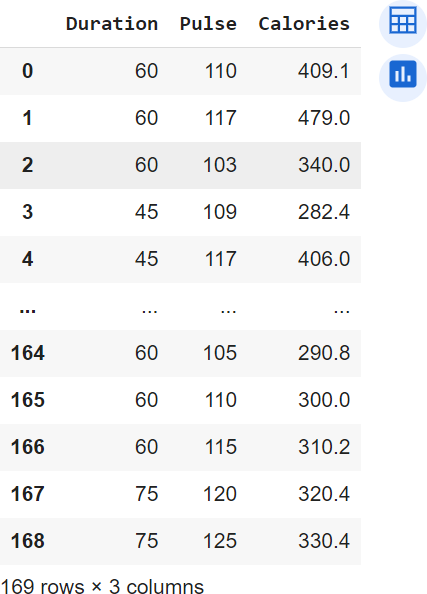


* 1. Filter the dataframe to select the rows with calories values > 500 and pulse <100.

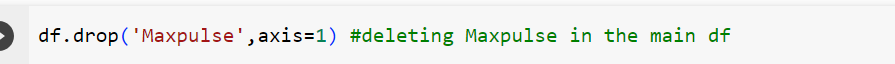


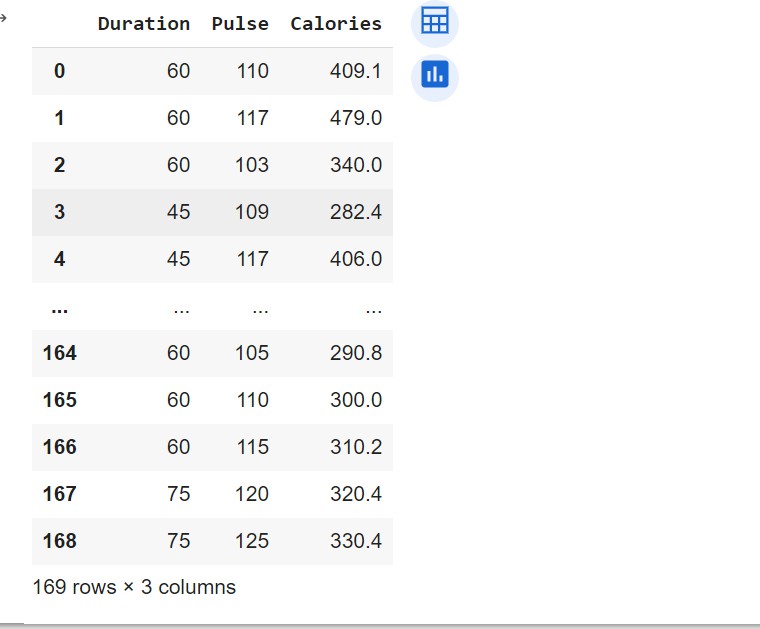


* 1. Create a new “df\_modified” dataframe that contains all the columns from df exceptfor “Maxpulse”.

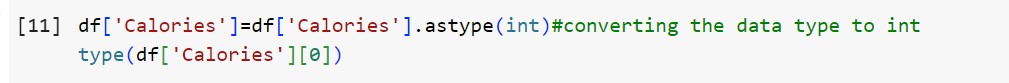


* 1. Delete the “Maxpulse” column from the main df dataframe



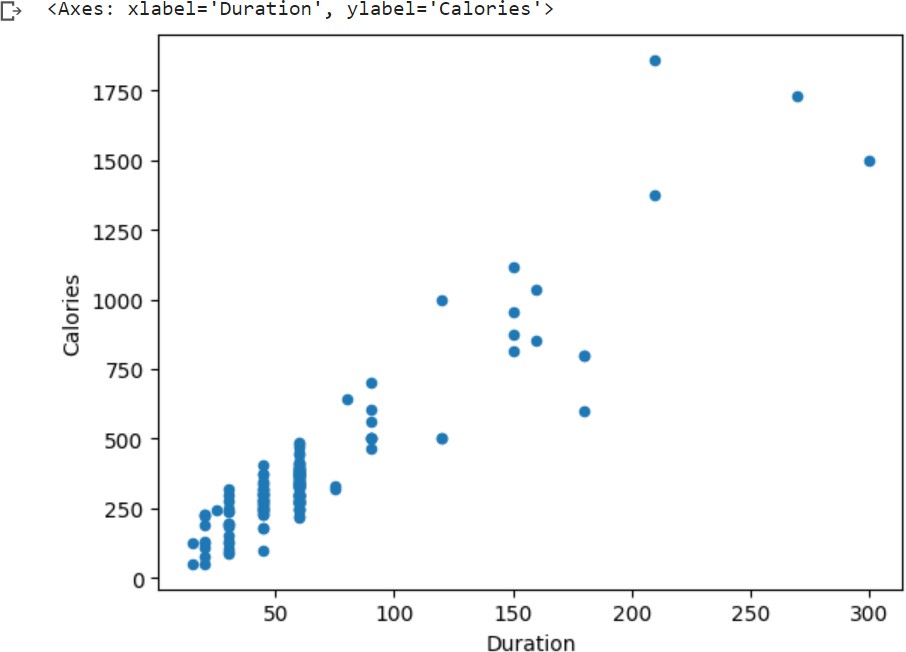


* 1. Convert the datatype of Calories column to int datatype.



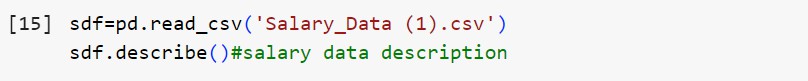
* 1. Using pandas create a scatter plot for the two columns (Duration andCalories).



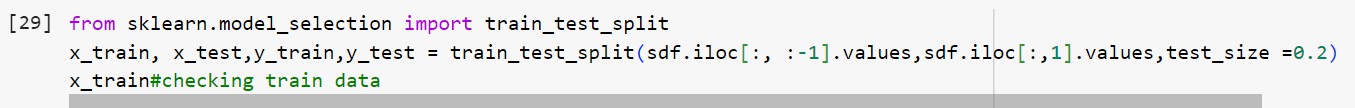


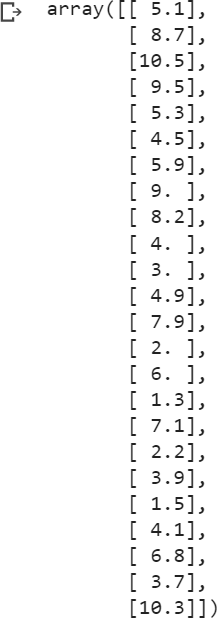
# Linear Regression

1. Import the given “Salary\_Data.csv”

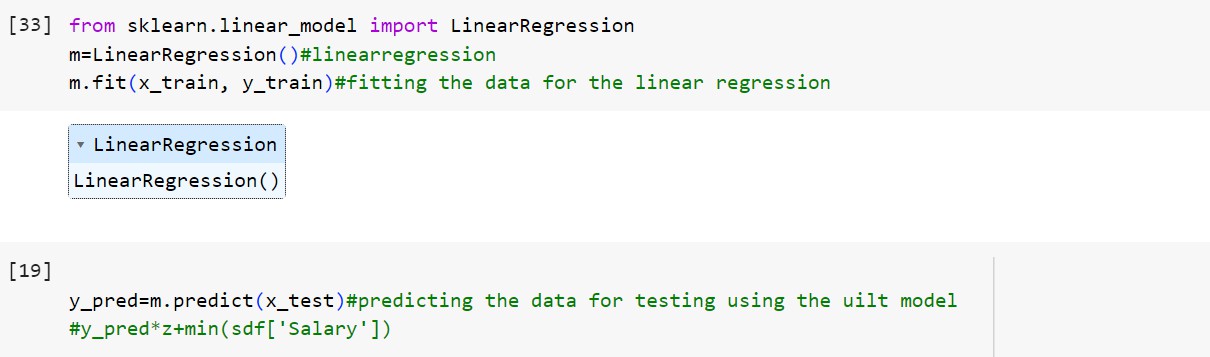


1. Split the data in train\_test partitions, such that 1/3 of the data is reserved as test subset.

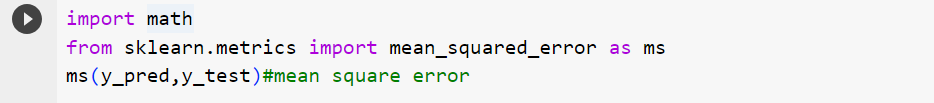




1. Train and predict the model.



1. Calculate the mean\_squared error



1. Visualize both train and test data using scatter plot.

