**Spring 2024: CS5720 – NN &DL**

**In-Class Programming Assignment-4**

**NAME: Bala Rishik Marneni**

**STUDENT ID: 700746746**

**Video Link:** [**https://drive.google.com/file/d/1f06k2SAZJge\_oxUNx71fBz7HWcBitjBF/view?usp=drive\_link**](https://drive.google.com/file/d/1f06k2SAZJge_oxUNx71fBz7HWcBitjBF/view?usp=drive_link) **Github Link: [https://github.com/BalaRishik001/Neural-Networks-and-Deep-Learning-Assignments](https://d.docs.live.net/a4167d02bd5720cb/Documents/700746746%20ICP4%20NNDL.docx)**

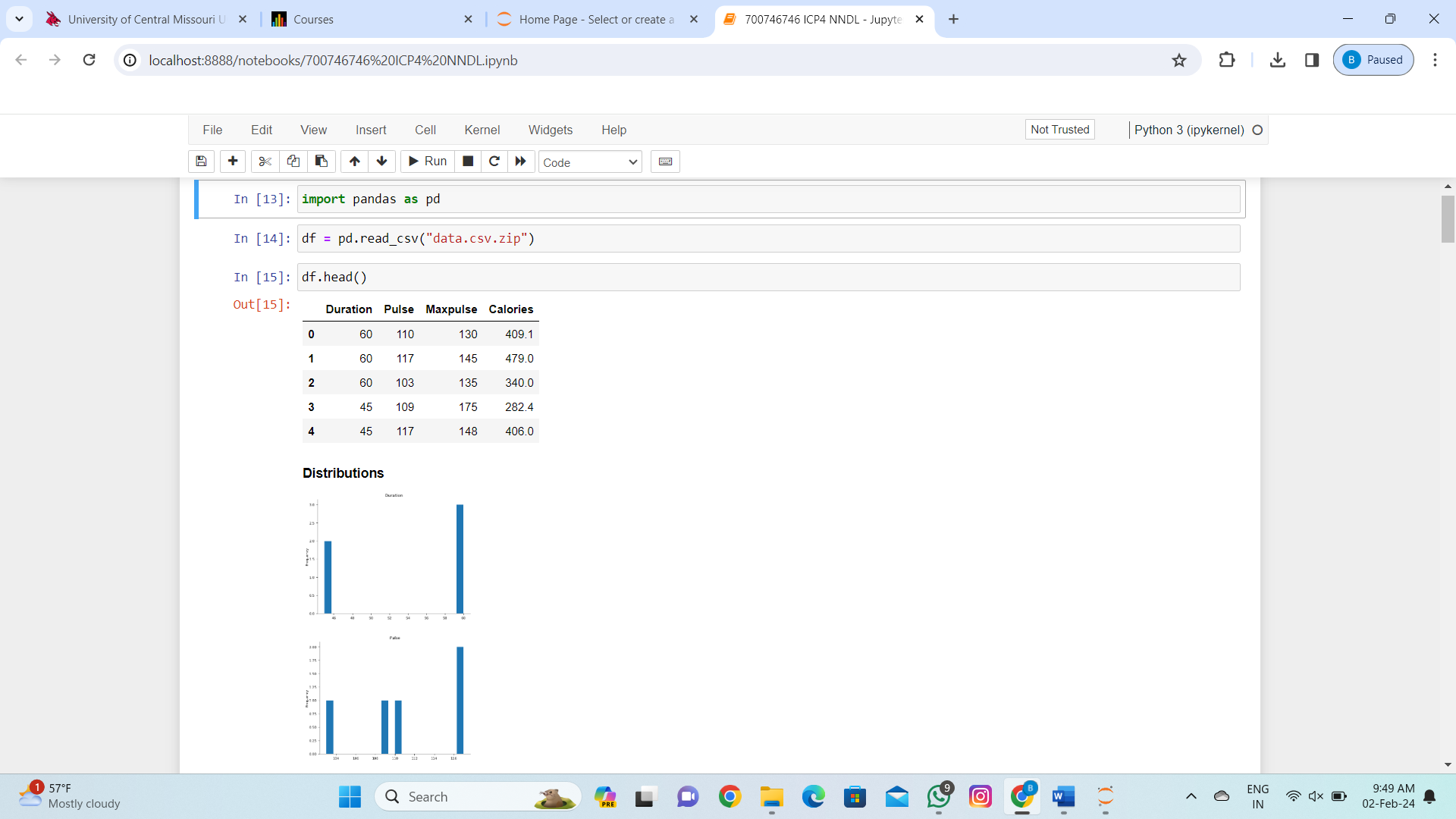
1. **Data Manipulation** 
   * 1. Read the provided CSV file ‘data.csv’.
     2. <https://drive.google.com/drive/folders/1h8C3mLsso-R-sIOLsvoYwPLzy2fJ4IOF?usp=sharing>
     3. Show the basic statistical description about the data.
     4. Check if the data has null values.

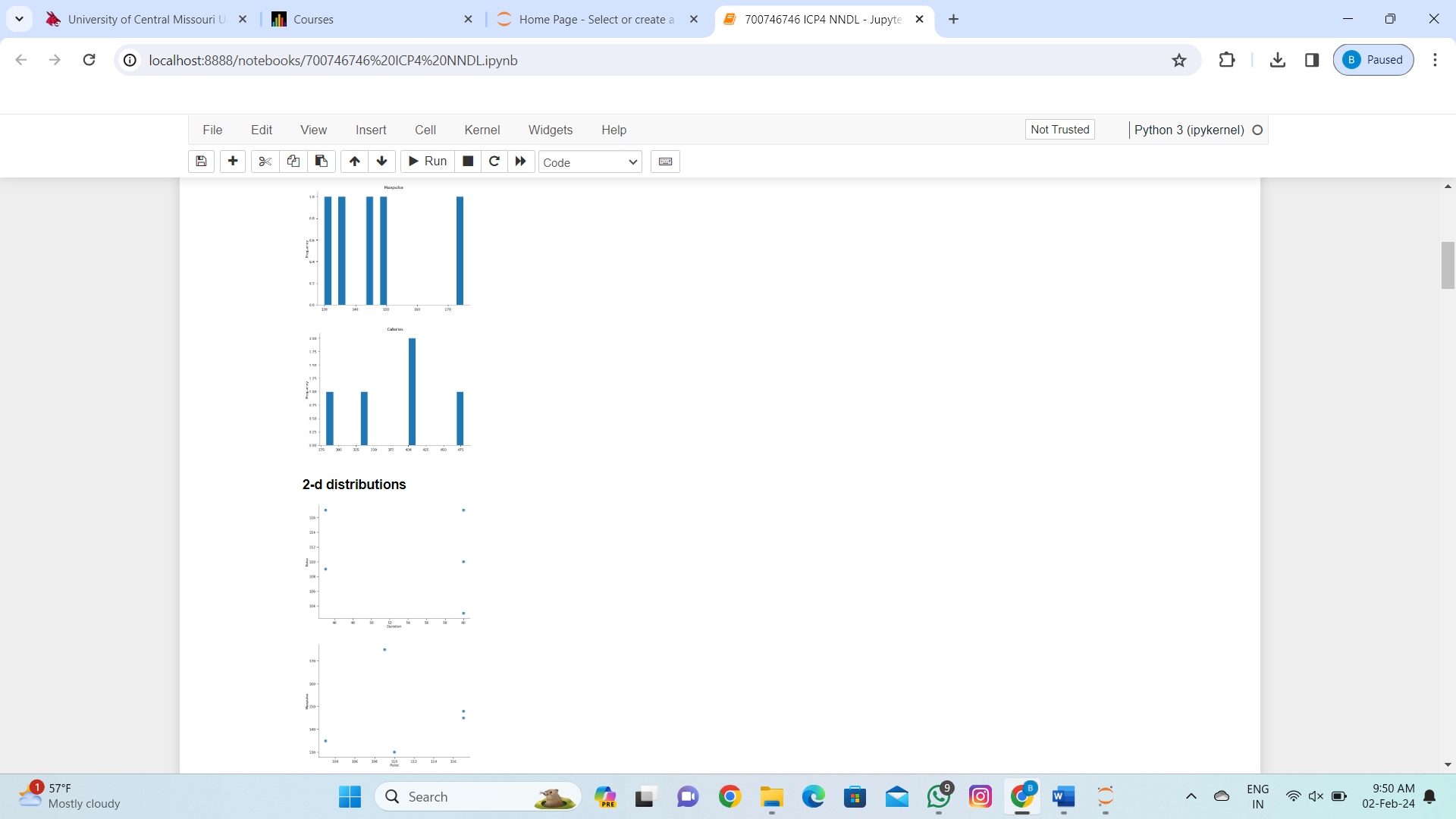
i. Replace the null values with the mean

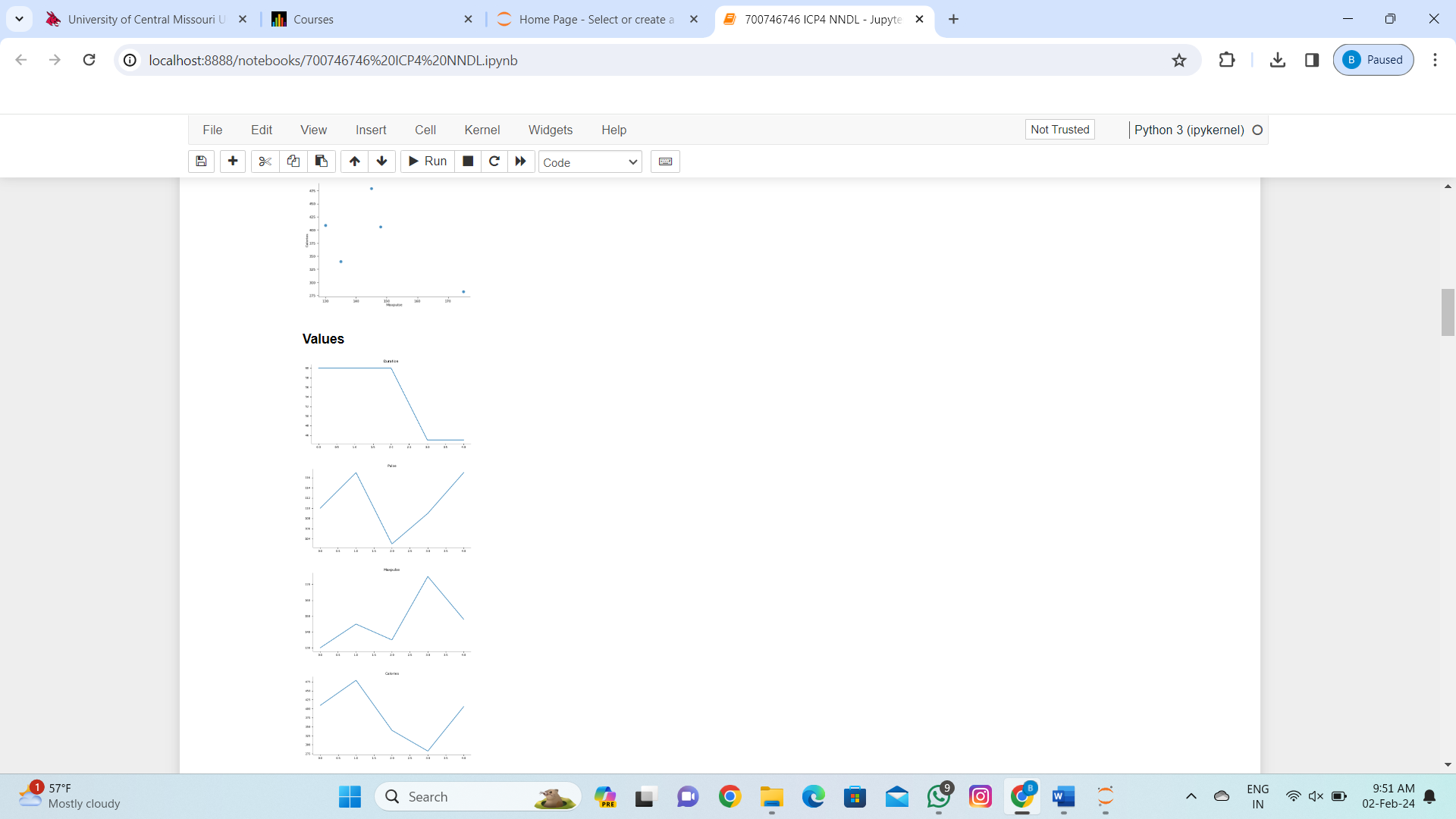
* + 1. Select at least two columns and aggregate the data using: min, max, count, mean.
    2. Filter the dataframe to select the rows with calories values between 500 and 1000.
    3. Filter the dataframe to select the rows with calories values > 500 and pulse < 100.
    4. Create a new “df\_modified” dataframe that contains all the columns from df except for

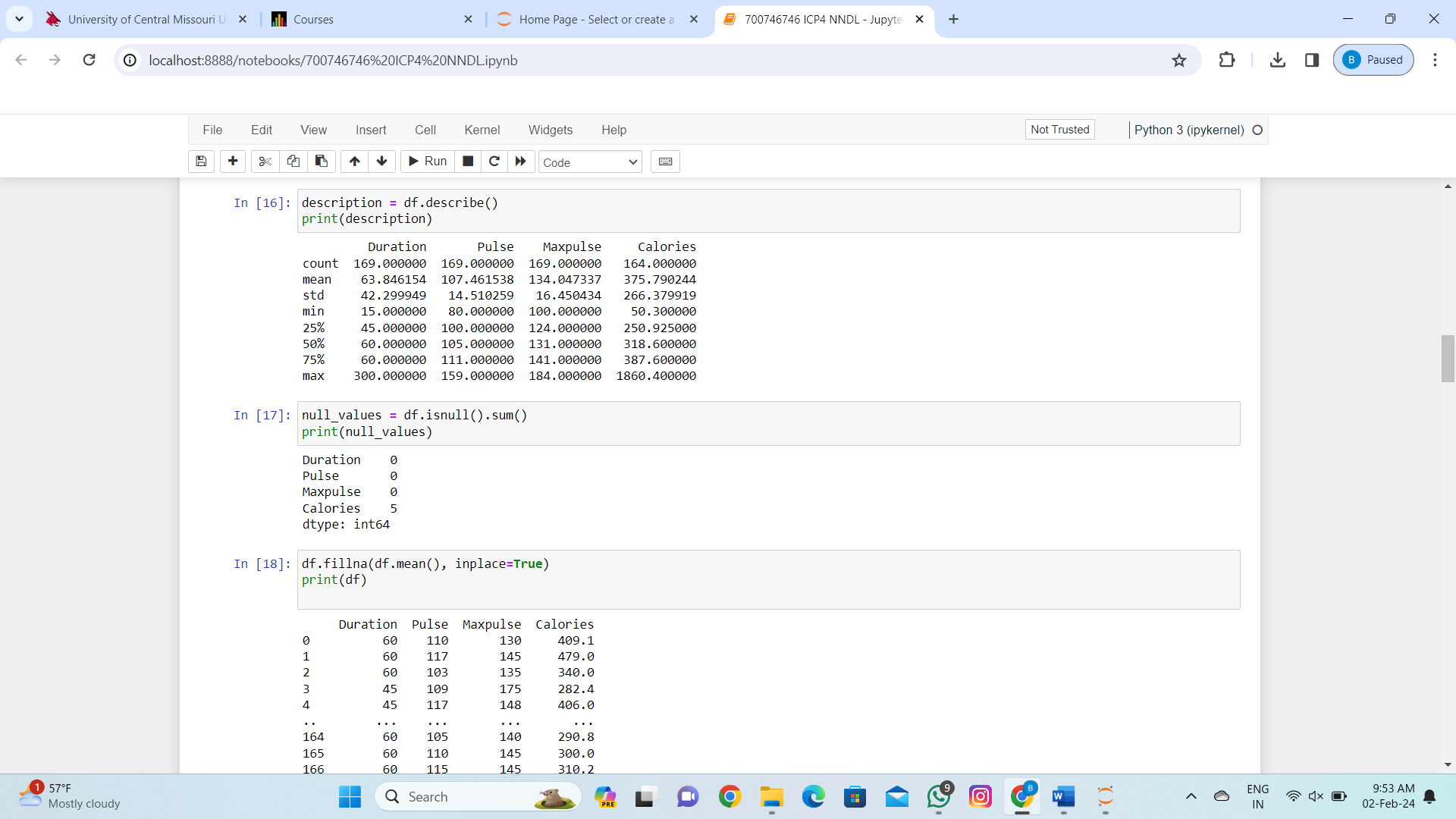
“Maxpulse”.

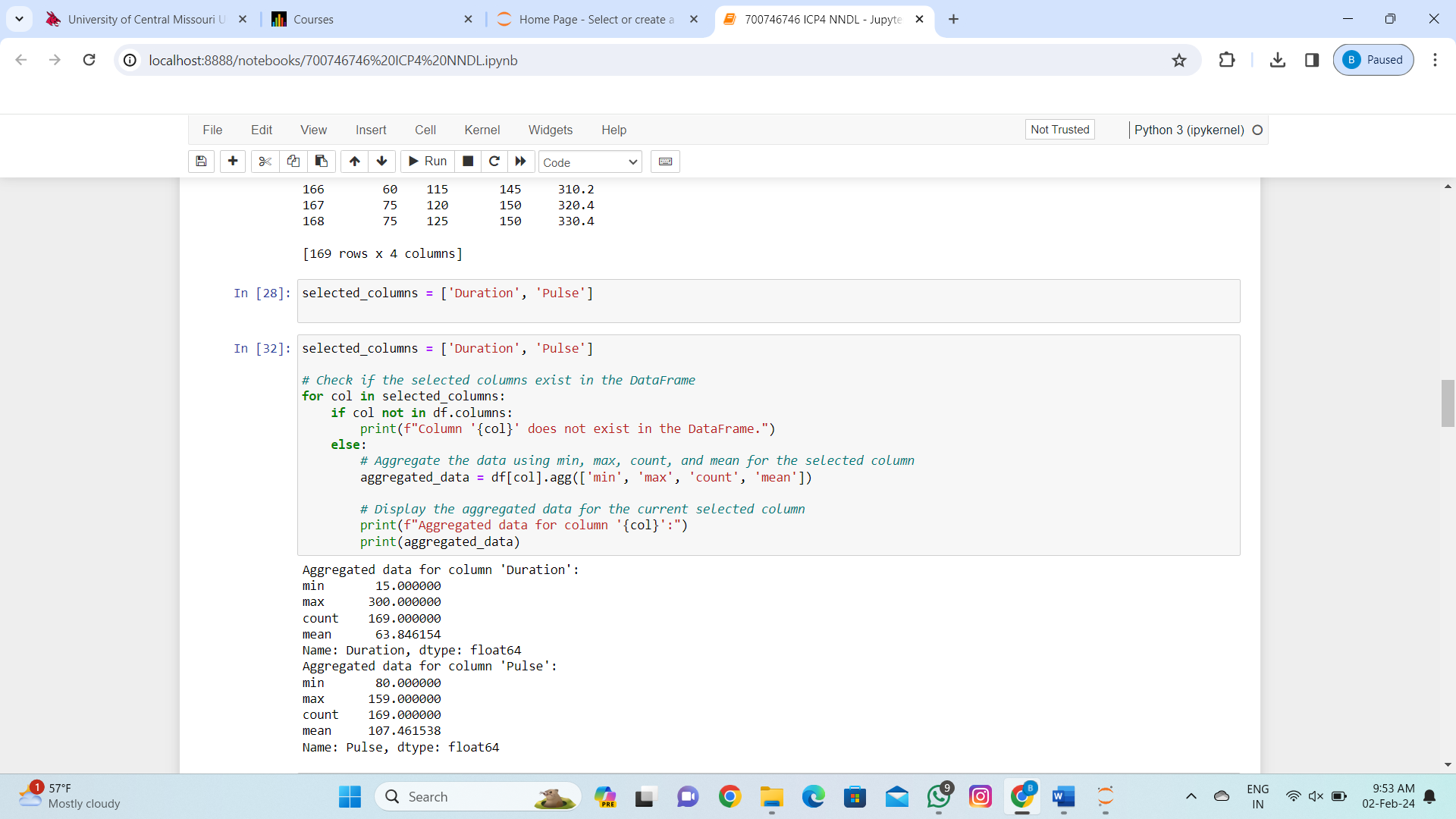
* + 1. Delete the “Maxpulse” column from the main df dataframe
    2. Convert the datatype of Calories column to int datatype.
    3. Using pandas create a scatter plot for the two columns (Duration and Calories). Example

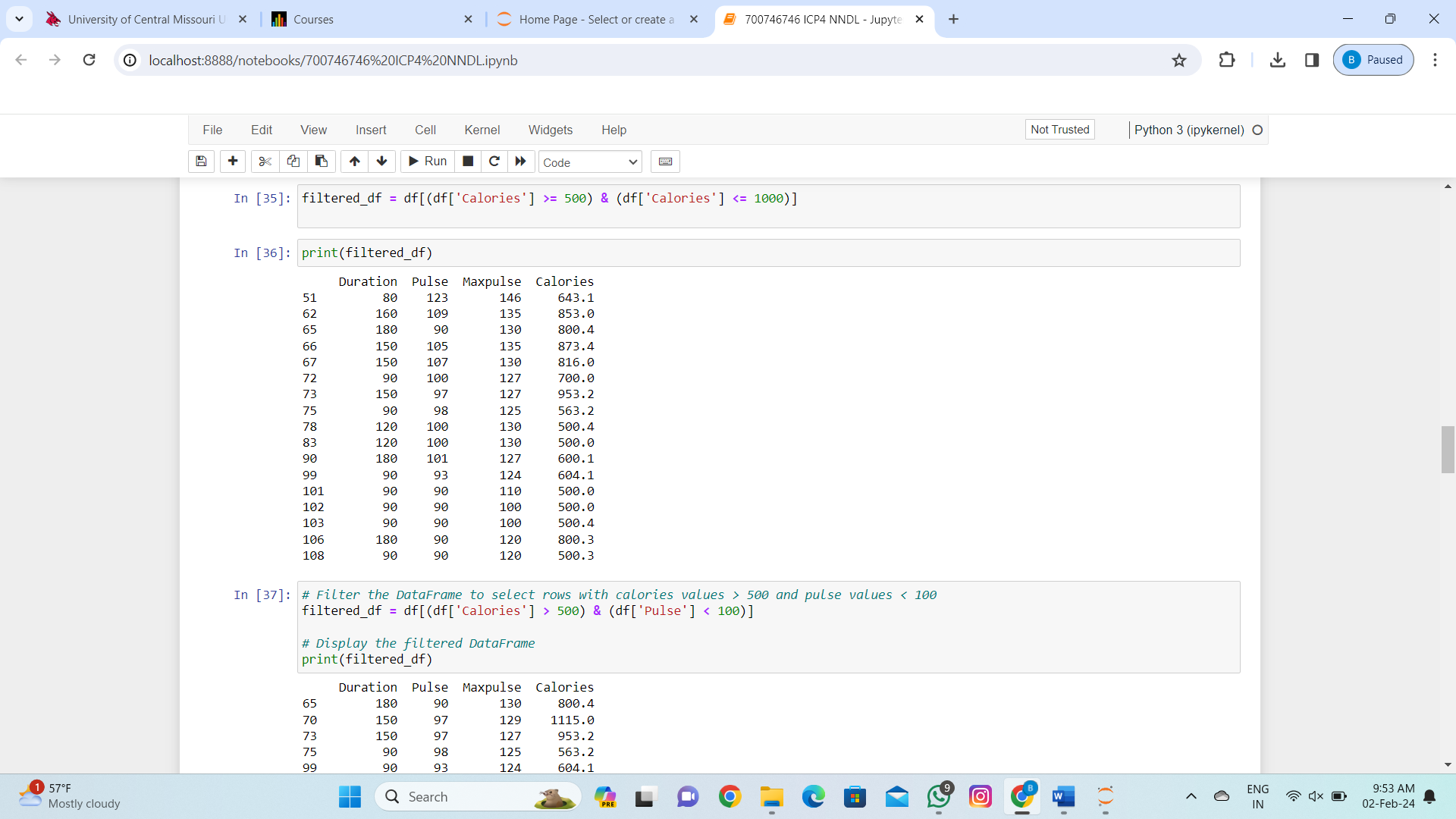


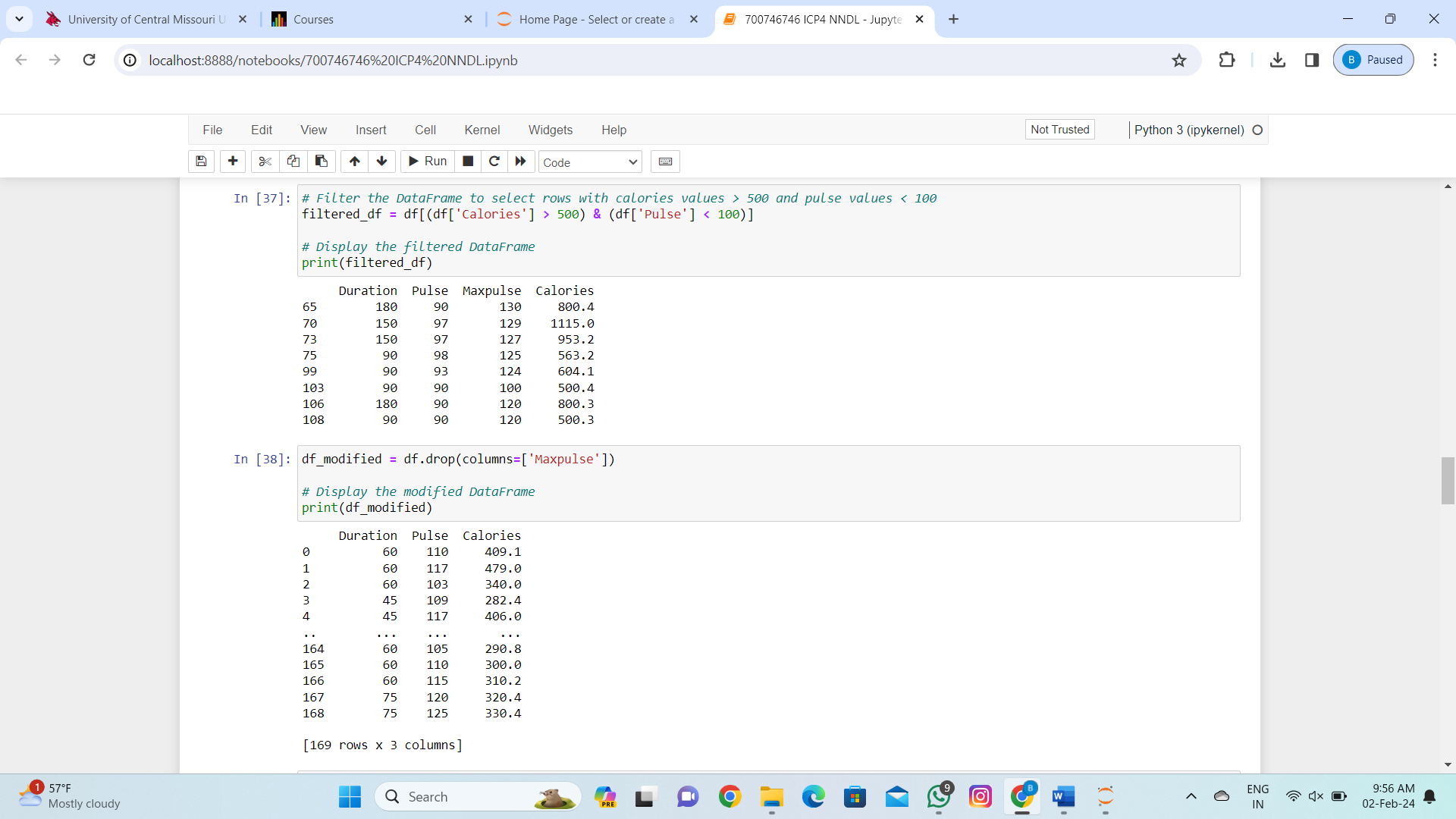


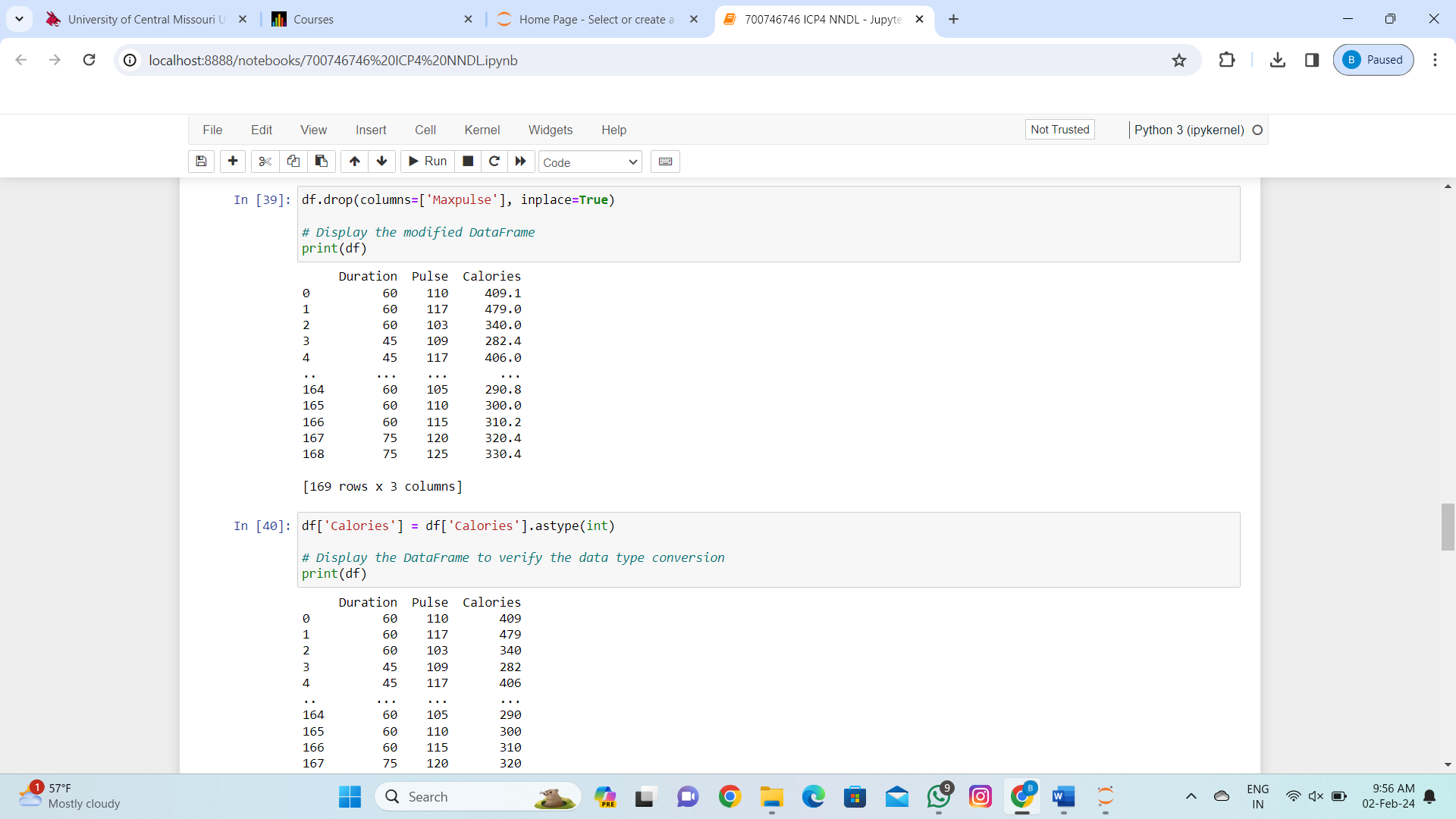


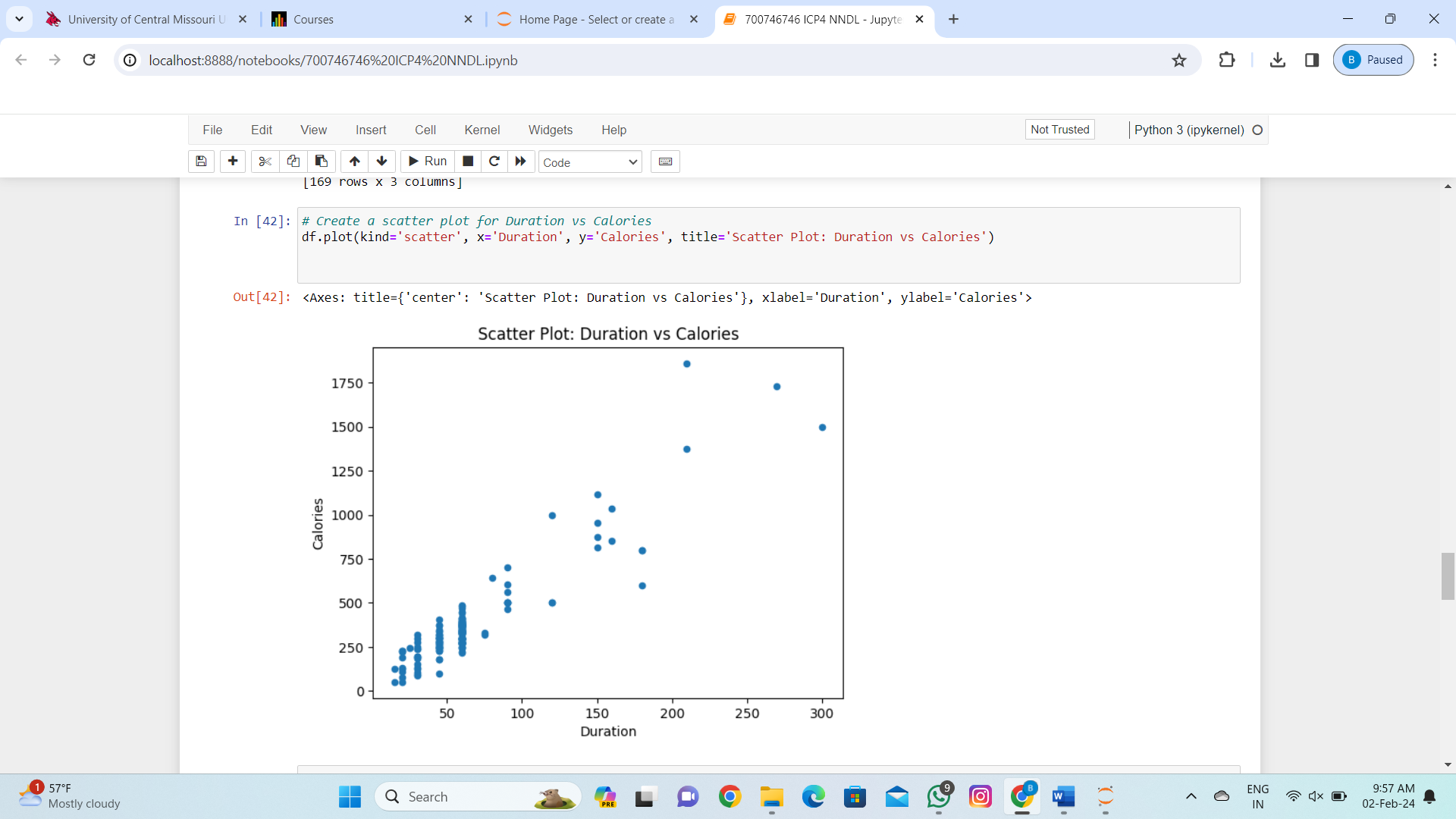






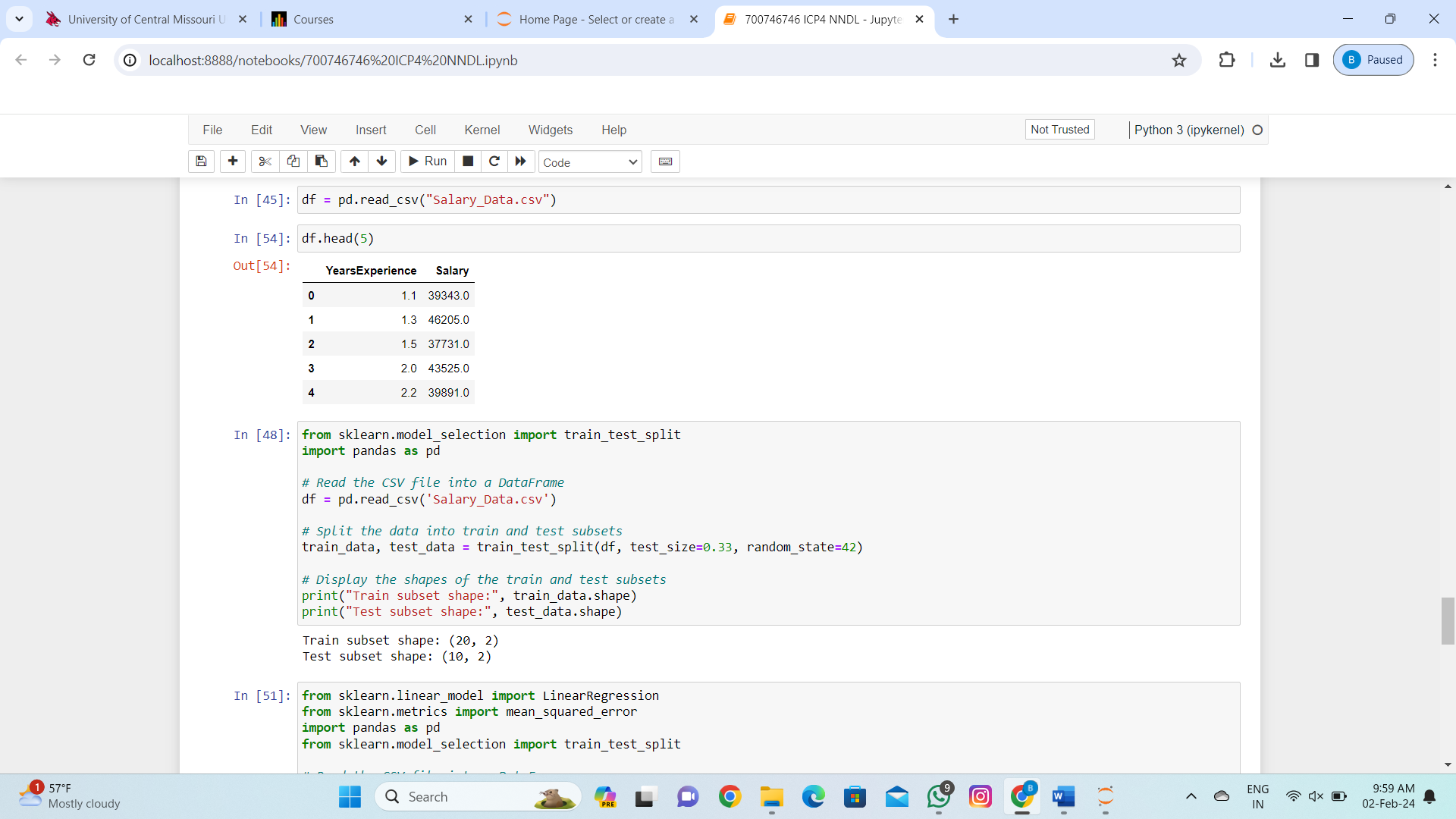


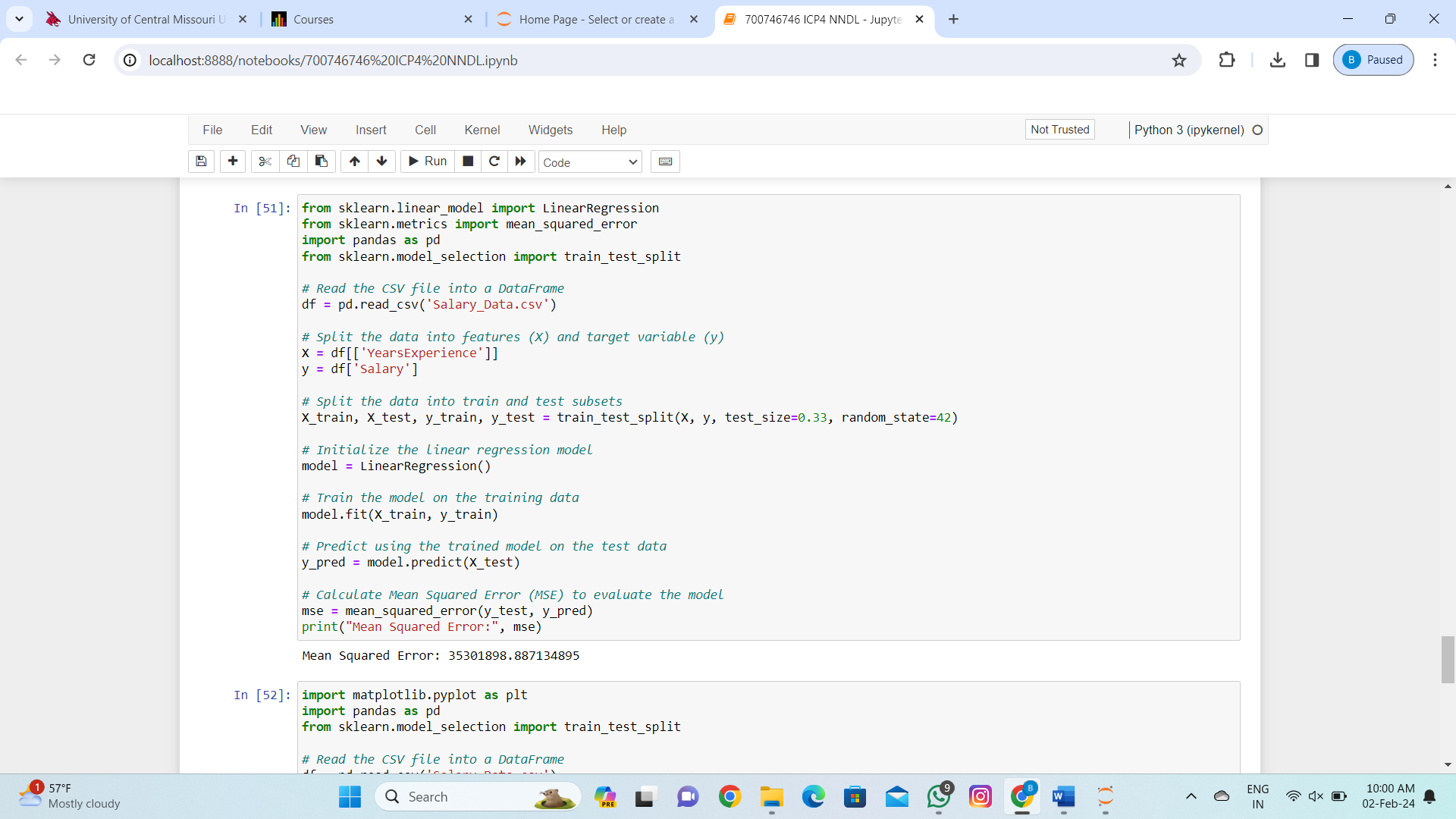


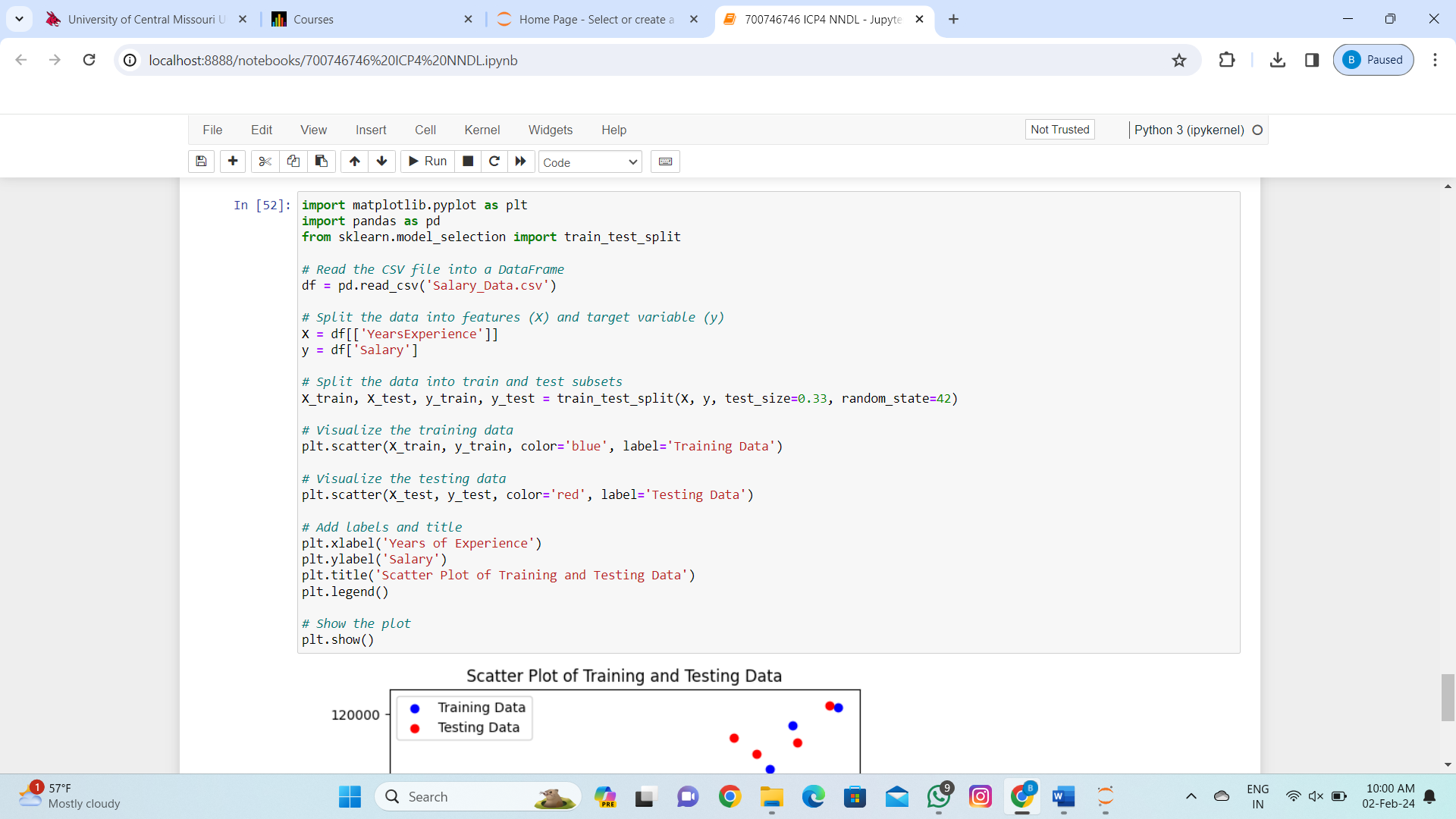


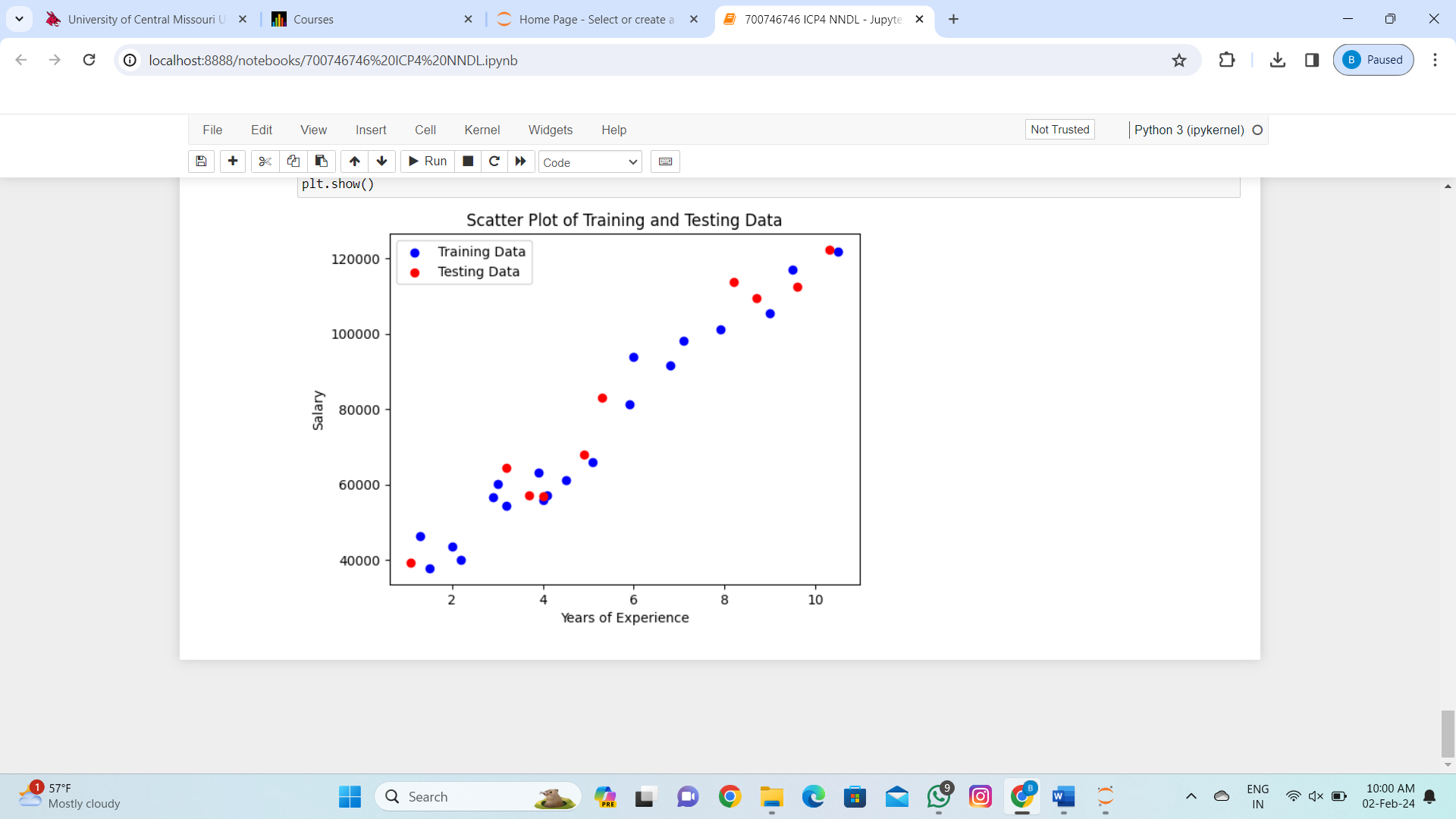
1. **Linear Regression**

* 1. Import the given “Salary\_Data.csv”
  2. 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.
  3. Calculate the mean\_squared error
  4. Visualize both train and test data using scatter plot.







  
  
  
 **THANK YOU**