**CS5542 Big Data Apps and Analytics**

**In Class Programming –9**

**Due 04/18/2022 (11:59 pm CST)**

**Regression:**

**Create a linear regression model in python using any dataset of your choice. For this model you can also create your own data. Find the best fit line in the data and calculate SSE (sum of square error) or MSE (Mean square error) , Y intercept, and Slope for the relationship in data. Explain your findings and understanding of these terms in detail in the report.**

ICP Requirements:

1. Successfully executing the code with linear regression model and calculating following:
   1. SSE or MSE
   2. Y intercept
   3. Slope
2. Detail explanation of each in report
3. overall code quality
4. Wiki Report quality, video explanation

Submission Guidelines:

1) Sign in into your GitHub account

2) Click on this link : <https://classroom.github.com/a/YfSYeINr>

3) Accept the ICP

4) Complete your ICP and create your wiki report (Pdf or Word doc).

5) Folders for ICP:

1. Create two folders for source code and documentation
2. Source code folder contains only the code and output file (if applicable)
3. Documentation folder contains wiki report and the images of your results

6) The wiki report should have at least the following:

a. What you learned in the ICP

b. ICP description what was the task you were performing

c. Challenges that you faced

d. Screen shots that shows the successful execution of each required step of your code

e. Out put file link if applicable

f. Video link (YouTube or any other publicly available video platform)

g. Any inside about the data or the ICP in general

7) Upload your ICP folders to your assignment GitHub repository

8) Click on the add a readme file on the next page and write done your and your partner name and email

You are all Done!!!!!!!