**Data Analysis - Final Project ( Deadline for Submission - March 25th 2021)**

**Submit Assignment**

* **Due** Thursday by 11:59pm

* **Points** 30

* **Submitting** a text entry box, a website url, or a file upload

* **Available** Feb 6 at 12am - Mar 25 at 11:59pm about 2 months

Prepare your Final Project on Data Analysis using dataset that you will bring from any sources from the web.

For Presentation, you can refer to the ”Data Analysis Presentation” template provided in the Files Section.

[https://classroom.ucsc-extension.edu/files/1086793/download?download\_frd=1](https://classroom.ucsc-extension.edu/courses/4494/files/1086793/download?download_frd=1)

If you would like to make changes to the presentation, please feel free

and enhance your presentation. The template attached is only to give an idea of how the presentation looks.

1.The dataset should have a minimum of 50,000 to 100,000 Records(100,000 Preferred)

2.Provide Description about your dataset

3.Describe the variables used in your dataset

4.Provide the summary of the dataset ( if you are using more than one dataset for comparison, provide the summary for both the datasets)

5.Provide the Descriptive Statistics for each variable

a)Min, Max, sum, range, mean, median, var, std. dev, NULLs, NAs( if any)

6.Provide the Descriptive Statistics of the dataset by:

     a)Histogram

     b)Box Plot

    c)Pie Chart

    d)Any other visualization

7.Find the overall mean for each variable

8.Predictive Modeling.

a)Use Linear Regression in two different models ( You can use any models that can be applied to your data)

9.Conclusion of your project

The points(30) will be given based on the presentation, code submission and the output.

During the final class on March 13th, you will be presenting the final draft of your presentation.

While submitting the final project, please include the following:

     - Final presentation ( PPTX or PDF or Word)

     - .RMD File

     -  & HTML File output

Here are the sample presentation templates that you can use a reference.

Please create the presentations on your own as it will help you in explaining.

[https://classroom.ucsc-extension.edu/files/1225714/download?download\_frd=1](https://classroom.ucsc-extension.edu/courses/5091/files/1225714/download)

[https://classroom.ucsc-extension.edu/files/1225715/download?download\_frd=1](https://classroom.ucsc-extension.edu/courses/5091/files/1225715/download)

[https://classroom.ucsc-extension.edu/files/1225716/download?download\_frd=1](https://classroom.ucsc-extension.edu/courses/5091/files/1225716/download)

[https://classroom.ucsc-extension.edu/files/1225717/download?download\_frd=1](https://classroom.ucsc-extension.edu/courses/5091/files/1225717/download)

Thanks

Partha