

Project 2: Midterm: Point 50 + 5 Extra point

Presentation Due by 3rd May

Note*: Don't use homework dataset/ or example.

Data analysis using **Panda, Scripy**, NumPy, and Mat plot

Two Options:

Either you create an entirely new project, or you can extend your midterm project

Point -5

Step:1

Choose your topic, generate dummy data, or download data if ANY

Tables: Minimum 4

Observations: 20

Columns: Minimum 5

Point -5

Step: 2

Read DATA

CLEAN AND VALIDATE YOUR DATA (For example, Handle missing duplicate data..)

Point -10

Step:3

Design your Functionality - Manipulation of data. (MERGE, APPEND, DUMMY VARIABLES, USE DIFF FUNCTIONS, ETC.

Minimum: 5 functionalities

For example: Using if /else or using formula generate new columns OR UPDATE COLUMNS

Point 10:

Step: 4 Design your objectives

What kind of reports you are going to create using statistic functions.

Min -5

Point 10: I may add here inferential statistics

Create correlation and Linear regression

T-test

Point 5

Step 6: Add Visualization

Point 5

Step 7 Presentation – 4

15 to 20 slides.

Extra point -5

Step 7: Additional point for out of box thinking

