

CMP 142 - Programming Assignment #3

Visualizing the Data and Building Machine-Learning Models

You are to work in your teams to solve the following problem.

PROBLEM DESCRIPTION

Over the course of the semester, your team will carry out a large-scale data science project. You will perform the following steps:

- 1) Analyze a real-world data science problem,
- 2) Collect and prepare data,
- 3) Explore the data using current tools and techniques,**
- 4) Build machine learning models to analyze the data,**
- 5) Analyze the models to select the best models, and
- 6) Draw conclusions about the problem based on the data model.

Your team will create a notebook (titled **Program03.ipynb**) that follows the steps outlined in Sections 15.2 – 15.3 (Classification) and 15.5 (Regression) to explore your team's dataset visually and to build a machine-learning model using your dataset to make predictions about new data.

ADDITIONAL NOTES

As always, specific instructions will be given in class. These instructions are required even if they do not appear in this description.

WHAT TO TURN IN

Upload copies of your program files (**Program03.ipynb**) using to the Program #3 - Drop Box on Moodle. Make sure you include the names of the files and the names of your team members in comments at the top of each of your files. Turn everything in by **5pm on Tuesday, April 19**.