11/19/19

Project: Data Science

Background

Our team is required to incorporate the Panda python package which would provide us with more adjustable and speedy data structures that are engineered to simplify our work with relational or labeled data in python.

Objectives

Our team's objective is to use the Panda python package tutorial provided by our instructor to submit a Google Doc/PDF/Word Doc file with cumulative results for Object Creation, Viewing Data, Selection, Operations, Merge, Grouping, Reshaping, Time Series, Plotting and Getting Data In Out (csv). To track our progress we will be using our time in class and Slack to discuss what work has been completed, what is currently being worked on and what else needs to be done. Our team will create a python file that will have the following packages installed; MySQL-connector, Numpy, Scipy, Pandas and matplot to make sure the file is running properly. We would also set up Github and utilize it for branching, pushing, and creating necessary pull requests to analyze, review and revise ours results. Branching would create and duplicate the settings of the python file for each of the team members. Pushing will create a perpetual snapshot of the work being done on the project. Pulling would ensure the work being done complies with the specifications given for this project. The Data Science project would be completed on November 20th, 2019. The verdict would be to make sure the python file is working properly and has results for all the previous conditions given. Once that happens the project would be considered complete.

Scope

The python file must be working properly and provide results for all the conditions listed previously. Those conditions include results for Object Creation, Viewing Data, Selection, Operations, Merge, Grouping, Reshaping, Time Series, Plotting and Getting Data In Out (csv). Each of the team members would be sticking to the time given in each scope and will not overcommit outside of the scope.

Timeframe

	Tasks	Start and End Dates
Sprint I	Statement of Work Set up Github Data Science Project Get team's forks set up Create a python file using IDE Install the necessary packages	11/13-11/16 10 hours to complete

Sprint II	Follow steps for panda provided from website Input results into python file Make sure python file is running correctly	11/17-11/18 10 hours to complete
Sprint III	Make sure python file is running correctly Turn in SoW Turn in Data Science project	11/19/-11/20 10 hours to complete

Monitoring and Evaluation

We will have our progress tracked via slack and sprint meetings in class. The code will be pushed, pulled and tested through our GitHub repository and Pycharm as our IDE.

Note: Semion Kalinin Team helped us

Team Members

Edison Dela Cruz
Mustafa Musa
Alan Abdullah

Approval Signatures

Edison Dela Cruz Alan Abdullah Mustafa Musa