INFM600 Information Environment

0201 - Team Brilliant

Work Plan

Submitted By:

Aniket Jadhav

Paridhi Mathur

Xinyun Zhang

Project Proposal

The New Coder dataset is a survey crafted for people who had recently started to learn to code through various platforms. The respondents were asked various questions such as their country of citizenship, the subjects they studied in university as well as their current and expected employment status. One interesting aspect of this dataset is that it caters to people from all backgrounds irrespective of their career field and grants researchers a glimpse of how adults learn to code. The research questions that we plan to work on are:

1. Do people who are younger learn relatively quickly and expect higher salaries eventually?
2. Do high investments in learning necessarily provide better return on investment than others?
3. Does a major in computer science or information technology or experience in related fields play an important role in learning and reduce the amount of time spent in learning and practicing?
4. Do people who stay far from learning centers spend less time learning?
5. Do people with programming experience spend more money on learning how to code?

Expected Challenges: The data has various anomalies that would challenge our analysis. For e.g. Responses to the question 'How old are you?' range from 0 to 120 which seems far from being realistic. The data is mostly nominal and contains free text field which would impose a fine challenge on us in order to derive the outcome. The data is messy, contains a lot of null values and would require extensive cleaning. Considering the measure of task in hand, we have allocated approximately two weeks for the process.

Project Timeline: The planned timeline for each of the sub-tasks is as shown in Table1:

Table 1: Project Timeline

Task Name	Start Date	End Date	Duration (days)
Work Plan	9-28	10-5	7
Studying the data set	10-5	10-12	7
Data cleaning	10-12	10-19	7
R tutorial	10-19	10-26	7
Dataset analysis	10-26	11-2	7
Analyze and plot graphs	11-2	11-9	7
Conclusion	11-9	11-16	7
Poster for presentation	11-16	11-23	7
Presentation rehearsal	11-23	11-30	7

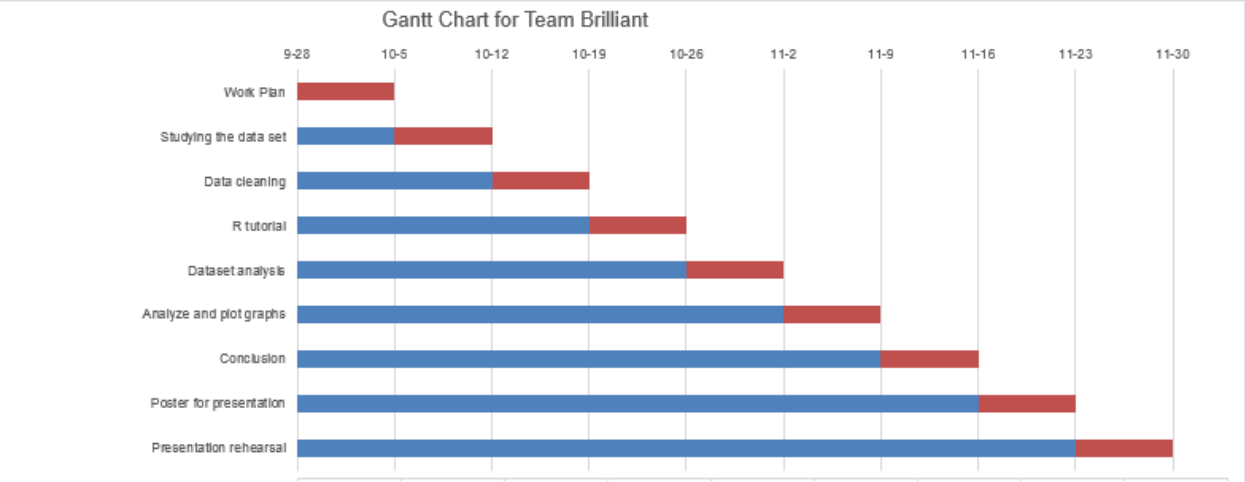

Fig. 1: Gantt chart

Table 1 and Fig.1 together provide the chronological order and the expected timeline of the tasks that would be followed. Starting with the work plan in Week 5 and ending in the presentation rehearsal in Week 15, the tasks in-between include data cleaning, dataset analysis, conclusion and presentation rehearsal.

Effort Allocation: The three of us have different strengths and various learning objectives from this project. The software that we plan to use for project implementation and analysis are MS Excel and R. We also plan to use Trello, Box and GitHub for project management, documentation and version control purposes respectively. Each one of us has different levels of expertise that enables us to strategize our project work in such a way that further develops our skillset rather than working on something we already know. Thus, we plan to meet, learn together and work equally on all aspects to maximize our learning process.

Target Audience for the analysis: The dataset and its analysis can prove useful for a number of organizations, MOOCs and educational institutes that are offer programming courses and related trainings. The answer to the research question 'Do people who are younger learn relatively quickly and expect higher salaries eventually?' can help the institutes to tailor their courses accordingly and perform target marketing to get a higher enrollment. The answer to the question 'Does a major in computer science or information technology or experience in related fields play an important role in learning and reduce the amount of time spent in learning and practicing?' can help people planning to take up a course see how their background would help them in the course. This study thereby enables a variety of audience to gain advantage from the results of our survey.

Word count: 635

References:

Dataset:

freeCodeCamp. (2016). 2016-FCC-New-Coders-Survey-Data [Data set CSV file]. Retrieved from https://github.com/FreeCodeCamp/2016-new-coder-survey September 13, 2016

Terms of Use:

Given data set is released under Open Database License (ODbL) v1.0 at http://opendatacommons.org/licenses/odbl/1.0/