

INFM 600 - Information Environments

Project Work Plan

Team Awesome

Binghuan Zhang

Rajat Aghi

Ruchira Kapoor

About the Data Set

Our data set is the one that is published by [Mashable](#)^[1] which is an online platform for media and entertainment. This particular dataset contains details of news articles published like their name, date published, category etc.

Proposed Research Questions

- Which is the best day to release a new article? Are people less likely to read about business on the weekend?
- What is the text to image ratio in articles and what effect does it have on the number of shares?
- What are people more likely to share: positive news or negative news?

Target Audience

- Individual Bloggers
- News Aggregating Websites
- Media Companies

Planned Timeline for Completing the Project

No.	Phase	Task Due	Starting Date	Ending Date
1	Understanding the data and coming up with research questions.	Work Plan	9/28/2016	10/5/2016
2	Data Cleaning and understanding Excel	-	10/5/2016	10/8/2016
3	Final Data Cleaning for the project	Data Cleaning and Documentation	10/9/2016	11/1/2016
4	Meeting with the instructor	Progress Meeting	11/1/2016	-

5	Playing with R - Analyzing Data	R Script (Draft)	11/2/2016	11/14/2016
6	Analysis of research questions using R	R Plot (Draft)	11/15/2016	11/30/2016
7	Final Analysis and Presentation	Git Package	12/1/2016	12/14/2016

The projected timeline is presented as a Gantt chart:

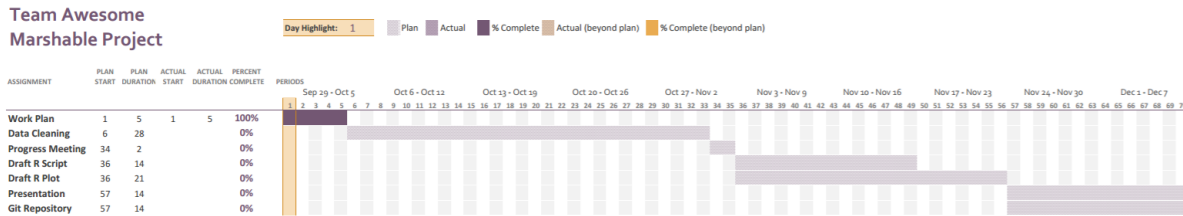


Figure 1 Gantt Chart - Projected Timeline

The excel file for this Gantt chart can be accessed from [here](#).

Projected Effort Allocation

We do not intend to restrict one person to one domain only as this will reduce the possible quality of the project. For instance, if only one person works on one of the research question the possibility of getting better results becomes less. Similarly, if only one person is entitled to work on one tool, for example R, there is too much load on that particular person and the others won't be able to learn new things. We will definitely work on our strengths and help each other to with the same. As per that following are the strengths of the people in our team:

- Binghuan Zhang - Tableau, Excel
- Rajat Aghi- Github, Excel, R
- Ruchira Kapoor- Documentation, R

However, we intend to work with equal diligence towards the successful completion of this project.

References

[1] UCI Machine Learning Repository: Data Set. (2015, May 31). *Online News Popularity Dataset*. Retrieved from [https://archive.ics.uci.edu/ml/datasets/Online News Popularity#](https://archive.ics.uci.edu/ml/datasets/Online+News+Popularity#) on Sep 26, 2016 11:15:23 AM.

Total Word Count- 388 words