Distributed Computing Project 1
Initial Planning Document

Group Members
- Joshua Ingram
- Sarah Nash
- Fehmi Neffati
- Raymond Blaha
- Paul Hwang

1) What is your communications strategy? What are your expectations/norms regarding team-member communication?

Communication Strategy:

Our team will maintain regular communication via Google Space, a Google Doc logbook, and comments for any commits on GitHub. As much as possible, progress on the project will be through decentralized individual or small team efforts. Progress made will be communicated regularly and when needed for major challenges, the full group meetings will be organized.

Communication Expectations:
- Each member must check the Google Space a minimum of 2 times per day, and more actively if working on a part of the project with another member.
- Each day, a standup thread will be created in the chat. Each member will insert what they have accomplished the last day and what is planned to be worked on next. It is okay if nothing has been accomplished for short stretches of time.
- Full group meetings will be scheduled on an as-needed basis and will be held in the library.

Group Member Availabilities:

Joshua Ingram - Flexible, but not available on Sundays.

Sarah Nash - Flexible, less available on weekends.

Fehmi Neffati - 8am - 6pm Monday-Friday.

Raymond Blaha - Flexible, less so Sunday / Monday. Any other day available from 10:30AM - 9PM

Paul Hwang - Flexible, but usually busy Thursday, Sunday, Monday nights

2) What tools/technologies are you using to facilitate group organization/communication?

Primary methods of communication (ranked):
1. Google Space Group Chat
2. Google Doc Logbook
3. Email

Technologies for Organization:
1. GitHub (https://github.com/FNeffati/DCP1)
2. Google Drive

3) What is your meeting strategy? How often will you meet? Virtual or in-person? What team or partial teams?

Meeting Strategy:

Each meeting that takes place will have an agenda organized by the project manager. Group members will present their individual progress since the last group meeting, their team progress, and will present any problems they have had that need to be addressed in the meeting. At the end of each meeting, a plan of action will be produced with a tentative next meeting date.

Frequency of Meetings:
- Full group meetings will be made on an as-needed basis (estimated 1-2 times per 2 weeks). These will be held in person.
- Partial group meetings will be up to the discretion of the members of the teams. These can be virtual or in person.

4) What is your high-level initial plan of attack for this project?

Plan of Action:
1. Topic: Analyze star wars fandom fanfiction
2. Web Scraping to collect data
3. Clean csv file
4. Break up into teams and look at:
   a. Calculate author metrics (determine metric)
   b. Post hits, kudos, bookmarks
   c. Popular characters
   d. Popular tags
   e. Popular character pairings
   f. Fanfiction lengths (words, sentences)
   g. Post dates and frequencies
5. Combine results and perform regression analysis to see what is associated with author popularity
6. (possibly) Look at author popularity and level of writing (complex language)
7. Create presentation

5) Initially, who is doing what?

Group Member Roles:

Joshua Ingram - Meeting agenda maker, post dates and frequencies

Sarah Nash - Web scraping for data, popular characters, popular character pairings

Fehmi Neffati - Web scraping for data, popular tags

Raymond Blaha - Fanfiction lengths (words, sentences)

Paul Hwang - author metric, post hits, kudos, bookmarks