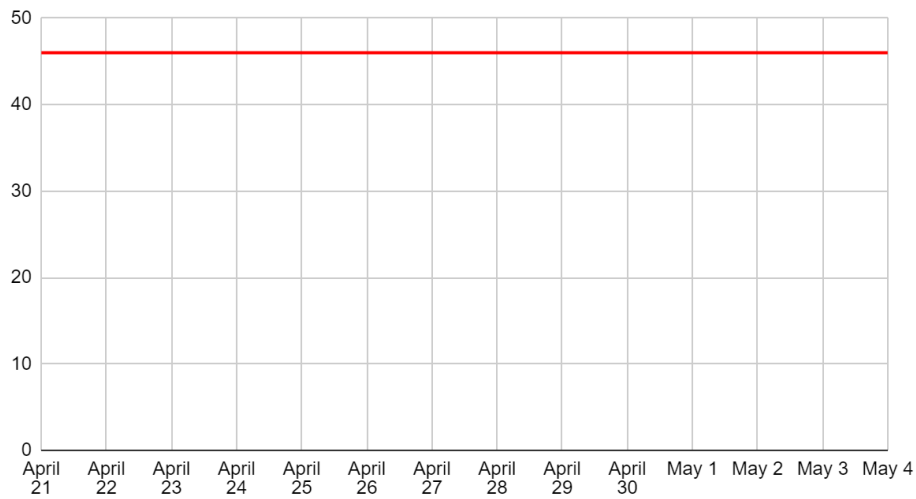


1. Heading: Sprint 2 plan, PathFinder, PathFinders, PathFinder 1.0, June 10th, Revision #1 & Revision Date: April 21th, 2021
2. Goal: Finish web scraping features, properly format data to be stored in SQL files and sent to the database. Work on data visualization and calculations on a small scale.
3. Task listing, organized by user story:
 1. As a developer, I want to be able to scrape data from a set of job postings: 5 story points
 - i. Combine the Web Scraper and the Web Crawler and test it on a small scale:
 1. 5 hours
 - ii. Set up proxy servers in Crawler
 1. 7 hours
 2. As a developer, I want to avoid scraping duplicate URLs
 - i. Use database to keep track of job IDs, then query it to avoid duplicates
 1. 4 hours
 3. As a developer, I want the data to be formatted correctly for the database
 - i. Text Preprocessing 4 hours
 - ii. Data identification 6-10 (Depending on method used)
 - iii. Data retrieval(with regular expressions) 5
 - iv. Generating Formatted SQL files 4-6 hours
 - v. Establishing Tables 6 hours
 4. As a user, I want to be able to have a Webapp to access the calculations on the scraped data:
 - i. Basic calculations 1 hour
 - ii. Single visualization 4 hours
4. Team roles:
 1. Product Owner: Ming Jeng
 2. Scrum Master: Andrew Craig
 3. Developers: Chainy Carson, Alice Liang, Eric Ngor
5. Initial task assignment:
 1. Andrew: Establishing Tables
 2. Chainy: Combine the Web Scraper and the Web Crawler and test it on a small scale
 3. Ming: Text Pre-Processing
 4. Alice: Establishing Tables
 5. Eric: Avoid duplicate webpages
6. Initial burnup chart:

Ideal Work Hours



7. Initial scrum board:Trello

8. Scrum times:

1. Tu/Thur/Sat after 5:30

9. Burn Up

Ideal Work Hours

