1. Data Set
    a. Reddit HyperLinks
        i. Source: http://snap.stanford.edu/data/soc-RedditHyperlinks.html
        ii. Summary
            1. The subreddit hyperlink network comes from the posts that create hyperlinks from one subreddit to another. Each hyperlink originates from a post in the source community and links to a post in the target community. Each hyperlink is annotated with three properties: the timestamp, the sentiment of the source community post towards the target community post, and the text property vector of the source post.
            2. Each post has a title and a body, therefore since the hyperlink can be present in either the title or the body of the post, there will be a network file for each.
            3. The total number of nodes we are given is 55,863 (these are the total number of subreddits), and the total number of edges 858,490 (these are the hyperlinks between the subreddits)
            4. We are given the following details about the data:
                a. SOURCE_SUBREDDIT: the subreddit where the link originates
                b. TARGET_SUBREDDIT: the subreddit where the link ends
                c. POST_ID: the post in the source subreddit that starts the link
                d. TIMESTAMP: time time of the post
                e. POST_LABEL: label indicating if the source post is explicitly negative towards the target post. The value is -1 if the source is negative towards the target, and 1 if it is neutral or positive. The label is created using crowd-sourcing and training a text based classifier, and is better than simple sentiment analysis of the posts. Please see the reference paper for details.
                f. POST_PROPERTIES: a vector representing the text properties of the source post, listed as a list of comma separated numbers.
2. Algorithms
    a. One of our goals is implement and algorithm to find the shortest path between 2 nodes/subreddits
        i. We plan to use Dijkstra's Algorithm
    b. One of our goals is also to find the shortest path between 2 nodes/subreddits through a specific third node/subreddit