Crawl.py

Libraries:

- ~praw
- ~sqlite3
- ~Beautiful Soup
- ~multiprocessing
- ~ison

Functionalities:

- ~ Database setup
- ~ uses multiple Reddit

API applications to crawl

- ~ inserts post and comment data into sqlite database
- ~incorporates some error handling

Setup_database.py

Libraries:

~SQLite3

Functionalities:

- ~Database connection
- ~ Table creation: post and comments table
- ~ id, post_id, url, permalink, subreddit, body, score

Index data.py

Libraries:

- ~ re (for tokenizing)
- ~sqlite3

Functionalities:

- ~Tokenization
- ~Normalization
- ~building inverted index
- ~Storing the inverted index
- ~Database interaction