Web Scrape Procedure

Generates json files with game
data contained within them

scrape_core.scrape

scrape_core.py

json data

scrape_core.load

Generates an up-to-date
dictionary from json files

scrape_cron.py

Runs scrape_core.py routine
every xxx hours/minutes

Notes:

1) The generation of json files should avoid redundancy as much as possible. Needs a separate function within scrape_core.py which checks for json files not already within the current iteration of the database structure.

2) Needs tests to make sure it is functioning properly. Create scrape_tests.py to run (which will test out the procedure on a small subset of the data; perhaps on a single season or something).

3) Are we okay not using a sql database? Do we really want/need it?
Perhaps, if we are going to attempt to select data using complex conditions (e.g. – I want all players between the years of X an Y, who scored Z points against west coast teams.). If we do want an sql database. We should create a function in scrape_core.py to create an sql database each time we scrape new data.