Big Data Chess Project – Planning

# Problem

User Story: To create an application which analysis your chess games, and tells you in which openings you strggle the most or are the strongest.  
  
What we planned so far is having an analysis only on the given player. But if this turns out to be less complicated than we thought it would be, we could extend the analysis.

# Relevant Data

## Chess games:

<https://www.reddit.com/r/chess/comments/9ifkaq/how_i_downloaded_all_my_chesscom_games_using/>  
  
Chess.com and Lichess API implemented in another github: <https://github.com/evamvid/chesscom-games/blob/master/chesscom-games.py>

## Chess Openings (For Clustering)

chessTempo: <https://old.chesstempo.com/chess-openings.html>  
has all chess openings, but I did not find an option to download, could be webscraped

Github eco.json : <https://github.com/hayatbiralem/eco.json>  
another Git Project , which calls itself Encyclopedia of Chess Openings. Have to look more into detail into it, but it seems to have all chess openings in json formats..

# Existing Applications

Opening Tree: Free Opensourece software which allows loading in all games from different chess sites, and lets you chose which time formats. Lets you discover the moves that you play the most just in general lets you explore your openings. But does nto show quality of opening only game win percentage at any given position.   
all on git hub as well.

<https://www.openingtree.com> github: <https://github.com/openingtree/openingtree>