**506 Final Project Plan**

Kehan Liao

My project plan is to extract data from my Facebook group “MarketNoire” which is an online market place where people sell and resell stuffs. Things on sale includes football ticket, used furniture, used textbook, used clothes, etc. To be specific I want to use the data of football ticket sales to analyze the relationship between the popularity of football game and the degree of competitiveness. So, I can extract the frequency of the resell and request of specific games from Facebook group “MarketNoire”, to compare with the game statistics extracted from Sportradar.com. The output from Facebook will be the number of times or frequency of each game appeared in the group, and the output from Sportradar will be the football statistics (name of the two football teams, Umich and the rival, and the scores).

API for Facebook: <https://developers.facebook.com/tools/explorer>

API for Sportradar: <http://developer.sportradar.com>

Since the Facebook require the access token, so the person whoever needs to run my program has to be in the Facebook group, and paste their access token to my python code. Or, I could also provide the cached data for him/her to see my source.

To access Facebook API needs the user to login to Facebook to get the access token, in which to get access to MarketNoire group ID. To access Sportradar API need to register for account first and get API keys to access the data.

I know both API is working, and I extracted the data successfully from the API, but to be honest it is not cached. I have some problem with the code and not be able to go to office hours during the break. But I think I can make it work, and I’ll ask questions in this week’s offices hours.

I will sort my data by frequency so it’s easy for human to read with the ranking. I would use *.sorted* to sort them from high to low.

The Class for Facebook is going to be called *Game*, with instances variables *competitorname*, *frequency* and *fullname*. The *competitorname* is the abbreviated team name of team UMICH Wolverine, the *frequency* is the times of the ticket of this specific appeared in the group and the fullname is the full name of the team. Instance might be Wisconsin Badgers, Purdue, Michigan State, etc. The class constructor requires a dictionary representing the frequency of each team appeared.

The Class for Sportradar is going to be called *GameSummary*, with instance variables *rivalname*, *home\_points*, and *away\_apoints*. Instances may be Badgers, Purdue, Michigan State, etc. So, the *rivalname* is the name of the rival, the *home\_points* are game points for Michigan Wolverine team for this specific game, and *away\_points* are game points for the rival (instance variable)’s points since they are playing the game in Umich. The class constructor requires a tuple representing the scores/points of each team earned.

I have summit 2 python files, one for extracting data from Facebook (506\_fb\_test.py) and another one from Sportradar (506\_sr\_test.py). One data file is the data extracted from Sportradar about the NCAA game summary in 2016. I spend a large amount of time figuring out the best API to use and don’t have enough time at the end to save the cached file, so the pythons files are not working as they supposed to be, but I’ll work on them more!