

Final Project

INFORMATION ARCHITECTURE, FALL 2019

WEB SCRAPPING AND VISUALIZATION

DUE : 11:59PM MONDAY, DECEMBER 16

1 Instruction

- In this project, you may learn how data scientists or data analysts collect, preprocess, and visualize the various sources of data.
- This project is **optional**.
- The entire project is worth 10% of your final grade.
- You can do the project by yourself or work in groups of two or three members.
- When working in groups, please clearly detail the contributions of the individual group members. Individual scores will be based on the contribution.
- You can use any inanimate source you like but you can only discuss the project with your team members or me.
- Submit only a single zipped file via eTL. **Do not send an email.**
- The file should contain the final report in a pdf form and the Python source codes.
- To submit your term project, you must include:
 - Project Name
 - Group Members
 - Project Description
 - Data Structures and Algorithms
 - Sample Results

2 Requirements

The following are all the requirements for your final project.

- Using `BeautifulSoup` or any alternative you want in Python, collect data from at least two different web sites.
- Using any data structures you learned in the lecture, combine the data sources.
- Visualize the combined results using `matplotlib` or any alternative you want.