# **Project Yasee Layout (Week 1)**

Overview:

In order to have a clear picture of how the 5-week internship should be planned out, we had a brief discussion on Monday to work out the preliminary structure of the project timeline and framework. As is the standard in the scientific community, Python is chosen as the primary programming language for development. We aim to produce a local deliverable Python program that will take an Excel form as input and outputs various data analysis charts and images.

Primary Progress:

* Finished configuring a virtual Windows 10 testing environment to test our software on native Windows machines.
* Developed a prototype program that fulfills the minimum requirement of the project, producing a “word cloud” from an Excel file, visualizing word frequency.
* Worked out a preliminary project framework:
  + main Module (Console-based user interface)
  + YaseeWordCloud module (A class that handles word cloud visualization)
  + ReportFile module (A class that manipulates Excel files and panda “DataFrame” data structure)
  + Unittest modules (with further explanation down below)
* Test-Driven Environment:

It has been decided that we will work on our project in a “test-driven” style. For each new module that we develop, we will write automatic tests for each newly developed functionality if possible. It may stall the project progress, but we hope that our project will have a solid foundation for further development and extension.

Primary Issues:

* Patching up our program:

Python is mostly used as a tool for “agile software development”. As a result, the language does not readily support developing user-facing applications on Windows PCs. We were having trouble patching up our python program into an EXE application that will directly run on a Windows computer without any Python environment configuration.

* Python development configuration issue on the Writing Center computers:

Due to a lack of administrator access, we were having trouble installing and configuring a Python environment on the client computer. Therefore, we can only run the program on a configured Windows virtual machine or a personal computer as of now.

* Manpower:

To develop a user-facing software with a friendly interactive interface, we realized that we might need a team of developers to deliver such a product, instead of one programmer doing both the “back-end” and “front-end” work.

Future Directions:

* Communicate with the UCI IT department to configure a working Python environment on the client machine.
* Work on the secondary requirements:
  + Visualize word frequency correlations (in relation to a chosen category/column in the Excel file)
  + Support analysis for word roots.
* Enlist outside help for project development:
  + We hope at least one developer can join to form a team. Their primary responsibility would be working on user-interface and all the “front-end” side of the development.
  + If things work out swiftly in Week 2, as the secondary requirements get developed and tested, we hope to possibly set up a small server with the help of team members. This will convert our program into a small web-app, bypassing the need of any environment on the client machine.