Project Plan

Analysis

1) IPython Notebook scripts

- Repositories with ".ipynb" scripts

- Number of cells for script

- Code vs Text(Markdown)

      → Some scripts have "worksheet" item, others have only "cells"

- Execution

      → Number of errors

      → Types of errors

- Types of Outputs from scripts

- Most common used library in scripts (e.g. pandas or matplotlib)

- Structure of scripts (e.g. first imports or they are inside)

-

2) Repositories

- Empty repositories

- License and Readme files

      → Words contained

      → Readability analysis

- Number of commits

- Number of contributors

- Starting date of project

-

3) Three "data_copy*.json" files analysis

- Difference in number of repositories with IPython

-

Unit Testing

TKinter