**Administration**

**Tasks:**

* Figure out Github
* Buildout hosting area / Set up file structure
* Setup app documentation (readme.md)
* Create project / milestone framework
* White board project
* Model the models

**Code naming convention**

* Variables (single value): camelCase
* Lists/Arrays: name\_name

Primary Datasets:

1. Crime\_Data \*Obtained from CoP Crime Datasets
2. Housing\_Data \*Obtain from CoP or Oregon State data website (API)

Secondary Datasets:

1. Zip-code info associated with unique Portland Areas
2. Zip-code info associated with housing data addresses

**Dataset Prep**

**Tasks:**

* Pull in datasets from City of Portland or figure out if an API call would work
* Pull in zip code dataset based on Portland neighborhoods
* Crime Data:
  + - Raw Crime Data -> Scrub for anomalies and null data -> merge data with zip code dataset -> Rank Crimes based on severity (i.e. arbitrary value) -> add up value of crimes based on zip codes
      * crime\_index = {‘zipcode’: #value}
    - Integrity and accuracy: Testing ongoing at each stage of manipulation to ensure data integrity.
    - Discussion: the crime\_index could be derived in a number of ways; however, for the purposes of this app/project, the value of the offense type will be given a value of 1-20 (i.e. murder = 20, parking violation = 1) and then summed by zip-code to develop an ‘index’. This methodology could potentially skew data, however, it appears to make the most sense over counting the number of crimes in an area (i.e. averageing data based on the severity of crime vs. treating all crimes as equal).

Crime Data:

Methodology: use a combination of Python and excel to scrub/sanitize data and add in zip code info based on