**Weekly Plan**

**Team Members**

* **Nancy Fujikado**
* **Alexei Mendoza**
* **Sarah Sanchez**
* **Kijahre Fikiri**

Progress: begin in-process done

**Rubric Segment 1** (19% of final grade) **Due 11/6** *note: content should be in README*

1. **Presentation Content: Team members have drafted their project, including the following:**

* Selected topic
* Reason why they selected their topic
* Description of their source of data
* Questions they hope to answer with the data

1. **GitHub**

* Main Branch Includes a README.md
* README.md must include: Description of the communication and protocols
* At least one branch for each team member
* Each team member has at least four commits from the duration of the first segment

1. **Machine Learning Model:** **Team members present a provisional model that stands in for the final machine learning model and accomplishes the following:**

* Takes data in from the provisional database
* Outputs label(s) for input data

1. **Database: Team members present a provisional database that stands in for the final database and accomplishes the following:**

* Sample data that mimics the expected final database structure or schema
* Draft machine learning module is connected to the provisional database
* Team members present a fully integrated database.

**~~From 10/4 to 10/11:~~**

* Meeting between team members (10/11 @2pm)
* Create code to grab data from Home Depot’s site

**~~From 10/11 to 10/18:~~**

* Refactor code to grab data from Lowe’s site
* Create database schema (See “Estimator\_Scehma.sql)
* Research web scraping a third site and get necessary data for third web scraping site
* Create Estimator Worksheet Framework
* Edit schema to include “link”

**From 10/18 to 10/23:**

* Fix “Multi-Vendor\_grab” so that it only grabs information where price is listed
* Make a better price variable (currently looks like this: ‘['rich\_snippet']['top']['detected\_extensions']['price']’
* Clean raw data:
  + Drop na’s
  + Drop vendors with only (1) value count
  + Return a data set of minimums
* Figure a way to download new data into an existing Excell worksheet

**From 10/24 to 10/30:**

* Meet new team members:
  + Explain project
  + Establish communication and GitHub protocols
  + Divide up remaining workload
  + Discuss presentation structure
* Current workload:
  + Create database with CSV file on PG Admin
  + Polish a deliverable “Segment 1” (and possibly Segment 2)
  + Create descriptive stats sheet=> average, max, and SD of prices
  + Create a git hub io or live host or website
  + Start on Power Point
  + Add VBA drop down menus to estimator sheet
  + Find a way to activate ‘Working\_Estimator\_Script’ once a day and save results each day a search is done.
  + Find a way to host daily results on a web page, git hub io or live html

**From 10/31 to 11/6:**

* Discuss presentation structure
* Get a web host domain (if we haven’t found a way around it)
* Start working on Power Point
* Start working on regression analysis
* Start working on Visualizations
* Polish a deliverable for ‘Segment 2’ or 3

**From 11/7 to 11/13:**

* Fine tune worksheet and website
* Practice and fine tune our presentation
* Polish remaining Segment deliverable(s) and turn them in
* Present