**Canon Camera Scraper**

**Macro Goals:**

Autogenerate pdf report:

\* Use ipython to write the python script **(Done)**

\* Build data frame to store all cameras **(Done)**

\* Use plotly to generate graphs and plots in iPython (Example Done)

\* Allow data frame to be updated (write to csv, check for duplicates using pandas, update day count, when disappears note selling price) **(Done)**

\* Use a task scheduler **(Done)**

\* Upload to git

\* Build Average Price Calculator

\* Web Facing HTML front-end

\* Automate with Server

\* Set Up Server to Run Automatically

**------------------------------------------------------------------------------------------------------------------------------**

Build Data Frame...

1. Add column for days **(Done)**

2. Check for items disappearing and at what price **(Done)**

3. Check for duplicates using pandas (if more than 90% description match change price same item)

Verify Task Scheduler

1. Automate Jupyter Notebook **(Done)**

Lens Removal

1. Using regex to identify lens only listings **(Done)**

Title

1. Using regex, determine the camera make and model **(Done)**

2. Build list of all makes and models

Find what listing comes with

\* Watch Safari Books online video on NLP

\* Use NPL to perform analytics on DLSR search

\* Use word matching to determine if it is body only **(Done)**

\* Use regex matching to determine lens it comes with **(Done)**

\* Use regex to determine number of actuations

\* Find if it comes with batteries/bag/more than one lens **(Done)**

Build average price model

\* Group by number of lens it comes with

\* Store average price based on group

Bonus: Amazon tie in

\* Find the value of the lens/accessories on Amazon

\* Find difference in price between craigslist and amazon listing