**Review List**

**Programming:**

Pandas

SQL

Numpy (repo)

Scikit Learn (repo)

Visualization

Data Cleaning

EDA

**Modeling:**

* Overview (Scratch 11), https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/
* Gradient Descent (Scratch 8)
* Logistic Regression (Scratch 16),
* kNN: python (Scratch 12), assignment: mine plus sklearn
* Decision Trees: python (Scratch 17), sklearn: http://scikit-learn.org/stable/modules/tree.html
* Linear Regression
* Regularized Regression:

Look through: <http://scikit-learn.org/stable/tutorial/statistical_inference/supervised_learning.html>

**Programming:**

Visualization:

<https://www.youtube.com/watch?v=q7Bo_J8x_dw&list=PLQVvvaa0QuDfefDfXb9Yf0la1fPDKluPF>

<https://github.com/gSchool/dsi-pandas-matplotlib>

<http://nbviewer.jupyter.org/github/jrjohansson/scientific-python-lectures/blob/master/Lecture-4-Matplotlib.ipynb>

<https://blog.modeanalytics.com/python-data-visualization-libraries/>

EDA:

LR EDA lecture

Unix:

<https://www.digitalocean.com/community/tutorials/using-grep-regular-expressions-to-search-for-text-patterns-in-linux>

**Statistics:**

Statistical power: http://rpsychologist.com/d3/NHST/