1. ## SAT Scores + Summary Statistics - *Week 1*
   - ○ Provided with a dataset of SAT scores from across the United States, students will perform exploratory analysis for their client the College Board, using NumPy, Matplotlib, and Tableau to apply basic summary statistics.
   - ○ **Format**: Jupyter Notebook

2. ## Billboard Hits + Data Munging - *Week 2*
   - ○ Using a dirty dataset of Billboard hits, students will use Pandas to munge data, create a problem statement, and perform exploratory analysis for a local music publisher.
   - ○ **Format**: Jupyter Notebook

3. ## Liquor Sales + Linear Regression - *Week 3*
   - ○ Given access to state liquor sales data, students will choose between performing market research or conducting tax audits, using Pandas, StatsModels, and SKLearn to transform data, perform linear regression, and plot results.
   - ○ **Format**: Jupyter Notebook

4. ## Web Scraping + Logistic Regression - *Week 4*
   - ○ Posing as a private contractor, students will scrape website data and use Pandas, Statsmodels, and NLTK to clean and analyze data, perform logistic regression, and evaluate correlation coefficients.
   - ○ **Format**: Jupyter Notebook

5. ## Disaster Relief + Classification - *Week 5*
   - ○ As a researcher for a disaster response agency, students will pull remote data on Titanic survivors in order to build a local database, run a logistic regression classification model, and validate results from test subsets.
   - ○ **Format**: Jupyter Notebook

6. ## IMDB API + Random Forests - *Week 6*
   - ○ Acting as a member of the Netflix data science team, students will collect data from IMDB's API and use SQL to join it with additional scraped website data, in order to construct a random forest model that identifies ratings indicators and correlates these findings with viewer sentiment analysis.
   - ○ **Format**: Jupyter Notebook

7. ## Airport Delays + Cluster Analysis - *Week 7*
   - ○ Working as an airport operations consultant, students will analyze plane delay data in US airports, creating a local PostgreSQL database, performing a principal component analysis, and writing up a technical report describing clustering methods and evaluation metrics.

- ○ **Format**: Jupyter Notebook