CAPSTONE-CANNABIS RECOMMENDATION SYSTEM

Task

Building a Content-Based recommendation system, focusing on the effects and flavors of the three different types of strains of Cannabis.

Business Value

- Showing the user different strains than what they are used to.
- Taking the customer out of their comfort zone by providing suggestions different from the type of strain they like.

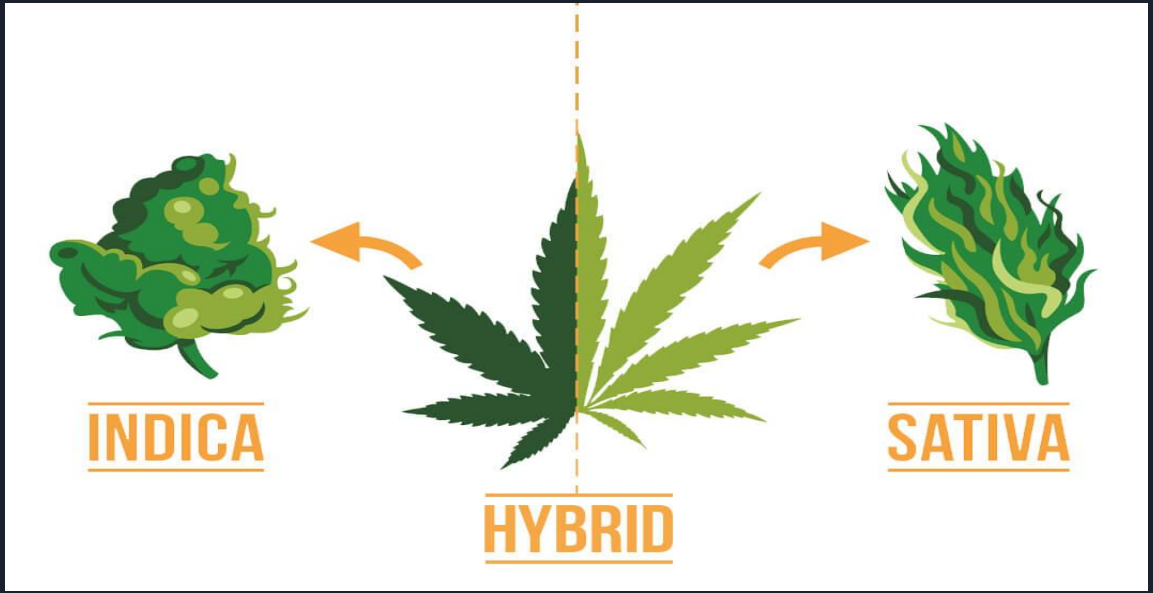
OSEMN Process

1. Obtain the Data
2. Scrub the Data
3. Explore the Data
4. Model the Data
5. Interpret the model

1. Obtain the Data

Data from Kaggle Datasets:

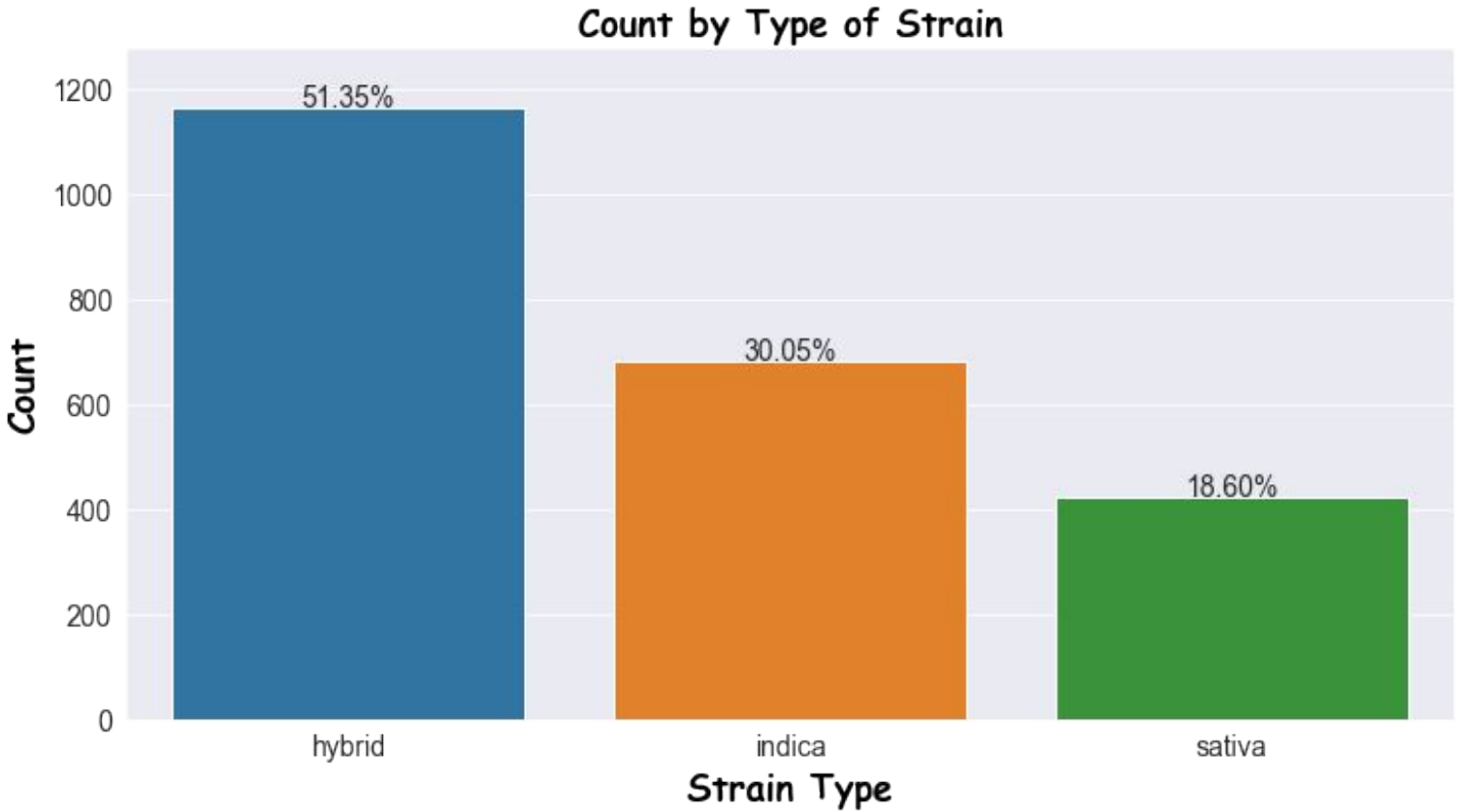
https://www.kaggle.com/kingburrito666/cannabis-strains



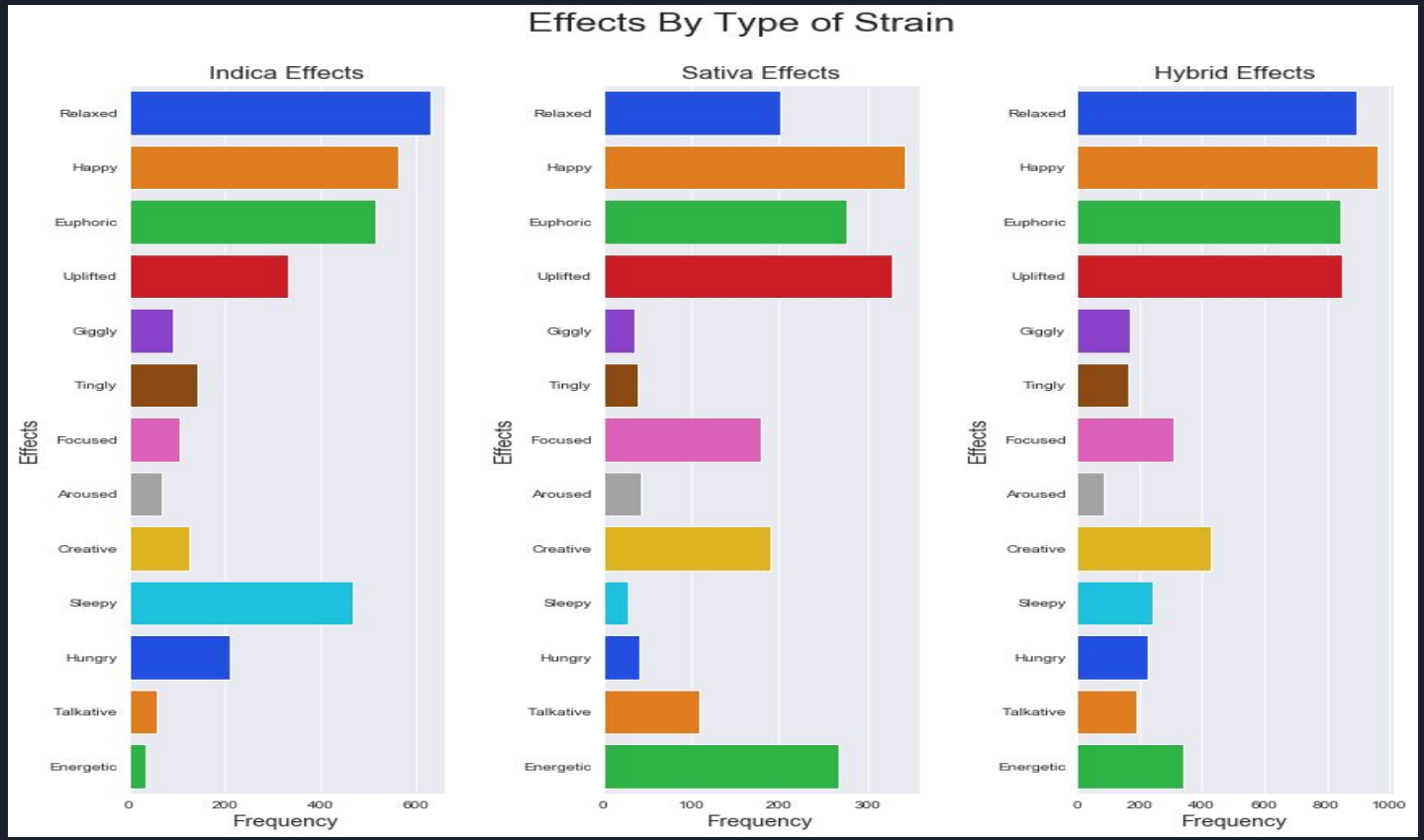
2. Scrub the Data

- Dropped missing values
- Dropped duplicated strains
- Dropped strains that had no reported effects and flavors

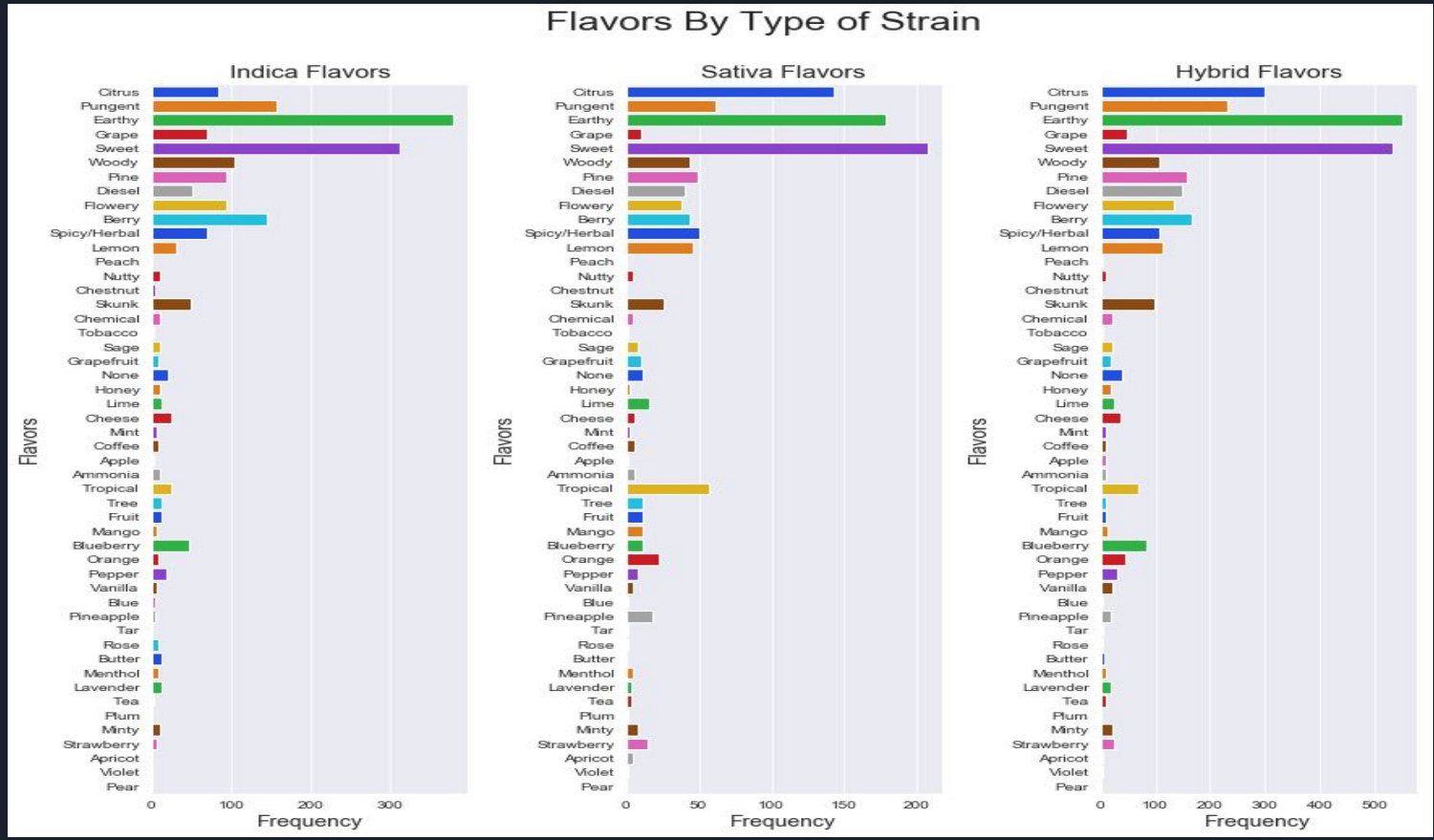
3. Explore the Data



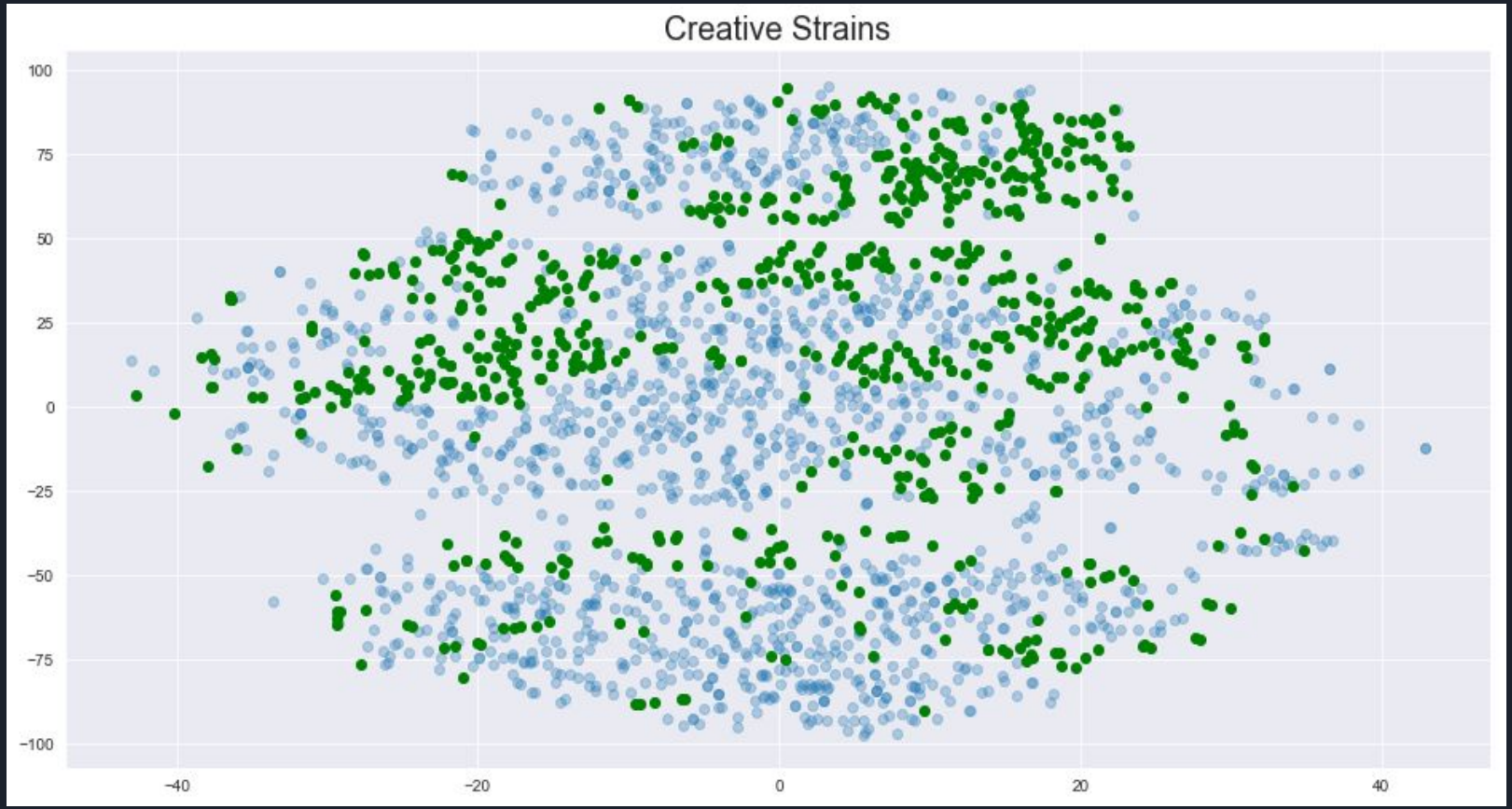
3. Explore the Data



3. Explore the Data



3. Explore the Data



4. Model the Data

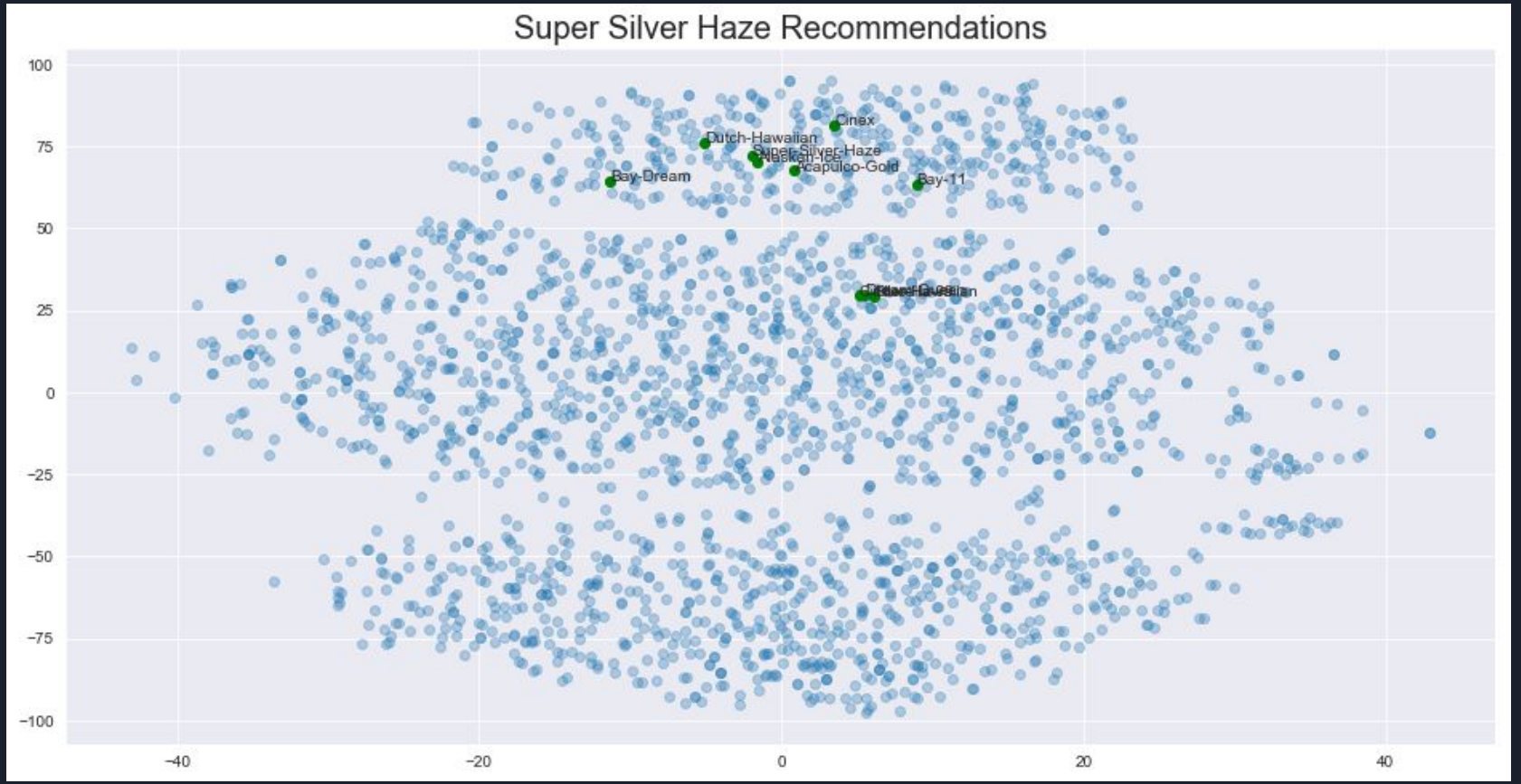
- One-Hot Encoding
- Cosine Similarity Distance Metric
- Using similarity scores recommendations were made

5. Interpret

Super SIlver Haze (Sativa) strain recommendations:

1. Acapulco Gold (.888)
2. Alaskan Ice (.888)
3. Bay 11 (.888)
4. Bay Dream (.888)
5. Blue Hawaiian (.888)
6. Cinderella 99 (.888)
7. Cinex (.888)
8. Dream Queen (.888)
9. Dutch Hawaiian (.888)

5. Interpret


Super Silver Haze Recommendations
Cinex
Dutch-Hawaiian
Super-Silver-Haze
Acapulco-Gold
Bay-Dream
Bay-11
100
75
50
25
0
−25
−50
−75
−100
−40
−20
0
20
40

Business Recommendations

- Collect more user-based data with ratings and average price.
- Use raw data available, provides more variety than reduced data

Future Work

- Use current ratings reported and implement into the current recommender.
- Gather more pricing data and new strains created, as well as user-based data to apply Collaborative Filtering.
- Use current descriptions reported to gain further insight and better suggestions.
- Apply other dimensionality reduction other than PCA, such as Autoencoders.

THANK YOU!

Jose J. Villalobos - Student Flatiron School/Data Science

E-mail: teacherjoe83@hotmail.com

GitHub: https://github.com/UberJoe83