Dashboard Blueprint

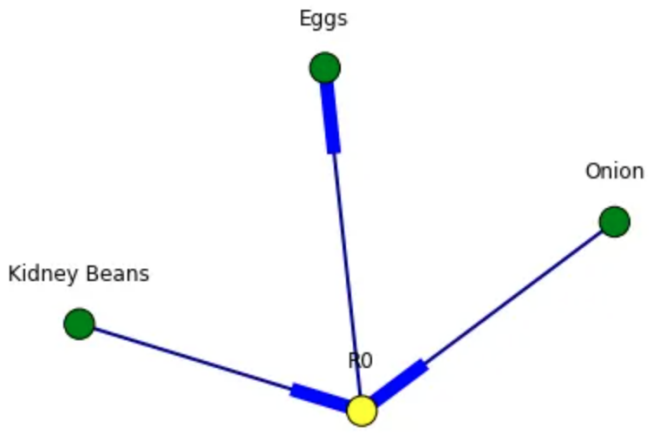
Research Question #1

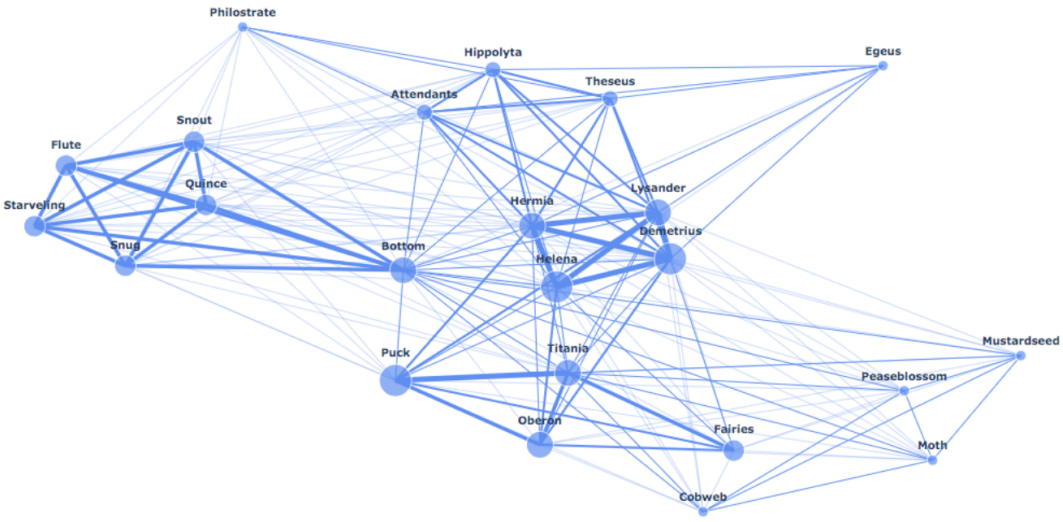
Can we predict which products a customer will most likely purchase together within various product segments?

Machine Learning: Apriori Algorithm

Visualization Plans:

- NetworkX module for charting association rules: https://intelligentonlinetools.com/blog/2018/02/10/how-to-create-data-visualization-for-association-rules-in-data-mining/
- NetworkX python module represents association rules through a diagram.
- Each diagram represents one rule association and provides arrows that connect the associative products together, please see image below for an example.
- The root node (R0) represents one association rule with incoming and outcoming edges attached to the root. The diagram represents the association between products.
- Example:





Visualization Tools:

- Python's NetworkX Library and Plotly
- https://towardsdatascience.com/tutorial-network-visualization-basics-with-networkx-and-plotly-and-a-little-nlp-57c9bbb55bb9

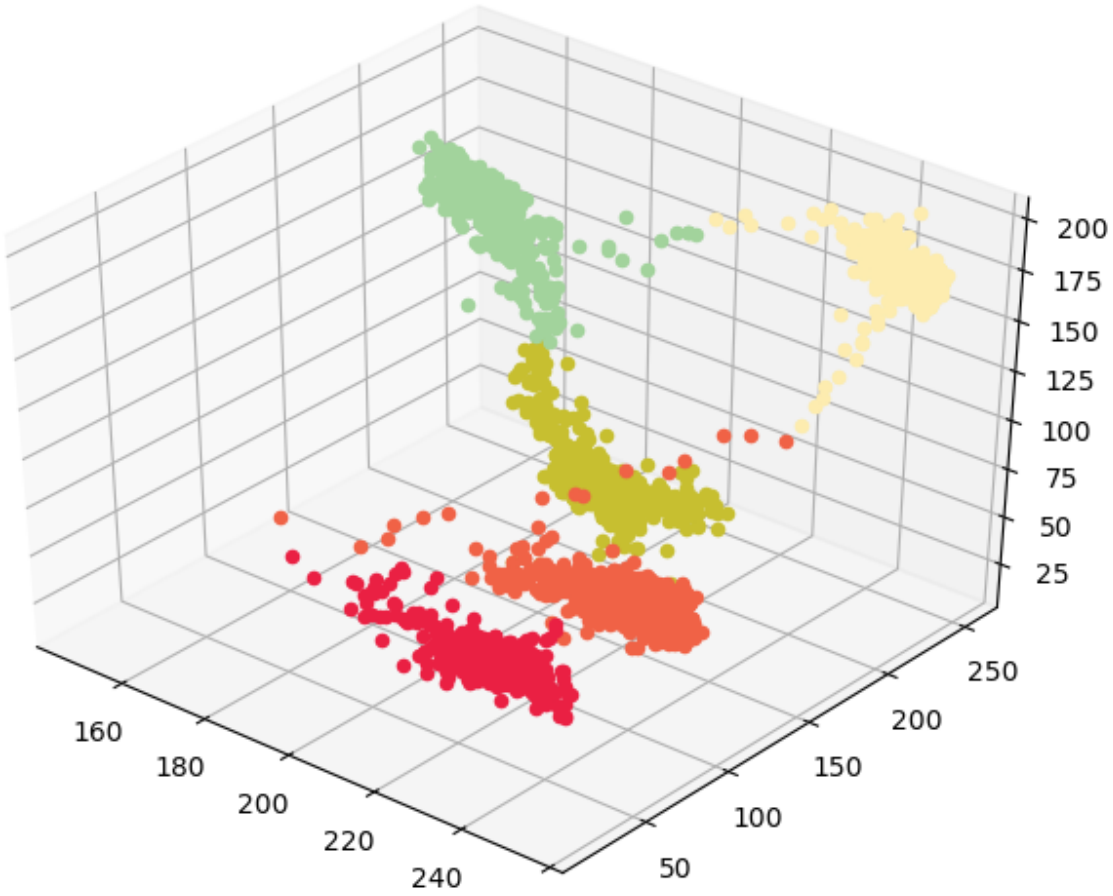
Research Question #2

Can we identify customer segments based on the purchased product categories to better target marketing campaigns?

Machine Learning: Unsupervised Learning K-Means Cluster Analysis

Visualization Plans:

- Visualize K-Means clusters through a 2D and 3D scatter plot with 3 product segments.
- The goal of this scatter plot is for the viewer to easily identify customer segments based on the purchased product categories. For example, a user can visually see high customer activity within a product category or view low customer activity within a product category.
- https://www.naftaliharris.com/blog/visualizing-k-means-clustering/
- Examples:
    -



Visualization Tools:

- Using Python, the 2D graph will be created using Hvplot.pandas dependency.
- Using Python, the 3D scatter graph will be created using Plotly.express
- https://hvplot.holoviz.org/
- https://plotly.com/python/plotly-express/

Question #3

Can we extract key topics within product reviews to help companies analyze and interpret customer feedback?

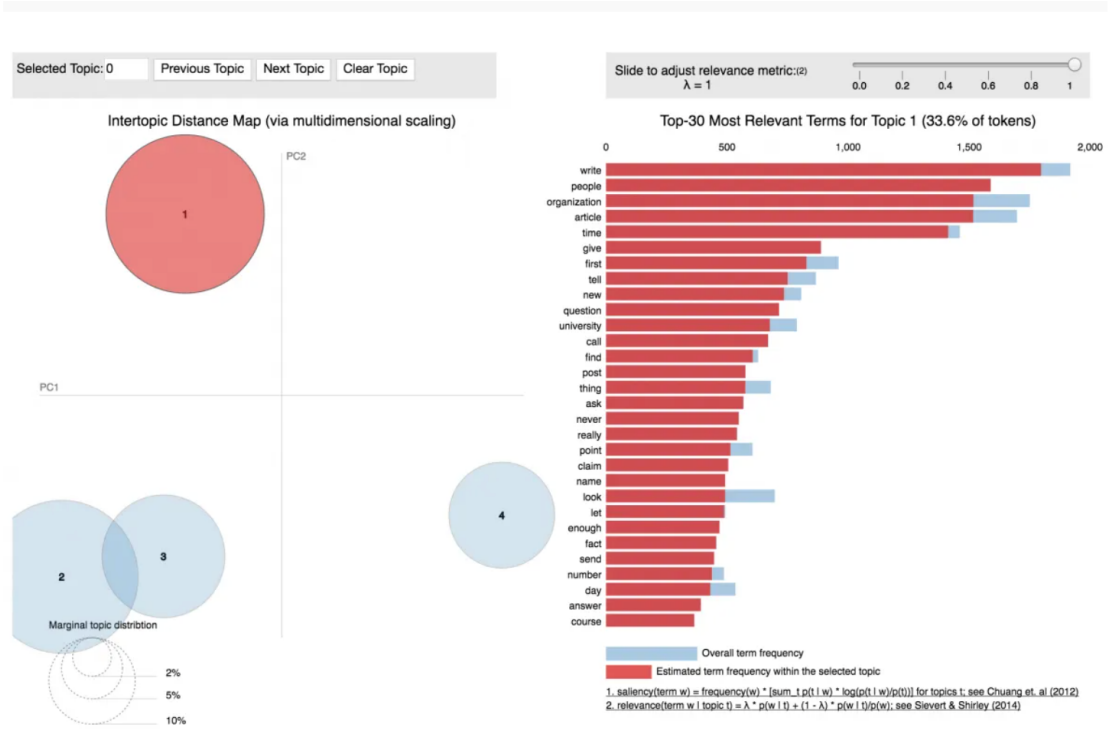
Machine Learning: Natural Language Processing Topic Analysis & Sentiment Analysis

Visualization Plans: Topic Analysis & Sentiment Analysis

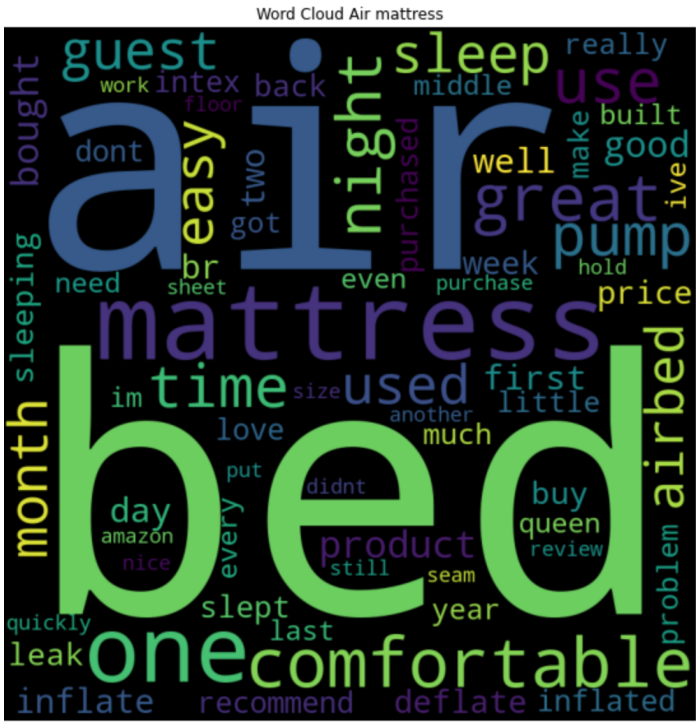
- In order to visualize highlighted topics within product reviews, a bar chart and a bubble chart can be used to display the weight and frequency of a word through a Latent

Dirichlet Allocation model (LDA).
https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/

- To display this visualization, the user can select a certain word from a dropdown. After selecting a word, a graph and a bubble chart would pop up showing the frequency and the weight of the word. A longer bar and a larger bubble would represent a word with high frequency and heavy weight. The color of the bubble and bar chart would also represent whether the sentiment of the world is negative or positive.
- Example:



- A word cloud can also be used to display the frequency and weight of the topics within review descriptions.
https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/
- Example:



Visualization Tools

- To visualize the bubble and bar charts, the Python dependency PyLDAvis.genism can be utilized to display the visualization.
    - https://pyldavis.readthedocs.io/en/latest/modules/API.html
- To create the dropdown menu, the Pyphi module and Jsonify dependency can be utilized to display the visualization through HTML.
    - https://pyphi.readthedocs.io/en/latest/api/jsonify.html
- The word cloud can be visualized through the Python Matplotlib dependency https://matplotlib.org/stable/contents.html