**Content-based Recommender System**

The RS would take item X, the item being promoted, and return the items most similar to it. The similarity would be measured in % and a list of the most similar items presented to the user.

The system would be trained on item features, e.g. description, name, store,..

* Is there any way to have labeled data here?
* Would different features be able to have different weights?
* How would the accuracy of the model be measured?

The model would then be able to take in one item, multiply it with the weight vector and return a value which would be flipped to a % (highest % is the most similar)

Then investigate how well this model predicts cannibalization items

* Would need to discuss some theory that states that cannibalization items are similar items
* Compare accuracy of the items based on features and the accuracy of predicting cannibalization items.
* This would be a classification problem so we would need labeled data.

\*Cannibalization is a known effect of recommender systems. Here we are however trying to predict what those cannibalization items will be

**Market Basket Analysis**

Similar.

Here the focus is on finding correlated items based on sale history, i.e. items that are often sold together.

This way it would be possible to estimate the sales increase of a set of items when item X is put on promotion.