**Wayfair Competitors Analytics Proposal**

Yilin Huang

1. **Problem Statement**

Wayfair Inc. is an American e-commerce company that sells home goods. the company was founded in 2002 and now sells many home furnishings and décor items and over ten million products from over 10,000 suppliers. Wayfair self-service bid management system on the Supplier Extranet allows suppliers to boost their current products on wayfair.com through an auction model.

Well Woven want to understand more about the products from competitors on Wayfair to support the decision-making process for the future campaigns.

1. **Business Goal Analysis:**

The goal for this project is to:

1. Better understand the products from our competitors
2. Build Model to predict the performance of products in new campaigns
3. **Dataset Description**

There are three types of data in this dataset:

a. Text Data: Product Description/Product Name

b. Numerical Data: Page/Size\_num/Sale\_Price/Original\_Price/Discount/Review\_Numbers/

Customer\_Rating\_Total/Size\_Rating/Material\_Rating/Color\_Rating

c. Categorical Data:Shop\_Name/Sponsor/Type

1. **Project Design**
2. Crawl competitors’ data from Wayfair
3. Preprocessing dataset
4. Conduct Exploratory Data Analytics
5. Conduct Text Analytics
6. Build Model to predict the rating and reviews for the products for future campaigns
7. **Questions Expected to Answer**

Exploratory Data Analytics

a. What kind types of rug have higher Rating/Review?

b. What kind of attributes of product will have higher Rating/Review?

c. Is Sponsor useful to generate higher Review?

d. Is Page Rank useful to generate higher Review and what page will have more Review?

e. Does more sizes has more higher Rating/Review?

f. What range of sale price/original price/Discount have higher review?

Text Analytics:

a. What keywords our competitors use very often to describe their product

b. What kind of keywords products have will generate higher Reviews

Model Building:

a. Could we use the existing dataset to build a higher accuracy model to predict the products in the future campaigns?

PS: There maybe more questions will be answered by the future analytics process if I find anything useful information from the dataset.