**SUMMARY**

**DATA PREPARATION-**

I considered *events* data for initial analysis

* Since the dataset was huge, I considered only the visitors who had viewed, added to cart and transacted the items
* Since the analysis was conducted to recommend properties of items in add to cart based on viewed items, I filtered for only “view” and “add to cart” events.
* For implementing recommender model, I created a single table by aggregating viewed and cart items by visitor-id
* The recommender model was trained on a random sample of 100 visitors and tested on a random sample of 100 visitors

**KEY DESCRIPTIVE ANALYSIS-**

Of 1,407,580 Unique Visitors…..

Over a Period of 12 years and 4 months

Only **10,228 (0.7%)** Visitors had viewed, added to cart and bought the item

Other analyses have been included with the source code

**MODEL SELECTION-**

Since the goal was to recommend items based on properties of the viewed items, I felt analysing item similarities would be more appropriate. Hence, I chose **Item based Collaborative filtering**. **Cosine similarity** was used to determine the similarity among items and top 10 similar items based on cosine similarity was selected for recommendation.

**MODEL PERFORMANCE-**

The performance was calculated as follows –

*No. of recommended items in cart*

*Total No. of unique items in cart*

While testing on a sample of 100 visitors, the performance was 88.5%