

Online Retail Giftware Dataset

The idea of this project is to perform market basket analysis and derive recommendations for the most frequently bought items together. Based on the time of the year we can also determine the demand and supply needs and what items to stock.

This Online Retail II data set contains all the transactions occurring for a UK-based and registered, non-store online retail between 01/12/2009 and 09/12/2011. The company mainly sells unique all-occasion giftware. Many customers of the company are wholesalers.

[UCI Machine Learning Repository: Online Retail II Data Set](#)

Column	Description
InvoiceNo	A 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.
StockCode	A 5-digit integral number uniquely assigned to each distinct product.
Description	Product Name
Quantity	The quantities of each product per transaction
InvoiceDate	The day and time when transaction was generated
UnitPrice	Product price per unit in sterling
CustomerID	A 5-digit integral number uniquely assigned to each customer
Country	The name of the country where a customer resides

Problem Statement

How can we give recommendations of the giftware items to the customers so that we help them shop better for every occasion and thereby improve the revenue of our company in the upcoming quarter.

Context

This Online Retail II data set contains all the transactions occurring for a UK-based and registered, non-store online retail between 01/12/2009 and 09/12/2011. The company mainly sells unique all-occasion giftware. Many customers of the company are wholesalers. We are trying to increase the revenue and help customers shop better and be ready for the gift occasions.

Success Criteria

Using the statistical inference we are planning expecting an increase in the revenue by +5%.

Scope

Our current scope is to create a recommendation engine that determines the frequently bought items together and compare it with before and after the study.

Constraints

Currently our dataset has only 2 years of data. It is offline and will can't be tested on the website using A/B testing.

Stakeholders

Raghunandan Patthar

Datasource

[UCI Machine Learning Repository: Online Retail II Data Set](#) - This contains 1067371 rows.