Prediction and analysis of online sales in retail industry Domain: Retail Industry – Online Sales

Introduction

Studying the behavior of the online retail sales is one of the key factor for the business to expand. Over a decade of time, the demand for online sales has seen a tremendous rise in terms of customers and business. When compared to traditional methods of retail, online retail stands out in various ways. Understanding the customer purchase history to help effectively promote future product offerings to customer-base is one of the key solution. It helps establish successful product types and improve future offerings.

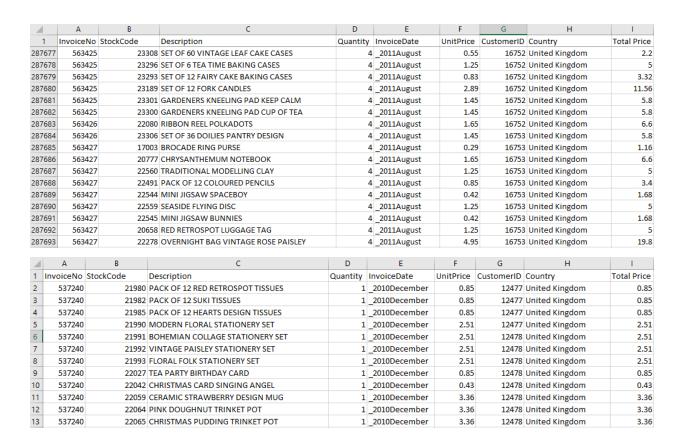
Background

A large variety of choices to choose from, firsthand information on price differences and discreetness of the activity are few of the main reasons that tempts people to buy things online. Eight in 10 Americans are now shopping online, according to a new study from Pew Research. More users these days prefer paying through their mobile phones than cash. Online sales is one such safe and convenient gateway that allows them to depend on this means of payment.

Dataset Description

The dataset focusses data for a period of two years. The data extracted maps to the online sales, predicting the change in sales. It analyzes which of the items are frequently purchased together amongst many other conclusions that can draw from the dataset. The preprocessed data consists of nearly 0.3 million rows without any null values.

Variable Name	Description	Data Type
InvoiceNo	Invoice Number	Nominal
StockCode	Product code	Nominal
Description	Description of the item	Nominal
Quantity	The number of products per transaction	Numeric
InvoiceDate	Invoice Data and Time	Numeric
UnitPrice	Unit price of each item	Numeric
CustomerID	Customer Number	Nominal
Country	Name of the country	Nominal
Total Price	Total Price of the order	Numeric



Outcomes of the Project

- 1) Items customers have purchased together frequently
- 2) Who are the most loyal customers?
- 3) Percentage change in total sales over a period of time
- 4) Number of sales and customers on a particular day.
- 5) Which item was widely sold in which country?
- 6) Predicting the sale of the most widely sold item in a particular location for next two months.

Resources Needed:

1. Weka

Approaches in the Project:

- Prediction Algorithm
- Frequent pattern tree Algorithm

Timeline

We would like to submit the entire work by end of March

References:

https://techcrunch.com/2016/12/19/79-percent-of-americans-now-shop-online-but-its-cost-more-than-convenience-that-sways-them/

http://archive.ics.uci.edu/ml/datasets/Online+Retail

 $\frac{https://www.yellowfinbi.com/blog/2010/10/yfcommunitynews-data-analysis-for-the-retail-industry-part-1-100068$