Online Retail Analytics

-A Big Data for Business Project

Group 3

Enock Chemochek

Kshitiz Kharel

Data Source and Description

Online Retail Data Set Sourced from UCI Machine Learning Repository, link to the dataset.

Transnational dataset containing transactions occurring between 01/12/2010 and 09/12/2011 for UK-based online retail.

Features Information

Feature	Description
InvoiveNo	Invoice Number, 6-digit unique number assigned to each transactions. InvoiceNo. containing 'C' indicates cancellation
StockCode	Product code
Description	Product name
Quantity	Quantity purchased per transaction
InvoiceDate	Date and Time of each transaction
UnitPrice	Product price per unit
CustomerID	5-digit number unique to each customer
Country	Country from where the products are purchased

Objective of the Project

- Label products purchased by customers

- Cluster and label customers based on their historical transactions

- Build & evaluate customer classification models for use in classifying new customers.