# **Title**: **Retail Transection Analysis**

|  |  |  |
| --- | --- | --- |
| First name | Last Name | IIT Email |
| Amrutha | Gowda | [alakshmanegowda@hawk.iit.edu](mailto:alakshmanegowda@hawk.iit.edu) |
| Vishnu | Pajjuri | [vpajjuri1@hawk.iit.edu](mailto:vpajjuri1@hawk.iit.edu) |
| Gauri | Khatri | [gkhatri@hawk.iit.edu](mailto:gkhatri@hawk.iit.edu) |
| Aditya | Dubey | [adubey6@hawk.iit.edu](mailto:adubey6@hawk.iit.edu) |

**Type of Your Projects (select one or more):** Data preprocessing, Multiple linear regression, ANOVA

1. **Introduction**

When consumer expectations change faster than the speed of business, you can’t afford to rely on fragmented data to make daily decisions that affect your bottom line of the business. But timely insights is easier than it sounds when you’re weighed down by system-crushing volumes of data, we have the power to create seamless customer experiences and deliver the personalized service that customers expects and demands. This analysis will make significant impact on the retail business as we will be having timely insights and structured analysis of the data.

1. **Data Sets**

Data describes the customers of different origin have showed interest in different products at different time.

This dataset consists of 8 columns which are mentioned as follow-:

Invoice No

Stock Code

Description

Quantity

Invoice Date

Unit Price

Customer

Country

Data is gathered from Kaggle website . we have around 4.7 lakh records and 8 features associated.

Data describes the customers of different origin have showed interest in different products at different time.

1. **Research Problems**

In order to grow retail business, improve efficiency and gain customer insights it is necessary to perform predictive analysis over the data. To grow retail business in a competitive market we need to know certain things in advance like customers interest, product’s demand in countries.

**Potential Solution**

We will be performing below analysis on the data:

1. Preprocessing of Data- cleaning of the data by removing unnecessary records.
2. Hypothesis Testing- Our hypothesis is based on comparison of average product prices among countries.

H0= average price of

1. ANOVA testing- With help of variance of product sold according to countries we will compare the average product sold out.

These solutions will help retail business to grow as it will lead to

* More accurate prediction.
* Product demand.
* Valuable insights of the products & customers.

1. **Evaluations**

Since our data set is large enough which is more than 5k hence we will go for N-fold evaluation. In which we randomly divide the dataset into k groups keeping first fold for testing and training the model on k-1 folds.

1. **Expected Outcomes**

By performing analysis our expected outcomes are follows-

* Comparison of Average price of product in different countries.
* Bottom and top most countries according to quantity sold out.
* Online Retail market of countries.
* Anticipating customer’s interest.
* Prediction for faster and more accurate management.

After you finished your proposal, you should ask yourself the following questions:

1. Is my objective/goal is clear in the proposal? Am I able to decompose the high-level objectives into some practical problems which can be solved by my proposed solutions?

Yes, here we are analyzing the product sold and prices based on countries. will be beneficial to high level problem’s solution with respect to countries. Yes, it can be solved using my proposed solution.

2. Can my solutions help me solve the proposed problems? why? are there any requirements on the data given by my solutions? Did I introduce the data? Can I used my solutions on the proposed problems?

Yes, while declaring the proposed solution different parameters are compared upon which retail stock will be analyzed.

3. Do I have a clear evaluation approach? Can I reasonably evaluate my solutions to tell that my solutions are good ones?

Well, we do have clear evaluation approach, as we will be able to predict the retail data as product prices and demand country wise.

4. Can the reader understand every detail in my proposal?

Yes, the proposed solution has been described in detail and simple manner so that the reader does not find difficulty to understand the concept.