**Development Part1**

|  |  |
| --- | --- |
| **Date** | **17-10-2023** |
| **Team ID** | **3909** |
| **Project Name** | **Product Demant using Machine learnings** |

**Table of Contents**

|  |  |
| --- | --- |
| **S.No** | **Title** |
| 1 | Problem statement |
| 2 | Data Collection |
| 3 | Data Preprocessing |
| 4 | Data visualization |
| 5 | Conclusion |

**Problem Statement:**

Predicting product demand is a critical task for businesses of all sizes. By accurately forecasting demand, businesses can optimize their inventory levels, pricing, and marketing campaigns. However, traditional demand forecasting methods are often inaccurate, especially in today's rapidly changing markets. Machine learning (ML) can be used to develop more accurate and robust demand forecasting models. ML models can learn from a variety of data sources, including historical sales data, marketing data, and social media data. This allows ML models to identify complex patterns and trends that traditional forecasting methods may miss.

Specific objectives:

1. The goal of this machine learning project is to develop a model that can accurately predict product demand.
2. The model should be able to take into account a variety of factors, including:
   1. Historical sales data Marketing data (e.g., ad spend, promotion spend)
   2. Social media data (e.g., customer sentiment, product reviews)
   3. Economic data (e.g., GDP growth, unemployment rate)
   4. Competitive data (e.g., competitor pricing, product launches)

**Data collection:**

Data collection for product demand is the process of gathering information about how much of a product people are likely to buy. This information can be used to make decisions about product development, marketing, and pricing.

There are a variety of ways to collect data for product demand. Some common methods include:

* **Sales data**: This is the most direct way to measure product demand. Sales data can be collected from point-of-sale systems, e-commerce platforms, and other sources.
* **Customer surveys**: Customer surveys can be used to collect information about customer preferences, purchase intentions, and other factors that can influence demand.
* **Market research:** Market research firms can collect data on product demand through a variety of methods, such as focus groups, interviews, and online polls.
* **Social media data:** Social media data, such as likes, shares, and comments, can be used to gauge interest in a product and predict demand.

When collecting data for product demand, it is important to consider the following factors:

* **Target market**: The target market is the group of people who are most likely to buy the product. It is important to collect data from a representative sample of the target market.
* **Time frame**: The time frame over which the data is collected is important. For example, data collected during the holiday season may not be representative of demand throughout the year.
* **Accuracy:** It is important to collect accurate data. This can be done by using reliable data sources and by carefully designing surveys and other research methods.

Once the data has been collected, it can be analysed to identify trends and patterns. This information can then be used to make predictions about future demand.

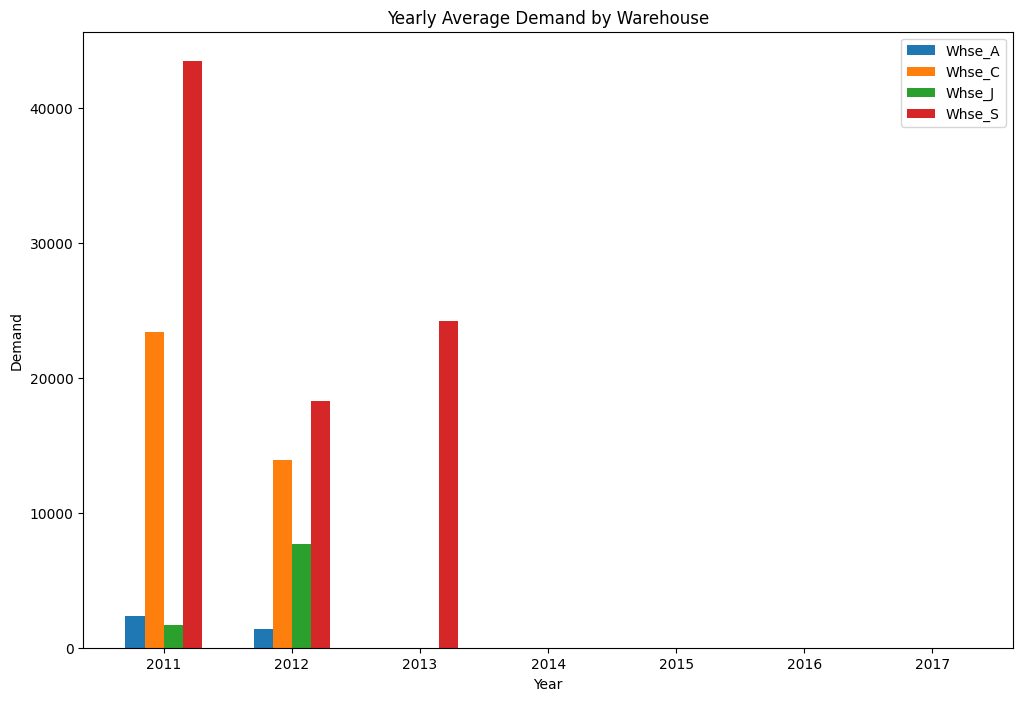
**Data Pre-processing:**

Data preprocessing is the process of cleaning, transforming, and integrating data in order to make it ready for analysis. This may involve removing errors and inconsistencies, handling missing values, transforming the data into a consistent format, and scaling the data to a suitable range.

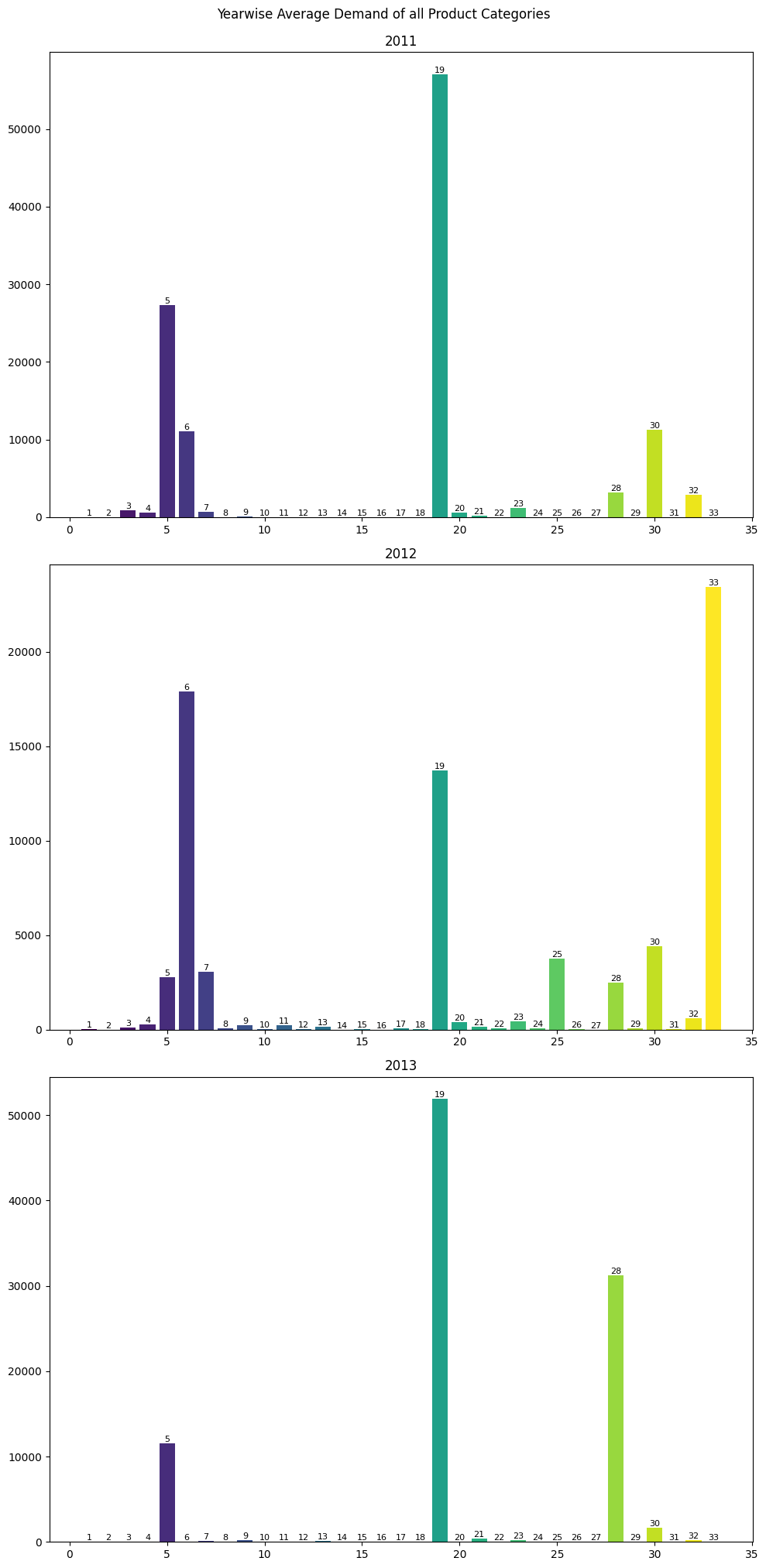
* **Removing errors and inconsistencies:** This may involve identifying and correcting typos, outliers, and other errors in the data. It is also important to ensure that the data is consistent in terms of its format and encoding.
* **Handling missing values:**Missing values can be handled in a variety of ways, depending on the nature of the data and the specific machine learning model that will be used. Some common methods for handling missing values include:
  + **Imputation:** This involves replacing missing values with estimated values. For example, the mean or median of the feature could be used to impute missing values.
  + **Deletion:** This involves removing rows or columns with missing values. This is a simple approach, but it can lead to a loss of data.
* **Transforming the data into a consistent format:** This may involve converting categorical data to numerical data, normalizing the data, and creating new features.
* **Scaling the data to a suitable range:** This is important because some machine learning models are sensitive to the scale of the data.

By carefully preprocessing the data, we can improve the accuracy and performance of our machine learning models.

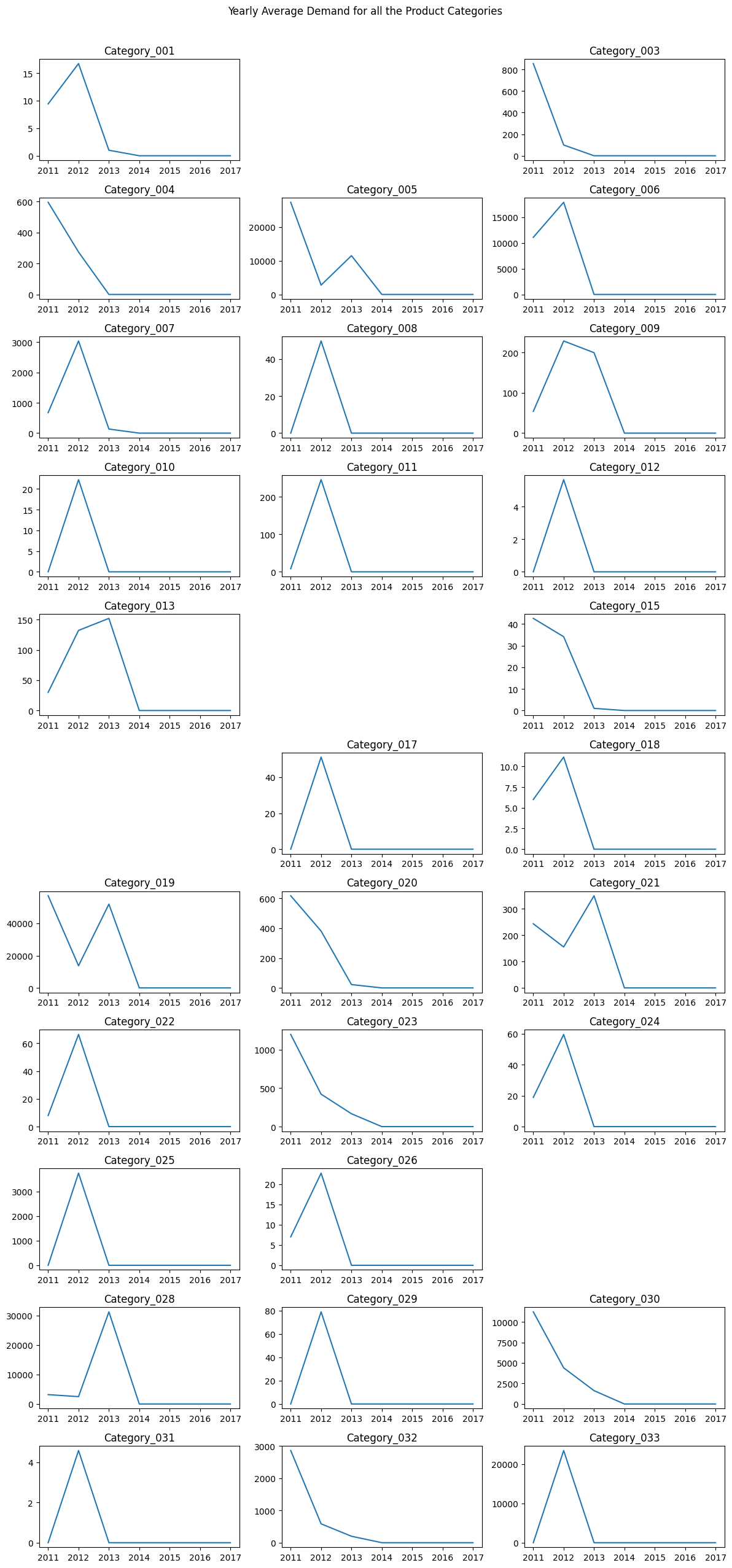
**Data visualization:**

****

**Yearly average demand by warehouse**

****

**Yearwise Average demand of all products**

****

**Yearwise average demand of product categories**

**Conclusion:**

Data preprocessing for product demand is a crucial step in demand forecasting, improving the accuracy, performance, and interpretability of machine learning models. It involves cleaning, transforming, and integrating data to make it ready for analysis.