DASHBOARD TO ANALYZE IMPORT AND EXPORT ANALYSIS OF INDIA

A PROJECT REPORT

Submitted by

MONISHA K (21ADR029)

MOHAMED RIZWAN M (21ADR028)

DINESH KUMAR B (21ADR013)

NAVEENAN M (21ADR031)

for

20ADC33 DATA ANALYSIS

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

KONGU ENGINEERING COLLEGE

(Autonomous)

PERUNDURAI ERODE–638060

DECEMBER 2022

KONGU ENGINEERING COLLEGE
(Autonomous)

PERUNDURAIERODE–638060
DECEMBER 2022

Department of Artificial Intelligence

20ADC33–DataAnalysis Project Report

Signature of course in-charge

Signature of the HOD

Submitted for the continuous Assessment viva voice examination held on _____

EXAMINER I

EXAMINER II

ABSTRACT

The import-export industry is important to the economy of the nation. A country participates in international trade when it purchases goods and services from another nation; by contrast, a nation participates in international trade when it sells goods and services to other nations. To balance and benefit one nation, it imports what it lacks and exports what it has in huge quantities. And that is how one way or the other, all the countries grow their economy**.** India is one of the growing countries and it's growing demand for consumer goods and services has also led to an increase in imports, while the continued growth of the Indian economy has allowed the country to become competitive in the global export market.

This report shows the clear understanding of India's import and export by establishing charts and dashboard using the power BI service. The dataset "EXPORTS AND IMPORTS ANALYSIS OF INDIA" is taken from kaggle. Before using the data it should be cleaned, then the data is visualized. Data modeling is the act of creating a visual representation of a full information system or specific parts of it in order to express relationships between data points and structures. After preprocessing and modeling the data, the pie charts, bar charts, tree map, slicer and some other charts are created to understand the import and export of India by using visualization field. Finally the charts are uploaded in the power BI service to create the dashboards with slicer under some category and to make the analyze clear.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

The logistics involved in international trade are referred to as import and export. An import is a good or service that is made overseas and bought in your native nation. When native industries are unable to provide comparable goods and services affordably or effectively, customers are attracted to imported goods and services. Products and resources are less expensive to import frequently depends on free trade agreements and tariff schedules. A good or service that is exported is one that is produced in one nation but is sold to a customer abroad. Export is one of the earliest forms of trade and occur often between nations. Exporting to new areas can increase sales and profits and even provide companies the possibility to capture a large portion of the global market. The Indian economy greatly benefits from international trade. International trade is important since the nation needs to import a variety of goods. India imports and exports the same number of goods in both ways. Let's look at the import and export data for India.

History of Indian trade

Trade agreements existed between ancient India and China, Europe, Iran, and Egypt. India joined the Commonwealth of Nations in 1947, after its freedom from the British. However, the economy of the nation was closed. The government kept a watchful eye on imports and exports. India's economy was modernized following that. No one has turned around since then. In this sense, the word "Indian Trade" refers to the participants of India in the trade. India showed improvement than other modern undeveloped nations in terms of trade volume and the variety of commodities joining trading lists.

Current Indian trade

India imported 367 million in thousands of dollars and exported 275 million in thousands of dollars. There are top fifteen trading partners of India account for 60% of its overall trade. US, China, Hong Kong, Singapore, UK, Bangladesh, Germany, Netherlands, Nepal, Malaysia, Italy, and Saudi

Arabia are a few of these. According to the trading partner agreements, all of these Indian trading partners import and export goods to India.

India’s major import and export

Crude oil, electronic equipment, gold, silver, electrical machinery and equipment, electronic equipment, including computers, iron and steel, organic and inorganic chemicals, and electronic products and machinery are India's top imports. Refined petroleum, jewellery, ready-to-wear clothing, ready-to-eat foods, packaged medications, vehicles and other forms of transportation, ready-to-eat foods, cotton, ready-to-eat foods, ready-to-eat foods, and ready-to-eat foods are among India's top exports. India's growing demand for consumer goods and services has also led to an increase in imports, while the continued growth of the Indian economy has allowed the country to become competitive in the global export market.

Analysis of import and export using power bi

Power BI is a software that is developed by Microsoft. It is used to visualize the data which mainly focus on business intelligence. Power BI is the cluster of number of computer program, applications, and connectors that analyze different and unrelated data sources and provide useful insights and conclusions.

Power BI desktop : It is a Microsoft Business Intelligence which helps the users to import, analyze and create report for the data which has more complex information and the data from number of sources. Small and medium-sized enterprises often utilise Power BI Desktop, which is a free Windows-based program for PCs and laptops.

Power BI service : Power BI Service is an online Software as a Service application for developing and visualizing dashboards and reports**.**

Here the Power BI Desktop is used to create the pie charts, bar charts, tree map etc… using the visualization field to make insight about the data. And the Power Bi Service is used to create dashboard to make them look better.

1.2 Data collection

Data collection is the procedure of collecting, measuring and Analyzing accurate insights for research using standard validated techniques.
https://www.kaggle.com/datasets/rajanand/import-and-export-by-india

The dataset was provided by the Kaggle challenge. Users of Kaggle have access to data sets that they may search and share, analyse, and build models in a web-based data science environment, collaborate with other data scientists and machine learning experts, and participate in competitions to solve data science objectives.

The dataset is available in Excel format. The data is directly downloaded from the Kaggle. For better understanding of the imports and exports by India, the data is available by principle commodity and country wise for 3 years from Apr'2014 to Mar'2017.There are 6 tables in which three of the tables are for import and the other three tables for export data.

1. PC_Export_2014_2015
2. PC_Export_2015_2016
3. PC_Export_2016_2017
4. PC_Import_2014_2015
5. PC_Import_2015_2016
6. PC_Import_2016_2017

Each tables contains seven attributes. The attributes are

1. pc_code **:** Integer, Principle Commodity Code
2. pc: String, Principal Commodity Name
3. unit: String, measurement of quantity
4. country_code: Integer, country code
5. country_name: String, country name
6. quantity: Integer, quantify of export or import
7. value: Integer, monetary valeu of the quantity (in million USD)

1.3 Problem statement

Maintaining the proper balance between import and export is essential for a country and to reduce the Indian import expenditure and raise its profit through exports. A country's import and export activities can have an impact on its GDP, exchange rate, degree of inflation, and interest rates. They will benefit from it as the economy of our nation grows. For that, the comparison of the products that India import from other countries and try to reduce the import of some products by producing that product on our own, but it can be done only with the help of clear knowledge about the past records of import and export of our country. So analyzing this data will give the clear insight about the products which includes in imports and exports, it will leads to improve the country’s economy.

1.4 Business objective

1. From the analysis the production of the highly exporting products increases.
2. From the analysis the imports from other countries decreases to increase our economy.
3. Introduce some new products according to the demand previous product.

.

CHAPTER 2

DATA PREPARATION AND MODELING

2.1 Data cleaning

Data cleansing is the process of deleting, altering, or manipulating data to enhance its quality and accuracy before it is used for analysis. Errors are fixed, duplication are eliminated, and missing data are filled in to achieve this. Data cleaning ensure that the data is correct and reliable before any analysis is done on it, making it a crucial step in the data analysis process. When it comes to data analysis, this information is typically not required or useful because it could impede the process or produce unreliable results. Data cleaning is not just about deleting data to create room for new data; rather, it is about figuring out how to increase an it set's accuracy without necessarily deleting data.

2.2 Data transformation

The process of data transformation is comprised of several sub-processes including cleaning, standardization, de-duplication, verification, sorting, and other duties. Transformation is generally considered to be the most important part of the ETL process. By deleting duplicates and ensuring that raw data arrives at its new location reasonably satisfied and ready for use, data transformation promotes data integrity.

ETL Process

Extraction - The majority of firms utilise a variety of data analysis techniques to develop business intelligence while working with data from numerous sources. The data needs to be free to move between platforms and apps in order to implement such a complicated data approach. Data must first be extracted from its original sources, such as a data warehouse or data lake, before it can be moved to a new location. Both structured and unstructured data are imported and combined into a single repository in this initial step of the ETL process..

Transformation - Sorting and normalising the data in the database or data warehouse keeps some or all of it on hand and available for customised reporting. Although there is more cost associated with holding this much data, there are more opportunities to mine it for useful business intelligence in close to real-time.

Loading- The final stage of the ETL procedure is data loading into the target data storage database. Typically, a large volume of data must be loaded into a database system in a little amount of time (nights). Therefore, performance optimization should be applied to the load process.In case of load failure, recovery mechanisms should be configured to restart from the point of failure without data integrity loss. Data Warehouse admins need to monitor, resume, and cancel loads as per prevailing server performance.

Steps involved in transformation are

1. Load the data into the Power BI Desktop by using get data in the top of the screen.

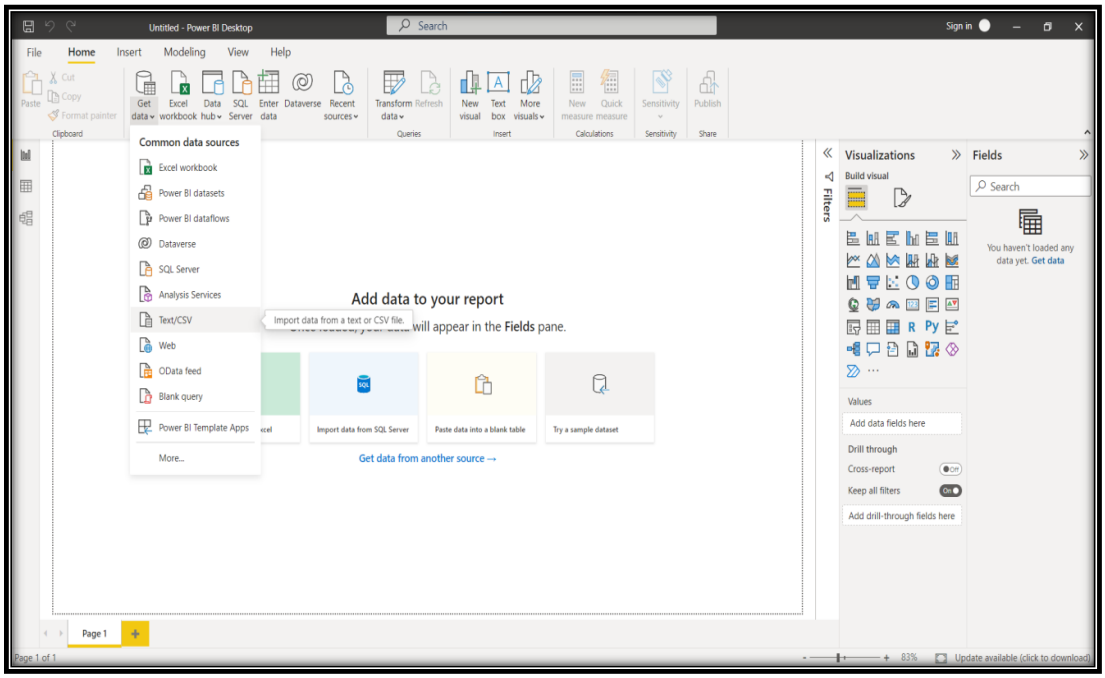


Figure: 2.1 Loading the data into Power BI

2. Then go to the Power Query Editor by clicking transform data in the ribbon at the home tab.

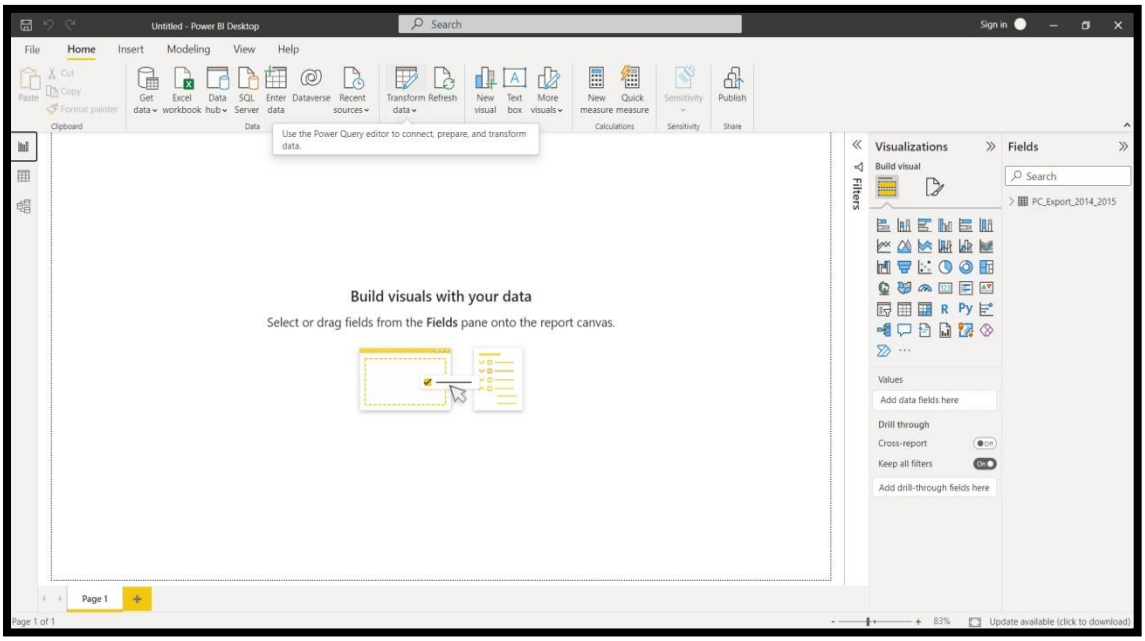
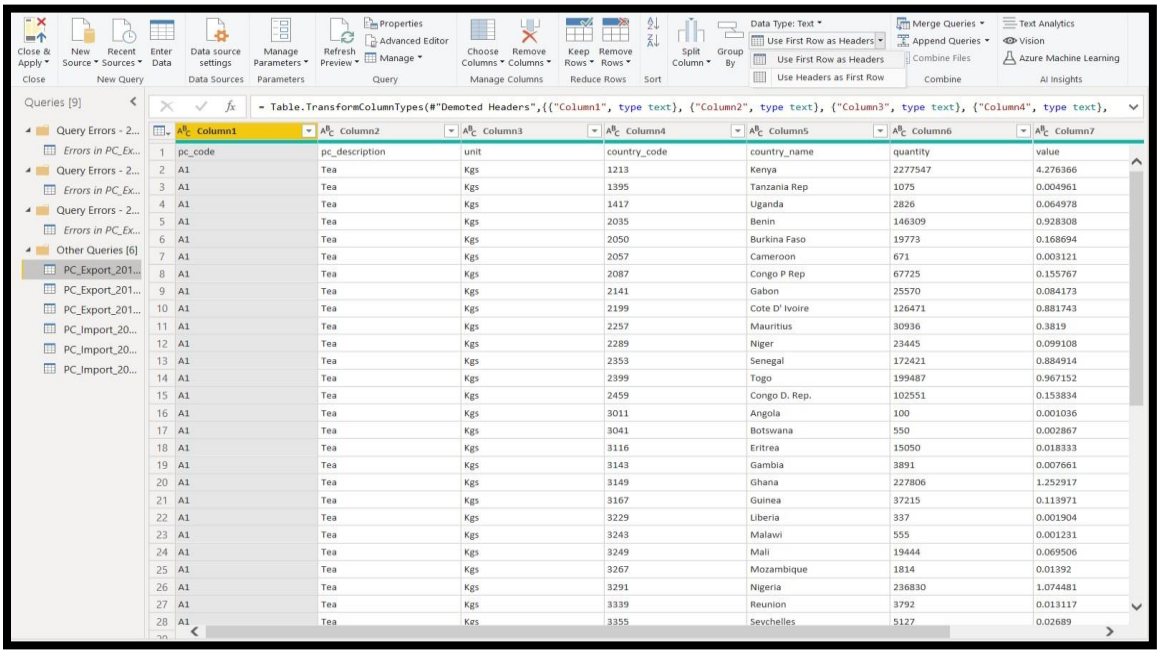


Figure: 2.2 Transformation of data

3. Make an insight of data and be sure that the column names are used as header, if not change the first row as header.



	Column1	Column2	Column3	Column4	Column5	Column6	Column7
1	pc_code	pc_description	unit	country_code	country_name	quantity	value
2	A1	Tea	Kgs	1213	Kenya	2277547	4.276366
3	A1	Tea	Kgs	1395	Tanzania Rep	1075	0.004961
4	A1	Tea	Kgs	1417	Uganda	2826	0.064978
5	A1	Tea	Kgs	2035	Benin	146309	0.928308
6	A1	Tea	Kgs	2050	Burkina Faso	19773	0.168694
7	A1	Tea	Kgs	2057	Cameroon	671	0.003121
8	A1	Tea	Kgs	2087	Congo P Rep	67725	0.155767
9	A1	Tea	Kgs	2141	Gabon	25570	0.084173
10	A1	Tea	Kgs	2199	Cote D' Ivoire	126471	0.881743
11	A1	Tea	Kgs	2257	Mauritius	30936	0.3819
12	A1	Tea	Kgs	2289	Niger	23445	0.099108
13	A1	Tea	Kgs	2353	Senegal	172421	0.884914
14	A1	Tea	Kgs	2399	Togo	199487	0.967152
15	A1	Tea	Kgs	2459	Congo D. Rep.	102551	0.153834
16	A1	Tea	Kgs	3011	Angola	100	0.001036
17	A1	Tea	Kgs	3041	Botswana	550	0.002867
18	A1	Tea	Kgs	3116	Eritrea	15050	0.018333
19	A1	Tea	Kgs	3143	Gambia	3891	0.007661
20	A1	Tea	Kgs	3149	Ghana	227806	1.252917
21	A1	Tea	Kgs	3167	Guinea	37215	0.113971
22	A1	Tea	Kgs	3229	Liberia	337	0.001904
23	A1	Tea	Kgs	3243	Malawi	555	0.001231
24	A1	Tea	Kgs	3249	Mali	19444	0.069506
25	A1	Tea	Kgs	3267	Mozambique	1814	0.01392
26	A1	Tea	Kgs	3291	Nigeria	236830	1.074481
27	A1	Tea	Kgs	3339	Reunion	3792	0.013117
28	A1	Tea	Kgs	3355	Seychelles	5127	0.02689

Figure: 2.3 Changing the first row as header

4. It is necessary to check all the fields are in the correct datatype. The datatype of pc_code is changed from number to integer, and the field quantity is converted into whole number for easy to use.

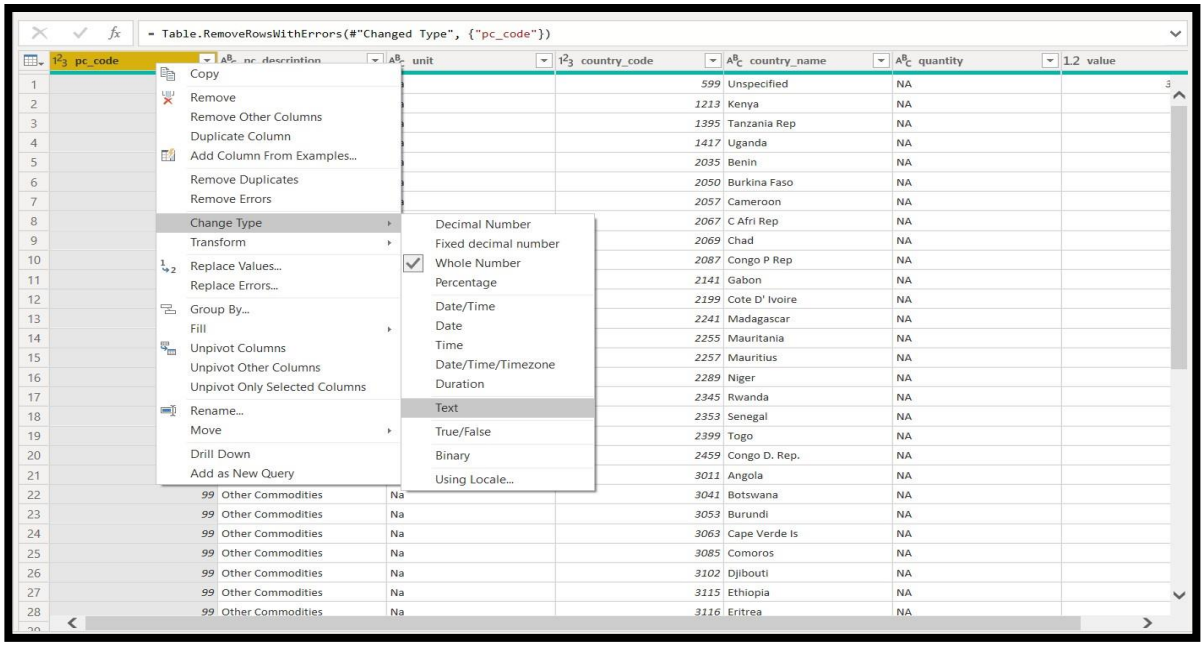


Figure: 2.4 changing the datatype

5. Removing of null values plays the import role in accuracy. The value NA is removed from the units.

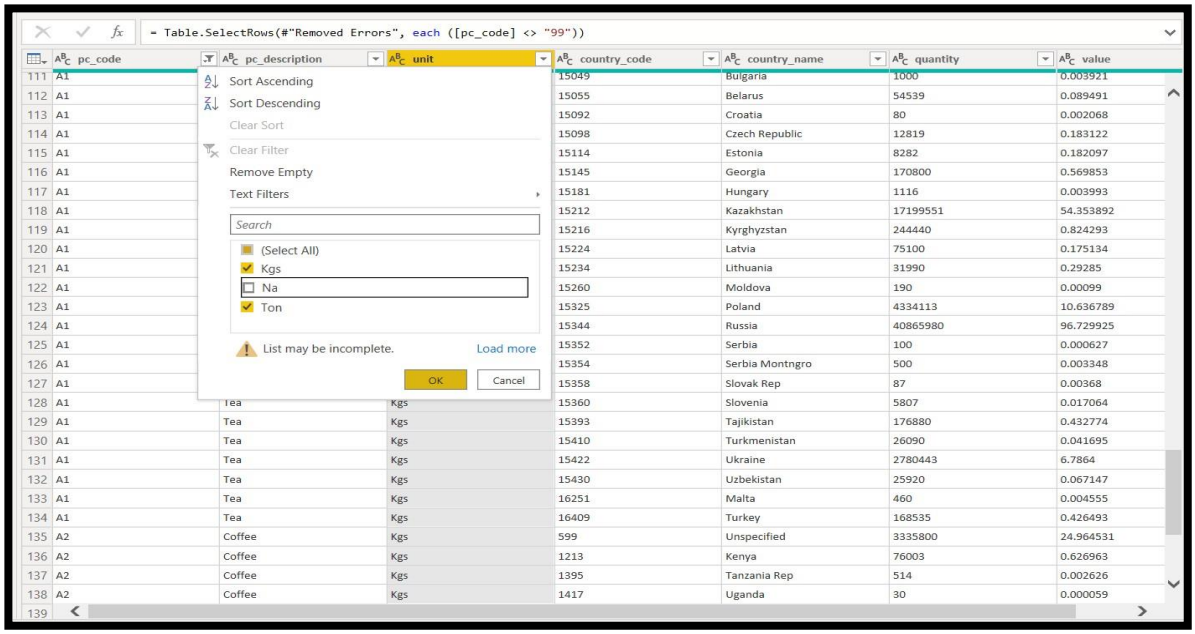


Figure: 2.5 Removal of null values

6. While changing from one type to other type or removing null values, there may be possibilities of occurring errors. And it is need to be removed for the better accuracy.

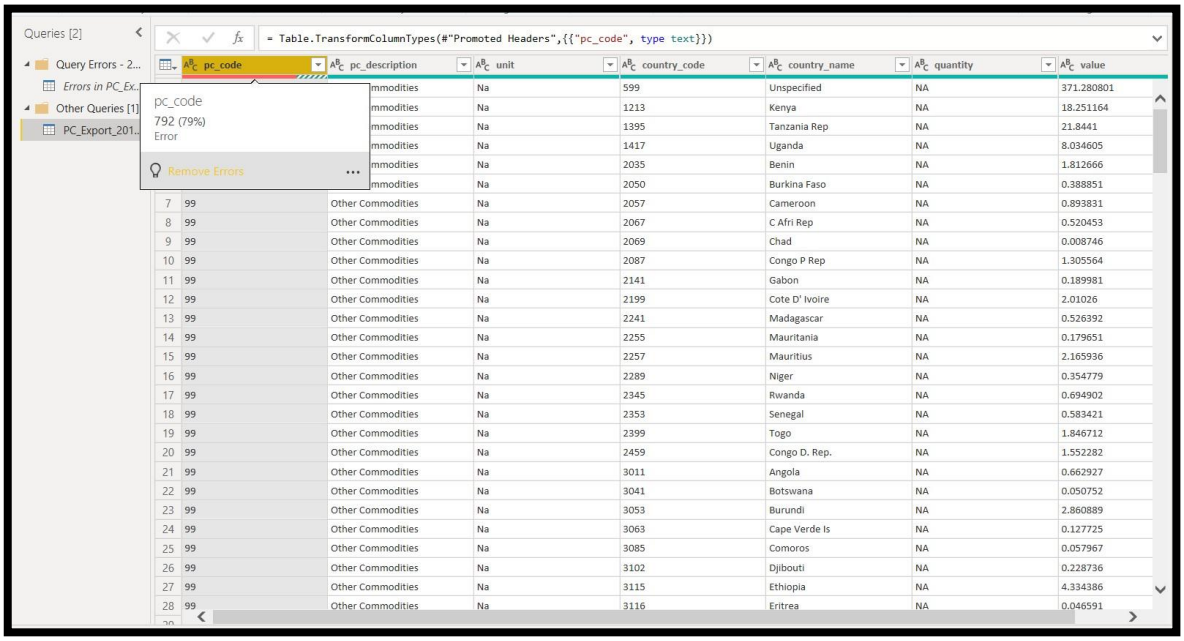


Figure: 2.6 Removing the errors

7. Then the important and required values should be selected for better understanding of that values. Here the required and essential commodities are taken.

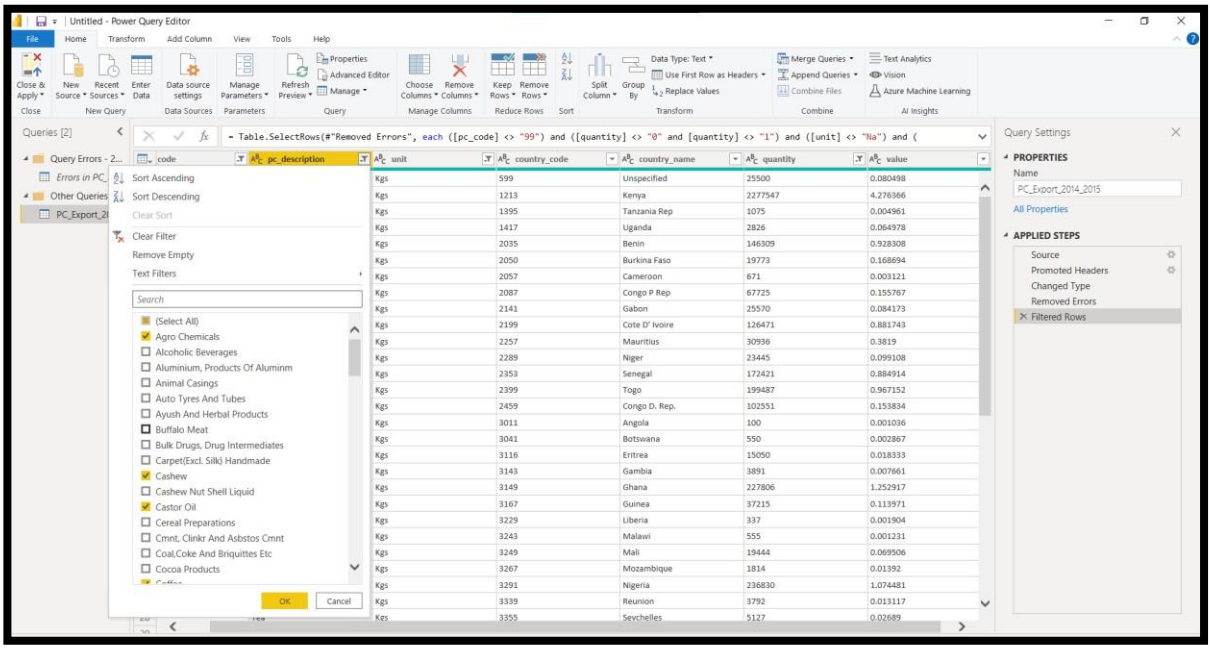


Figure:2.7 Selecting the important commodities

8. Finally, close the Power Query Editor and apply the changes to use the data for analyzing the import and export of India.

DAX function

Data Analysis Expressions (DAX), a package of expressions, functions, operators, and constants, that helps user to create tables, columns, measurements. DAX function is used to yield one or more values that make a relationship between the variables of the data and used for data analyze process. The existing model is used to create the new information about the data. For this import and export dataset, some of the measures are

Measure : 2014 to 2015 max sale = MAX(PC_Export_2014_2015[Total sales])

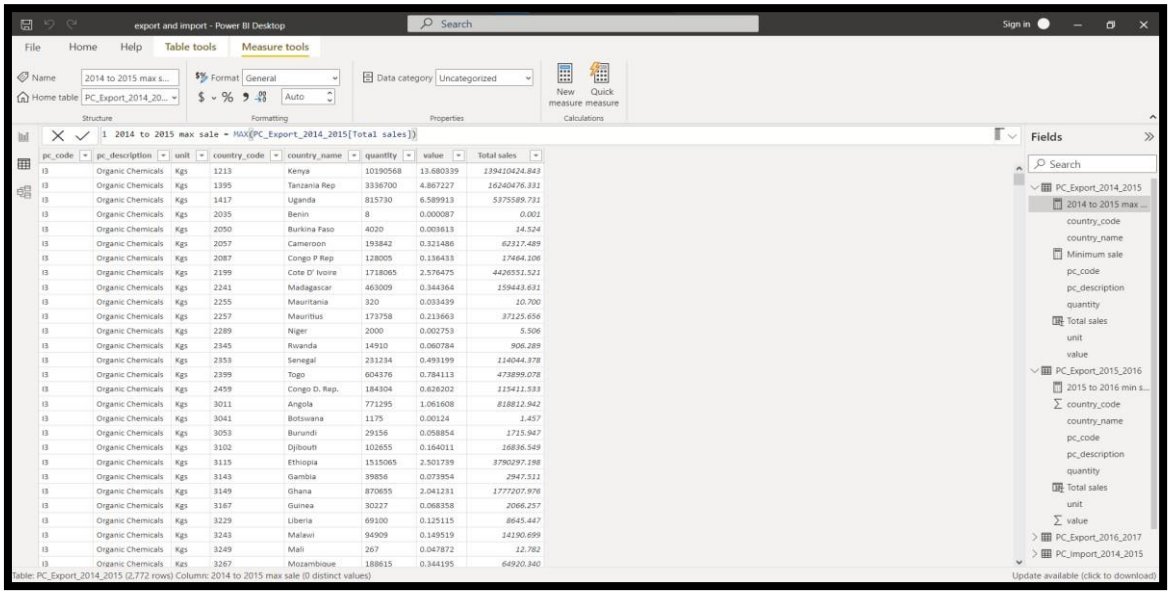


Figure: 2.8 DAX function for maximum total sales

Measure : 2014 to 2015 min sale = MIN(PC_Export_2014_2015[Total sales])

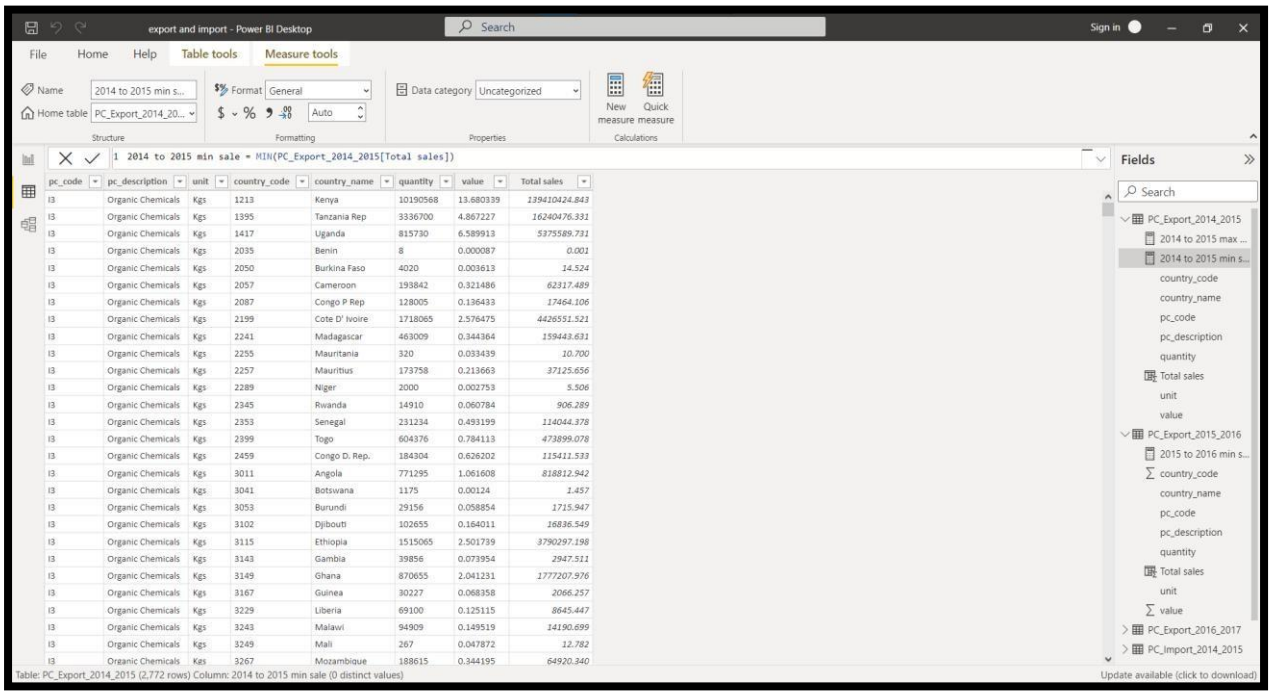


Figure: 2.9 DAX function for minimum total sales

Measure : avg of export from 2014 to 2015 = AVERAGE(PC_Export_2014_2015[Total sales])

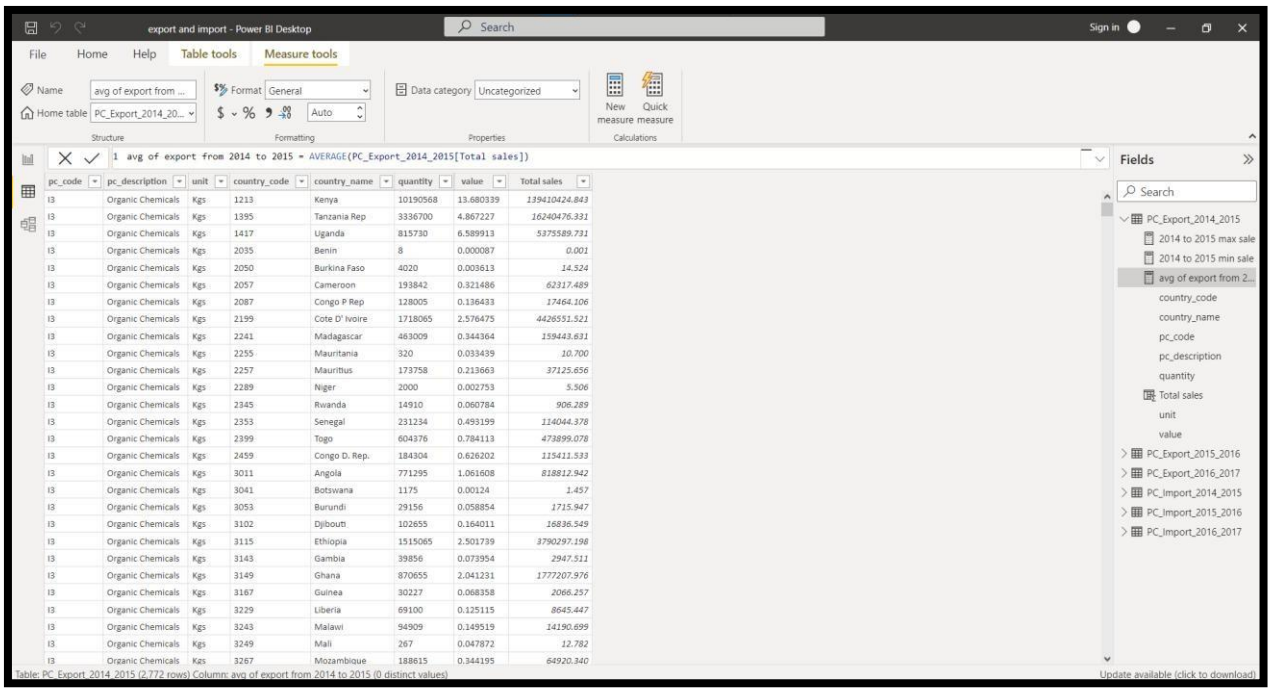


Figure: 2.10 DAX function for average total sales

Calculated columns

Calculated columns are the columns that are created to add new information about the data in the existing dataset by using DAX function. Unlike DAX function calculated column returns separate value for each rows

Columns : Total sales = (PC_Export_2014_2015[quantity])*(PC_Export_2014_2015[value])

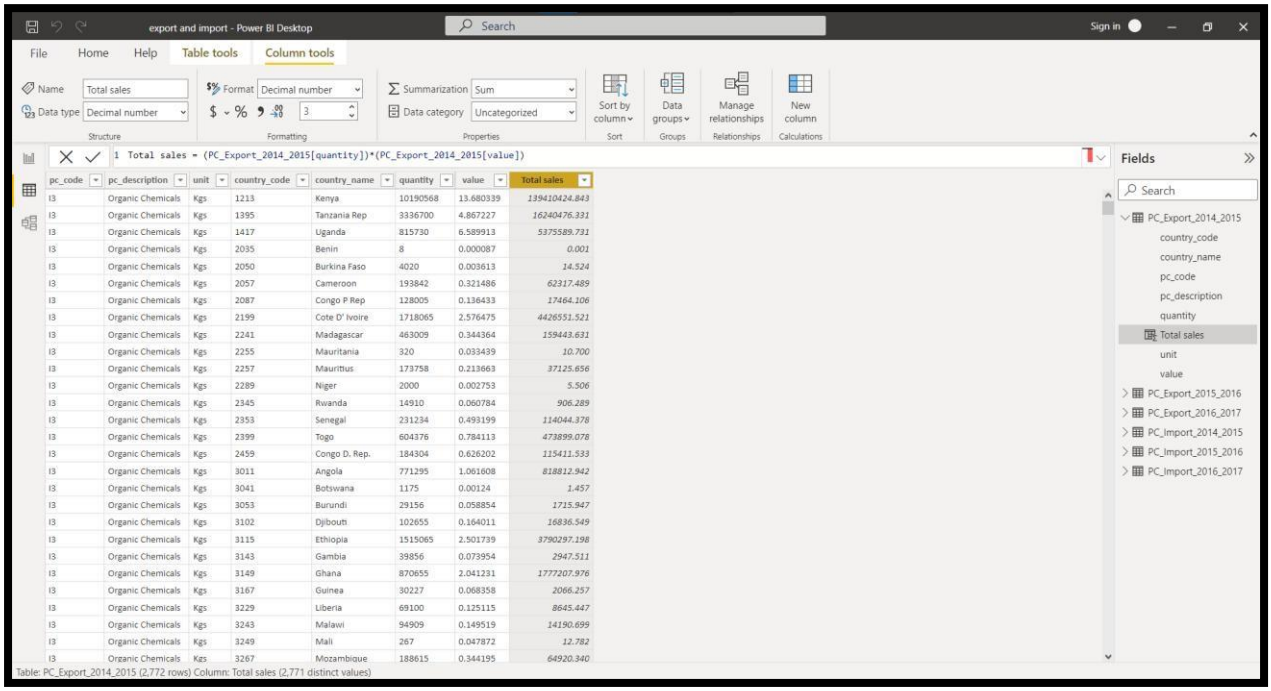


Figure: 2.11 Calculated column for total sales

Calculated tables

The flexibility to add new tables based on data that are already entered into the model is given by calculated tables. In order to set the values of the table, there is a need to construct a Data Analysis Expressions (DAX) formula rather than querying and loading data into the columns of your new table from a data source. Like wise, the measures and calculated columns are created for all the six tables.

2.4 Data modeling

Data modeling is the act of creating a visual representation of a full information system or specific parts of it in order to express relationships between data points and structures. The goal is to show how various types of data are used and stored within the system, the connections between these various types of data, the various ways that data can be organized and classified, as well as its formats and features. Data models are created to meet business requirements. Prior to the design of a new system or during an iteration of an existing one, rules and requirements are developed through input from business stakeholders.

In the import and export dataset which contains six tables, where all the attributes of each tables are interrelated with each other. The main attributes are pc_code and country_code. The attribute pc_code says that how other countries import and export commodities to and from India. It is used to analyze the top commodities imported and exported from India. Similarly, the country_code represents the countries who are the partners of Indian trade.

Figure: 2.12 Data modelling

CHAPTER 3

DATA ANALYSIS AND INTERPRETATION

3.1 Data analysis

Data analysis is the process of modifying, processing, and cleaning raw data in order to obtain useful, relevant information that supports business decision-making. The process offers helpful insights and statistics, frequently presented in charts, graphics, tables, and graphs, which minimize the risks involved in making decisions.

Power Bi tool is used to make the charts in order to analyze India's import and export, estimate the value, and find key facts about the trade. For the analysis, the data are divided into 3 categories.

1. Total sales

- Which country is most profitable for the year 2014 to 2015?
- Compare the total sales of import for all the three years.
- Give the maximum sales for import and export based on the year.
- Compare the average sales for the export between the year 2014 to 2015 and 2015 to 2016.
- Which product has most profit based on the total sales?

2. Products

- List the top 3 countries which export marine products for the year 2016 to 2017.
- How many countries export tea and sugar based on the year.
- What are the countries that export petroleum product?

3. Country

- Which country import large amount of vegetable oil?
- What are the countries that import minimum products from India?
- Which country export mass quantity of product?
- Which unit is mostly transported to a country?

1. Total sales

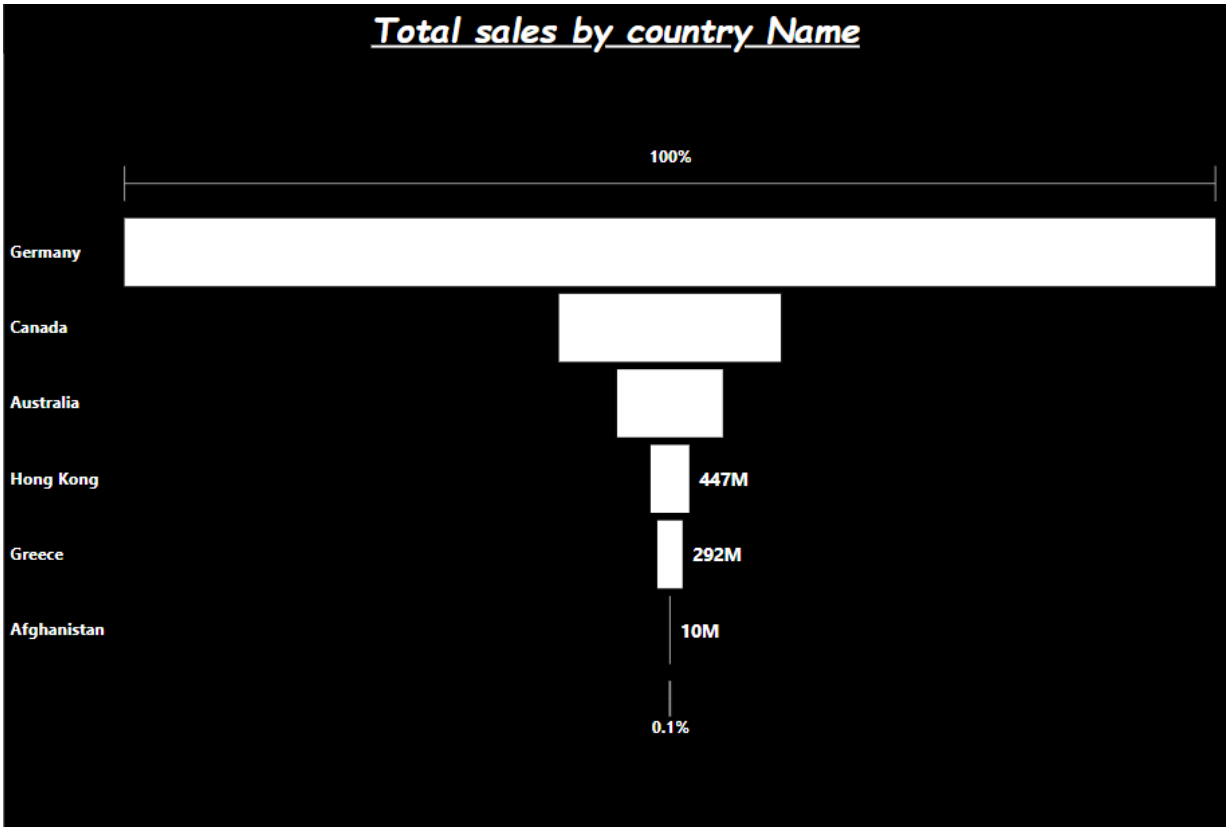


Figure: 3.1 Country based in total sales

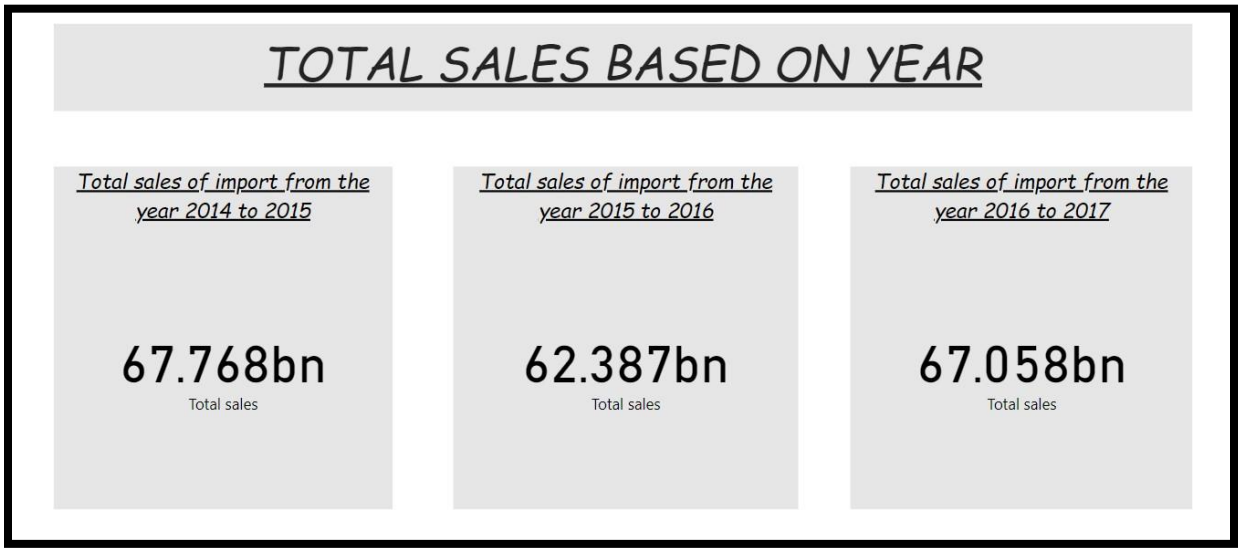


Figure: 3.2 Total sales by year

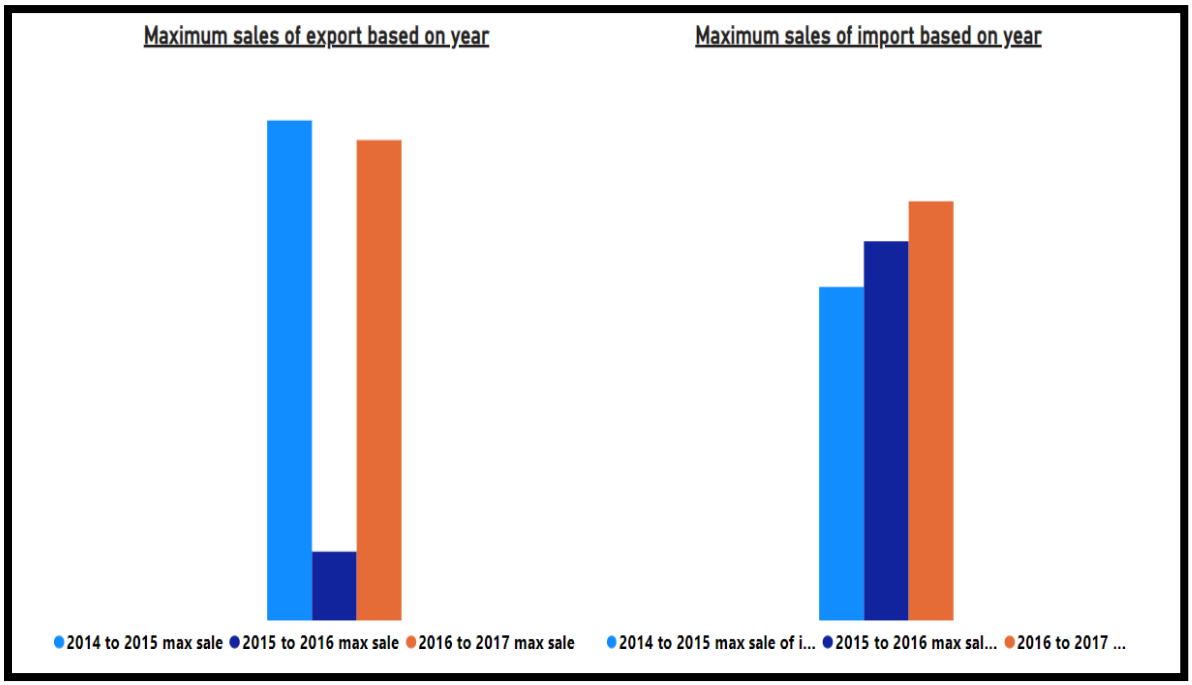


Figure: 3.3 Maximum sales for import

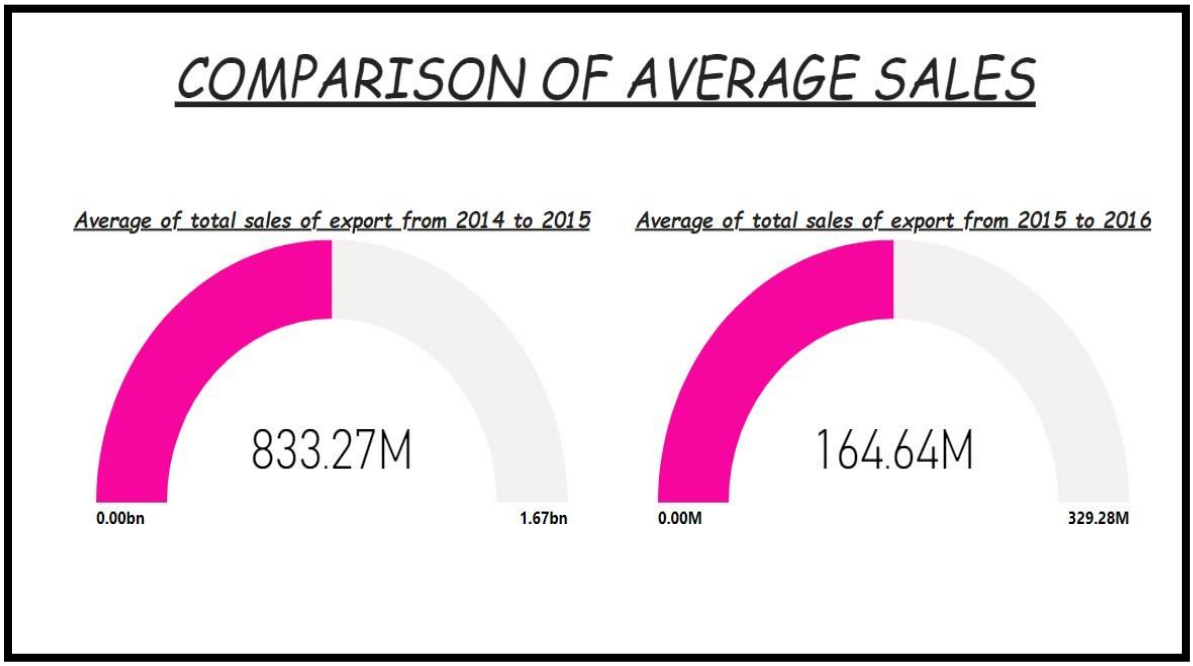


Figure: 3.4 Comparison of average sales

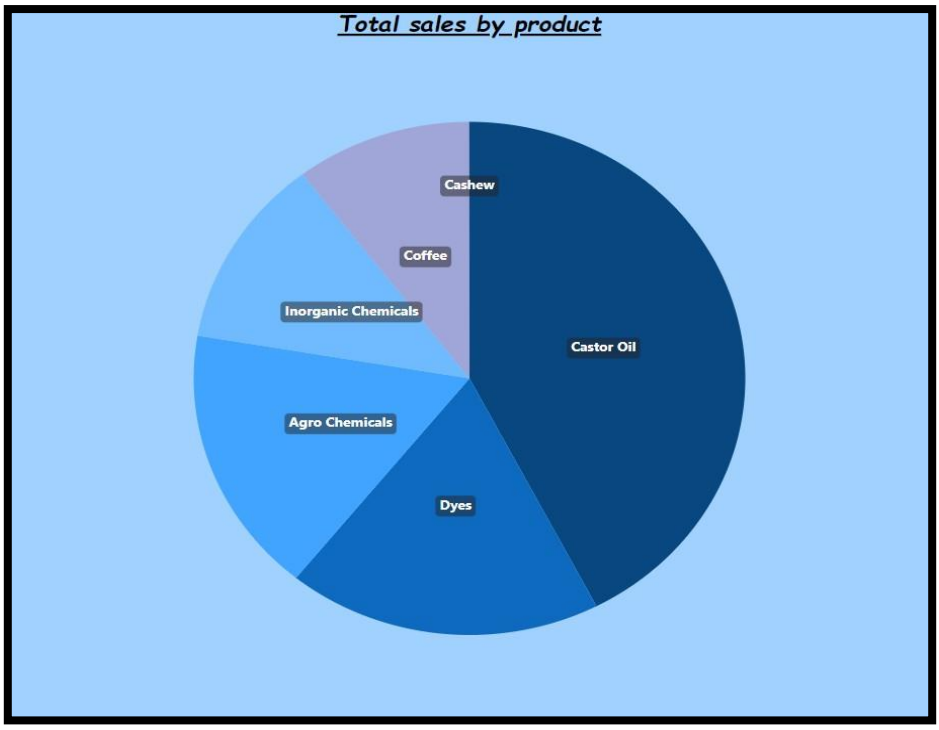


Figure: 3.5 Total sales of products

2. Products

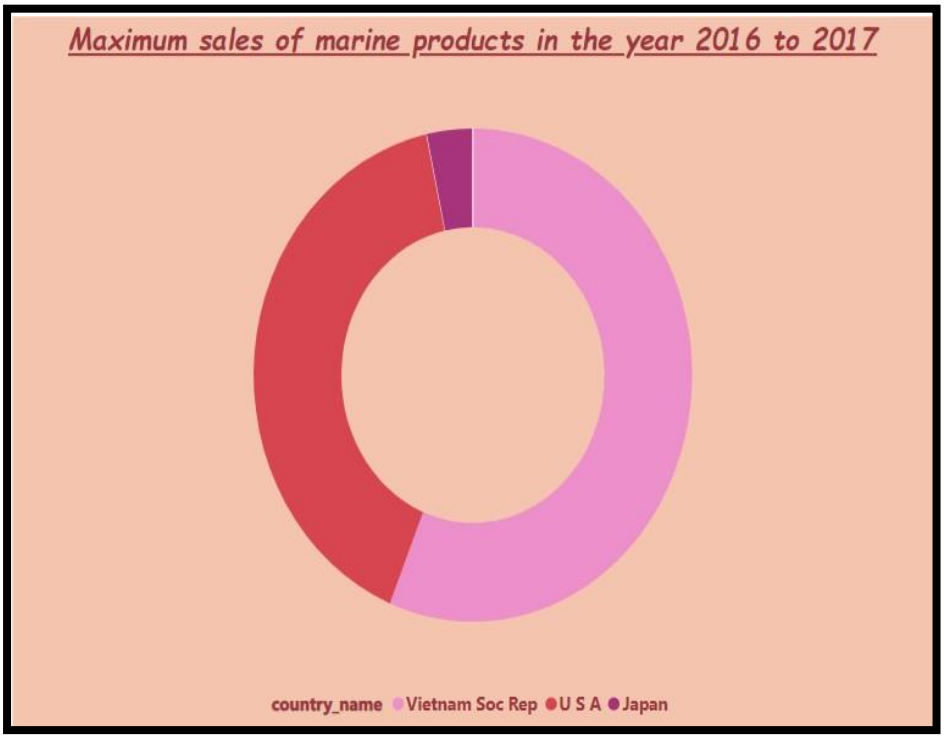


Figure: 3.6 Maximum sales of marine products

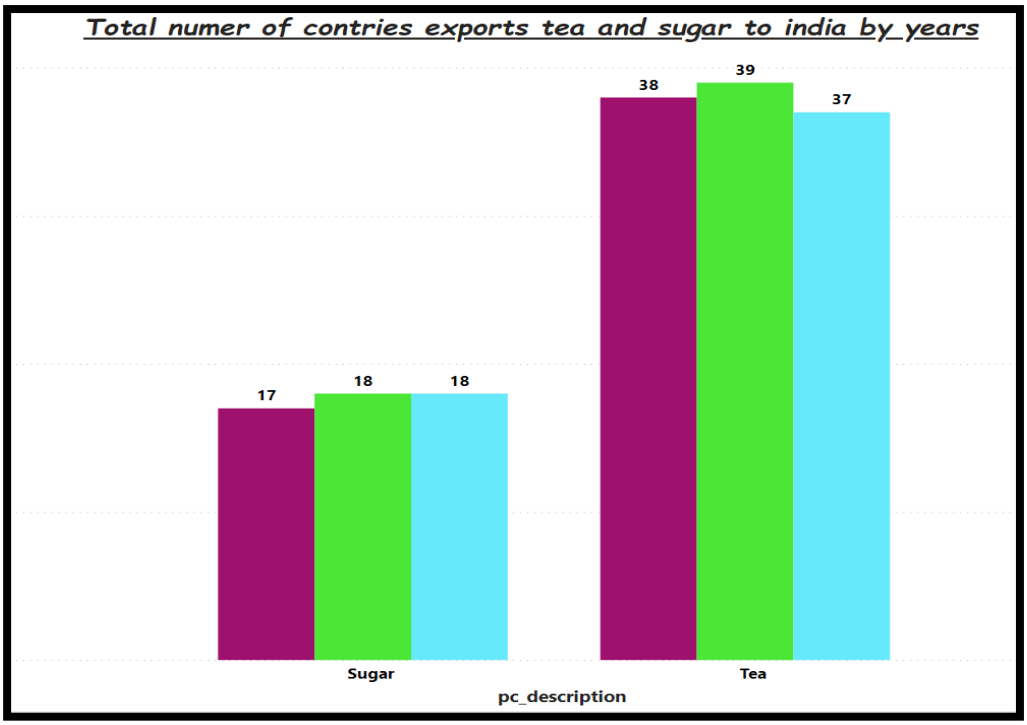


Figure: 3.7 Number of countries export tea and sugar

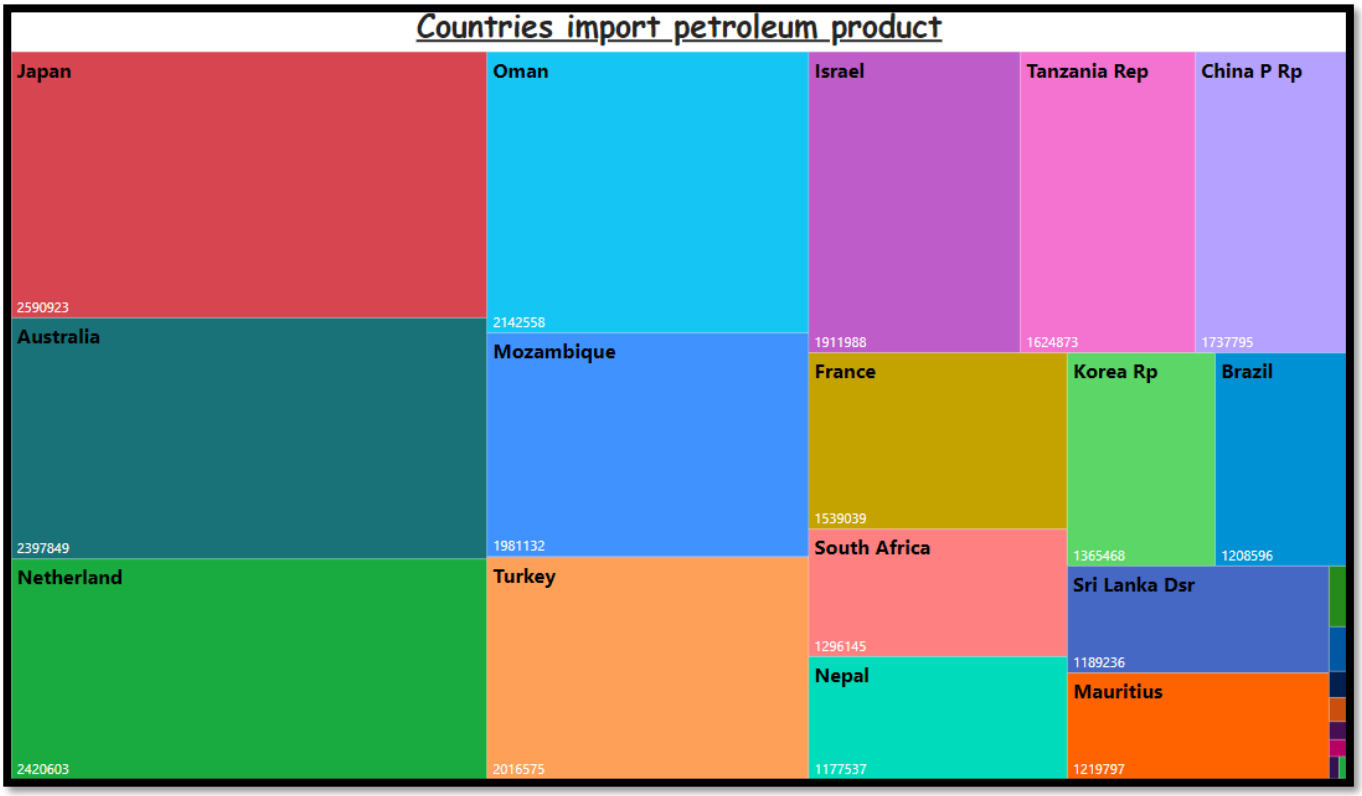


Figure: 3.8 petroleum product importing country

3. Country

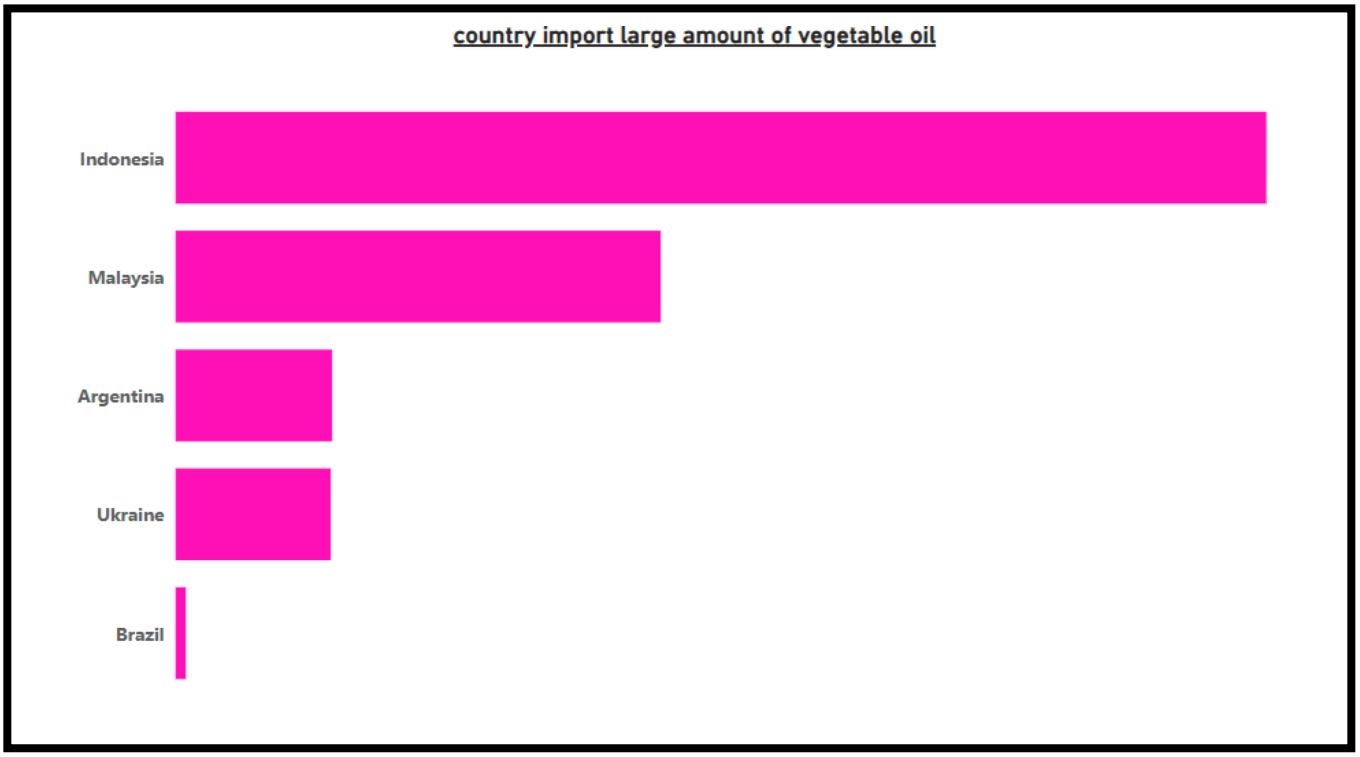


Figure: 3.9 country import vegetable oil

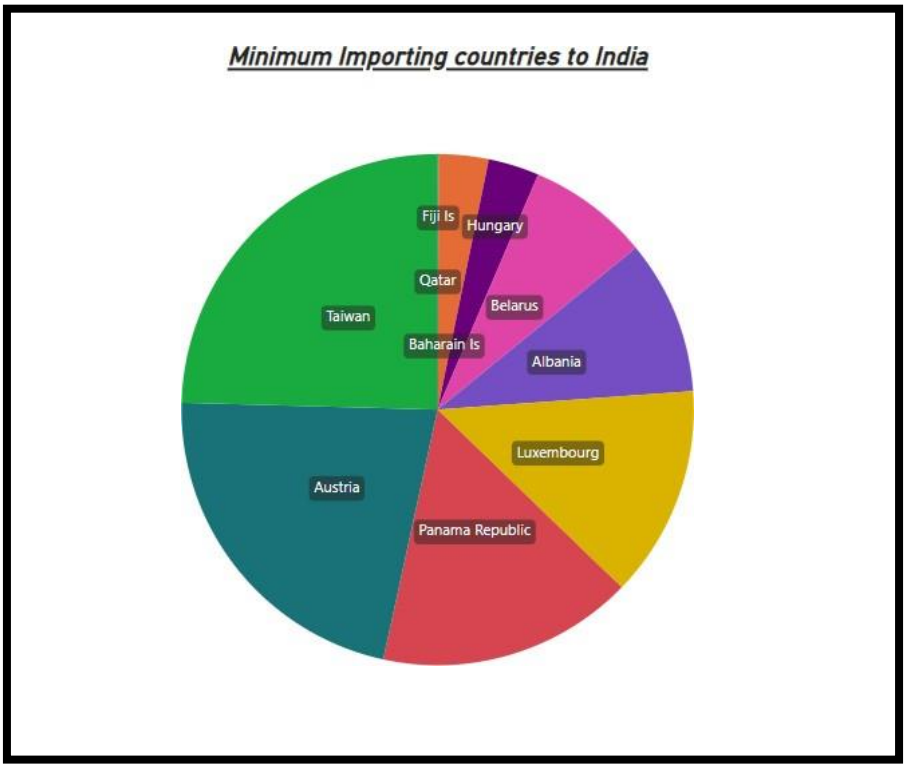


Figure: 3.10 Maximum importing countries

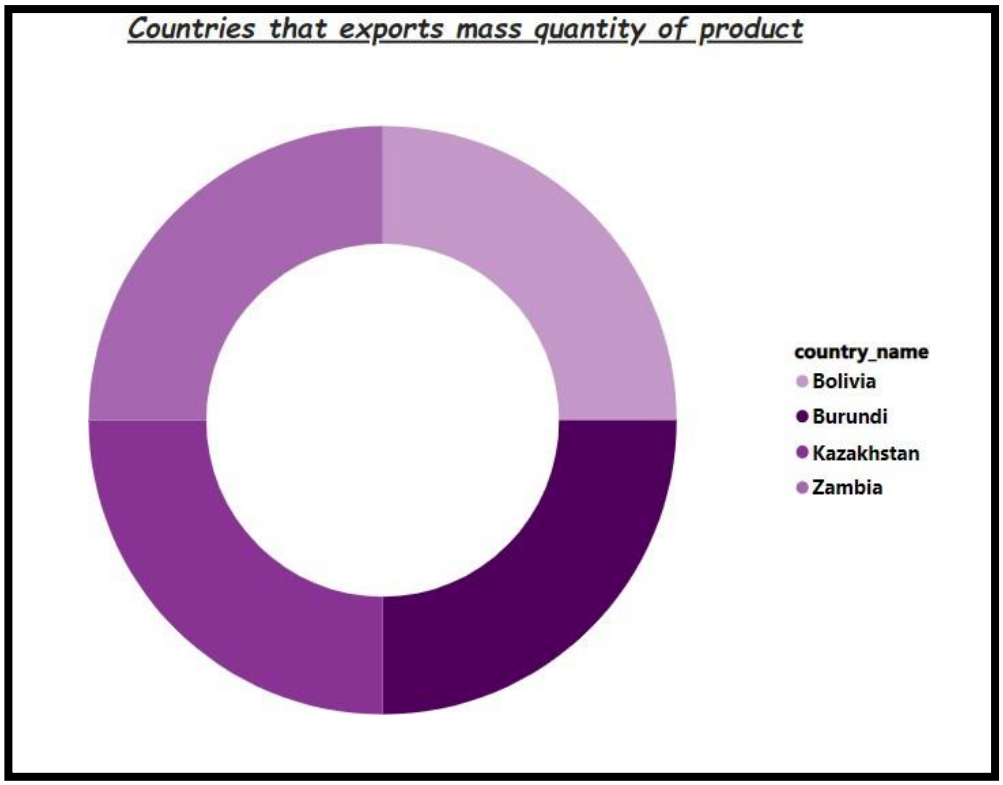


Figure: 3.11 Mass quantity of country's export

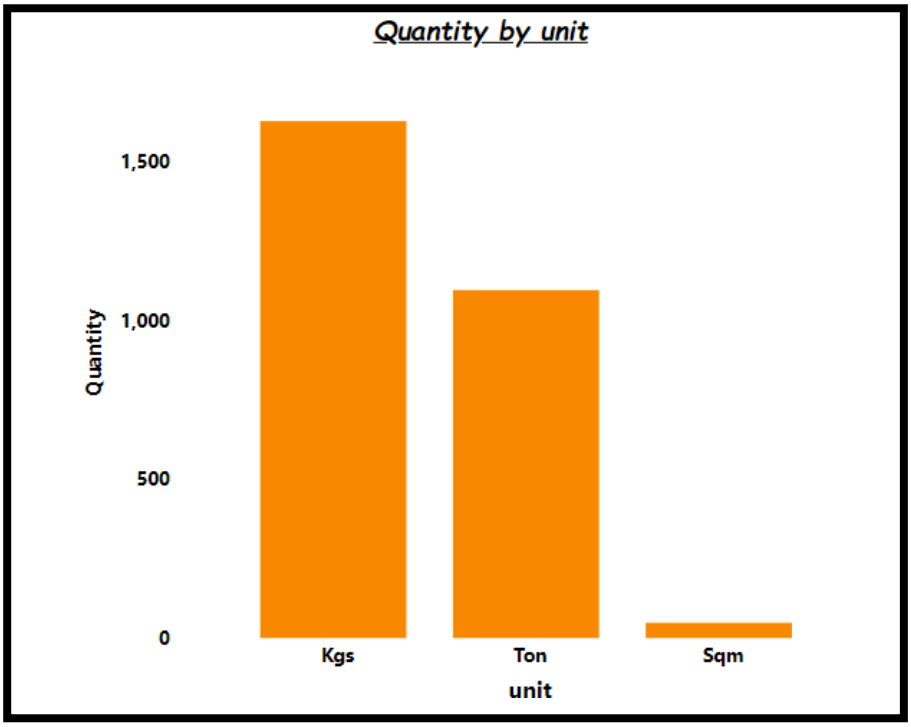


Figure: 3.12 Unit preferred by countries

Finally the slicers are created for all the six tables which give insight about the total sales and units imported by other countries form India by taking country name, quantity and pc_description as the value.

Figure: 3. 13 slicer for import of India

3.2 Publishing dashboard

The Power BI publishing dashboard is a powerful tool that enables users to easily create, share and publish interactive Power BI dashboards with their colleagues, friends and family. With the dashboard, you can quickly create visualizations and insights from your data, as well as gain access to a range of advanced features, such as automatic refresh, data alert notifications, security and privacy settings, and more. Power BI publishing dashboard also provides an easy-to-use interface for users to customize the look and feel of their dashboards, and allows for simple sharing and collaboration with others.

Figure: 3.14 Dashboard for import and export analysis

3.3 Inferences

- Among Germany, Canada, Australia, Hong Kong, Greece and Afghanistan, Germany is the most profitable country and the next country is Canada.
- The total cost of imports in the year 2014-2015 is the highest among the three years.
- The maximum sales were achieved in imports in 2016-2017 and exports in 2014-2015.
- The average export sales from India is highest in 2014-2015.
- Castor oil and dyes were the most profitable products.
- Most of the marine products were exported to Vietnam, USA and Japan from India.
- While comparing the export of tea from India is higher in 2015-2016 as 39 countries and the export of sugar from India is higher in the year 2015-2016 and 2016-2017 as 18 countries.
- Japan is the country which imports large amounts of petroleum.
- A large amount of vegetable oil is imported to Indonesia.
- The minimum products were imported to India from Qatar and Fiji.
- The countries which export in mass quantities are Bolivia, Burundi, Kazakhstan and Zambia.
- Maximum products were transported in quantities of Kilograms.

CHAPTER 4

CONCLUSION ANDFUTURE WORK

4.1 Recommendations

1. It can be used to calculate the number of countries importing or exporting that particular products to India.
2. It is used to calculate the which type of goods are imported and exported to the specific country.
3. It is used to calculate a particular product's minimum export or import.
4. It is used to calculate a particular product's maximum export or import.
5. Year wise analysis can be done with this dashboard.

REFERENCES

[1] Dutta, M. (1965). Import structure of India. The Review of Economics and Statistics, 295-300.

[2] Bhattacharyya, M. (1989). Import intensity of exports: A case study of Indian economy. Indian Economic Journal, 36(3), 94.

[3] Goldar, B. (2017). Determinants of import intensity of India's manufactured exports under the new policy regime. In Perspectives on Economic Development and Policy in India (pp. 217-232). Springer, Singapore.

[4] Kumar, M. R., & Janagam, D. (2011). Export and import pattern of medicinal plants in India. Indian Journal of Science and Technology, 4(3), 245-248.