

# **DATA SCIENCE MINOR PROJECT REPORT**

(Project Third Semester August-December 2020)

## **SALES DASHBOARD**

Submitted by

**Name:** D Balaji

**Registration No:** 11805130

**Programme and Section:** Bachelor of Technology (CSE) & KM073

**Corse Code:** INT217

**Under the Guidance of**

**Miss. Ashu (23631) , Assistant Professor**

**Discipline of CSE**

**Lovely School of Computer Science**

**Lovely Professional University, Phagwara**



## **CERTIFICATE**

This is to certify that **D Balaji** bearing Registration no. **11805130** has completed **INT217** project titled, “**Sales Dashboard**” under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**Designation of the Supervisor**

**School of Computer Science**

Lovely Professional University

Phagwara, Punjab.

Date: 5<sup>th</sup> December 2020

## **DECLARATION**

I, **D Balaji** student of **Bachelor of Technology** under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 5<sup>th</sup> December 2020

Signature : ***D Balaji***

**Registration No.** 11805130

D Balaji

## **ACKNOWLEDGEMENT**

In the accomplishment of this project successfully, many people have best owned upon their blessings and the heart pledged support this time I am utilizing to thank all the people who have been concerned with this project.

Primarily, I would thank god for being able to complete this project with success. Then I would like to thank my Instructor “Ashu” whose valuable guidance has been the once that helped me to patch this project and make it full proof success. Her suggestions and instructions have served as the major contributions towards the completion of this project.

Then I would like to Thank my parents and the friends who have helped me with their valuable suggestions and guidance has been very helpful in various phase of the completion of the project.

I would like to express my special thanks of gratitude to my teacher **Ashu, UID: 23631** as well as our University: **Lovely Professional University** who gave me the golden opportunity to do this wonderful project on the topic “**Sales Dashboard**”, which also helped me in doing a lot of Research and I came to know about so many new things I am really thankful to them.

Last but not least I would like to thank my classmates who have helped me a lot.

D Balaji.

## **Table of Contents**

<b>Introduction</b>	<b>6</b>
<b>Objectives/Scope of the Analysis</b>	<b>8</b>
<b>Source of dataset</b>	<b>9</b>
<b>ETL process</b>	<b>10</b>
<b>Analysis on dataset</b>	<b>23</b>
<b>List of Analysis with results</b>	<b>29</b>
<b>Conclusion</b>	<b>30</b>
<b>Bibliography</b>	<b>31</b>

## **INTRODUCTION**

Sales Dashboard: Sales teams operate in a fast-paced, target-oriented environment. Ask any sales person and she'll know exactly how she's performing that month – and whether she'll hit her sales target. Sales Dashboards provide an at-a-glance window into your team's performance so you can help your organization crush its revenue targets. Sales dashboards translate data into easily consumable pieces of information, providing you with instant insight about your business. A sales leader board allows the entire team to monitor their performance on a daily, weekly, or monthly basis. It's used as a motivational tool to foster friendly competition among internal sales teams. Sales dashboards are a great venue for leader boards because they provide a way to pull the data in, and a graphical interface to display the data.

Provide sales reps with a dashboard that allows them to track their individual performance. Include reports for key metrics like meetings booked, open opportunities, the number of deals in their pipeline, forecasted revenue, and any other performance indicators your team uses. In our dashboard we used various data collected from the sources and it helps us in tracking the sales done in last two to three years in all the months. We used pivot tables and pivot charts, slicers to create a dynamic sales dashboard and the column headings of the dashboard include Brand, country, Salesman, customer industry, sales channel, Sales value, region, products, Units, Imports and the date. Pivot tables are used as the most important and served as the major part in the creation of the dashboard. If you have your data in a single Excel table, you can summarize the data in the way that is required using Excel PivotTables. A PivotTable is an extremely powerful tool that you can use to slice and dice data.

Slicers are visual filters. You can filter a pivot table or chart by adding a slicer on any field. The dynamic dashboard allows the user to select a report from the range of reports within the model and decide where to position it on the page. The user can select "hide" to hide a report that they do not want to see or select "view" to preview it prior to choosing its position. It is one that updates automatically with real-time data. Whenever a change is made to a project schedule or record, any dashboard that draws on that information is updated. You get real-time information, every time. A dynamic dashboard helps the user to understand the things in an easier way and they can estimate the sales in a particular year. A pivot table is a program tool, that is a built-in feature of MS Excel,

allowing you to organize and summarize selected columns and rows of data in a spreadsheet to obtain a desired report, helping in visualization of data. A pivot chart is especially useful for user when dealing with tremendous amounts of data. A PivotTable is an interactive way to quickly summarize large amounts of data. You can use a PivotTable to analyse numerical data in detail, and answer unanticipated questions about your data.

A PivotTable is especially designed for: Querying large amounts of data in many user-friendly ways. A PivotTable is an interactive way to quickly summarize large amounts of data. You can use a PivotTable to analyse numerical data in detail, and answer unanticipated questions about your data. A PivotTable is especially designed for: Querying large amounts of data in many user-friendly ways. Subtotalling and aggregating numeric data, summarizing data by categories and subcategories, and creating custom calculations and formulas. Expanding and collapsing levels of data to focus your results, and drilling down to details from the summary data for areas of interest to you. Moving rows to columns or columns to rows (or "pivoting") to see different summaries of the source data. Filtering, sorting, grouping, and conditionally formatting the most useful and interesting subset of data enabling you to focus on just the information you want. Presenting concise, attractive, and annotated online or printed reports.

Slicers provide buttons that you can click to filter tables, or PivotTables. In addition to quick filtering, slicers also indicate the current filtering state, which makes it easy to understand what exactly is currently displayed. Slicers are visual filters. Using a slicer, you can filter your data (or pivot table, pivot chart) by clicking on the type of data you want.

Slicers are a powerful new way to filter pivot table data. It's easy to add a Slicer: Select a cell in the pivot table. Excel offers many tools and charts, timelines being one class of these. A timeline is a type of chart which visually shows a series of events in chronological order over a linear timescale. The power of a timeline is that it is graphical, which makes it easy to understand critical milestones, such as the progress of a project schedule.

## **Objectives/Scope of the Analysis:**

**“Sales Performance Analysis”** identifies the difference between desired and actual **performance**. Understanding the different enhances employee handling and morale. Better focus and decision making as a result of **sales performance analysis** boosts customer relations.

A Sales Performance Analysis is a way of assessing where your business currently stands compared to where it wants to be in future. It uses industry businesses standards, performance, and other elements for comparison. It shows the gaps between the current state of your business and how it should be performing and what are the causes for the gap in between.

### **Need for Sales Performance Analysis:**

A Sales Performance Analysis addresses the questions “Where are we?” and “Where do we need to be?” clearly. As an organization grows, it loses track of some objectives and even relationships with customers and peers. That is why businesses need gap analysis: to put them back on track.

### **Importance of Sales Performance Analysis:**

A gap exists in three categories: business goals, people’s performance, and organizational capability. Although they can be differently prioritized with different approaches, conducting an analysis on them bears the following significance:

- Provides comprehensive overview of a business goals (actual and desired)
- Analyzing all categories allows directors and executives to measure their resources in their bid to meet missions, goals, and objectives.
- Decision making
- A business needs sales performance analysis to focus its efforts and make informed decisions especially in resource allocation.
- Stakeholder satisfaction
- Measuring a gap realistically builds stakeholder confidence.
- Employee motivation
- Sales performance analysis identifies the difference between desired and actual performance. Understanding the different enhances employee handling and morale.
- Improved customer relations



## **SOURCE OF DATASET**

This dataset is obtained from the link:

<https://drive.google.com/file/d/1UDEsCqzaOQ2K9AXBUaruT48hxVucG3fM/view?usp=sharing>.

This dataset is all about the sales of the electronic products along with their brands and the customer industry and it also describes about the customer industry that the specific industry has chosen for the selling of their products and shows the values of units left and sales value.

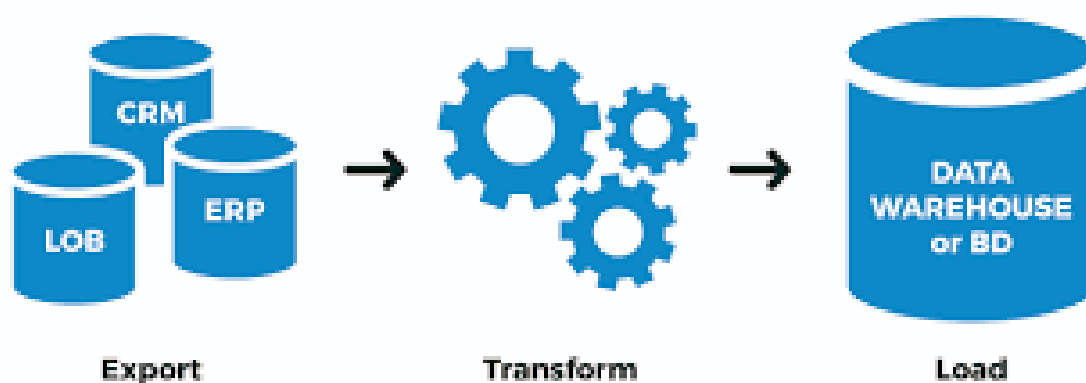
### **Google's Datasets Search Engine:**

Google has been the search engine giant, and they helped all the ML practitioners out there by doing what they are legends at, helping us find datasets. The search engine does a fabulous job at getting datasets related to the keywords from various sources, including government websites, Kaggle, and other open-source repositories.

## ETL PROCESS

In computing, extract, transform, load (ETL) is the general procedure of copying data from one or more sources into a destination system which represents the data differently from the source(s) or in a different context than the source(s). The ETL process became a popular concept in the 1970s and is often used in data warehousing. Data extraction involves extracting data from homogeneous or heterogeneous sources; data transformation processes data by data cleaning and transforming them into a proper storage format/structure for the purposes of querying and analysis; finally, data loading describes the insertion of data into the final target database such as an operational data store, a data mart, data lake or a data warehouse.

A properly designed ETL system extracts data from the source systems, enforces data quality and consistency standards, conforms data so that separate sources can be used together, and finally delivers data in a presentation-ready format so that application developers can build applications and end users can make decisions. Since the data extraction takes time, it is common to execute the three phases in pipeline. While the data is being extracted, another transformation process executes while processing the data already received and prepares it for loading while the data loading begins without waiting for the completion of the previous phases. ETL systems commonly integrate data from multiple applications (systems), typically developed and supported by different vendors or hosted on separate computer hardware. ETL is a process in Data Warehousing and it stands for **Extract**, **Transform** and **Load**.



## **Why Do We Need ETL?**

It is essential to properly format and prepare data in order to load it in the data storage system of your choice. The triple combination of ETL provides crucial functions that are many times combined into a single application or suite of tools that help in the following areas:

- Offers deep historical context for business.
- Enhances Business Intelligence solutions for decision making.
- Enables context and data aggregations so that business can generate higher revenue and/or save money.
- Enables a common data repository.
- Allows verification of data transformation, aggregation and calculations rules.
- Allows sample data comparison between source and target system.
- Helps to improve productivity as it codifies and reuses without additional technical skills.
- A basic ETL process can be categorized in the below stages:

Data Extraction

Data Cleansing

Transformation

Load

A viable approach should not only match with your organization's need and business requirements but also performing on all the above stages.

## **Traversing the Four Stages of ETL:**

1. Know and understand your data source — where you need to extract data
2. Audit your data source

3. Study your approach for optimal data extraction
4. Choose a suitable cleansing mechanism according to the extracted data
5. Once the source data has been cleansed, perform the required transformations accordingly
6. Know and understand your end destination for the data — where is it going to ultimately reside
7. Load the data

### **Data Extraction and Data Cleaning:**

#### **Data Source:**

It is very important to understand the business requirements for ETL processing. The source will be the very first stage to interact with the available data which needs to be extracted.

Organizations evaluate data through business intelligence tools which can leverage a diverse range of data types and sources.

The most common of these data types are:

1. Databases, Flat files, Web Services, Other sources such as RSS Feeds etc.

First, analyse how the source data is produced and in what format it needs to be stored. Traditional data sources for BI applications include Oracle, SQL Server, MySQL, DB2, Hana, etc.

Evaluate any transactional databases (ERP, HR, CRM, etc.) closely as they store an organization's daily transactions and can be limiting for BI for two key reasons:

1. Querying directly in the database for a large amount of data may slow down the source system and prevent the database from recording transactions in real time.
2. Data in the source system may not be optimized for reporting and analysis.

### **Usage and Latency:**

Another consideration is how the data is going to be loaded and how will it be consumed at the destination.

Let's say the data is going to be used by the BI team for reporting purposes, so you'd certainly want to know how frequently they need the data. Further, if the frequency of retrieving the data is very high but volume is low then a traditional RDBMS might suffice for storing your data as it will be cost effective. If the frequency of retrieving the data is high, and the volume is the same, then a traditional RDBMS could in fact be a bottleneck for your BI team. That type of situation could be well served by a more fit for purpose data warehouse such as Snowflake or Big Data platforms that leverage Hive, Druid, Impala, HBase, etc. in a very efficient manner.

There are many other considerations as well including current tools available in house, SQL compatibility (especially related to end user tools), management overhead, support for a wide variety of data, among other things.

### **Auditing Source Data:**

Data auditing refers to assessing the data quality and utility for a specific purpose. Data auditing also means looking at key metrics, other than quantity, to create a conclusion about the properties of the data set. In short, data audit is dependent on a registry, which is a storage space for data assets. So, ensure that your data source is analysed according to your different organization's fields and then move forward based on prioritizing the fields.

### **Analysing Data Extraction:**

The main objective of the extraction process in ETL is to retrieve all the required data from the source with ease. Therefore, care should be taken to design the extraction process to avoid adverse effects on the source system in terms of performance, response time, and locking.

### **Steps to Perform Extraction:**

- **Push Notification:** It's always nice if the source system is able to provide a notification that the records have been modified and provide the details of changes.
- **Incremental/Full Extract:** Some systems may not provide the push notification service, but may be able to provide the detail of updated records and provide an extract of such records. During further ETL processing, the system needs to identify changes and propagate it down. There are times where a system may not be able to provide the modified records detail, so in that case, full extraction is the only choice to extract the data. Make sure that full extract requires keeping a copy of the last extracted data in the same format to identify the changes. While using Full or Incremental Extract, the extracted frequency is critical to keep in mind.
- **Challenges Faced During the Extraction Process:**

One of the challenges that we typically face early on with many customers is extracting data from unstructured data sources, e.g. text, emails and web pages and in some cases custom apps are required depending on ETL tool that has been selected by your organization. This can and will increase the overhead cost of maintenance for the ETL process.

Second, the implementation of a CDC (Change Data Capture) strategy is a challenge as it has the potential for disrupting the transaction process during extraction. Many times the extraction schedule would be an incremental extract followed by daily, weekly and monthly to bring the warehouse in sync with the source. Extraction of data from the transactional database

has significant overhead as the transactional database is designed for efficient insert and updates rather than reads and executing a large query.

- Change in data formats over time, Increase in data velocity and volume, Rapid changes on data velocity and volume, Rapid changes on data source credentials, Null issues, Change requests for new columns, dimensions, derivatives and features, Writing source specific code which tends to create overhead to future maintenance of ETL flows.

### **Data Cleansing Requirements:**

Data cleaning, cleansing, and scrubbing approaches deal with detection and separation of invalid, duplicate, or inconsistent data to improve the quality and utility of data that is extracted before it is transferred to a target database or Data Warehouse. With the significant increase in data volumes and data variety across all channels and sources, the data cleansing process plays an increasingly vital role in ETL to ensure that clean, accurate data will be used in downstream decision making and data analysis.

A solid data cleansing approach should satisfy a number of requirements:

- Detection and removal of all major errors and inconsistencies in data either dealing with a single source or while integrating multiple sources.
- Correcting of mismatches and ensuring that columns are in the same order while also checking that the data is in the same format (such as date and currency).
- Enriching or improving data by merging in additional information (such as adding data to assets detail by combining data from Purchasing, Sales and Marketing databases) if required.

- Data cleaning should not be performed in isolation but together with schema-related data transformations based on comprehensive metadata.
- Mapping functions for data cleaning should be specified in a declarative way and be reusable for other data sources as well as for query processing.

### **Data Cleansing Problems:**

Data quality problems that can be addressed by data cleansing originate as single source or multi-source challenges as listed below:

#### **Potential Problems with Data from a Single Source**

- Uniqueness
- Misspelling
- Redundancy/Duplicates
- Outside domain range
- Data entry errors
- Referential integrity
- Contradictory values

#### **Potential Problems with Data from Multiple Sources:**

- Naming conflicts at the schema level — using the same name for different things or using a different name for the same things

**Data Cleansing Approach:** While there are a number of suitable approaches for data cleansing, in general, the phases below will apply:



**Data Analysis:**

In order to know the types of errors and inconsistent data that need to be addressed, the data must be analysed in detail. For data analysis, metadata can be analysed that will provide insight into the data properties and help detect data quality problems. There are two related approaches to data analysis.

**Data Profiling:**

As data gets bigger and infrastructure moves to the cloud, data profiling is increasingly important. Data profiling, data assessment, data discovery, data quality analysis is a process through which data is examined from an existing data source in order to collect statistics and information about it. In this step, a systematic up-front analysis of the content of the data sources is required.

Data profiling requires that a wide variety of factoring are understood including the scope of the data, variation of data patterns and formats in the database, identifying multiple coding, redundant values, duplicates, nulls values, missing values and other anomalies that appear in the data source, checking of relationships between primary and foreign key plus the need to discover how this relationship influences the data extraction, and analysing business rules.

**Data Mining:**

Data mining, data discovery, knowledge discovery (KDD) refers to the process of analysing data from many dimensions, perspectives and then summarizing into useful information. It also refers to the nontrivial extraction of implicit, previously unknown, and potentially useful information from data in databases.

In actual practice, data mining is a part of knowledge discovery although data mining and knowledge discovery can be considered synonyms. Through a defined approach and algorithms, investigation and analysis can occur on both current and historical data to predict future trends so that organizations' will be enabled for proactive and knowledge-driven decisions.

**Defining Transformations and Mapping Rules:**

Many transformations and cleaning steps need to be executed, depending upon the number of data sources, the degree of heterogeneity, and the errors in the data. Sometimes, a schema translation is used to map a source to a common data model for a Data Warehouse, where typically a relational representation is used. First, data cleaning steps could be used to correct single-source instance problems and prepare the data for integration. Later in the process, schema/data integration and cleaning multi-source instance problems, e.g., duplicates, data mismatch and nulls are dealt with. Declarative query and a mapping language should be used to specify schema related data transformations and a cleaning process to enable automatic generation of the transformation code.

### **Verification:**

The transformation workflow and transformation definition should be tested and evaluated for correctness and effectiveness. Improving the sample or source data or improving the definition may be necessary. Multiple repetitions of analysis, verification and design steps are needed as well because some errors only become important after applying a particular transformation.

### **Transformation:**

Execution of transformational steps is required either by running the ETL workflow for loading and by refreshing the data in a data warehouse or during the period of answering the queries on multiple sources.

### **Backflow and Cleaned Data:**

After removal of errors, the cleaned data should also be used to replace on the source side in order improve the data quality of the source database. This process will avoid the re-work of future data extraction.

### **Data Transformation:**

Once data cleansing is complete, the data needs to be moved to a target system or to an intermediate system for further processing. The transformation step in ETL will help to create a structured data

warehouse. Transformation refers to the data cleansing and aggregation that prepares it for analysis. There are two approaches for data transformation in the ETL process.

1. **Multistage Data Transformation:** In this process, extracted data is moved to an intermediate area (staging) where transformation occurs prior to loading the data into the final target area (data warehouse).
2. **In-Warehouse Data Transformation:** In this process, the flow would be ELT (Extract, Load and then Transform). The extracted data will be loaded into the data warehouse and there the transformation will occur.

Below, aspects of both basic and advanced transformations are reviewed.

#### **Basic Transformation:**

- **Format Standardization:** Standardize the data type and length according to field format to make it easy for end user to retrieve data.
- **Cleaning:** Includes mapping of values to some derived/short meaning like mapping 'Male' to 'M', null to '0', etc.
- **Deduplication:** Involves removing of duplicate values.
- **Constraints Implementation:** Establishment of key relationships across tables.
- **Advanced Transformation:**

- **Decoding of Fields:** Data coming from multiple sources many times will be described by varying field values and often times legacy source systems use fairly cryptic codes to represent business values making it necessary to remove fields having similar information and or changing obscure codes into values that make business sense to users that consume the data.
- **Merging of Information:** It's common to merge related fields together and view the merged fields as a single entity, e.g. product, product price, product type, description, etc.
- **Splitting single fields:** Splitting a large text field into a single field for easier consumption, e.g. splitting full name into first name, middle name and last name.
- **Calculated and Derived Values:** At times, an aggregation can be required on the dataset before loading it to a Data Warehouse, e.g. calculating total cost and profit margin.
- **Summarization:** Values are summarized to obtain a total figure which is subsequently calculated and stored at multiple levels as business fact in multidimensional tables.
- **Data Loading:**

In this phase, extracted and transformed data is loaded into the end target source which may be a simple delimited flat file or a Data Warehouse depending on the requirement of the organization.

There are two types of tables in Data Warehouse: Fact Tables and Dimension Tables. Once the data is loaded into fact and dimension tables, it's time to improve performance for BI data by creating aggregates.

#### **Aggregation:**

In order to design an effective aggregate, some basic requirements should be met. First, aggregates should be stored in their own fact table. Next, all dimensions that are related should be a compacted version of dimensions associated with base-level data. Finally, affiliate the base fact tables in one family and force SQL to invoke it. Aggregation helps to improve performance and speed up query time for analytics related to business decisions.

### **An Effective Loading Process:**

#### **Referential Integrity Constraints:**

Referential integrity constraints will check if a value for a foreign key column is present in the parent table from which the foreign key is derived. This constraint is applied when new rows are inserted or the foreign key column is updated.

While inserting or loading a large amount of data, this constraint can pose a performance bottleneck. Hence, it's imperative to disable the foreign key constraint on tables dealing with large amounts of data, especially fact tables. Make sure that the purpose for referential integrity is maintained by the ETL process that is being used. Other Considerations for Effective Loading:

- Indexes should be removed before loading data into the target. They may be rebuilt after loading.
- Manage partitions. The most recommended strategy is to partition tables by date interval such as a year, month, quarter, some identical status, department, etc.
- In the case of incremental loading, the database needs to synchronize with the source system. The incremental load will be a more complex task in comparison with full load/historical load.

## Challenges with Incremental Loads:

- **Ordering:** To handle large amounts of data with high availability, data pipelines often leverage a distributed systems approach which implies that data may be processed in a different order than when it was received. If data is deleted or updated, then processing in the wrong order will lead to data errors, therefore maintaining and ordering is crucial for keeping data accurate.
- **Schema Evaluation:** It is necessary to evaluate the source schema at the time of loading the data to ensure data consistency.
- **Monitoring Capability:** Data coming from a variety of sources presents complexities, and potentially failures due to an API being unavailable, network congestion or failure, API credential expiration, data incompleteness or inaccuracy — monitoring is critical as recovering from these issues can be complex.
- A final note that there are three modes of data loading: APPEND, INSERT and REPLACE, and precautions must be taken while performing data loading with different modes as that can cause data loss as well.

## Best practices ETL process

### Never try to cleanse all the data:

Every organization would like to have all the data clean, but most of them are not ready to pay to wait or not ready to wait. To clean it all would simply take too long, so it is better not to try to cleanse all the data.

### Never cleanse Anything:

Always plan to clean something because the biggest reason for building the Data Warehouse is to offer cleaner and more reliable data.

## Analysis on the dataset

Objective 1: Displaying top 5 products sold considering the sales value (line chart)

Pivot table:

	A	B	C	D
1	Displaying top 5 products sold considering the sales value (line chart)			
2				
3	Product name	Sum of Sales Value		
4	sandals	₹ 12,323		
5	laptop	₹ 11,505		
6	shoes	₹ 11,476		
7	printer	₹ 10,825		
8	scanner	₹ 10,793		
9	Grand Total	₹ 56,922		
10				
11				

Pivot chart:



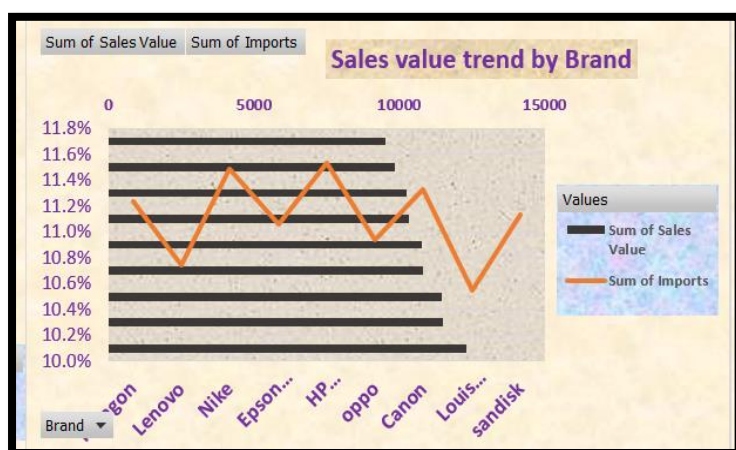
This pivot table and the pivot chart analysis the data and it displays the top five products that are sold in different regions along with the product names. The line chart here display the data precisely.

Objective 2: Displaying sales value trend by brand and showing percentage of each products import with the total imports (combo chart- Bar and line)

Pivot table :

	A	B	C	D	E	F	G
1	Displaying sales value trend by brand and showing percentage of each products im						
2							
3	Brand	Sum of Sales Value	Sum of Imports				
4	Paragon	12323	11.2%				
5	Lenovo	11505	10.7%				
6	Nike	11476	11.5%				
7	Epson Work	10825	11.1%				
8	HP Scanjet	10793	11.5%				
9	oppo	10337	10.9%				
10	Canon	10234	11.3%				
11	Louis Phillipe	9829	10.5%				
12	sandisk	9524	11.1%				
13	Grand Total	96846	100.0%				
14							
15							

Pivot chart:



Sales Trend Analysis is a concise process with results that are of critical importance for the business. We can plot sales trends from historical data, and use them to estimate future performance. Keep in mind that these can end up being highly incorrect. Therefore, it is better to forecast based on more drilled-down data, and perform additional analysis, to better estimate sales over future periods. Sales Trend Analysis gives valuable insights into how our business operates. We gather actionable information on the performance of our business, which can significantly improve our decision-making process. Employing Sales Trend Analysis leads to decisions based on the actual data of the company.

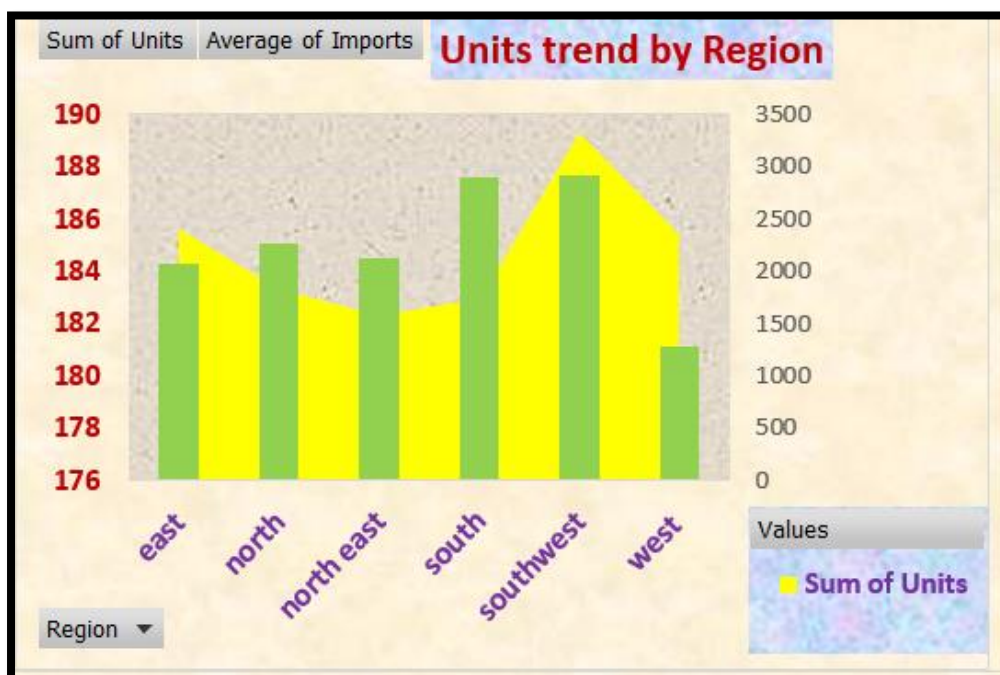


Objective 3: Displaying units value trend by region and showing the average of imports in those region (combo chart-column and area)

Pivot table:

1	<b>Displaying units value trend by region and showing th</b>		
2			
3	<b>Region</b>	<b>Sum of Units</b>	<b>Average of Imports</b>
4	east	2401	184.3103448
5	north	1819	185.0689655
6	north east	1588	184.5357143
7	south	1742	187.6206897
8	southwest	3315	187.6428571
9	west	2352	181.137931
10	<b>Grand Total</b>	<b>13217</b>	<b>185.0406977</b>
11			
12			
13			

Pivot chart:



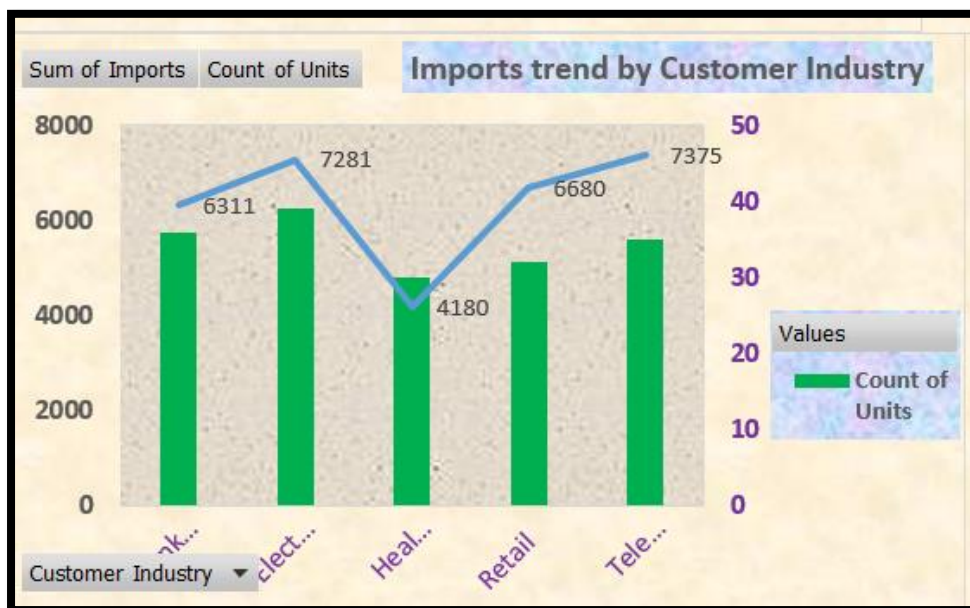
Sales Region analysis highlights the rate of sales in the market of company's sales. The below insights shows the distribution of products sold in the firm across the country.

Objective 4: Displaying imports trend by customer Industry and showing count of unit's value  
(combo chart- line and column)

Pivot table:

	A	B	C	D
1	<b>Displaying imports trend by customer Industry and showing co</b>			
2				
3	<b>Customer Industry</b>	<b>Sum of Imports</b>	<b>Count of Units</b>	
4	Banking	6311	36	
5	Electronics	7281	39	
6	Healthcare	4180	30	
7	Retail	6680	32	
8	Telecom	7375	35	
9	<b>Grand Total</b>	<b>31827</b>	<b>172</b>	
10				
11				

Pivot chart:



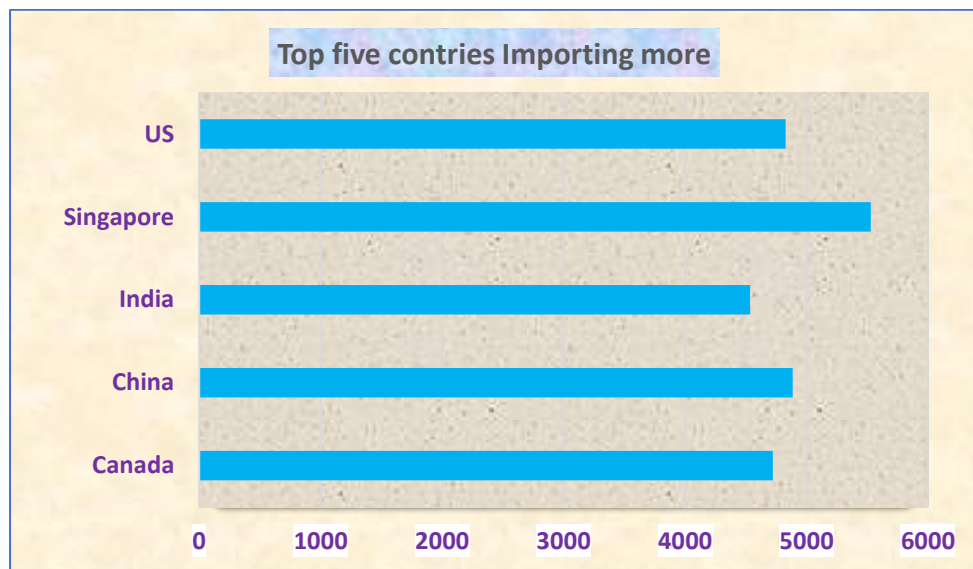
This chart displays the data about the Imports based on the customer industry. And also displays the count of units in those respective customer industry and displays the sum of imports and count of units of different customer industries.

Objective 5: Displaying top 5 countries that are importing more (Bar chart)

Pivot table:

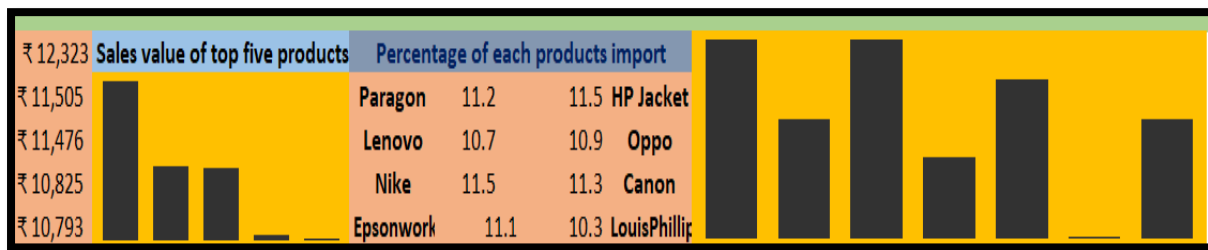
	A	B	C	D	E	F
1	Displaying top 5 countries that are importing more (Bar chart)					
2						
3	Country	Sum of Imports				
4	Canada	4722				
5	China	4887				
6	India	4535				
7	Korea	2888				
8	Malaysia	4438				
9	Singapore	5528				
10	US	4829				
11	Grand Total	31827				
12						
13						

Pivot chart:



This Pivot chart displays the bar chart of values of the imports made by the various countries and also displays the names of top five countries importing more.

## Spark lines Implementation:



A sparkline is a tiny chart in a worksheet cell that provides a visual representation of data. Use sparklines to show trends in a series of values, such as seasonal increases or decreases, economic cycles, or to highlight maximum and minimum values. Position a sparkline near its data for greatest impact.

## Slicers Implementation:



Slicers provide buttons that you can click to filter tables, or PivotTables. In addition to quick filtering, slicers also indicate the current filtering state, which makes it easy to understand what exactly is currently displayed.

## Timeline:



A timeline is a type of chart which visually shows a series of events in chronological order over a linear timescale. The power of a timeline is that it is graphical, which makes it easy to understand critical milestones, such as the progress of a project schedule.

## Dashboard



A final dashboard where all the KPIs are given the functionality to filter out objectives based on the need of user, managers, decision makers in the firm.

### Key Performance Indicators in the below Analysis:

- Sales Units by region
- Imports trend by Customer Industry
- Sales of top five products
- Top five countries who are importing more
- Sales trend by Brand

## **Conclusion:**

Dashboards might be implemented at all the levels of the enterprise, as a tool that changes the business culture. One of the most important benefit of using Dashboards is the fact that managers can analyse a single screen where are monitored the key risk indicators, or the key performance indicators and can make decisions and undertake actions to mitigate the risks and improve the performance of the enterprise. MS Excel is a useful tool for creating interactive Dashboards, that analyse millions of data in seconds. A Dashboard that monitors the evolution of sales, enables the manager to make decisions regarding: the market position, the potential customers, the performance of employees, who are the most efficient suppliers and collaborators etc.

An Excel dashboard is one-pager (mostly, but not always necessary) that helps managers and business leaders in tracking key KPIs or metrics and take a decision based on it. It contains charts/tables/views that are backed by data. A dashboard is often called a report, however, not all reports are dashboards. IP Professionals use dashboards to quickly gain insights into the most important aspects of their data. They get real time insights and competitive analyses, and use them to identify items that require urgent action, streamlining workflows and properly purposing resources.

### **Bibliography**

1. <https://medium.com/hashmapinc/etl-understanding-it-and-effectively-using-it-f827a5b3e54d>
2. <https://www.guru99.com/etl-extract-load-process.html>
3. MICROSOFT EXCEL 2016 BIBLE: THE COMPREHENSIVE TUTORIAL RESOURCE  
by JOHN WALKENBACH, WILEY.
4. FUNDAMENTALS OF BUSINESS ANALYTICS by R.N. PRASAD, SEEMA ACHARYA,  
WILEY.
5. [https://www.sas.com/en\\_in/insights/data-management/what-is-etl.html](https://www.sas.com/en_in/insights/data-management/what-is-etl.html)