iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44132

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Data Analysis and Visualization of Sales Dataset using Power BI

Ms. Sarika Singh[1], Ms. Lavina Jadhav[2]
[1]*MCA Student, MET-Institute of Computer Science, Mumbai, India*
[2]*Assistant Professor, MET-Institute of Computer Science, Mumbai, India*

Abstract: ***Power BI has completely revolutionized the worlds of business intelligence, data visualization, and analytics. Power BI is an online service that allows users to search for data, transform it, visualize it, and share the reports and dashboards they create with other users in the same or different departments/organizations, as well as the general public. Power BI is used by over 200,000 organizations in 205 countries as of February 2017. Power BI's Quick Insights feature is an innovative feature that is built on a growing set of advanced analytical algorithms. This function may be activated with a single click after uploading a data set to Power BI, and it generates a number of reports depending on the data's analysis without any need for human participation. This also aids in reducing human errors in calculations and statistical techniques, which can lead to research that isn't verifiable. Power BI is simple to use and ready for adoption as a platform for Research Data Analysis, accepting even Excel spreadsheets as input. The pandemic of the coronavirus has hastened the adoption of mobile-friendly analytics and business intelligence platforms. Companies all over the world now want their employees and customers to have access to data and analytics from any location. The mobile BI market will be worth more than $20 billion by 2024, according to Mordor Intelligence. Furthermore, mobile devices such as smartphones and tablets are expected to account for more than 72 percent of all internet traffic. As a result, implementing mobile-friendly BI platforms across your organization in 2022 makes even more sense from a business perspective. The purpose of this paper is to show how Power BI can quickly transform a piece of research data into a set of shareable analytical reports and dashboards.***
Keywords: Data Analysis, Data Visualization, Power BI, Sales Dataset

I. INTRODUCTION

We are all always thinking about the future and what is expected to happen in the coming weeks, months, and even years, and in order to do so, a look into the past is required. Businesses must be able to see their development and the factors influencing their sales [1]. In this technological era of large-scale data, organizations must reconsider current techniques to better understand clients in order to achieve a competitive advantage in the market. Data is useless if it cannot be analyzed, comprehended, and applied in context [2]. In this effort, we leveraged AtliQ sales data to create business value. A picture is worth a thousand words, and business analytics may assist in creating a picture by visualizing data to provide retailers with business insights. With these insights, firms can make meaningful modifications to their future plan to maximize profitability and success. Most raw data, particularly large scale databases, are worthless in their unprocessed state. We can extract valuable insights from this stockpile of bits by using the Power BI tools [3]. The primary goal here is to read and analyze the available datasets from AtliQ in order to generate insights and an overall view of the firm. Any organization, business, or company's success is dependent on its sales department, because it is the sole element of the organization that earns revenue and money and brings it to the table. The significance of sales is as follows:
Sales data is a broad word that encompasses a wide range of indicators, but in general, if you can measure something in relation to the sales process, it's viable sales data. With such a broad definition of sales data, it might be difficult to decide where to focus your efforts

II. WHAT IS DATA ANALYSIS?

Data analysis entails establishing a dataset, studying it, cleaning it by removing any Na values or outliers, and converting it to generate a useful result.
Power BI is a cloud-based data analysis tool that extracts, visualizes, and provides real-time insight. Power BI brings together data from a range of sources to provide you a complete picture of your company's data assets. It also takes a 'big picture' approach, using current knowledge to analyse what has happened in the past to help you make better decisions in the future.

III. PROBLEM STATEMENT

AtliQ Hardware is an Indian firm that only sells computer hardware and peripherals in India. Surge stores, Nomad stores, and other stores can be found all over India. The company's headquarters are located in Delhi.

A. Scenario

The company's sales manager is up against a number of obstacles. He's having trouble keeping track of sales in a rapidly expanding market. He is having difficulties with his company's insights. To do so, he has some of the company's regional managers working for him in North, South, and Central India. So he dials their number and inquires about the information he seeks. They inform him of the sales and increase from the previous quarter. As a result, the issue is that the conversations are verbal. As a result, regional managers sugarcoat the truth, and the company's CEO is unaware of the facts. He is unable to act despite knowing that sales are dropping because he lacks a clear picture of the sales. The regional manager sends him with excel files after he requests the records. However, he is unable to make modest decisions using this method. All the management wants is a picture of the company's weakest sector, where it needs to focus in order to boost sales and improve performance. He's looking for easy-to-understand, digestible information. As a result, he's more interested in a dashboard where he can look at real data, because data speaks the truth. He only needs a simple data visualization tool that he can use on a regular basis.

As a result, by utilizing such tools and technology, one may make data-driven decisions that aid in increasing a company's sales.

As a result, in this project, we will assist a corporation in creating its own sales dashboard using Power BI.

IV. DATA DISCOVERY

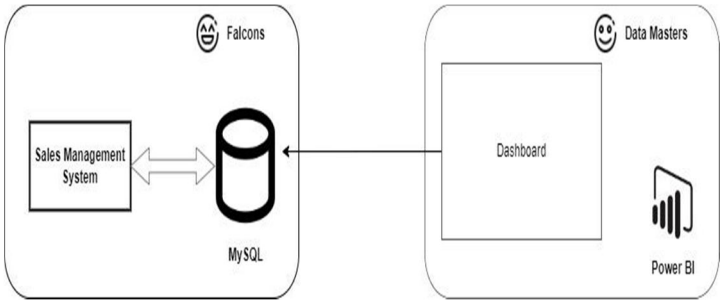
A. AIMS Grid for Project Planning

AIMS grid is a four-part project management tool.

1) Purpose (what to do exactly)
2) Stakeholders (who will be involved)
3) End Result (What are your objectives)
4) Success Criteria (cost optimization and time save)

In our example, the final result will be the creation of a dashboard, and the success criteria will be increasing sales through cost optimization while saving time for the company's manager.

B. Flowchart



C. How will the Firms Operate?

The sales management system is owned by a group of software engineers (falcons). This system's entries are kept in a MySQL database.

The Data Analyst (Data Master) team contacts the software engineers to request access to the database, which they will use to generate the PowerBI dashboard.

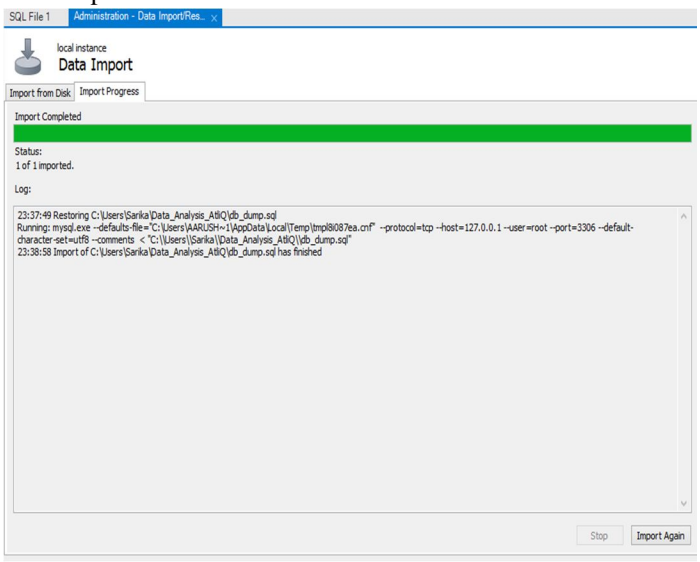
In the same way, our project will be carried out. We'll get the data from the database via the company's website, then transform and import it into PowerBI to create the dashboard.

V. DATA ANALYSIS USING SQL

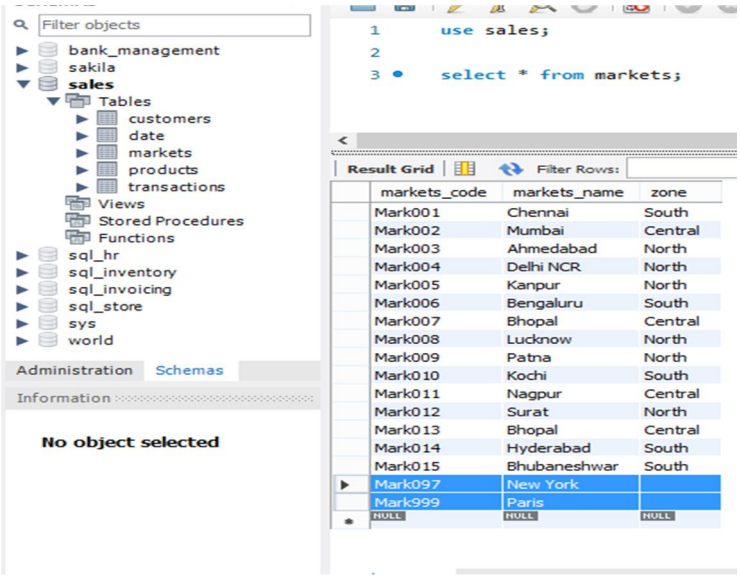
1) Step 1: Using MySQL Workbench to Import Data

The data is imported from a MySQL file that already exists. For further data analysis, this file must be loaded into MySQL Workbench.

The photo below demonstrates that the import was successful.



2) Step 2: Perform a basic data analysis by looking at various tables and reflecting erroneous values. We can observe that the table market contains some inaccurate data. The AtliQ hardware company is solely active in India, yet there are records of various non-existent cities in India.



We can observe that some table transactions have a negative amount, which is not possible.

```
1
2 •  select * from sales.transactions;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency
Prod001	Cus001	Mark001	2017-10-10	100	41241	INR
Prod001	Cus002	Mark002	2018-05-08	3	-1	INR
Prod002	Cus003	Mark003	2018-04-06	1	875	INR
Prod002	Cus003	Mark003	2018-04-11	1	583	INR
Prod002	Cus004	Mark003	2018-06-18	6	7176	INR
Prod003	Cus005	Mark004	2017-11-20	59	500	USD
Prod003	Cus005	Mark004	2017-11-22	36	250	USD
Prod003	Cus005	Mark004	2017-11-23	39	21412	INR
Prod003	Cus005	Mark004	2017-11-27	35	19213	INR
Prod003	Cus005	Mark004	2017-11-28	310	170185	INR
Prod003	Cus005	Mark004	2017-11-29	184	101194	INR
Prod003	Cus005	Mark004	2017-11-30	35	19213	INR
Prod004	Cus005	Mark004	2017-11-29	17	9426	INR
Prod004	Cus005	Mark004	2017-12-19	1	218	INR
Prod005	Cus005	Mark004	2018-08-07	5	3093	INR
Prod003	Cus006	Mark004	2017-12-04	58	30306	INR
Prod005	Cus006	Mark004	2018-06-29	38	52319	INR
Prod005	Cus006	Mark004	2018-07-02	93	126296	INR
Prod005	Cus006	Mark004	2018-07-03	79	107500	INR
Prod005	Cus006	Mark004	2018-07-04	1	273	INR

transactions 3

You can see that some of the transactions are in US dollars. As a result, filtration is also required by converting into INR.

```
2 •  select * from sales.transactions;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency
Prod001	Cus001	Mark001	2017-10-10	100	41241	INR
Prod001	Cus002	Mark002	2018-05-08	3	-1	INR
Prod002	Cus003	Mark003	2018-04-06	1	875	INR
Prod002	Cus003	Mark003	2018-04-11	1	583	INR
Prod002	Cus004	Mark003	2018-06-18	6	7176	INR
Prod003	Cus005	Mark004	2017-11-20	59	500	USD
Prod003	Cus005	Mark004	2017-11-22	36	250	USD
Prod003	Cus005	Mark004	2017-11-23	39	21412	INR
Prod003	Cus005	Mark004	2017-11-27	35	19213	INR
Prod003	Cus005	Mark004	2017-11-28	310	170185	INR
Prod003	Cus005	Mark004	2017-11-29	184	101194	INR
Prod003	Cus005	Mark004	2017-11-30	35	19213	INR
Prod004	Cus005	Mark004	2017-11-29	17	9426	INR
Prod004	Cus005	Mark004	2017-12-19	1	218	INR
Prod005	Cus005	Mark004	2018-08-07	5	3093	INR
Prod003	Cus006	Mark004	2017-12-04	58	30306	INR
Prod005	Cus006	Mark004	2018-06-29	38	52319	INR
Prod005	Cus006	Mark004	2018-07-02	93	126296	INR
Prod005	Cus006	Mark004	2018-07-03	79	107500	INR
Prod005	Cus006	Mark004	2018-07-04	1	273	INR

transactions 3

3) Step 3: Perform a primary analysis of the database by executing various SQL commands.

a) Determine the total number of records in the transaction table.

```
2 •  select count(*) from sales.transactions;
```

Result Grid | Filter Rows: | Export:

count(*)
150283

b) Determine the total number of records in the customer table.

```
2 •  select count(*) from sales.customers;
```

Result Grid | Filter Rows: | Export:

count(*)
38

c) To retrieve entries from the transaction table that include a specified market code.

```
1 •  select * from sales.transactions where market_code = 'Mark002';
2
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

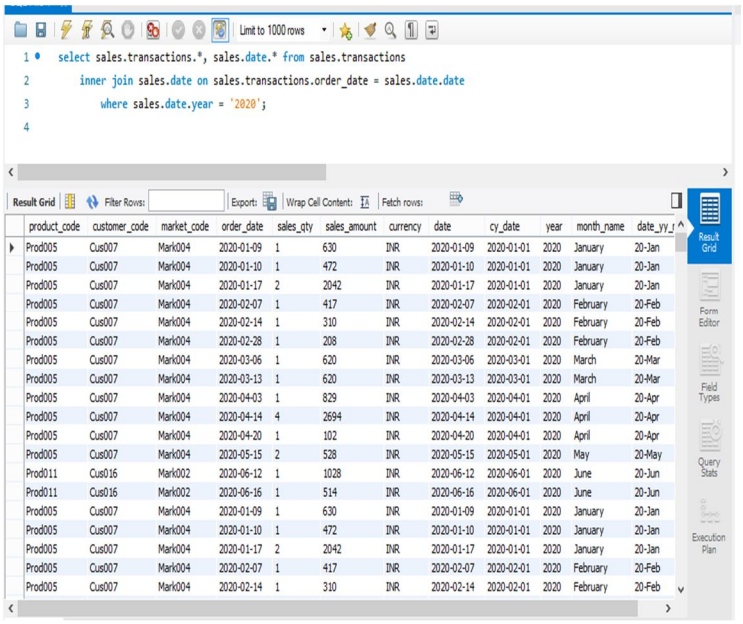
product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency
Prod001	Cus002	Mark002	2018-05-08	3	-1	INR
Prod009	Cus009	Mark002	2018-11-14	3	3032	INR
Prod011	Cus016	Mark002	2017-10-06	1	505	INR
Prod011	Cus016	Mark002	2017-10-13	1	1514	INR
Prod011	Cus016	Mark002	2017-10-20	1	2023	INR
Prod011	Cus016	Mark002	2017-10-27	25	37889	INR
Prod011	Cus016	Mark002	2018-02-08	1	593	INR
Prod011	Cus016	Mark002	2018-03-01	1	593	INR
Prod011	Cus016	Mark002	2018-10-12	1	514	INR
Prod011	Cus016	Mark002	2018-10-31	3	4106	INR
Prod011	Cus016	Mark002	2018-12-07	1	514	INR
Prod011	Cus016	Mark002	2018-12-14	1	514	INR
Prod011	Cus016	Mark002	2018-12-21	1	514	INR
Prod011	Cus018	Mark002	2018-12-28	1	0	INR
Prod011	Cus016	Mark002	2019-01-28	1	514	INR
Prod011	Cus016	Mark002	2019-02-01	1	514	INR
Prod011	Cus016	Mark002	2019-02-08	1	1028	INR
Prod011	Cus016	Mark002	2019-02-15	1	2056	INR
Prod011	Cus016	Mark002	2019-02-22	1	1542	INR
Prod011	Cus016	Mark002	2019-03-01	1	514	INR

d) To locate transactions from a specific year that are linked by a date table.

```
1 •  select sales.transactions.*, sales.date.* from sales.transactions
2        inner join sales.date on sales.transactions.order_date = sales.date.date;
3
```

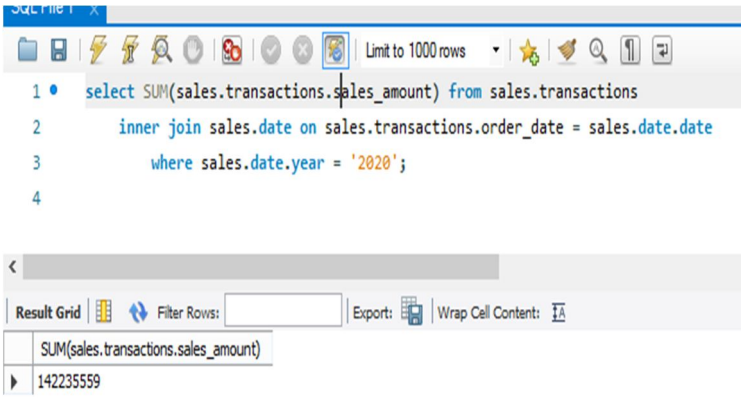
Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency	date	cy_date	year	month_name	date_yy_
Prod001	Cus001	Mark001	2017-10-10	100	41241	INR	2017-10-10	2017-10-01	2017	October	17-Oct
Prod001	Cus002	Mark002	2018-05-08	3	-1	INR	2018-05-08	2018-05-01	2018	May	18-May
Prod002	Cus003	Mark003	2018-04-06	1	875	INR	2018-04-06	2018-04-01	2018	April	18-Apr
Prod002	Cus003	Mark003	2018-04-11	1	583	INR	2018-04-11	2018-04-01	2018	April	18-Apr
Prod002	Cus004	Mark003	2018-06-18	6	7176	INR	2018-06-18	2018-06-01	2018	June	18-Jun
Prod003	Cus005	Mark004	2017-11-20	59	500	USD	2017-11-20	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-22	36	250	USD	2017-11-22	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-23	39	21412	INR	2017-11-23	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-27	35	19213	INR	2017-11-27	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-28	310	170185	INR	2017-11-28	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-29	184	101194	INR	2017-11-29	2017-11-01	2017	November	17-Nov
Prod003	Cus005	Mark004	2017-11-30	35	19213	INR	2017-11-30	2017-11-01	2017	November	17-Nov
Prod004	Cus005	Mark004	2017-11-29	17	9426	INR	2017-11-29	2017-11-01	2017	November	17-Nov
Prod004	Cus005	Mark004	2017-12-19	1	218	INR	2017-12-19	2017-12-01	2017	December	17-Dec
Prod005	Cus005	Mark004	2018-08-07	5	3093	INR	2018-08-07	2018-08-01	2018	August	18-Aug
Prod003	Cus006	Mark004	2017-12-04	58	30306	INR	2017-12-04	2017-12-01	2017	December	17-Dec
Prod005	Cus006	Mark004	2018-06-29	38	52319	INR	2018-06-29	2018-06-01	2018	June	18-Jun
Prod005	Cus006	Mark004	2018-07-02	93	126296	INR	2018-07-02	2018-07-01	2018	July	18-Jul
Prod005	Cus006	Mark004	2018-07-03	79	107500	INR	2018-07-03	2018-07-01	2018	July	18-Jul

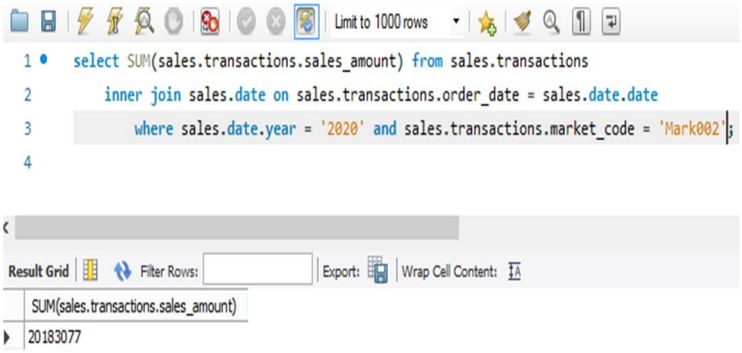


We're doing an inner join here by connecting the date and year together, which only reveals records for the year 2020.

e) To determine the overall revenue for a given year.



f) To learn more about the work you made in a certain city (eg. Mumbai).

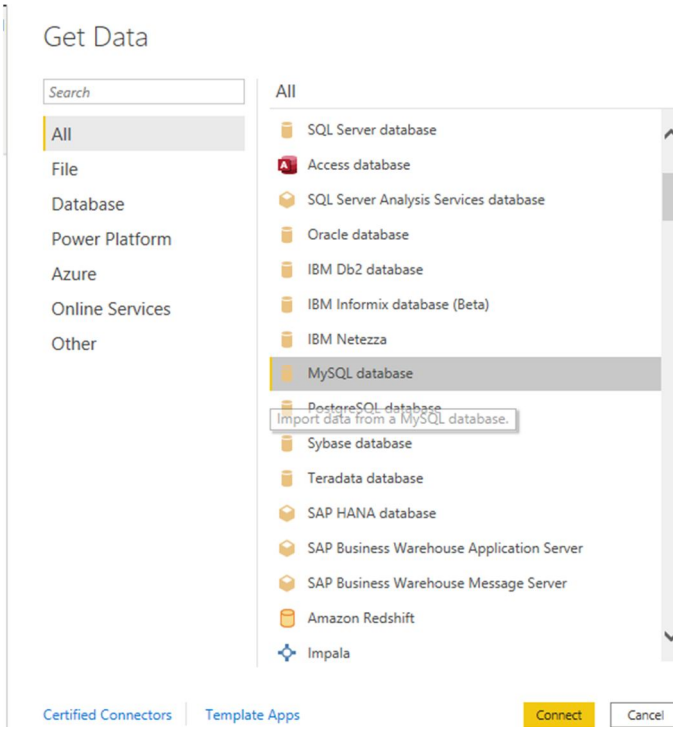


To obtain all of the records, the market code of Mumbai is used. Similarly, if we want something distinct from any other city, we use the market code for that city.

VI. DATA CLEANING AND ETL

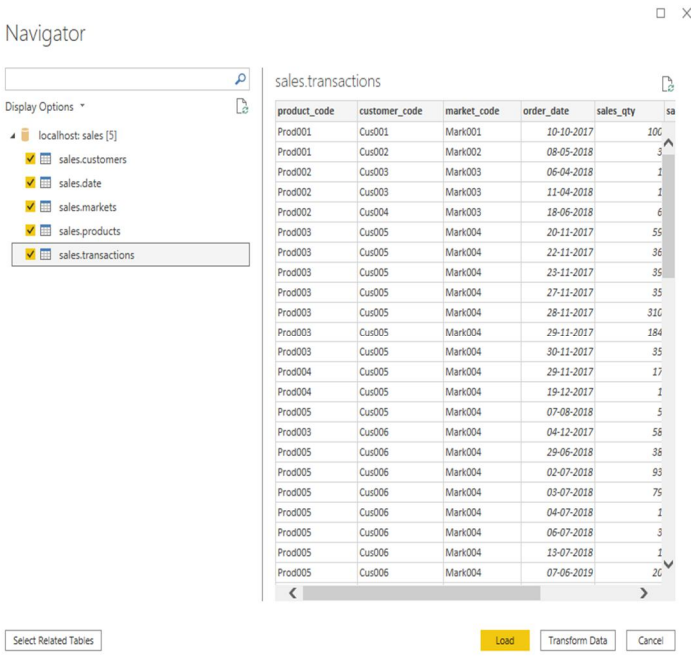
(EXTRACT TRANSFORM LOAD)

1) *Step 1:* We'll connect MySQL to the PowerBI desktop application.

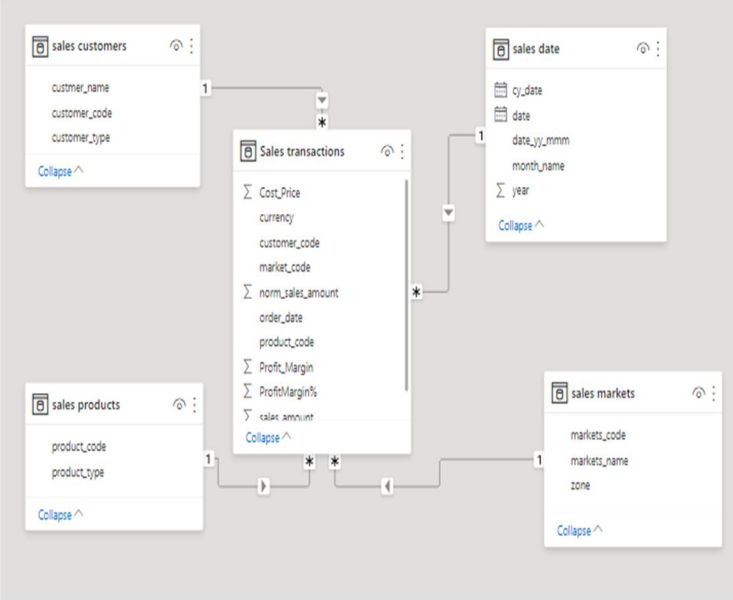


To connect to the desktop, we're going to use a MySQL database.

2) *Step 2:* Loading data into PowerBI



All of the tables we've constructed in the database will be loaded here. This load option connects to SQL and pulls all of the data into the Power BI environment.



3) Step 3: Use Power Query to transform data.

a) Filtering the Market Table

We are sent to the Power query editor when we select the transform data option. The ETL is carried out in the Power Query Editor. We can clean, wrangle, and manage data here.

We need to filter out the rows with null values that aren't appropriate.

As a result, we're filtering the data and un-checking the blank box.

After filtration, this is the final result.

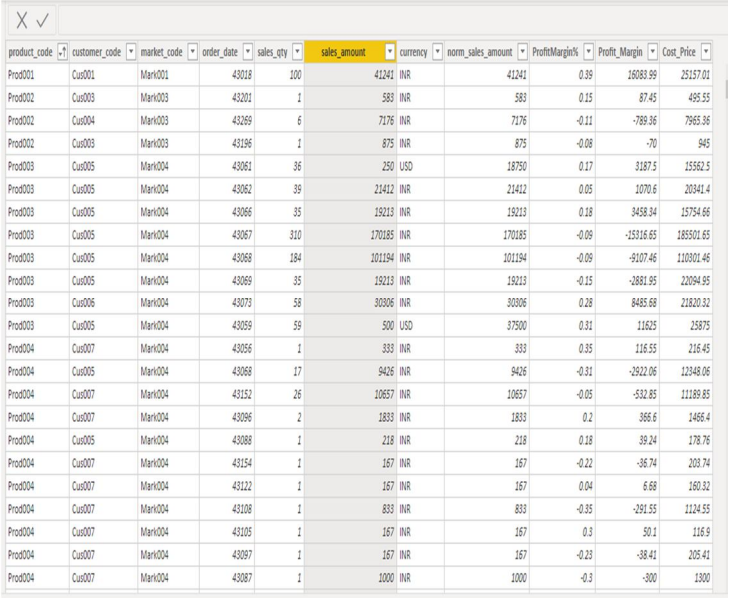
markets_code	markets_name	zone
Mark001	Chennai	South
Mark002	Mumbai	Central
Mark003	Ahmedabad	North
Mark004	Delhi NCR	North
Mark005	Kanpur	North
Mark006	Bengaluru	South
Mark007	Bhopal	Central
Mark008	Lucknow	North
Mark009	Patna	North
Mark010	Kochi	South
Mark011	Nagpur	Central
Mark012	Surat	North
Mark013	Bhopal	Central
Mark014	Hyderabad	South
Mark015	Bhubaneshwar	South

b) Filtering data in the Transactions Table

The intended output is produced when we perform the query in MySQL to filter some negative and zero values that appear in the table. Now we'll use PowerBI to accomplish the same filters.

We can remove the values we don't want from the table by deselecting them.

After filtration, the result. We need to sanitize that data because the zero values represent certain junk values that aren't possible.



product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency	norm_sales_amount	ProfitMargin%	Profit_Margin	Cost_Price
Prod001	Cus001	Mark001	43018	100	41241	INR	41241	0.39	16083.99	25157.01
Prod002	Cus003	Mark003	43201	1	583	INR	583	0.15	87.45	495.55
Prod002	Cus004	Mark003	43269	6	7176	INR	7176	-0.11	-789.36	7965.36
Prod002	Cus003	Mark003	43196	1	875	INR	875	-0.08	-70	945
Prod003	Cus005	Mark004	43061	36	250	USD	18750	0.17	3187.5	15562.5
Prod003	Cus005	Mark004	43062	39	21412	INR	21412	0.05	1070.6	20341.4
Prod003	Cus005	Mark004	43066	35	19213	INR	19213	0.18	3458.34	15754.66
Prod003	Cus005	Mark004	43067	310	170185	INR	170185	-0.09	-15316.65	185501.65
Prod003	Cus005	Mark004	43068	184	101194	INR	101194	-0.09	-9107.46	110301.46
Prod003	Cus005	Mark004	43069	35	19213	INR	19213	-0.15	-2881.95	22094.95
Prod003	Cus006	Mark004	43073	58	30306	INR	30306	0.28	8485.68	21820.32
Prod003	Cus005	Mark004	43059	59	500	USD	37500	0.31	11625	25875
Prod004	Cus007	Mark004	43056	1	333	INR	333	0.35	116.55	216.45
Prod004	Cus005	Mark004	43068	17	9426	INR	9426	-0.31	-2922.06	12348.06
Prod004	Cus007	Mark004	43152	26	10657	INR	10657	-0.05	-532.85	11189.85
Prod004	Cus007	Mark004	43096	2	1833	INR	1833	0.2	366.6	1466.4
Prod004	Cus005	Mark004	43088	1	218	INR	218	0.18	39.24	178.76
Prod004	Cus007	Mark004	43154	1	167	INR	167	-0.22	-36.74	203.74
Prod004	Cus007	Mark004	43122	1	167	INR	167	0.04	6.68	160.32
Prod004	Cus007	Mark004	43108	1	833	INR	833	-0.35	-291.55	1124.55
Prod004	Cus007	Mark004	43105	1	167	INR	167	0.3	50.1	116.9
Prod004	Cus007	Mark004	43097	1	167	INR	167	-0.23	-38.41	205.41
Prod004	Cus007	Mark004	43087	1	1000	INR	1000	-0.3	-300	1300

c) Converting the Currency in the Transaction table from USD to INR

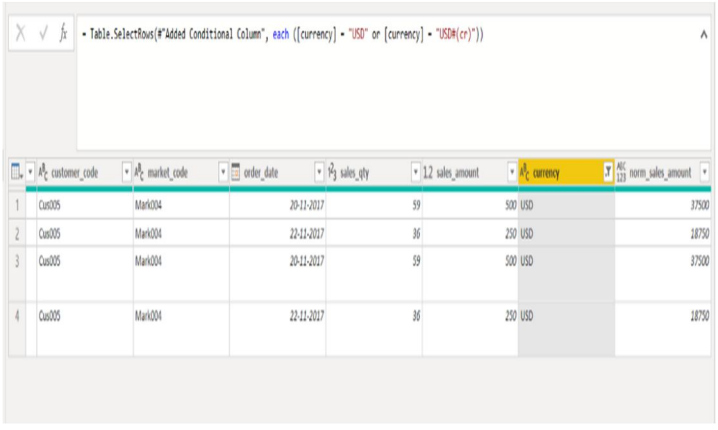
Because the company solely works in India, no USD figures can be calculated. As a result, we'll need to use some algorithms to convert those USD values to INR. Here, we'll create a new column and populate it with the translated currency value.

Obtaining the total values using USD as the currency.

We converted the USD currency into INR by applying the correct conversion procedure.

When we checked the filter, we came across something.. The USD is showing up twice for some reason. As a result, we'll now filter those values as well.

As a result, all USD values have been converted to INR.



VII. CREATING A REPORT OR A DASHBOARD

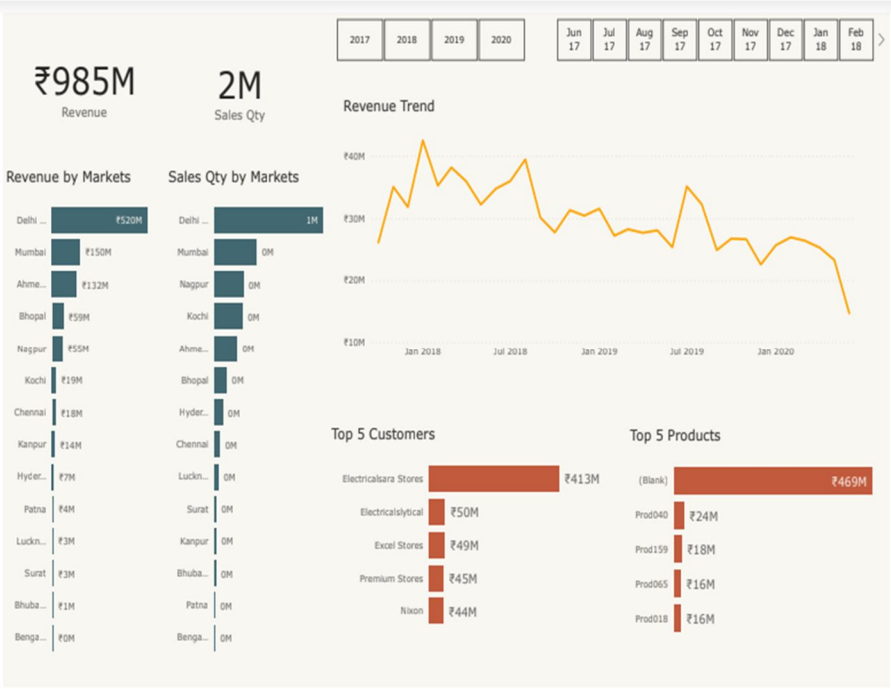
Dashboards and reports are built based on the requirements. After that, the dashboard is designed based on what the organization genuinely wants to search for and what is more vital to the company. There are many different ways to make a dashboard. In general, the dashboard should be simple to comprehend and use.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

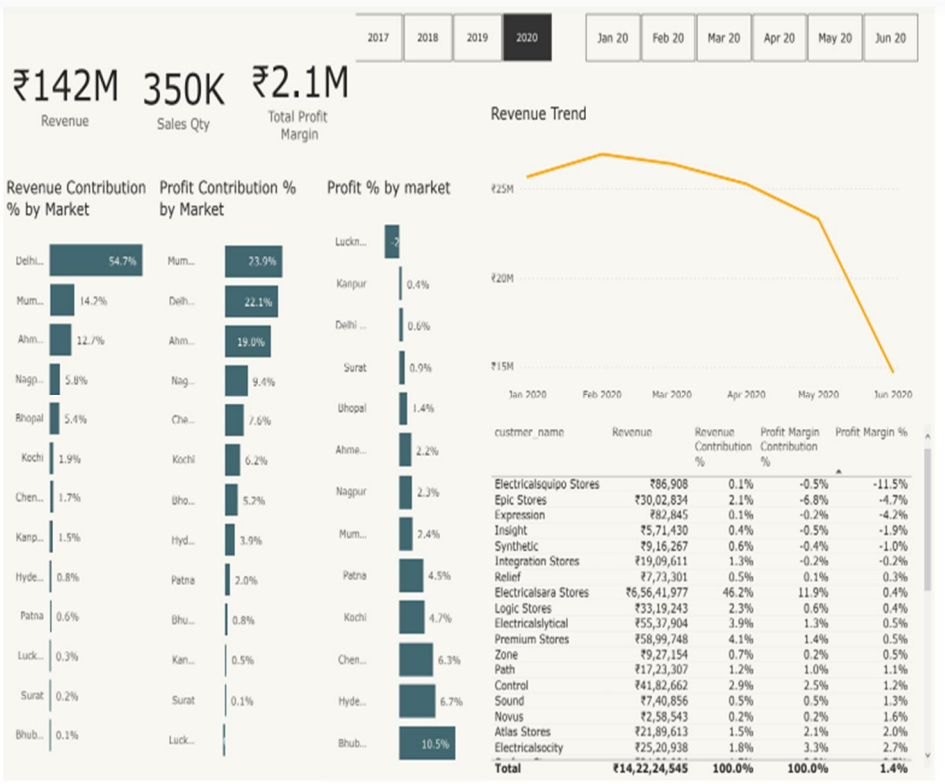
This is the dashboard once it has been entirely formatted.

A. Key Insights

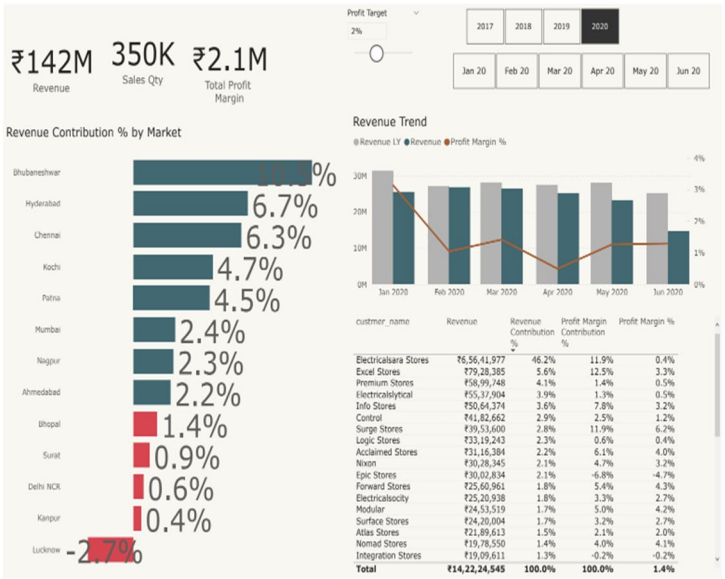
Every year's overall performance and sales profit and loss are included in the key insights dashboard.



B. Profit Analysis



C. Performance Insights



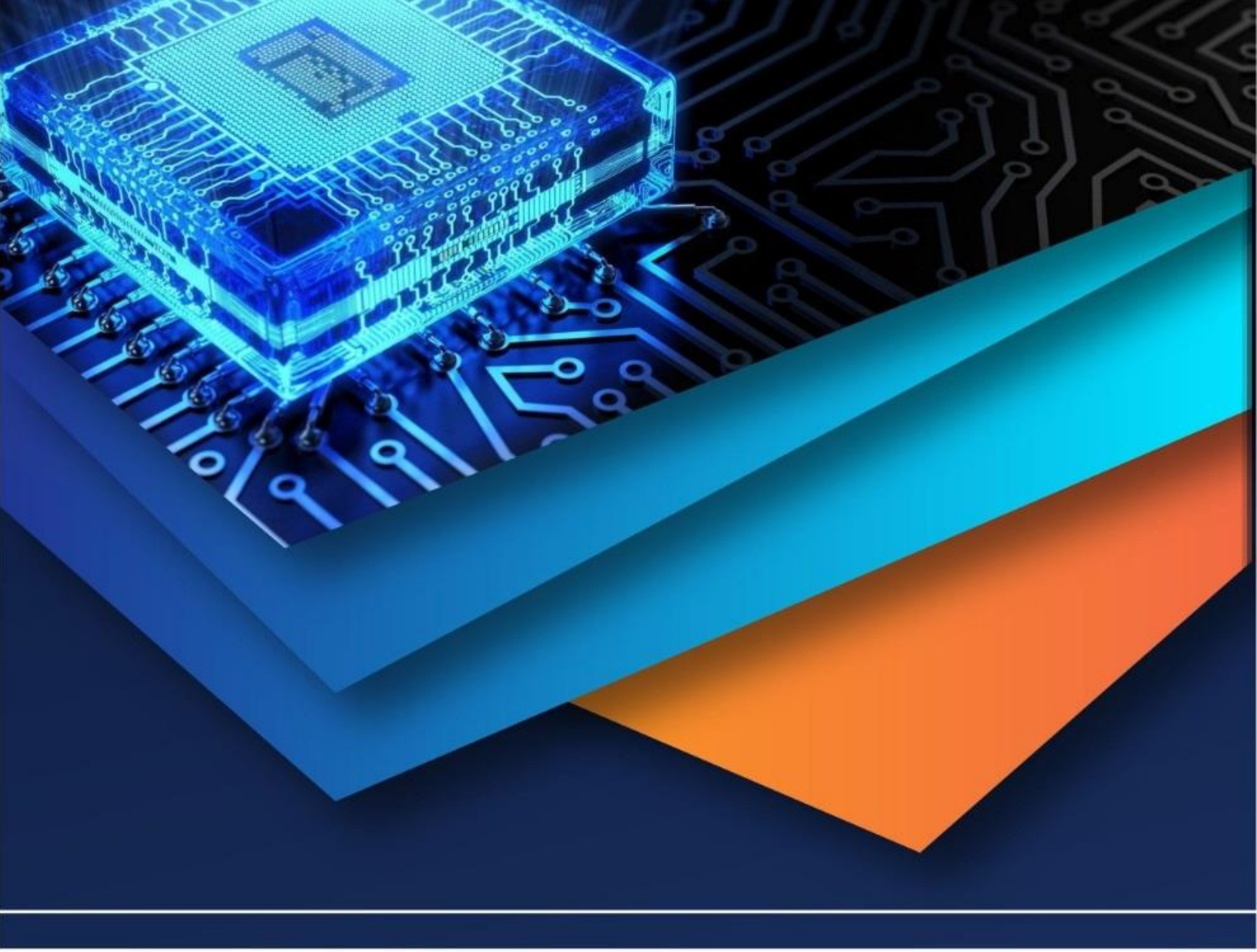
custmer_name	Revenue	Revenue Contribution %	Profit Margin Contribution %	Profit Margin %
Electricalsara Stores	₹6,56,41,977	46.2%	11.9%	0.4%
Excel Stores	₹79,28,385	5.6%	12.5%	3.3%
Premium Stores	₹58,99,748	4.1%	1.4%	0.5%
Electricalslytical	₹55,37,904	3.9%	1.3%	0.5%
Info Stores	₹50,64,374	3.6%	7.8%	3.2%
Control	₹41,82,662	2.9%	2.5%	1.2%
Surge Stores	₹39,53,600	2.8%	11.9%	6.2%
Logic Stores	₹33,19,243	2.3%	0.6%	0.4%
Acclaimed Stores	₹31,16,384	2.2%	6.1%	4.0%
Nixon	₹30,28,345	2.1%	4.7%	3.2%
Epic Stores	₹30,02,834	2.1%	-6.8%	-4.7%
Forward Stores	₹25,60,961	1.8%	5.4%	4.3%
Electricalsocity	₹25,20,938	1.8%	3.3%	2.7%
Modular	₹24,53,519	1.7%	5.0%	4.2%
Surface Stores	₹24,20,004	1.7%	3.2%	2.7%
Atlas Stores	₹21,89,613	1.5%	2.1%	2.0%
Nomad Stores	₹19,78,550	1.4%	4.0%	4.1%
Integration Stores	₹19,09,611	1.3%	-0.2%	-0.2%
Total	**₹14,22,24,545**	**100.0%**	**100.0%**	**1.4%**

VIII. CONCLUSION

Overall, the objectives of this research were met. Through visualization, data analysis aids pupils in comprehending concepts. Various technologies are available to undertake corporate data analysis, however the Power BI visualization technique is the most popular technique for learning the fundamentals of data analysis. With the help of the visualization technique, data interpretation and data representation may be done quickly and easily. This strategy is highly useful for stronger conceptual design.

REFERENCES

[1] Manpreet Singh, Bhawick Ghutla, Reuben Lilo Jnr, Aesaan F S Mohammedand Mahmood A Rashid, "Walmart's Sales Data Analysis - A Big Data Analytics Perspective" [Online]. Available: https://www.researchgate.net/publication/328246040_Walmart's_Sales_Data_Analysis_-_A_Big_Data_Analytics_Perspective
[2] Sonia Singla, "Getting Started with Data Analysis using Power BI" [Online]. Available:
[3] https://www.analyticsvidhya.com/blog/2021/11/getting-started-with-data-analysis-using-power-bi/Effective Businesss Intelligence Solutions, "BUSINESS INTELLIGENCE TRENDS IN 2022" [Online]. Available: https://ebisgroup.com/business-intelligence-trends-in-2022/
[4] Maggie Sparkman, David Iseminger, Michele Hart, Kent Sharkey, "Sales and Marketing sample for Power BI: Take a tour" [Online]. Available: https://docs.microsoft.com/en-us/power-bi/create-reports/sample-sales-and-marketing
[5] Orina Mark , "Data Analytics and Visualization: A Comprehensive Guide" [Online]. Available: https://hevodata.com/learn/data-analytics-and-visualization/
[6] https://www.tutorialgateway.org/quick-insights-in-power-bi/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY