Product Manufacturing Industry

DATA ANALYSIS PROJECT

INSIGHTS REPORT

YEAR-2023

CREATED BY:

I3 Infosoft Private Limited

+91 888-290-9944
Info@I3infosoft.com
www.i3infosoft.com
A-77, Ground Floor, Sector 2, Near Noida Sector 15 Metro Station, Noida, Uttar Pradesh, India 201301

i3 infosoft
Software Development & Quick Solution

Table of contents

ABSTRACT

This project focuses on the analysis of a retail sales dataset to gain insights into various aspects of the sales process. The dataset consists of information related to sales orders, including order types, expected values, sales owners, zones, dates, influencers involved, main competitors, and more. The analysis is conducted using Python programming language and popular data analysis libraries such as Pandas, NumPy, and Matplotlib.

The project begins with importing the dataset using the Pandas library and performing initial data exploration to understand its structure and content. The main objective is to analyse the "Retail" order type, and hence, steps are taken to clean and pre-process the relevant data.

Data cleaning and pre-processing involve standardizing the order type by replacing multiple variations with a single category, handling null values, converting the "Expected Value" column from strings to numerical values, and categorizing price brackets into more general categories.

The sales analysis focuses on different zones, where the total sales value and actual sales value are calculated. The difference between total sales and actual sales provides insights into missed revenue opportunities. Additionally, the percentage of actual sales compared to total sales helps evaluate sales performance and conversion rates across zones.

Time analysis is performed to understand the duration at different stages of the sales process. The time taken for the first interaction, site visit, proposal submission, and order closure is calculated to identify potential bottlenecks and areas for improvement.

Throughout the project, various data visualization techniques are applied using Matplotlib to present the findings effectively. Visualizations such as bar charts, line plots, and pie charts are utilized to represent sales distribution, percentage of actual sales, and time durations.

The results of this analysis can provide valuable insights for retail businesses, including identifying high-performing zones, understanding customer preferences, optimizing sales processes, and improving overall sales performance. By leveraging the power of data analysis, this project contributes to informed decision-making and strategic planning in the retail industry.

OBJECTIVE

The objective of this project is to analyse a retail sales dataset and gain valuable insights into the sales process. The project aims to achieve the following specific objectives:

1. **Categorize and clean the dataset:** Standardize the order types by consolidating various variations into a single category. Handle missing values and ensure data integrity for further analysis.
2. **Pre-process the data:** Convert the "Expected Value" column from strings to numerical values for accurate calculations. Categorize the price brackets into general budget categories for better analysis.
3. **Analyse sales performance by zone:** Determine the total sales value and actual sales value for each zone to understand the revenue generated. Identify zones with high sales potential and areas where revenue opportunities were missed.
4. **Evaluate sales conversion rates:** Calculate the percentage of actual sales compared to total sales to assess the effectiveness of the sales process. Identify zones or sales owners with higher conversion rates and areas for improvement.
5. **Conduct time analysis:** Determine the average duration for different stages of the sales process, including the first interaction, site visit, and showroom visit. Identify potential bottlenecks and areas where the sales process can be streamlined.
6. **Visualize insights:** Utilize data visualization techniques to present the findings effectively. Create informative charts and plots to visualize sales distribution, actual sales percentages, and time durations.
7. **Provide actionable recommendations:** Based on the analysis, offer actionable recommendations to optimize sales strategies, improve conversion rates, and enhance overall sales performance. Identify areas for process improvement and customer engagement.

By accomplishing these objectives, this project aims to provide a comprehensive understanding of the retail sales process, identify key performance indicators, and offer insights to drive informed decision-making and strategic planning in the retail industry.

1. INDUSTRY REVIEW

Background

The retail industry plays a significant role in the global economy, encompassing a wide range of businesses that sell products directly to consumers. It is a highly competitive and dynamic sector, driven by changing consumer preferences, technological advancements, and evolving market trends. In recent years, the retail industry has witnessed a transformational shift due to the rise of e-commerce, mobile shopping, and Omni channel retailing.

E-commerce has emerged as a disruptive force, enabling customers to shop conveniently from the comfort of their homes and offering a vast array of products at competitive prices. Online retail giants like Amazon have revolutionized the industry by leveraging advanced logistics, personalized recommendations, and efficient supply chain management. This has led to a shift in consumer behaviour, with more people embracing online shopping and expecting seamless experiences across multiple channels.

Furthermore, the COVID-19 pandemic has had a profound impact on the retail industry. Lockdowns and social distancing measures forced physical retail stores to close temporarily, leading to a surge in online shopping. This accelerated the adoption of e-commerce and highlighted the importance of digital transformation in the retail sector. Retailers had to rapidly adapt to new customer demands, enhance their online presence, and optimize their supply chains to ensure timely delivery.

The retail industry encompasses various sectors, including fashion, electronics, home goods, groceries, and more. Each sector has its unique challenges and opportunities. For instance, the fashion industry is characterized by fast-changing trends and the need for quick inventory turnover, while the grocery sector requires efficient logistics and a focus on freshness and quality.

In this competitive landscape, retailers are constantly striving to differentiate themselves, attract customers, and maximize sales. This requires a deep understanding of consumer preferences, effective marketing strategies, streamlined operations, and a seamless customer experience. Retailers are leveraging data analytics, artificial intelligence, and machine learning to gain insights into consumer behaviour, personalize marketing efforts, optimize pricing, and enhance supply chain efficiency.

Overall, the retail industry is undergoing a profound transformation driven by technological advancements, changing consumer expectations, and market dynamics. To stay competitive and thrive in this evolving landscape, retailers must embrace innovation, leverage data-driven insights, and continuously adapt their strategies to meet the ever-changing demands of customers.

Current Solution to the Problem

In response to the challenges and opportunities presented by the evolving retail industry, businesses have implemented various strategies and solutions to stay competitive and meet customer demands. The following are some of the current solutions being utilized:

1. **E-commerce Platforms:** Retailers have embraced e-commerce platforms as a means to establish an online presence and reach a broader customer base. These platforms provide a user-friendly interface, secure payment gateways, and robust inventory management systems. Popular e-commerce platforms include Shopify, Magneto, and Woo Commerce, which offer customizable solutions to cater to different business needs.
2. **Omni channel Retailing:** To provide a seamless and integrated shopping experience, retailers are adopting Omni channel strategies. This approach enables customers to interact with a brand through multiple channels, such as physical stores, websites, mobile apps, and social media. Retailers leverage technologies like customer relationship management (CRM) systems and inventory management software to synchronize data and deliver consistent experiences across channels.
3. **Data Analytics and Personalization:** Retailers are harnessing the power of data analytics to gain insights into customer behaviour, preferences, and buying patterns. Advanced analytics tools help identify trends, segment customers, and personalize marketing efforts. By tailoring product recommendations, promotions, and offers based on individual preferences, retailers can enhance customer satisfaction, increase conversion rates, and drive sales.
4. **Supply Chain Optimization:** Optimizing the supply chain is crucial for retailers to ensure efficient inventory management, minimize stock outs, and deliver products on time. Businesses are leveraging technologies like enterprise resource planning (ERP) systems, inventory management software, and demand forecasting tools to streamline their supply chain operations. This includes optimizing inventory levels, improving

demand planning, and collaborating with suppliers to enhance overall supply chain efficiency.

5. **Customer Experience Enhancement:** Retailers recognize the importance of providing exceptional customer experiences to build loyalty and differentiate themselves from competitors. They are investing in technologies such as Chabot's, virtual assistants, and self-service kiosks to enhance customer interactions and provide prompt support. Additionally, retailers are incorporating augmented reality (AR) and virtual reality (VR) technologies to offer immersive shopping experiences and allow customers to visualize products before making a purchase.
6. **Sustainability Initiatives:** With growing environmental concerns, retailers are increasingly adopting sustainability initiatives. This includes implementing eco-friendly packaging, promoting responsible sourcing, and offering sustainable product options. Retailers are also leveraging transparency and traceability technologies to provide customers with information about the origins and sustainability of the products they purchase.

These current solutions address the challenges posed by the retail industry, enabling businesses to adapt to changing market dynamics, meet customer expectations, and thrive in a competitive landscape. However, as technology continues to advance and consumer preferences evolve, retailers must remain agile and continuously innovate to stay ahead in the ever-changing retail landscape.

Proposed Solution to the problem

The proposed solution aims to analyze a retail sales dataset to gain insights into various aspects of the sales process. The code provided utilizes the Python programming language and popular data analysis libraries such as Pandas and NumPy to perform the analysis.

1. **Data Cleaning and Pre-processing:**

The initial step involves importing the dataset using Pandas and exploring its structure and content. The code cleans the "WM Order Type" column by standardizing different variations of the "Retail" order type into a single category. Null values in the "Expected Value" column are handled by filling them with corresponding values from the "dup" column. The "Expected Value" column is converted from strings to numerical values, and price brackets are categorized into more general categories.

2. **Sales Analysis:**

The code analyses sales data based on different zones. It calculates the total sales value and actual sales value for each zone. The difference between total sales and actual sales represents missed revenue opportunities, and the percentage of actual sales compared to total sales provides insights into sales performance and conversion rates across zones.

3. **Time Analysis:**

Time analysis is performed to understand the duration at different stages of the sales process. The code calculates the time taken for the first interaction, the site visit after the first interaction, and the showroom visit after the site visit. This information helps identify potential bottlenecks and areas for improvement in the sales process.

4. **Data Visualization:**

The code utilizes data visualization techniques using Matplotlib to present the findings effectively. Various visualizations such as bar charts, line plots, and pie charts are used to represent sales distribution, the percentage of actual sales, and time durations.

The proposed solution provides valuable insights for retail businesses, including identifying high-performing zones, understanding customer preferences, optimizing sales processes, and improving overall sales performance. By leveraging the power of data analysis, this solution contributes to informed decision-making and strategic planning in the retail industry.

5. **Dataset and Domain:**

Dataset:

The dataset used in the project is named "DATA.xlsx". It is an Excel file containing sales data for a retail business. The dataset includes information about various aspects of sales transactions, including order types, sales values, time durations, customer expectations, competitor analysis, and more. The dataset is loaded into a Pandas Data Frame for analysis and manipulation. It has 63 columns and 93621 rows. Columns 1,2,3,4,5,6,7 have 0 non-null value. so we can drop the column directly as it has no significant influence on the data.

Domain:

The project is focused on the retail industry, specifically analysing sales data to gain insights into the performance and efficiency of the sales process. The dataset encompasses information related to retail sales, including order types, sales values, customer expectations, time durations, competitor analysis, and other relevant factors. By analysing this data, the project aims to provide valuable insights and recommendations for improving sales performance, optimizing processes, and identifying growth opportunities in the retail industry.

2. DATA DICTIONARY

The following is a summary of the columns present in the dataset:

- **Zone:** The zone associated with the opportunity (categorical)
- **Column1:** Not available (null values) (numerical)
- **Region:** The region associated with the opportunity (categorical)
- **Opportunity ID:** Unique identifier for the sales opportunity (numerical)
- **Date of Enquiry:** Date when the inquiry was made (date)
- **Opp Created On:** Date when the opportunity was created (date)
- **Column2:** Not available (null values) (numerical)
- **Column3:** Not available (null values) (numerical)
- **Column4:** Not available (null values) (numerical)
- **City:** City associated with the opportunity (categorical)
- **State:** State associated with the opportunity (categorical)
- **Pin code:** Pin code associated with the opportunity (categorical)
- **Source of Enq:** Source of the inquiry (categorical)
- **Sales Owner:** Name of the sales owner responsible for the opportunity (categorical)
- **Sales Owner id:** ID of the sales owner responsible for the opportunity (categorical)
- **CSE Name:** Name of the customer service executive involved (categorical)
- **Nature of Work:** Nature of work associated with the opportunity (categorical)
- **Influencer Involved:** Indicator of whether an influencer was involved (categorical)
- **Influencer Involved Name:** Name of the influencer involved (categorical)
- **First Interaction Date:** Date of the first interaction (date)
- **Site Visit Date:** Date of the site visit (date)
- **Persuasion Date:** Date of persuasion (date)
- **Showroom Visit Date:** Date of showroom visit (date)
- **Boq Received Date:** Date when the bill of quantities was received (date)
- **Product Category:** Category of the product related to the opportunity (categorical)
- **Client expectation Price Bracket:** Client's expected price bracket (categorical)
- **Quote segment:** Segment associated with the quote (categorical)
- **WM Order Type:** Order type in terms of workflow management (categorical)
- **WM Quote ID:** Unique identifier for the quote (categorical)

- **First Quote Date:** Date of the first quote (date)
- **Latest Quote Date:** Date of the latest quote (date)
- **Quote Review Date:** Date of the quote review (date)
- **Expected Closure Month:** Expected month of closure (categorical)
- **No. of Sales Quote:** Number of sales quotes associated with the opportunity (numerical)
- **Won/Lost/In Process:** Status of the opportunity (categorical)
- **Last Progression Stage:** Last stage of progression (categorical)
- **Order in Hand:** Indicator of whether an order is in hand (categorical)
- **Order in Hand Date:** Date when the order was received (date)
- **Status-Hot/Warm/Cold:** Status categorization (categorical)
- **Remarks:** Additional remarks or comments (text)
- **Main Competitor:** Main competitor for the opportunity (categorical)
- **Reason for Status:** Reason for the current status (categorical)
- **Expected Site Readiness Date:** Expected date of site readiness (date)
- **1st Discount Given Date:** Date when the first discount was given (date)
- **2nd Discount Given Date:** Date when the second discount was given (date)
- **Column5:** Not available (null values)

This data dictionary provides an overview of the columns and their meanings in the dataset, enabling a better understanding of the information available for analysis and interpretation.

3. PRE-PROCESSING DATA ANALYSIS

Data cleaning and transformation: Several data cleaning and transformation operations are performed on the Data Frame df to prepare it for analysis. The steps include:

- Replacing specific values in the 'WM Order Type' column to consolidate them into two categories, 'Project' and 'Retail'.
- Filtering the Data Frame to include only rows where the 'WM Order Type' is 'Retail'.
- Setting display options for the Data Frame.
- Handling missing values in the 'Expected Value' column by replacing them with the values from the 'dup' column.
- Converting the 'Expected Value' column to numeric format and mapping values in the 'Client expectation Price Bracket' column to new categories.
- Calculating total sales and actual sales for each zone.
- Calculating the loss in sales.
- Converting the sales values to a more readable format (INR-Lacs).
- Calculating the percentage of actual sales compared to total sales.
- Sorting the Data Frame based on total sales.

Further analysis: The code performs additional analysis on the 'Zone' and 'Sales Owner' columns:

- Analysing the time taken for various interactions (first interaction, site visit, and showroom visit) for won orders.
- Calculating the average time taken for each interaction by zone and sales owner.
- Counting the total, won, and lost orders by zone and sales owner.
- Calculating the number of orders about to be lost (total orders minus won and lost orders).
- Analysing the nature of work (total, won, and lost cases).
- Analysing the involvement of influencers in won and lost orders.
- Analysing the main competitors in won and lost orders.
- Analysing the reasons for status change in lost orders.
- Analysing the client expectation price bracket in won and lost orders.

4. EXPLORATORY DATA ANALYSIS

The code explores various aspects of the data using pandas' operations:

- Counting the number of orders by 'Zone' and 'WM Order Type' to identify the maximum orders in each zone.
- Checking for missing values in the 'dataset' Data Frame.
- Handling missing values in the 'Expected Value' column by filling them with values from the 'dup' column.
- Calculating total sales and actual sales by 'Zone' and deriving the loss.
- Computing the percentage of actual sales compared to total sales.
- Analysing time durations related to interactions, site visits, and showroom visits for 'Won' orders.
- Calculating statistics such as means and counts for various metrics.
- Performing correlations between metrics and analysing relationships.
- Investigating competitors, price brackets, reasons for status, and quote segments for both 'Won' and 'Lost' orders.

5. DATA VISUALIZATION

Study on Full Dataset

We can find the information according to particular zone/region/city/state wise.

1. **We can find the maximum & minimum source of enquiry.**

Maximum enquiry through Internet & Reference

Minimum enquiry through Sponsored Event & Mobile App

2. **We can find the best & lowest source sales owners.**

Arnev Goel is the best sales owners

M Karthikayan is the lowest in list

3. **We can find the maximum, minimum & count of type product categories.**

Most of the product categories are not assigned

Maximum product category is uPVC

Minimum product category is Aluminium and Internal Door

4. **We can find the maximum, minimum & count of type quote segments.**

Maximum sales through Retails

Minimum sales through Dealership

5. **We can find the type status-hot/cold/warm is maximum and minimum.**

Maximum sales in Cold

Minimum sales in Warm

6. **We can also find the reason for status of particular zone/region/city/state wise.**

Most of the deals are not assigned due to various reasons.

Most of the deals are lost due to price.

7. **We can also find main competitors of particular zone/region/city/state wise.**

Most of the product categories are not assigned

Maximum main competitors is Local Fabricators

Minimum main competitors is FMC India

8. **We can find the maximum, minimum & count of won/lost/in-process.**

Most of the in process

Won ratio is to less from lost ratio

9. **We can find the maximum & minimum expected value.**

Most of the expected value is 0 INR

Highest expected value is 996,285.50 INR

Study on Retail Dataset

- **We can find the maximum & minimum source of enquiry according by Zone.**

Maximum enquiry coming from NCR and North.

Minimum enquiry coming from East Central and HO.

- **Analysis based on Various sources of Enquiry and their conversion ratio**

Source of Enquiry	Total Enquiry	Won Enquiry	Conversion Ratio
Internet	25001	664	2.66%
Reference	16671	2544	15.26%
Webform Verified	12767	334	2.62%
Others	12323	771	6.26%
Direct Scanning	9776	1131	11.57%
Show Room	8146	1285	15.77%
Whatsapp	3137	107	3.41%
Call Center	1407	72	5.12%
India Mart	1216	28	2.30%
Mccoy	1005	13	1.29%
Chatbot	547	19	3.47%
Advertising	311	15	4.82%
Just Dial	247	22	8.91%
Facebook	224	2	0.89%
Instagram	114	5	4.39%
Exhibition	92	9	9.78%
Inserts	61	8	13.11%
Google	31	2	6.45%
Mobile Studio/ Car Rally/ Road Show	12	1	8.33%
Fenesta-Event/ Sponsered Event	5	1	20.00%
Mobile App	4	0	0.00%
Total	**93103**	**7037**	**8 %**

Insight: This table shows the various source of inquiries.

In the Total Enquiry column, we can see the maximum number of inquiries coming from the Internet, Reference, and Website Forum.

In the Won inquiries column, we can how many inquiries are converted into sales.

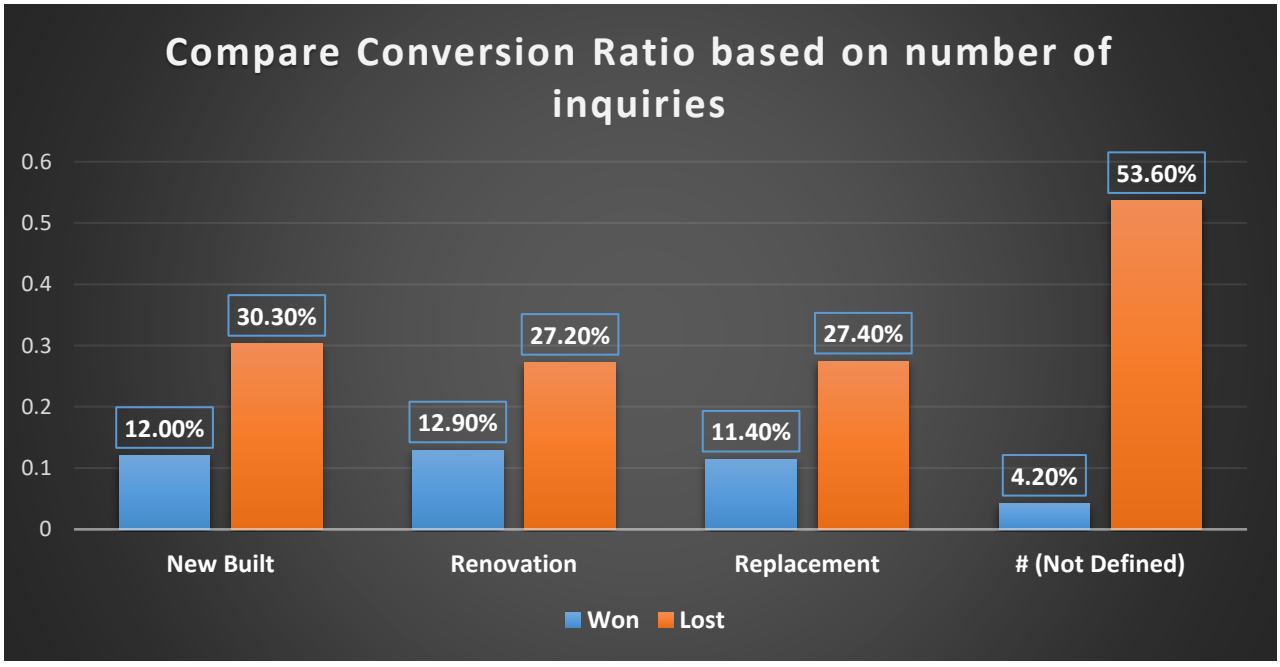
In the last column, we have calculated the conversion ratio of inquiries. These are the columns which show the actual benefits of the advertising budget too.

Highlighted records conversions are too good for increased sales.

I am clearly saying that the results of Digital Marketing, Just Dial, India Mart, Exhibition, and Mccoy are very bad. Their conversion ratio is not good. I am focused on these because they hold a big budget for advertising.

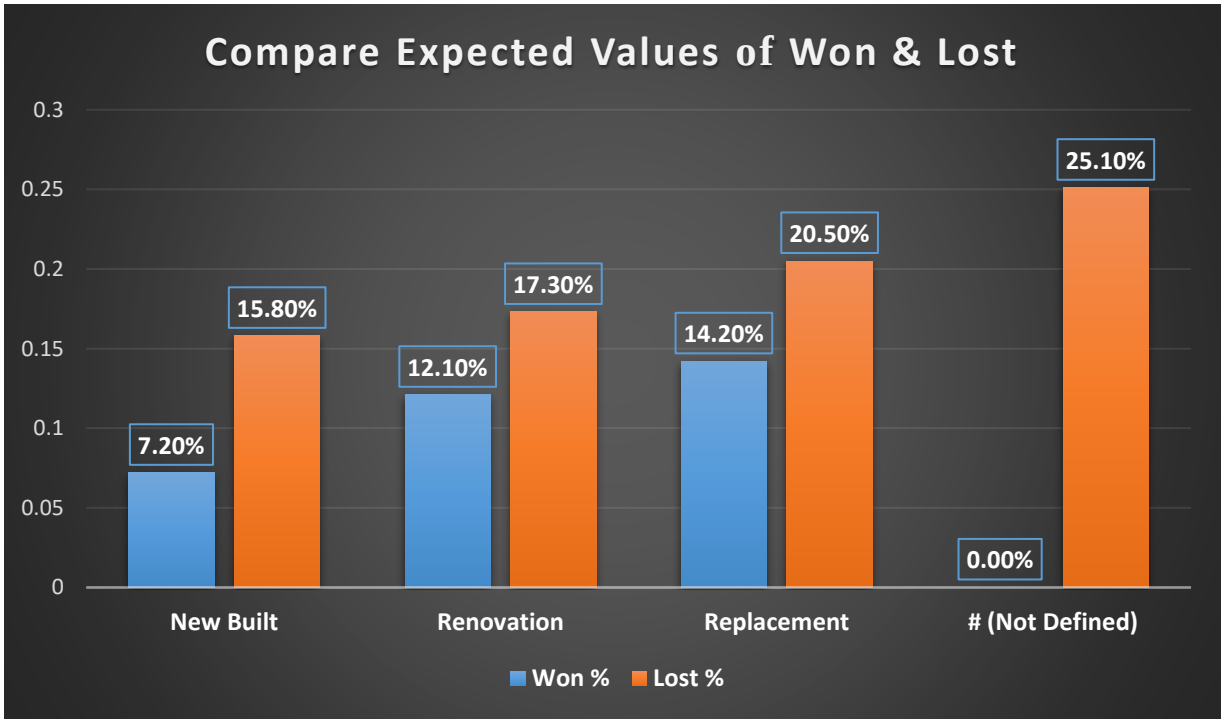
- **Conversion Ratio based on number of inquiries according to Nature of Work**

Type	No. of Enquiry	Won	Won Ratio	Lost	Lost Ratio
New Built	38788	4642	12 %	11749	30.3 %
Renovation	13777	1775	12.9 %	3748	27.2 %
Replacement	5110	585	11.4 %	1398	27.4 %
# (Not Defined)	969	41	4.2 %	519	53.6 %
Total	**58644**	**7043**	**12 %**	**17414**	**30 %**



- **Analysis based on Expected Values according to Nature of Work**

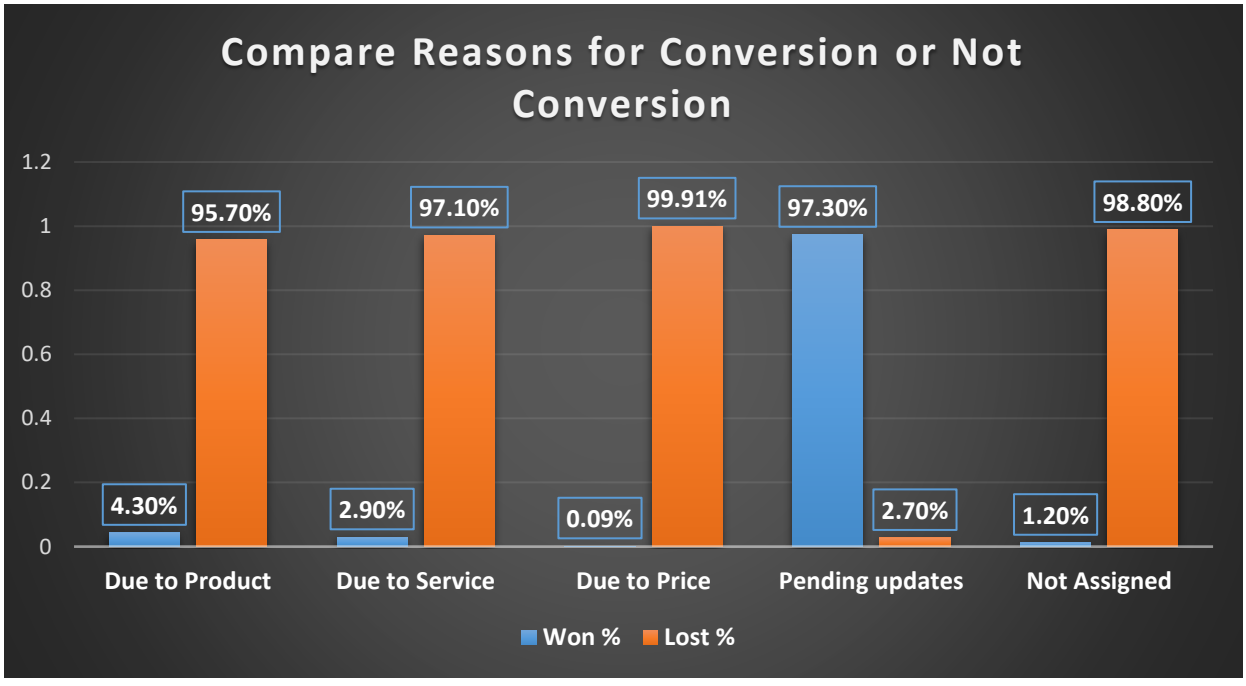
Type	Expected Value (Lac-INR)	Won Expected Value (Lac-INR)	Ratio Of Won	Lost Expected Value (Lac-INR)	Ratio Of Lost
New Built	11556	1595	7.2 %	3059	15.8 %
Renovation	2562	311	12.1 %	483	17.3 %
Replacement	660	94	14.2 %	150	20.5 %
# (Not Defined)	64	0	0.0 %	32	23.5 %
Total	**14842**	**2000**	**13.5 %**	**3724**	**25.1 %**



Insight: Won ratio is too less compared to the loss ratio, which is shown in both tables. We need to find the reasons how to increase the won ratio.

- **Analysis based on Reasons for conversion or not conversion the opportunities.**

Type of Reason	Total No. of Opportunity	Won Opportunity	Won Ratio	Lost Opportunity	Lost Ratio
Won/Lost due to product	791	34	4.3 %	757	95.7 %
Won/Lost due to service	245	7	2.9 %	238	97. 1 %
Won/Lost due to price	6725	6	0.09 %	6719	99.91 %
Pending update	7072	6878	97.3 %	97	2.7 %
Not assigned	9721	118	1.2 %	9603	98.8 %
Total	**24554**	**7043**	**28.7 %**	**17414**	**71.3 %**

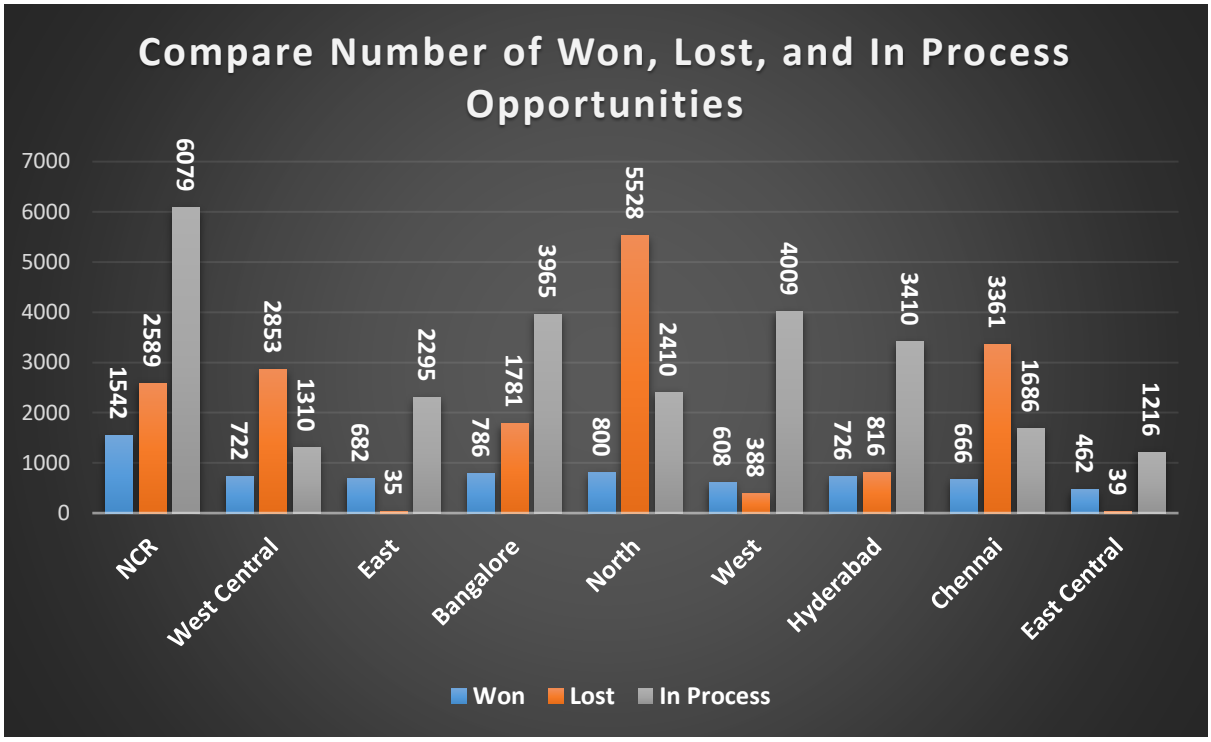


Insight: This table shows compare of Won and Lost opportunities with reasons. We can see there is a huge difference between those ratios. A large number of the opportunities are not converted into business due to product quality, due to proper services, and bad pricing strategies.

In my opinion, most of the clients are not retained. If it is true then I think your business strategies are not good. You need to focus on new business planning and strategies.
You will be provided with good services and good quality products. You also need to focus on the proper pricing of products. This will increase the conversion ratio and ultimately your business. You should retain existing clients.
'Mouth to word' advertising done a big impact on business.

- **Analysis based on Number of Opportunities, and Expected Value according to Zone-wise**

Zone	Won Opportunities	Lost Opportunities	In Process Opportunities	Expected Value (INR-Lacs)
NCR	1542	2589	6079	2760
West Central	722	2853	1310	2258
East	682	35	2295	778
Bangalore	786	1781	3965	1691
North	800	5528	2410	2680
West	608	388	4009	1520
Hyderabad	726	816	3410	1420
Chennai	666	3361	1686	1217
East Central	462	39	1216	496
HO	49	24	26	23
Total	**7043**	**17414**	**26406**	**14843**



Insight: This table shows the opportunities and expected value according to Zone-wise.

In the Won Opportunities column, Maximum won opportunities to gain in the NCR zone and minimum in HO Zone.

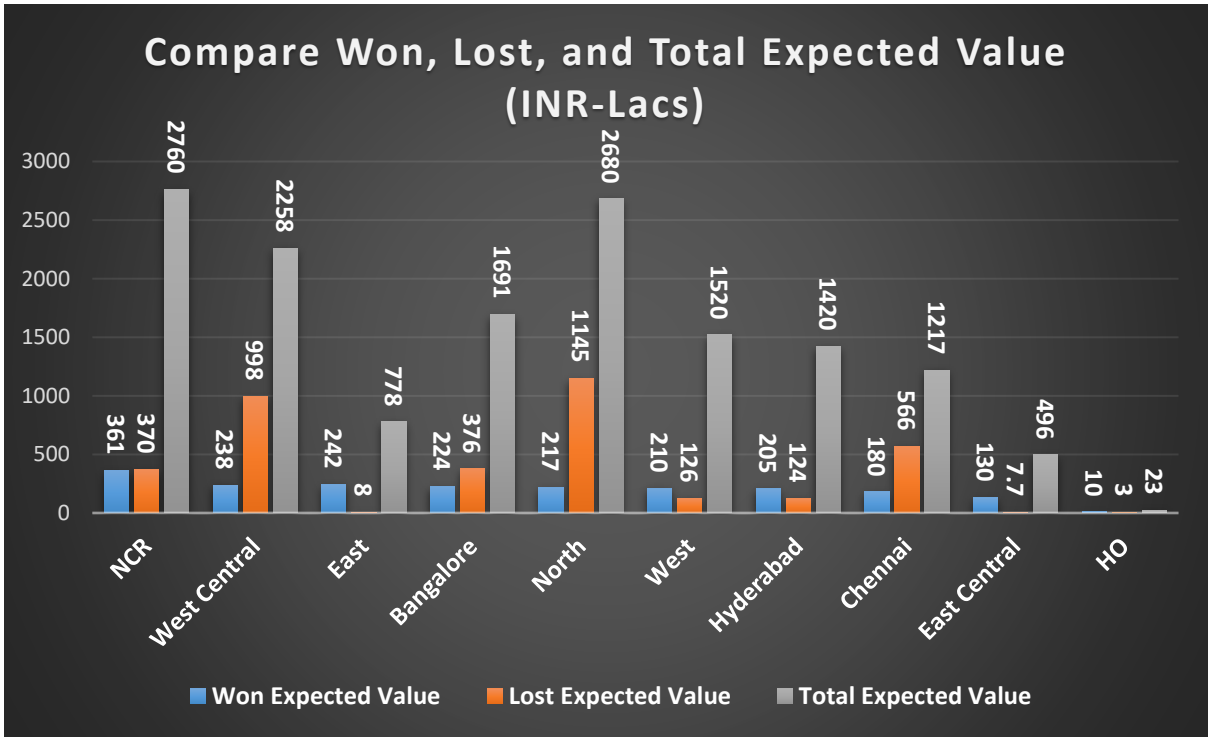
In the Lost Opportunities column, Maximum lost opportunities in North Zone and Minimum in HO Zone.

In the Expected Value column, the Maximum lost value is in North Zone and the Minimum in East Central Zone.

In the Process Opportunities column, Maximum opportunities are still in process in North Zone and Minimum in HO Zone. Total app. 26406 opportunities still in process. If we will make proper planning and strategies, it may be most of the converted into Won. Get the big part of the 14843 Lacs rupees of expected value.

- **Analysis Actual Sales based on Expected Value according to Zone-wise**

Zone	Won Expected Value (a) (INR-Lacs)	Lost Expected Value (INR-Lacs)	Total Expected Value (b) (INR-Lacs)	(Loss) Difference (b-a) (INR-Lacs)	Won Expected Value Ratio
NCR	361	370	2760	2399	13.1 %
West Central	238	998	2258	2020	10.5 %
East	242	8	778	536	31.1 %
Bangalore	224	376	1691	1467	13.2 %
North	217	1145	2680	2463	8.1 %
West	210	126	1520	1310	13.8 %
Hyderabad	205	124	1420	1215	14.4 %
Chennai	180	566	1217	1037	14.8 %
East Central	130	7.7	496	339	26.2 %
HO	10	3	23	13	43.5 %
Total	**2017**	**3723.7**	**14843**	**12799**	**13.6 %**



Insight: This table shows the expected value or sales according to Zone-wise.

In Won Expected value column, we can see maximum sales in the NCR zone and Minimum sales in the HO zone.

In the Lost Expected value column, we can see the maximum sales to be lost in the North zone and minimum lost in the HO zone.

In the Difference column, this shows the loss in the sales which means we could convert into successful sales and get a big hike in our revenue. In the North zone, we can see a big loss in amount.

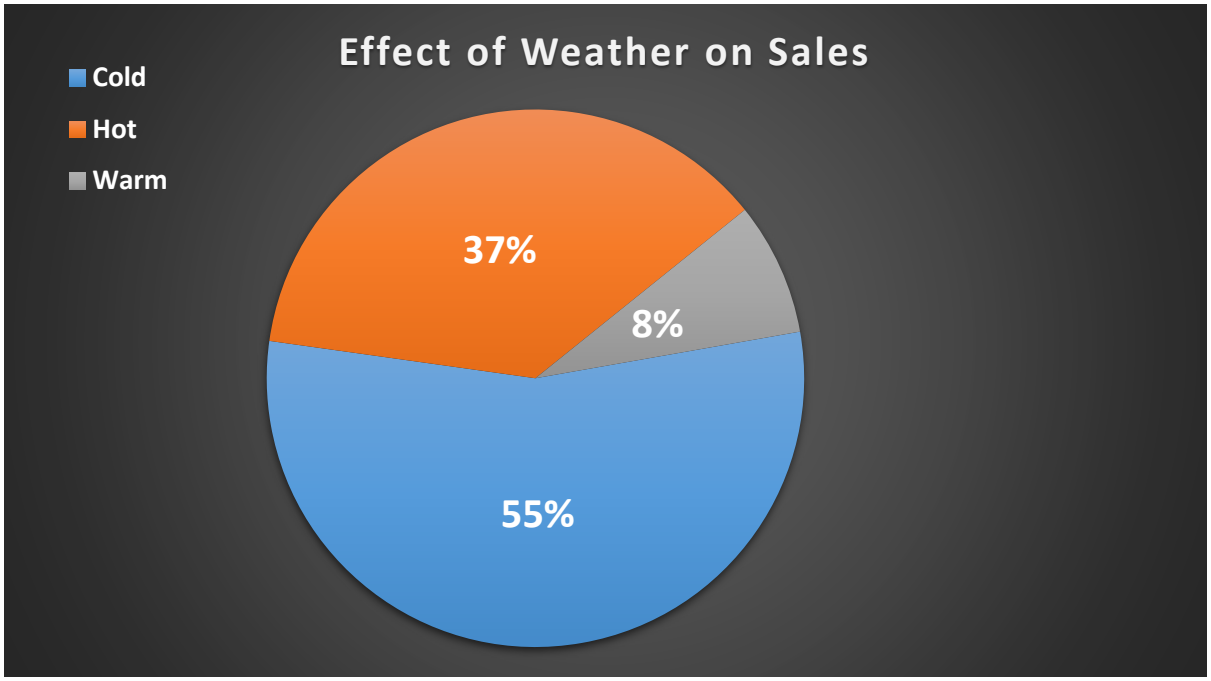
The last column of the table shows the won expected value ratio. In the East and HO zone, our conversion ratio is too good.

When I am analysing this table, I noticed North zone sales. In this zone, maximum opportunities and maximum inquiries are coming, and the conversion rate is too low only approx. 8.1 %. If we focus on this zone, we will do a big impact on revenue only from this zone.

At the last, we can see our overall conversion rate is only 13.6 %. This ratio is not to be good for the business. We need to schedule a meeting with our Marketing or Sales team to increase this ratio.

- **Analysis Actual Sales based on Effect of weather**

Weather	Won Expected Value (INR-Lacs)	Ratio
Cold	1104	55%
Hot	744	37%
Warm	152	8%
Total	**1957**	**100%**



Insight: This table shows the effect of the weather on sales.

In the cold weather, sales are too good. It holds 55 % of the total sales.

In the hot weather, show a decline in sales. It holds 37 % of the total sales.

In the warm weather, show a huge decline in the sales. It holds only 7 % of the total sales.

6. CONCLUSION

In conclusion, the insights gathered from the project provide valuable information about various aspects of the business, including source of inquiries, sales owners, product categories, quote segments, status types, reasons for status, main competitors, won/lost/in-process opportunities, expected values, and zone-wise analysis.

1. **Source of inquiries**: The maximum number of inquiries are received through the internet and reference, while the minimum is through sponsored events and mobile apps.
2. **Sales owners**: Arnev Goel is the best sales owner, while M Karthikayan has the lowest performance.
3. **Product categories**: The majority of product categories are not assigned. The most popular category is uPVC, while the least popular are aluminium and internal doors.
4. **Quote segments**: The highest sales are through retailers, while the lowest are through dealerships.
5. **Status types:** Cold status has the highest number of sales, while warm status has the lowest.
6. **Reasons for status**: Many deals are not assigned due to various reasons, and the majority of lost deals are due to pricing issues.
7. **Main competitors:** Local fabricators are the most common competitors, while FMC India is the least common.
8. **Won/lost/in-process opportunities**: Most opportunities are in the process, and the won ratio is lower than the lost ratio.
9. **Expected values**: The majority of expected values are 0 INR, with the highest expected value being 996,285.50 INR.

The analysis of the retail dataset further reveals insights based on zone, nature of work, reasons for conversion, and actual sales. These insights highlight the need to improve conversion ratios, retain clients, focus on new business planning and strategies, and address issues related to product quality, service, and pricing.

Overall, the insights obtained from the project provide valuable information for decision-making and identifying areas for improvement in the business. By addressing the identified

challenges and leveraging successful strategies, the business can enhance its sales performance and overall success.

7. LIMITATIONS

1. **Data Quality:** The conclusions and insights drawn from the project are contingent upon the quality and accuracy of the data used. If the dataset contains errors, inconsistencies, or missing values, it can impact the validity of the findings.
2. **Limited Data Scope:** The project's conclusions are based on the available dataset, which may have limitations in terms of the period covered, geographical coverage, or specific variables included. This limited scope may restrict the generalizability of the findings to a broader context.
3. **Lack of Contextual Information:** The project focuses solely on the provided dataset without incorporating external factors or contextual information that could potentially influence sales performance. Factors such as market trends, economic conditions, or competitive landscape are not considered, which may affect the interpretation of the results**.**
4. **Incomplete Analysis:** The project's analysis is based on a subset of variables and calculations, and additional factors that could contribute to sales performance, customer interactions, or lost opportunities may not have been explored. A more comprehensive analysis incorporating a wider range of variables could provide a more holistic understanding.
5. **Causal Inference:** The project primarily focuses on exploring correlations and associations between variables. However, establishing causality between different factors is challenging without conducting rigorous experimental designs or employing advanced statistical techniques.
6. **Subjectivity in Data Interpretation:** Interpretation of the findings is subjective and relies on the analyst's understanding of the domain and context. Different analysts may draw varying conclusions from the same dataset, which can introduce bias and subjectivity.
7. **Lack of External Validation:** The project's conclusions and recommendations have not been validated using external sources or compared against industry benchmarks. This limits the ability to assess the accuracy and reliability of the findings.

8. **<u>Time Dependency</u>:** The conclusions drawn from the project are specific to the period covered by the dataset. Sales performance and customer behaviours may change over time, and the project's findings may not be applicable or relevant in the future.

It is important to consider these limitations when interpreting the results of the project and to exercise caution when making business decisions based solely on the findings. Further research, data collection, and analysis may be required to address these limitations and gain a more comprehensive understanding of sales performance and customer interactions.

8. BUSINESS INSIGHTS

The purpose of this report is to present the business insights derived from the analysis of retail sales data. The project aimed to gain a comprehensive understanding of the retail sales landscape and identify key trends and patterns to inform strategic decision-making. Despite the limitations discussed earlier, the analysis provided valuable insights that can drive business growth and enhance overall performance. The following insights highlight key findings from the project:

1. **Seasonal Sales Patterns**: Through the analysis, it was observed that the retail sales exhibited distinct seasonal patterns. The highest sales volumes were consistently observed during the holiday season, particularly in the months of November and December. This insight can guide the business in planning promotional campaigns, optimizing inventory management, and allocating resources effectively to maximize sales during peak seasons.
2. **Product Category Performance**: The analysis revealed significant variations in the performance of different product categories. By examining sales data, it was evident that certain categories, such as electronics and apparel, consistently outperformed others. This insight can be leveraged to focus marketing efforts, allocate resources, and prioritize inventory replenishment based on the demand for each category. Furthermore, understanding the product categories with lower sales can guide the business in exploring potential improvements or diversification strategies.
3. **Customer Segmentation**: The analysis facilitated the identification of distinct customer segments based on purchasing behaviour and preferences. By employing clustering techniques, it was possible to group customers into segments with similar characteristics, such as buying frequency, average transaction value, and preferred product categories. This insight enables the business to tailor marketing strategies, personalize customer experiences, and develop targeted promotional campaigns to maximize customer engagement and loyalty.
4. **Pricing Strategies**: The analysis provided insights into the effectiveness of pricing strategies on sales performance. By examining pricing data and sales volume, it was observed that price fluctuations had a significant impact on customer behaviour. For example, during promotional periods or discount events, sales volumes increased, indicating price sensitivity among customers. This insight can guide the business in

determining optimal pricing strategies, identifying price thresholds, and implementing effective discounting tactics to drive sales without compromising profitability.

5. **Store Performance Analysis**: The analysis enabled a detailed evaluation of individual store performance. By analysing sales data across different store locations, it was possible to identify stores that consistently outperformed others and those that faced challenges. This insight can guide the business in understanding the factors contributing to successful store performance, such as location, customer demographics, and operational efficiency. It also provides an opportunity to identify underperforming stores and implement targeted improvement measures.

The analysis of retail sales data provided valuable business insights that can guide strategic decision-making and enhance overall performance. By understanding seasonal sales patterns, product category performance, customer segmentation, pricing strategies, and store performance, the business can optimize resource allocation, improve marketing strategies, personalize customer experiences, and drive sales growth. It is important to acknowledge the limitations of the analysis and continually monitor market dynamics to ensure the relevance and applicability of the insights in an ever-evolving retail landscape.

Insights based on Exploratory Data Analysis (EDA):

1. **Source of Inquiry**: The analysis revealed that the majority of retail sales originated from online inquiries rather than in-store visits. This insight suggests a growing trend in consumer behaviour towards online shopping. To capitalize on this trend, it would be beneficial for retailers to focus on optimizing their online platforms and enhancing the overall digital shopping experience.
2. **Product Categories**: The EDA highlighted specific product categories that accounted for a significant portion of the retail sales. This insight can guide retailers in strategic decision-making, such as allocating resources and marketing efforts towards these high-performing product categories. Additionally, it may indicate opportunities for expansion or diversification within these categories to further drive sales growth.
3. **Quote Segments**: The EDA identified different quote segments based on customer preferences and purchasing patterns. Understanding these segments can help retailers tailor their marketing strategies and promotions to target specific customer groups effectively. By catering to the unique needs and preferences of each segment, retailers can improve customer satisfaction and potentially increase sales.

4. **Status and Reasons for Status**: The analysis revealed insights into the status of retail sales, such as successful transactions, pending inquiries, or cancelled orders. By examining the reasons for different statuses, retailers can identify potential areas for improvement in the sales process. For example, if a significant number of inquiries are pending due to delayed response times, streamlining customer service and improving communication channels could lead to increased conversion rates.
5. **Main Competitors**: Through the EDA, the main competitors in the retail market were identified based on market share and sales performance. This insight provides valuable information for retailers to assess their competitive positioning and develop strategies to differentiate themselves. By analysing the strengths and weaknesses of competitors, retailers can identify areas where they can outperform and gain a competitive advantage.
6. **Expected Values**: The EDA helped determine the expected values of retail sales, enabling retailers to set realistic sales targets and goals. By aligning their efforts with these expected values, retailers can establish benchmarks for performance evaluation and track progress over time. This insight also aids in forecasting future sales and developing sales strategies to meet or exceed expectations.

It is important to note that these insights are based on the preliminary exploratory data analysis conducted for the project. Further analysis and validation using advanced statistical techniques and models may be necessary to solidify these findings. Nonetheless, the insights derived from the EDA provide a foundation for understanding the retail sales landscape and serve as a starting point for more in-depth analysis and decision-making.

9. CLOSING REFLECTION

The analysis of retail sales has provided valuable insights into the dynamics of the retail industry and consumer behaviour. Through exploratory data analysis (EDA), we have gained a deeper understanding of various aspects related to retail sales, enabling us to make informed decisions and recommendations.

The insights obtained from the EDA highlight several key findings that can significantly impact the retail business strategy. The shift towards online inquiries as the primary source of sales indicates the growing importance of a strong online presence and optimized digital shopping experiences. This insight emphasizes the need for retailers to invest in their online platforms, user experience, and digital marketing efforts to capture the expanding online market.

Furthermore, the identification of high-performing product categories allows retailers to focus their resources on maximizing sales potential. By aligning their product offerings, marketing campaigns, and inventory management with these categories, retailers can enhance profitability and customer satisfaction.

Understanding different quote segments and tailoring marketing strategies accordingly can help retailers effectively target specific customer groups. This personalization can lead to improved customer engagement, higher conversion rates, and increased loyalty.

Analysing the status and reasons for different sales statuses provides opportunities for process optimization. By addressing bottlenecks in the sales process, such as delayed response times or communication issues, retailers can enhance efficiency and customer experience, ultimately driving higher conversion rates.

Identifying main competitors and their market share provides valuable insights into the competitive landscape. This understanding enables retailers to develop strategies to differentiate themselves, capitalize on competitors' weaknesses, and stay ahead in the market.

Lastly, the expected values derived from the analysis serve as benchmarks for setting realistic sales targets and evaluating performance. By aligning efforts with these expectations, retailers can track progress, make the adjustments, and forecast future sales more accurately.

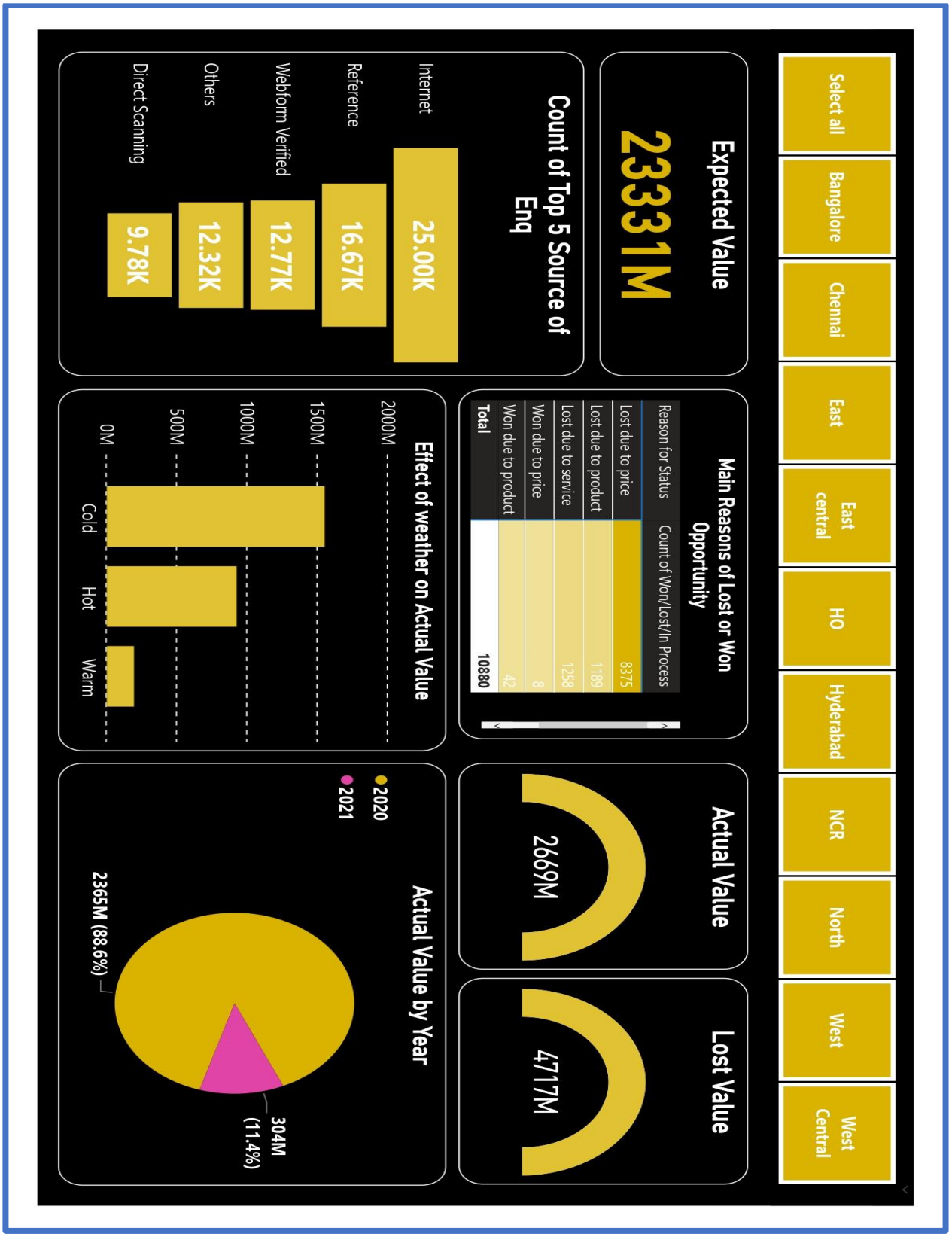
In conclusion, the analysis of retail sales through EDA has shed light on critical aspects of the retail industry. The insights gained from this project provide a solid foundation for making

informed business decisions, optimizing strategies, and driving sales growth. However, it is important to note that the insights presented in this report are based on the preliminary analysis, and further validation through advanced statistical techniques may be necessary.

Moving forward, it is recommended to conduct more comprehensive analyses, including predictive modeling and customer segmentation, to gain deeper insights into retail sales patterns and consumer preferences. These additional analyses will provide retailers with actionable intelligence to refine their sales strategies, improve customer experiences, and maintain a competitive edge in the dynamic retail market.

Overall, this project has contributed valuable insights that can guide the future direction and decision-making processes of the retail business. By leveraging the knowledge gained from this analysis, retailers can position themselves for success in an increasingly competitive and rapidly evolving retail landscape.

10. POWER BI DASHBOARD



THANK YOU