Professional Background

Currently I completed my B.E-ECE and I have secured 8.70 CGPA and have several skills including MS Excel, SQL, Tableau, Data Analytics and Data visualization

I have completed 8 projects in SQL, Excel, and Tableau for data analytics. These projects showcase my expertise in data analysis and interpretation. With strong analytical skills and proficiency in SQL, Excel, and Tableau, I bring valuable insights to the field of data analytic.

I have worked with several companies as an intern like Cognizant, Dagslore Technology etc, as a Data Analyst. I have worked on their Data Analytics course from scratch and managed different teams.

As a newcomer to the corporate world, I am eager to face its authentic challenges and gain insights into its operations. With my adaptable nature and openness to learning, I am confident in my ability to apply my theoretical knowledge in practical scenarios. I am enthusiastic about leveraging my skills and putting in the necessary effort to learn and grow in this dynamic environment.

Table of Contents

Instagram User Analytics

Description:

The **"Instagram User Analytics"** aims to analyse user behaviour, engagement patterns, and content preferences on the Instagram platform. The purpose is to derive insights that can help businesses, marketers, and influencers understand their audience better and optimize their content strategy accordingly.

The Problem :

You are working with the product team of Instagram and the product manager has asked you to provide insights on the questions asked by the management team.

Design:

1.Steps taken to clean the data:

First importing the datasets provided and merging them. Then removing the duplicates and the blank cells.

2.Improving the headers of each column with proper values.

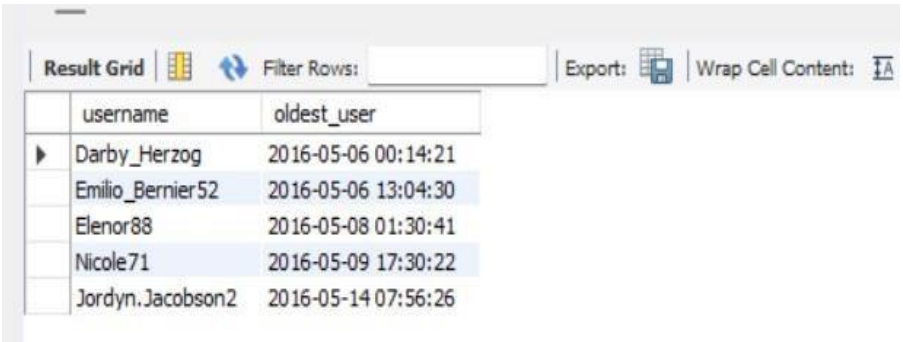
• Finding and replacing the data with correct values.

• Tools used for visualization: SQL

FINDINGS:

1.Rewarding Most Loyal Users: People who have been using the platform for the longest time.

These are the oldest user of Instagram.



username	oldest_user
Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

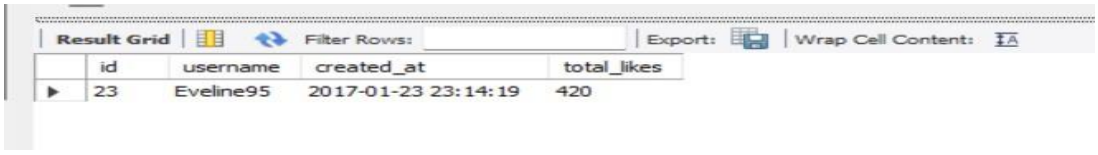
2.Remind Inactive User Engagement: By sending them promotional email to post their 1st photo

This user was inactive after their first photo

username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39

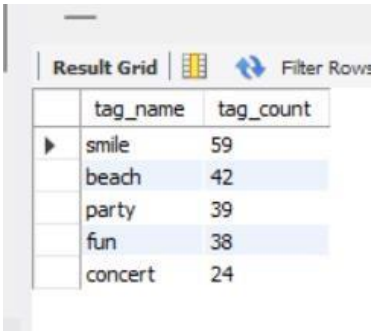
username
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67
Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard
Jessyca_West
Esmeralda.Mraz57
Bethany20

3.Contest Winner Declaration: Determine the winner of the contest and provide their details to theteam.



id	username	created_at	total_likes
23	Eveline95	2017-01-23 23:14:19	420

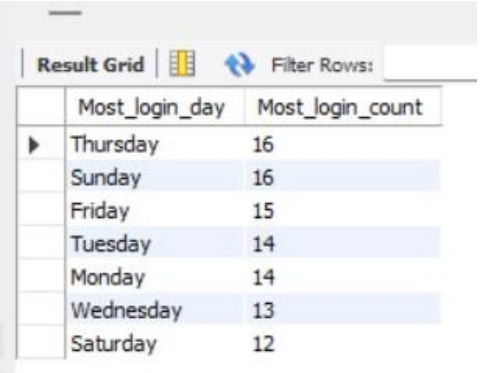
4. **Hashtag Research**: Identify and suggest the top five most commonly used hashtags on theplatform.



Result Grid | Filter Rows

tag_name	tag_count
smile	59
beach	42
party	39
fun	38
concert	24

5.Ad Campaign Launch: The team wants to know the best day of the week to launch ads.



Result Grid | Filter Rows:

Most_login_day	Most_login_count
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

ANALYSIS:

The analysis of the Instagram User Analytic project yielded several key insights:

1.User Demographics: The analysis revealed the age groups, gender distribution, and geographical locations of Instagram users. This information provides a clear understanding of the target audience and allows for targeted marketing campaigns.

2.Engagement Metrics: The project examined engagement metrics such as likes, comments, and shares. It identified the most engaging types of content and helped determine the factors that drive user engagement. This insight can be used to create compelling content and increase user interaction.

3.Content Performance: The analysis assessed the performance of different content formats, such as photos, videos, and stories. It identified which types of content generate the highest engagement and resonate most with the target audience. This information can guide content creation strategies maximum impact.

CONCLUSION :

In conclusion, the Instagram User Analytics project involved analysing and interpreting data related to Instagram users. Through the application of various analytical techniques.

Operation & Metric Analytics

Description:

The "Operation Analytics and Investigating Metric Spike" project aims to analyse operational data and investigate spikes or anomalies in key metrics. The purpose of this project is to gain insights into the performance and behaviour of systems or processes, identify any unusual patterns or trends, and take appropriate actions to address them.

Problem:

The organization faces inefficiencies and lacks visibility into its operational processes and metrics. This lack of insight hampers decision-making, affects productivity, and potentially leads to missed opportunities for improvement. There is a need to analyse operational data and metrics comprehensively to identify areas for optimization and enhance overall performance.

Findings:

1. **Process Inefficiencies:** Through data analysis, it was found that certain operational processes were inefficient, leading to bottlenecks and delays in task completion.

2. **Underutilized Resources:** Analysis revealed instances where resources were not being utilized optimally, resulting in wasted capacity and increased costs.
3. **Inconsistent Performance:** Metrics showed inconsistency in performance across different departments or teams, indicating the need for standardization and improvement.
4. **Impact of External Factors:** External factors such as market trends, seasonality, or regulatory changes were found to significantly affect operational performance and metrics.
5. **Lack of Data Integration:** Fragmented data sources and systems hindered the ability to gain a holistic view of operations, making it difficult to identify correlations and trends accurately.

Analysis:

1. **Root Cause Analysis:** Identified underlying causes of inefficiencies and discrepancies in operational performance.
2. **Trend Analysis:** Examined trends over time to understand patterns and fluctuations in metrics, allowing for proactive decision-making.
3. **Benchmarking:** Compared internal metrics against industry benchmarks to assess competitiveness and identify areas for improvement.
4. **Data Integration:** Integrated disparate data sources to create a unified dataset for comprehensive analysis, enabling a more accurate understanding of operations.

Conclusion:

Based on the findings, it's recommended to streamline processes, optimize resource allocation, and standardize performance metrics. Implementation will involve process redesign, resource reallocation,

and system integration. Continuous monitoring is essential to track effectiveness and identify new optimization areas. Improved data governance ensures data accuracy, integrity, and security, aligning operational analytics with organizational goals for sustainable performance improvement.

Hiring Process Analytics

Description:

This project is about Hiring Process Analytics where we will do our analysis over the data set provided to get insights about the data. Basically, we will perform Exploratory Data Analysis (EDA) where we will try to summarize, visualize, and identify missing values of the data which will help us to answer the various questions related to data which is being asked.

Design:

Steps taken to clean the data:

First importing the datasets provided and merging them. Then removing the duplicates and the blank cells. Improving the headers of each column with proper values. Finding and replacing the data with correct values.

Tools used for visualization: Excel

Tech-stack: MS Excel 2022

Findings:

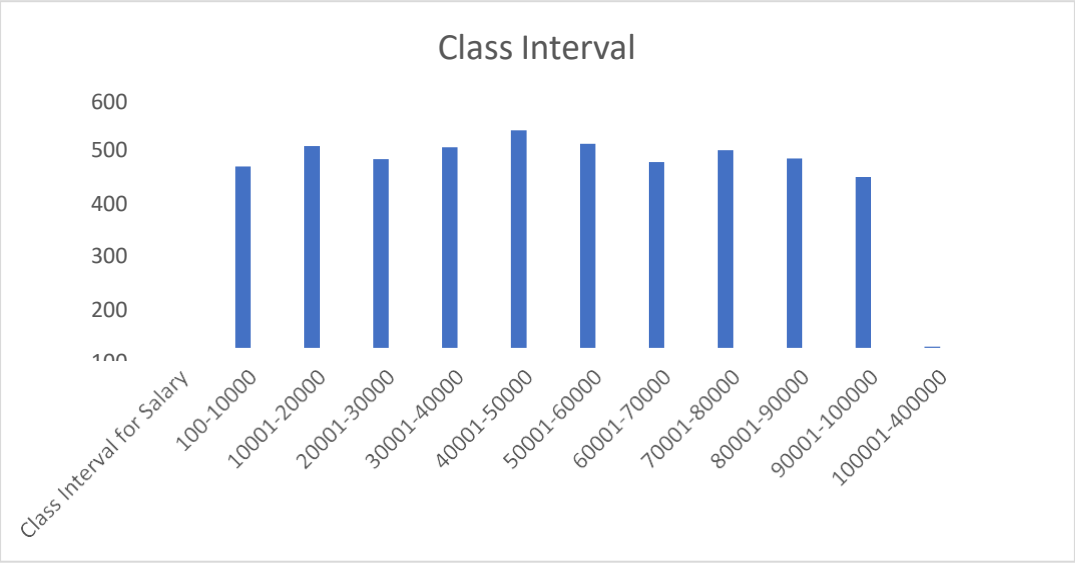
1.How many Males and Females are hired:

Event Name	Status	No Of Male and Female Hired
Male	Hired	2563
Female	Hired	1856

2.What is the average salary offered by this company?

	Average of Offered Salary
Finance Department	49628.00694
General Management	58722.09302
Human Resource Department	49002.27835
Marketing Department	48489.93538
Operations Department	49151.35438
Production Department	49448.48421
Purchase Department	52564.77477
Sales Department	49310.3807
Service Department	50629.88418
Total Average Salary	**49983.02902**

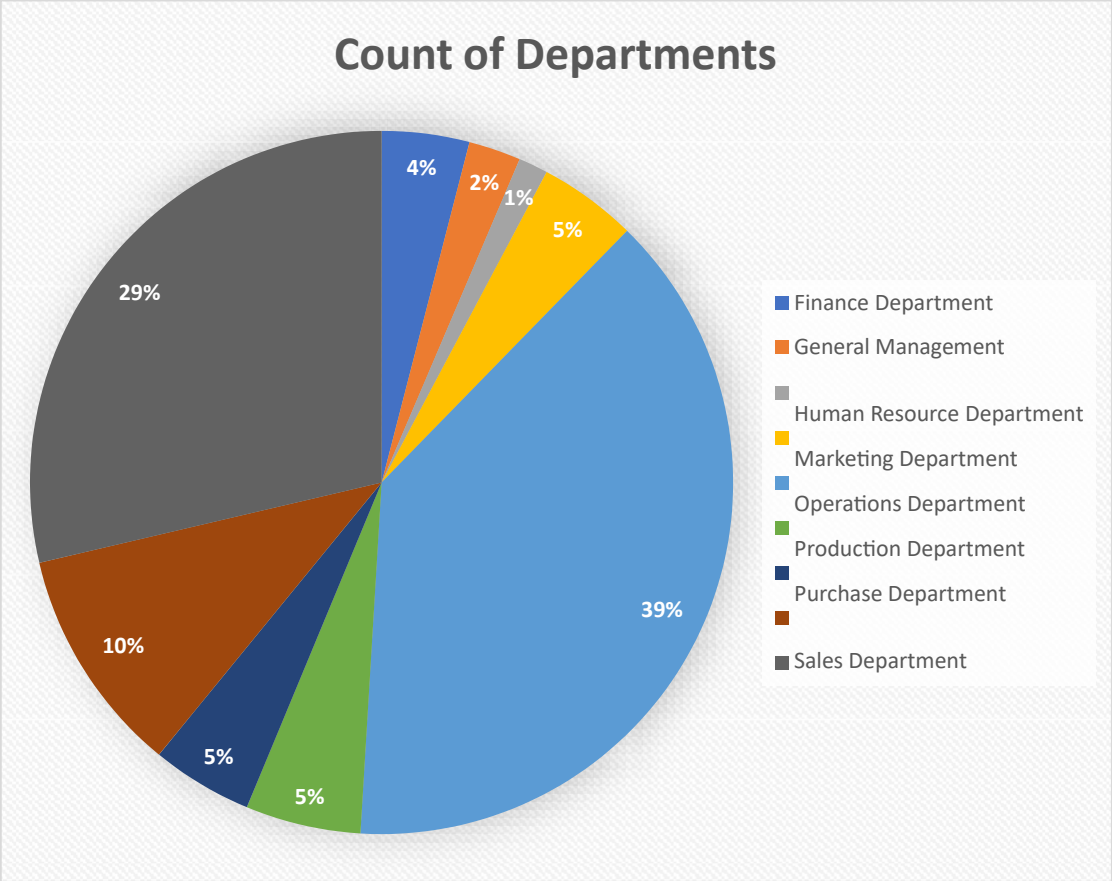
3.Create class intervals for the salaries in the company?



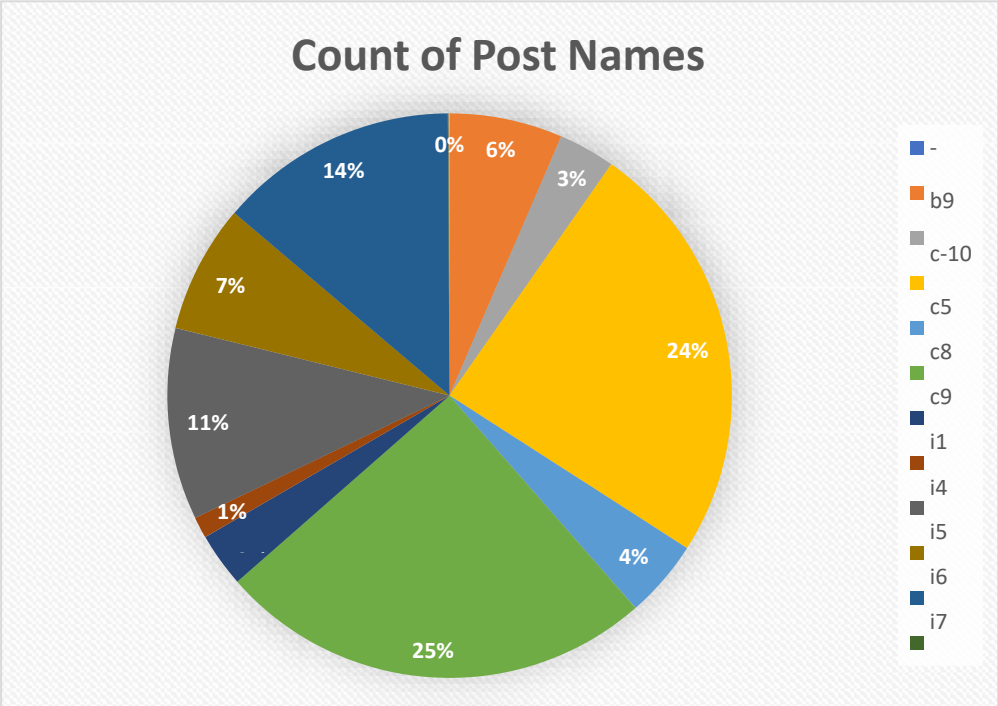
4.Departmental Analysis: Visualizing data through charts and plots is a crucial part of data analysis

Departments	Count of Department

Finance Department	288
General Management	172
Human Resource Department	97
Marketing Department	325
Operations Department	2771
Production Department	380
Purchase Department	333
Sales Department	747
Service Department	2055



5.Position Tier Analysis: Different positions within a company often have different tiers or levels.



Analysis:

The analysis of the Hiring Process Analytics involved examining the various stages and components of the hiring process within an organization. Here are the key findings:

1. **Time-to-Hire:** The analysis revealed the average time taken to complete the hiring process, from posting a job opening to making a final offer. It identified bottlenecks and delays in the process, such as lengthy candidate screening or interview scheduling, which impacted the overall time-to-hire.
2. **Sourcing Channels:** The analysis assessed the effectiveness of different sourcing channels, such as job boards, employee referrals, and social media platforms. It identified the most successful channels in attracting qualified candidates, helping

optimize recruitment strategies and allocate resources more effectively.

3. **Candidate Evaluation:** The analysis evaluated the methods and criteria used to assess candidates' qualifications, skills, and cultural fit. It identified areas for improvement in the evaluation process, such as the need for standardized interview questions or additional assessment tools.
4. **Cost Analysis:** The analysis also included an assessment of the costs associated with the hiring process, including advertising, recruitment agencies, and internal resources. It identified opportunities to optimize costs without compromising the quality of hires

Conclusion:

This project has deepened my understanding of how visualizations can be utilized in the **hiring process analytics** to gain insights into workforce distribution and make data-driven decisions to optimize organizational performance. It highlights the importance of data visualization in conveying complex information effectively and facilitating informed decision-making inhuman resources management.

IMDB Movie Analysis

Description:

IMDB Movie Analysis involved analysing data from the Internet Movie Database (IMDB) to gain insights into movie genres, ratings, box office performance, and the impact of directors and actors. The analysis provided valuable information for decision-making in the film industry, such as genre selection, talent acquisition, and release date planning. It offered a deeper understanding of the movie landscape and helped drive informed decision-making in the creation and distribution of films.

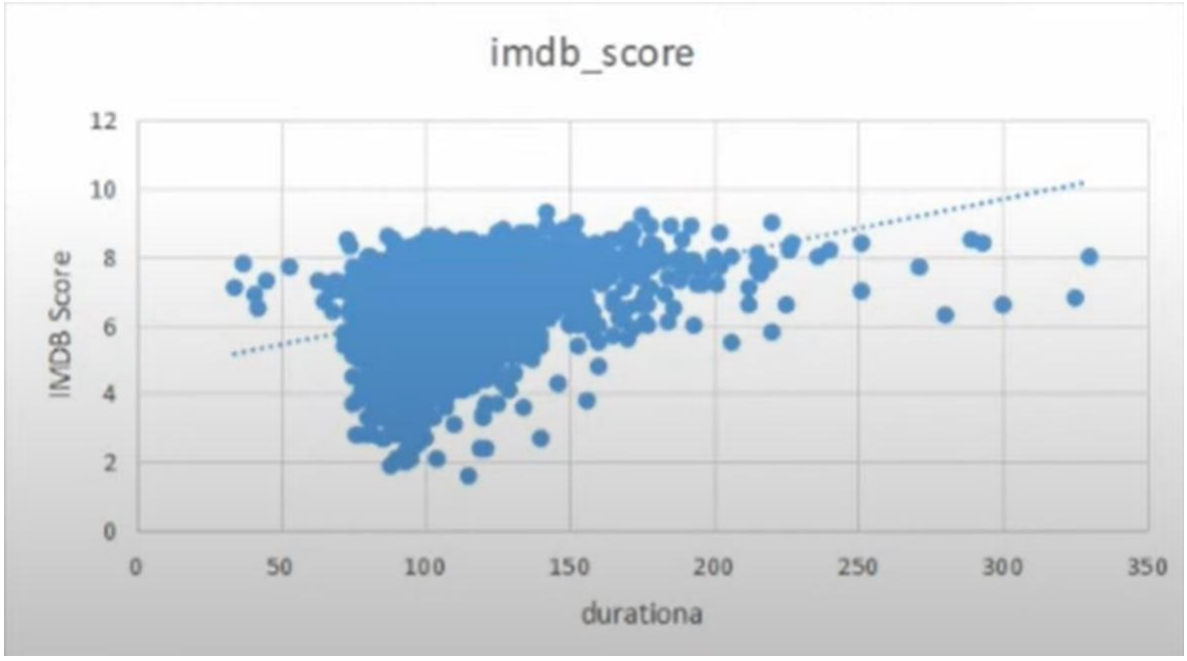
Problem:

The problem addressed in the IMDB Movie Analysis project was the lack of comprehensive insights into movie genres, ratings, box office performance, and the impact of directors and actors. The goal was to analyse the data from the Internet Movie Database (IMDB) to provide valuable information and address the need for data-driven decision-making in the film industry.

Findings:

Cleaning the data: This is one of the most important steps to perform before moving forward with the analysis. Use your knowledge learned till now to do this. (Dropping columns, removing null values, etc.)

1.Movie Duration Analysis: Analyse the distribution of movie durations and itsimpact on the IMDB score.



2.Movie Genre Analysis: Analyse the distribution of movie genres and their impact on the IMDB score.

Descriptive statistics for genres:

Average	230.5294118
Median	43
Mode	3
Max	1041
Min	2
var	120073.8
SD	346.5

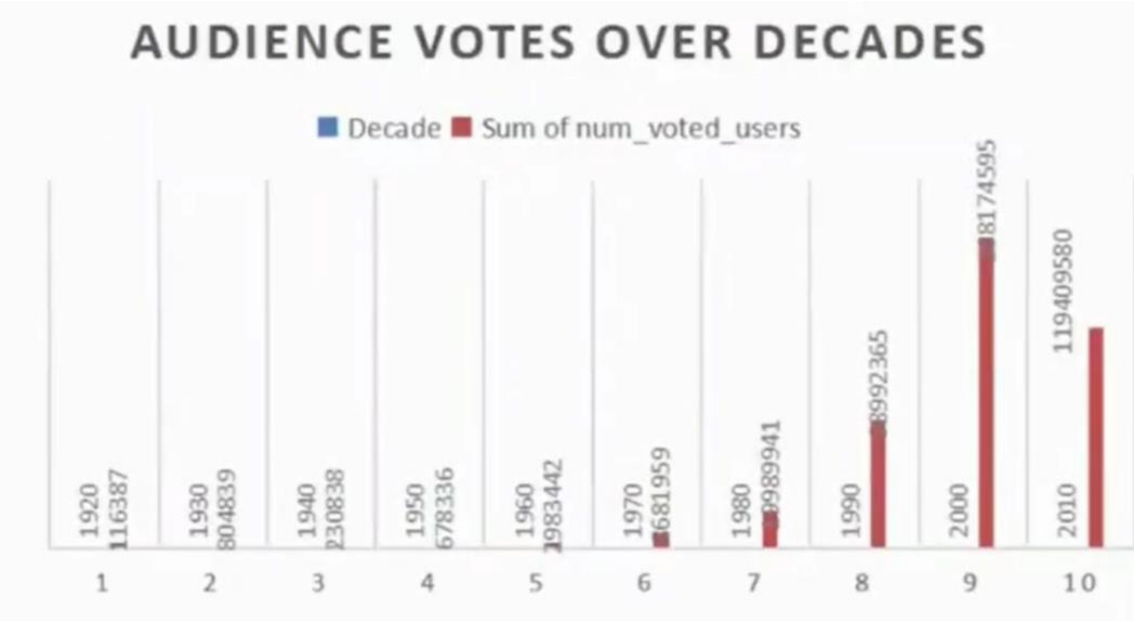
3.Director Analysis: Top directors based on their average IMDB score

Top 10 Directors	Average of imdb_score
Charles Chaplin	8.60
Tony Kaye	8.60
Alfred Hitchcock	8.50
Damien Chazelle	8.50
Majid Majidi	8.50
Ron Fricke	8.50
Sergio Leone	8.43
Christopher Nolan	8.43
Asghar Farhadi	8.40
Marius A.Markevicius	8.40

4.Budget Analysis: Explore the relationship between movie budgets and theirfinancial success.

Highest Profit Margin	523505847

5.critic-favourite and Audience-favourites Actors.



Result: Audience are getting more interest in movies and all by passing every decade.

Conclusion:

Through the IMBD Movie Analysis project, I have honed my logical, statistical, and technical skills to extract meaningful insights from the dataset. Concepts such as averages, frequency tables, and identifying outliers have enabled me to establish a deeper connection with the data and conduct thorough analyses.

By applying statistical techniques and leveraging the technical capabilities of Microsoft Excel, I have been able to streamline data analytics tasks and simplify complex calculations. Additionally, the data visualization functionalities of Excel have allowed me to effectively communicate insights by presenting data in a visually appealing manner. I have also gained an understanding of how to choose the appropriate visualization techniques based on the nature of the data and the desire outcome.

Bank loan case study

Description:

This case study attempts to demonstrate the application of EDA in a real-world business environment. In this case study, in addition to using the techniques learned in the EDA module, you will gain a basic grasp of risk analytics in banking and financial services, as well as how data is utilized to reduce the risk of lending money when lending to consumers.

Approach:

This case study has two enormous data sets: the current application and the previous application. Each included several unneeded columns that would be useless for risk assessments, as well as many blank data. I started by cleaning. To evaluate this enormous set of data, I first cleaned the data, located some outliers and deleted them, and then began performing univariate and bivariate analysis using pivot tables and charts.

DESIGN:

Following the data cleaning procedure, I split columns in the dataset based on two categories of variables.

1. Categorical variables
2. Numerical variables

Findings:

1.EDA

Categorical variables (non-numerical variables): person's occupation, education status.

Numerical variables: income, credit etc.,

The following are some of the categorical and numerical variables from the provided data set.

Categorical variables	**Numerical variables:**
Gender	Age
Number contract type	Days employed
Income Type	Amount Income

Education	Amount Annuity
Housing Type	Amount Credit

I completed full EDA on the present application and then on the previous application. Then, in this report, I summarised the results of both applications and provided business insights.

2.Missing Data: The existing application sheet included 161 columns.

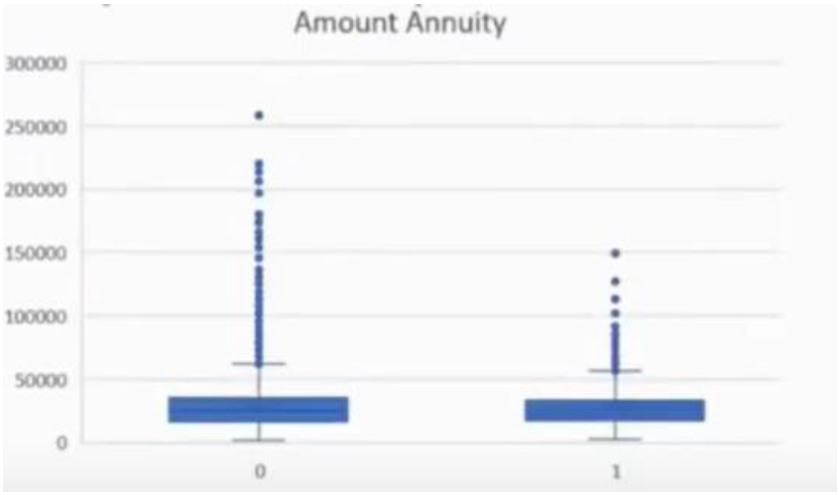
1. I deleted columns with more than 5% blank data.
2. I deleted a large number of useless columns

To eliminate blank values, I used the COUNTBLANK function

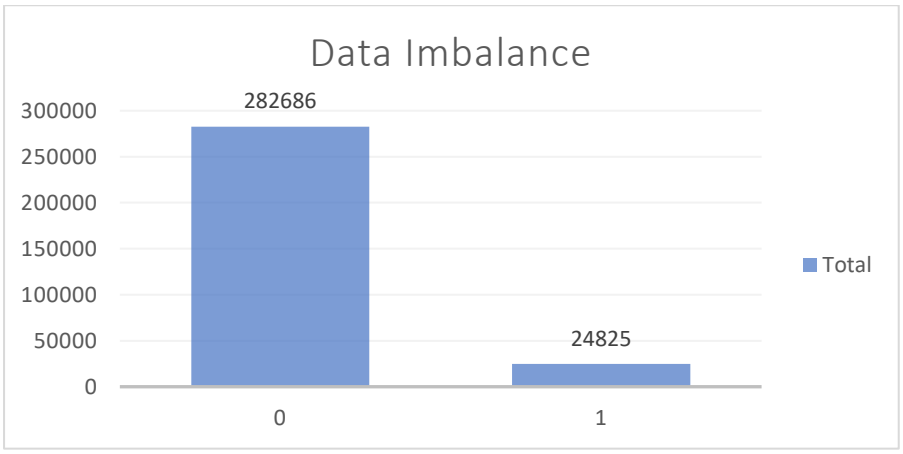
3.Outliers can only be identified on Numeric variables.

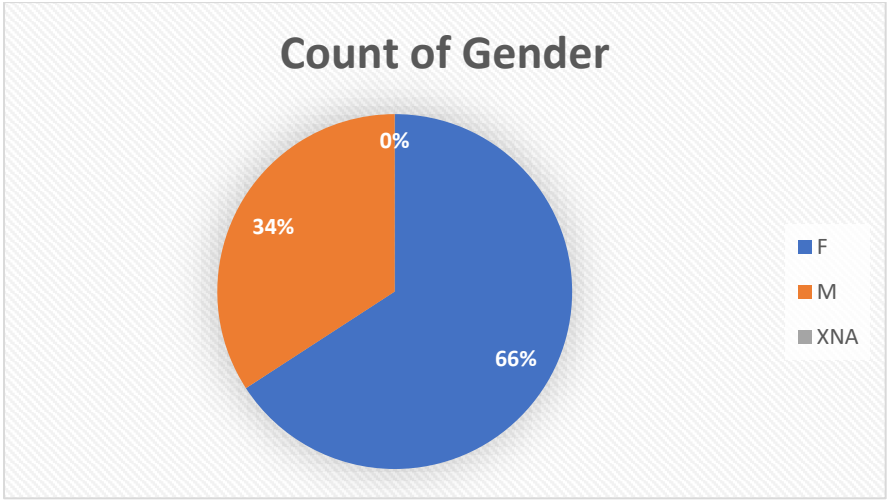
Box plotted Target column vs

1. Amount credit
2. Amount Income
3. Amount Annuity



4.Data imbalance: Data imbalance occurs when data is disseminated in an unequal manner. I plotted data imbalance using Pivot charts



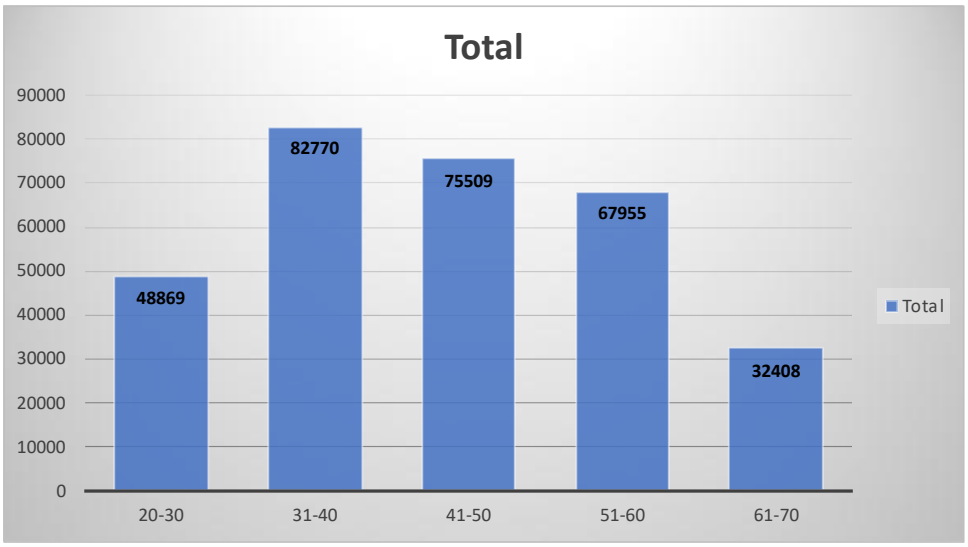


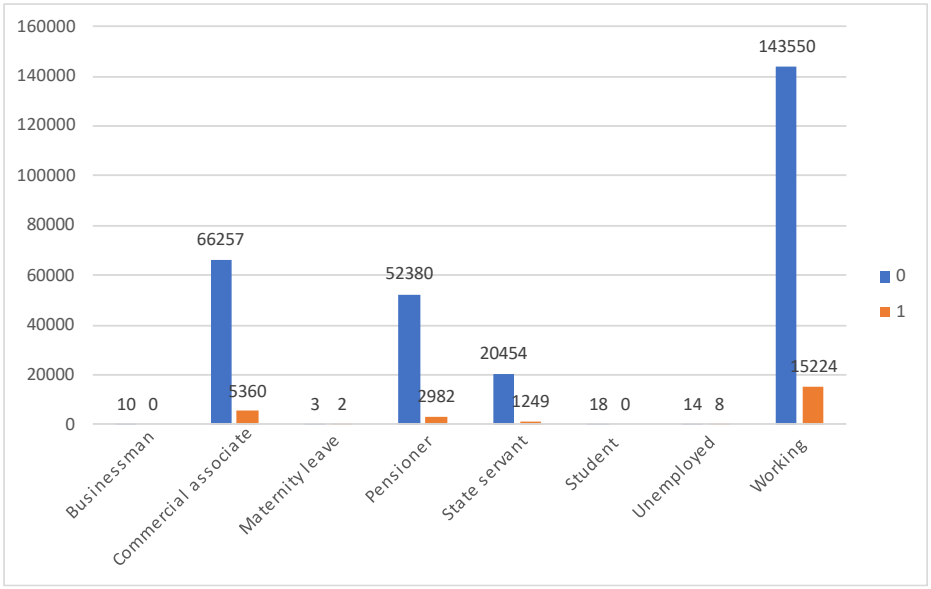
5.EDA

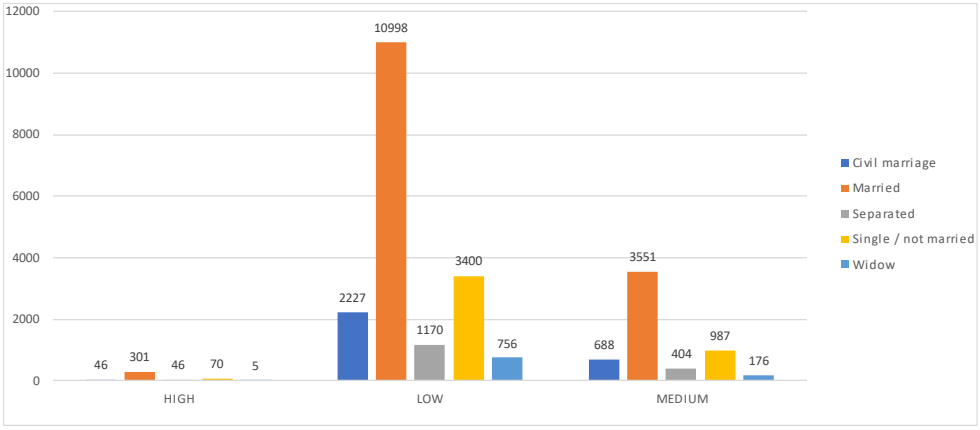
Univariate Analysis:

Customers have largely chosen cash and consumer loans. The majority of our clients are repeat customers.

The majority of current loan applicants are individuals who applied for loans less than ten months ago. More loans have been requested for consumer gadgets.





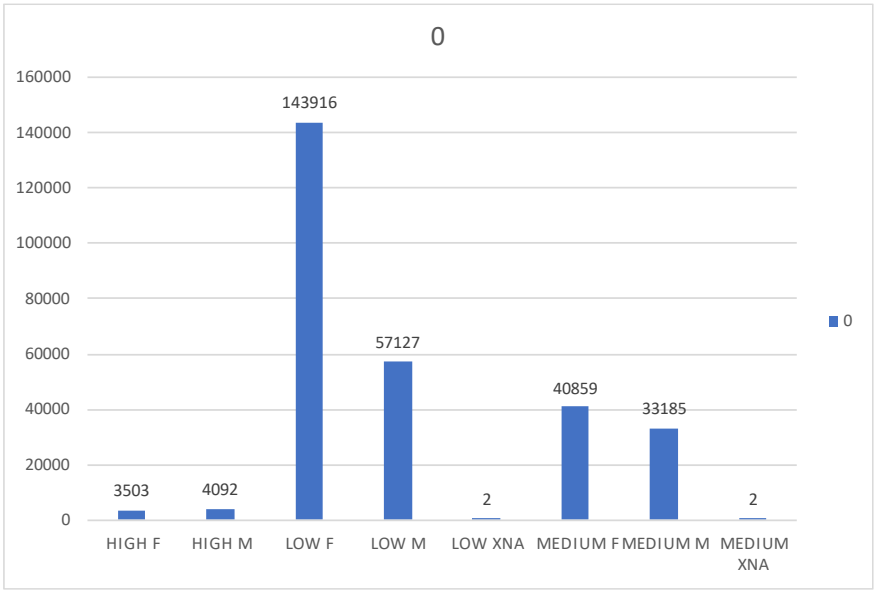


Bivariate Analysis:

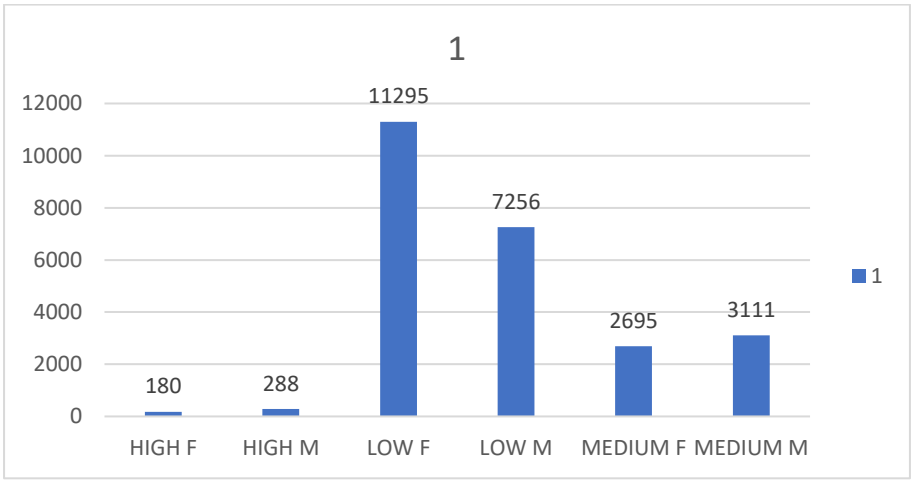
Inference

1. Customers who applied for more than Rs. 350,000 will most likely be denied. The majority of loans sought for through Credit and Cash agencies are cancelled.
2. New clients are overjoyed because the majority of their loans were approved. Thus far, car loans have been denied. Loans made to MLM partner clients are likely to be cancelled. Virtually 80% of the loans were authorised, with a steady stream of rejections.
3. Consumer loans have nearly no cancellations and the greatest approval rate. Several loans for the first Selling place area group were cancelled.
4. Clients who apply for another loan within 10 months of their previous loan are more likely to have it cancelled. Walk-in loans have a higher refusal rate.

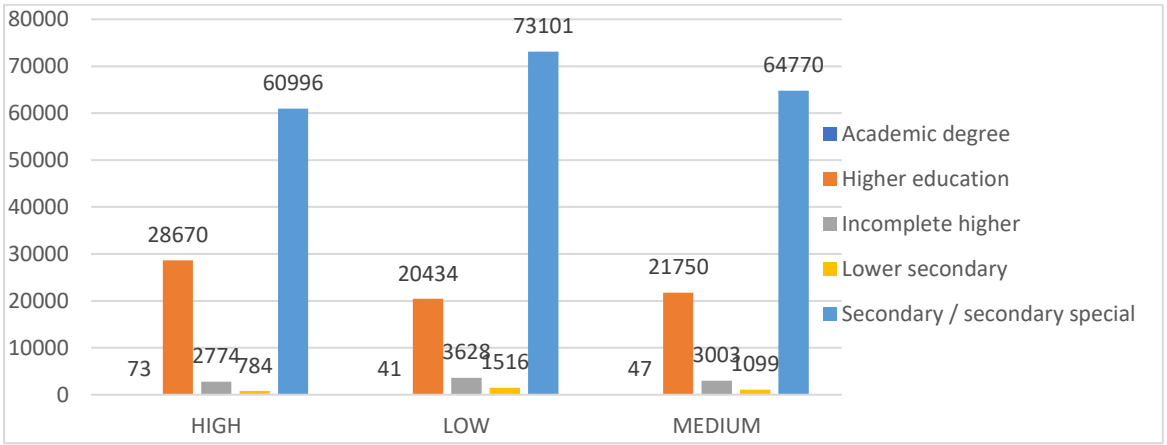
Target 0: Total_income_range vs Code_gender:



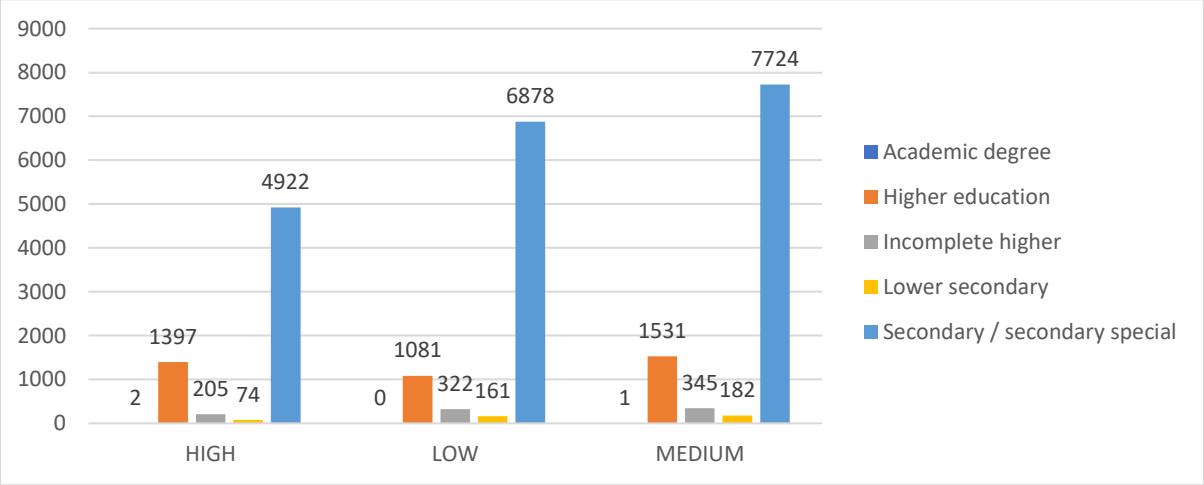
Target 1: Total_income_range vs Code_gender:



Target 0: Credit Amt vs Education status:



Target 1: Credit Amt vs Education status:



Conclusion:

In conclusion, the bank loan case study highlights the importance of robust risk assessment methodologies and stringent eligibility criteria in mitigating loan defaults. By leveraging advanced analytics banks can enhance decision-making processes, minimize risks, and improve overall performance.

Analyzing the Impact of Car Features on Price and Profitability

Description:

The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars.

This problem could be approached by analyzing the relationship between a car's features, market category, and pricing, and identifying which features and categories are most popular among consumers and most profitable for the manufacturer. By using data analysis techniques such as regression analysis and market segmentation, the manufacturer could develop a pricing strategy that balances consumer demand with profitability, and identify which product features to focus on in future product development efforts. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

Problem:

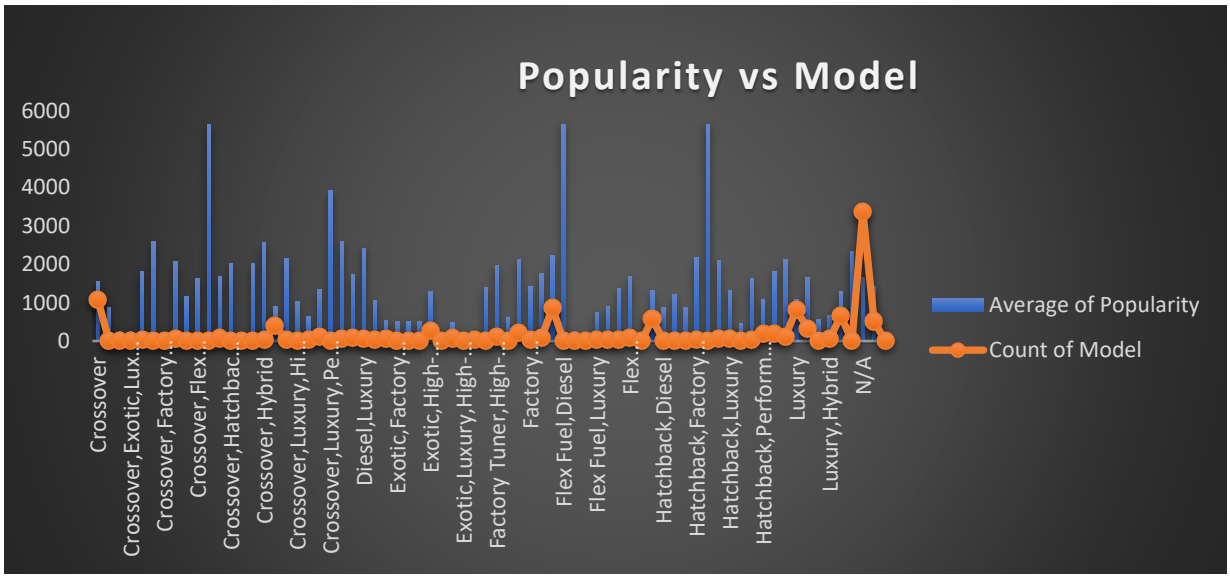
Investigating the relationship between a car's features and its popularity: By examining the popularity variable in the dataset, a data analyst could identify which features are most popular among consumers and how they affect a car's popularity. This could help manufacturers make informed decisions about product development and marketing.

Predicting the price of a car based on its features and market category: By using the various features and market category variables in the dataset, a data analyst could develop a model to predict the price of a car. This could help manufacturers and consumers understand how different features affect the price of a car and make informed decisions about pricing and purchasing.

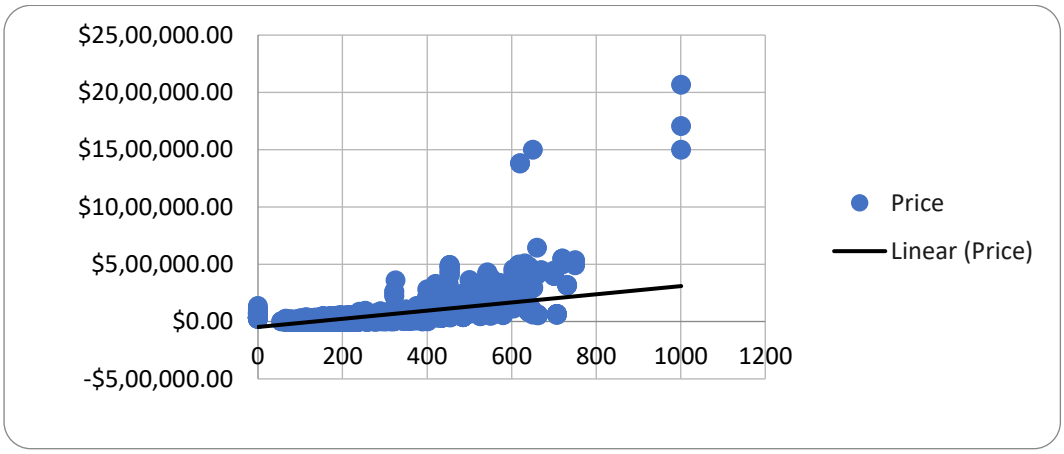
Overall, this dataset could be a valuable resource for data analysts interested in exploring various aspects of the automotive industry and could provide insights that could inform decisions related to product development, marketing, and pricing

Findings:

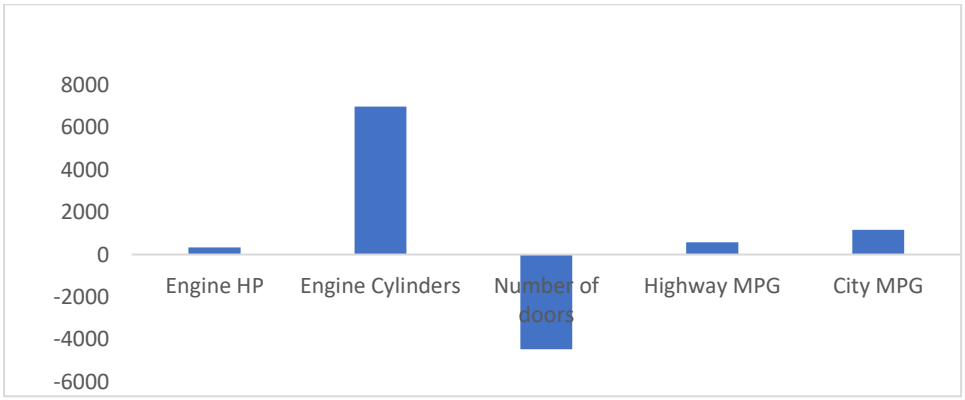
1.How Popularity of a car model vary across different market categories?



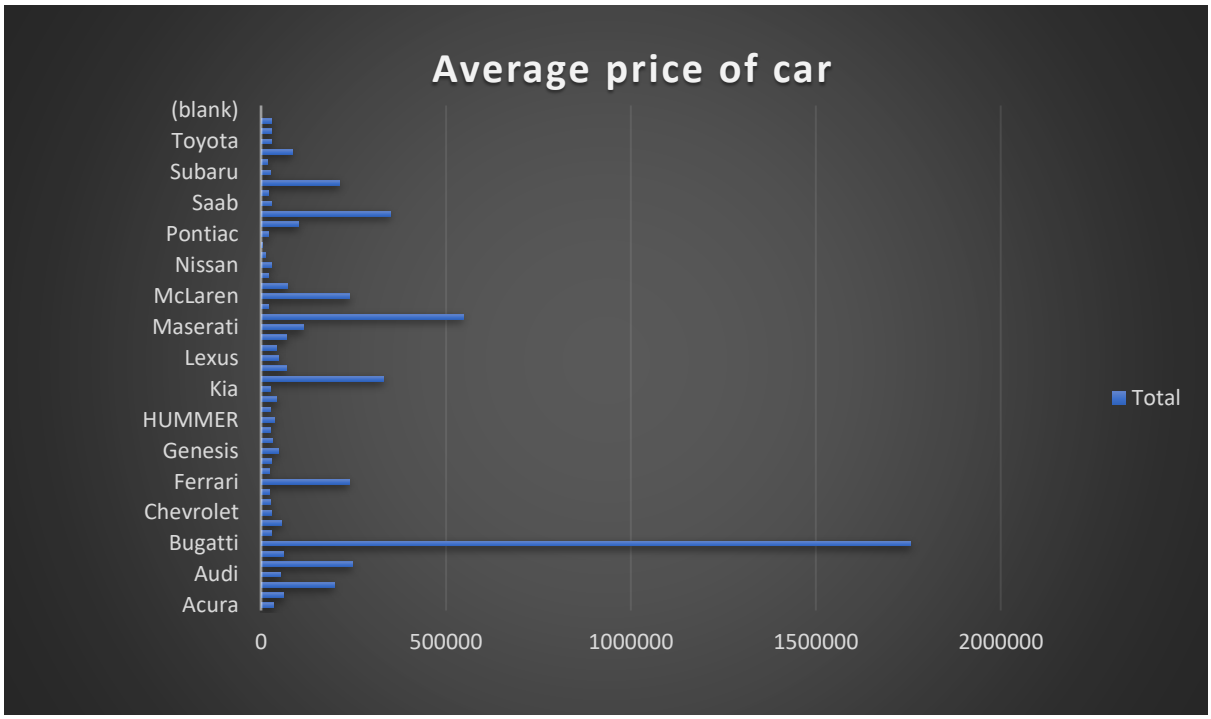
2. What is the relationship between a car's engine power and its price?



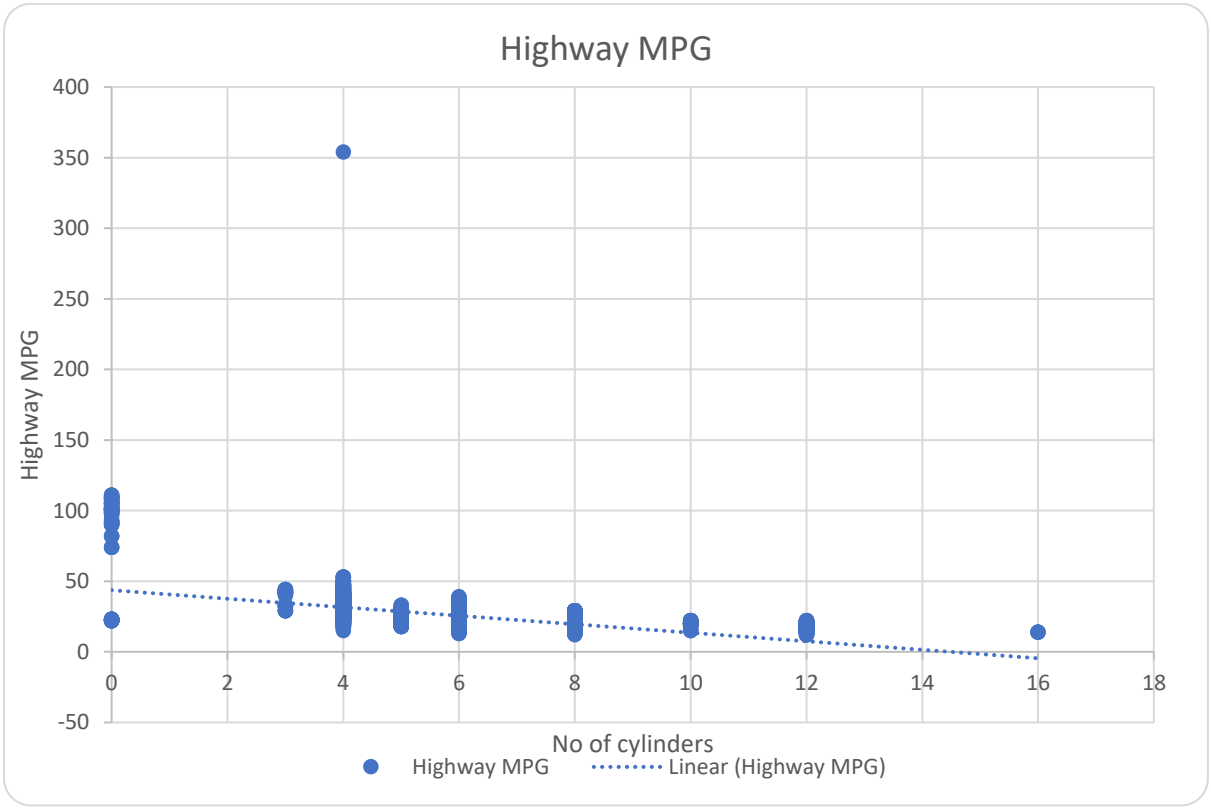
3. Which car features are most important in determining a car's price?



4. How does the average price of a car vary across different manufacturers?

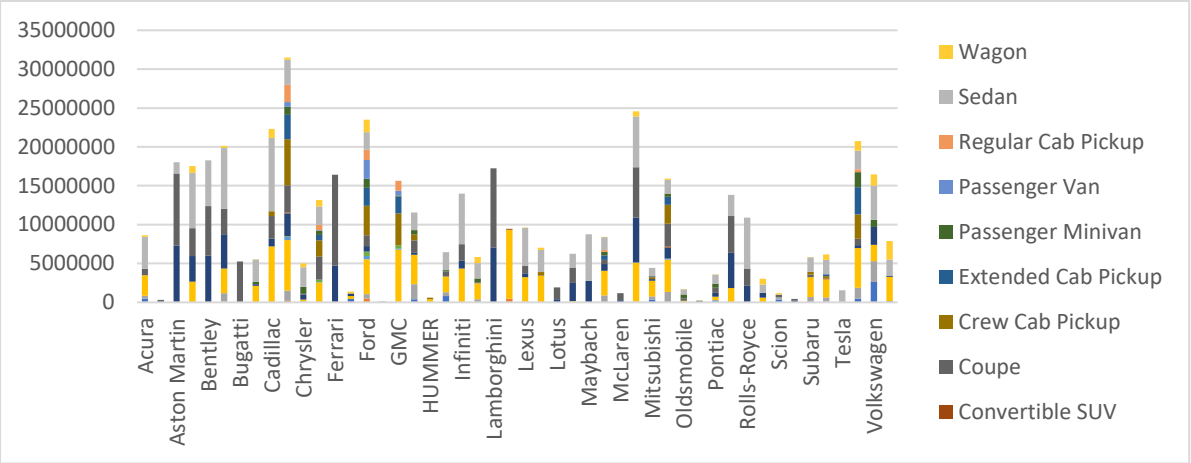


5.What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

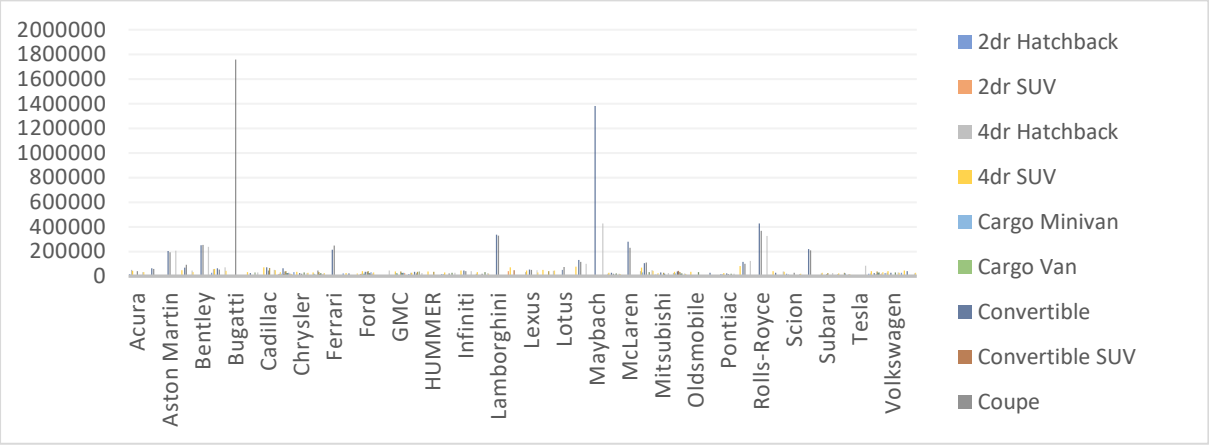


Building the Dashboard:

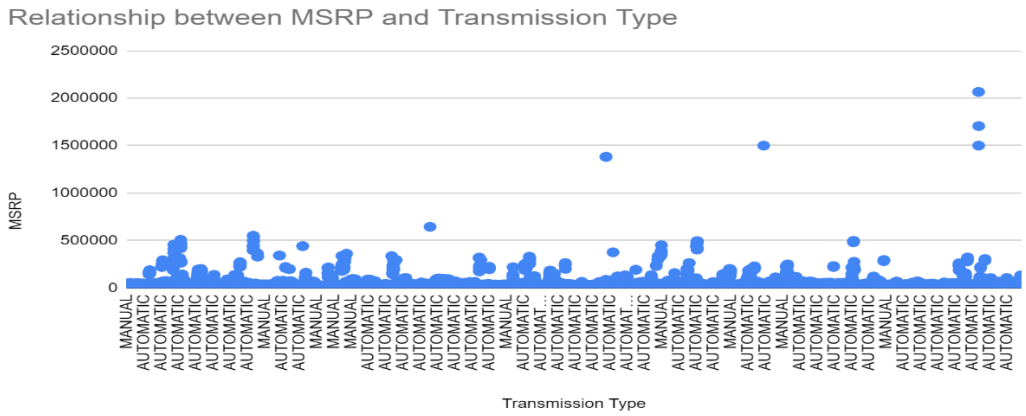
1. How does the distribution of car prices vary by brand and body style?



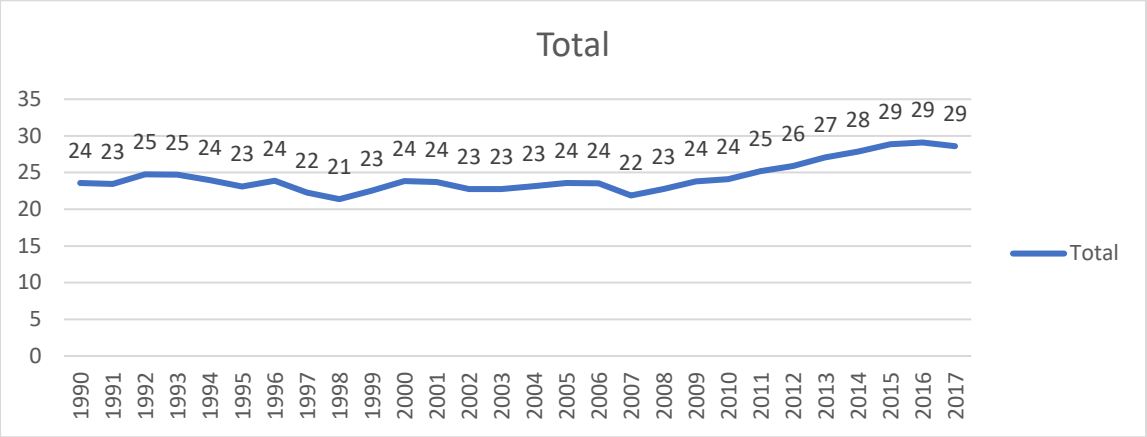
2.Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?



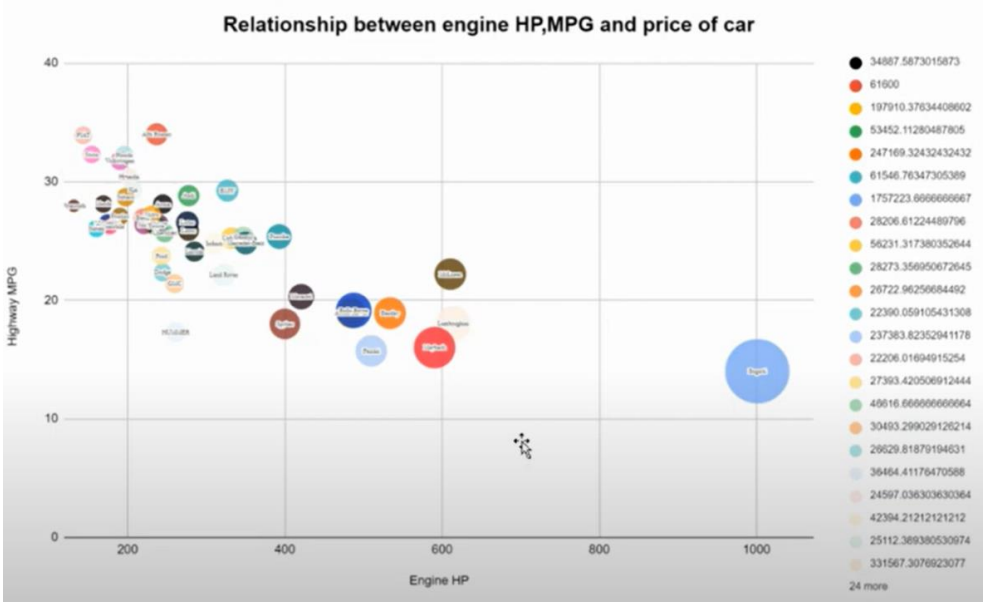
3.How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?



4.How does the fuel efficiency of cars vary across different body styles and model years?



5. How does the car's horsepower, MPG, and price vary across different Brands?



Conclusion:

The analysis of car features on price and profitability reveals that certain features have a significant impact on both. Advanced safety technologies, luxurious interior amenities, and cutting-edge infotainment systems contribute to higher prices and potentially

higher profitability. Understanding consumer preferences, market trends, and cost-effectiveness is crucial for optimizing pricing and profitability. Ongoing research and development are essential to stay competitive in the ever-evolving automotive industry. Manufacturers must carefully evaluate and prioritize features to enhance customer value and drive business profitability.

ABC Call volume trend analysis

Project Description:

The objective of the ABC Call Volume Trend Analysis project is to analyse the trends in call volume received by the ABC company over a specific period. Understanding call volume trends can provide valuable insights into customer behaviour, operational efficiency, and resource allocation. By identifying patterns and fluctuations in call volume, ABC can optimize its customer service strategies, staffing levels, and infrastructure to meet customer needs effectively.

Design:

Steps taken to clean the data:

First importing the datasets provided and merging them. Then removing the duplicates and the blank cells. Improving the headers of each column with proper values. Finding and replacing the data with correct values.

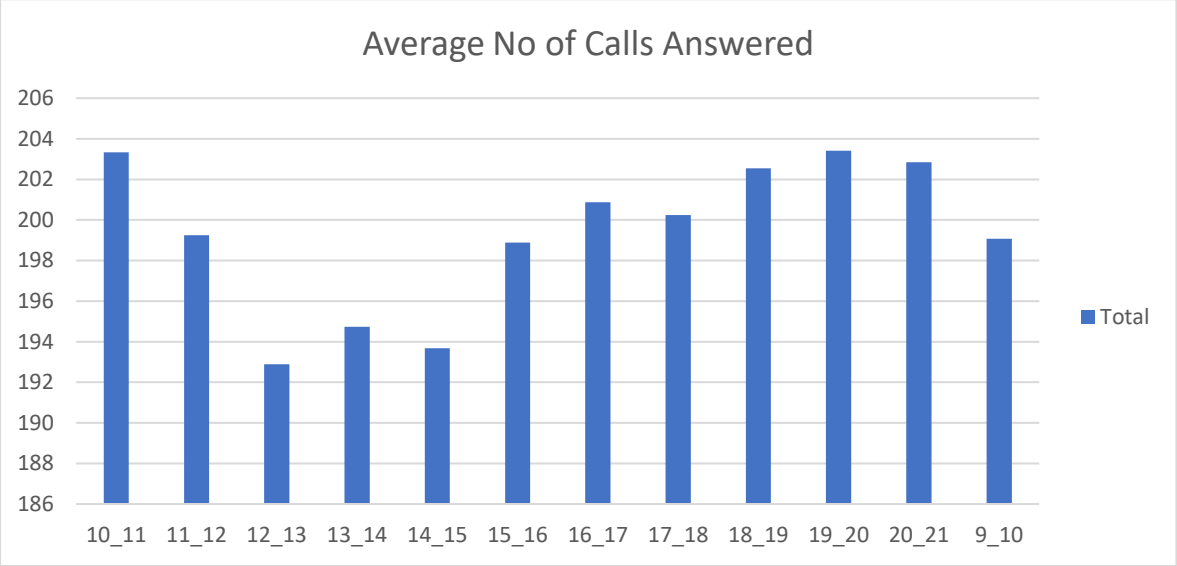
Tools used for visualization: Excel

Tech Stack Used

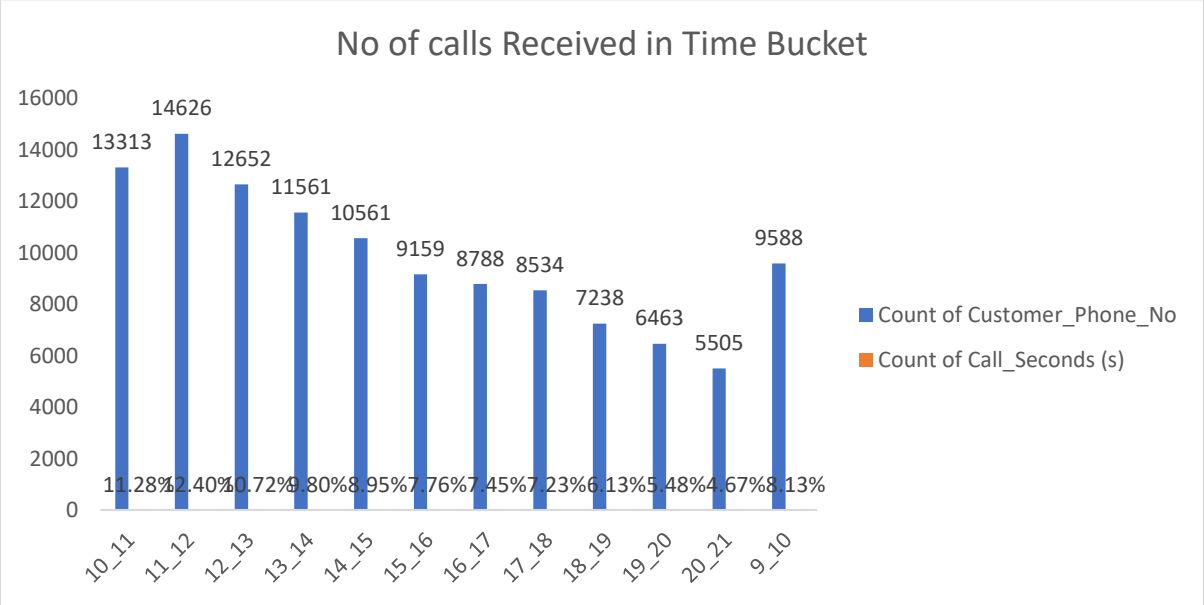
- Microsoft Excel

Findings:

1.What is the average duration of calls for each time bucket?



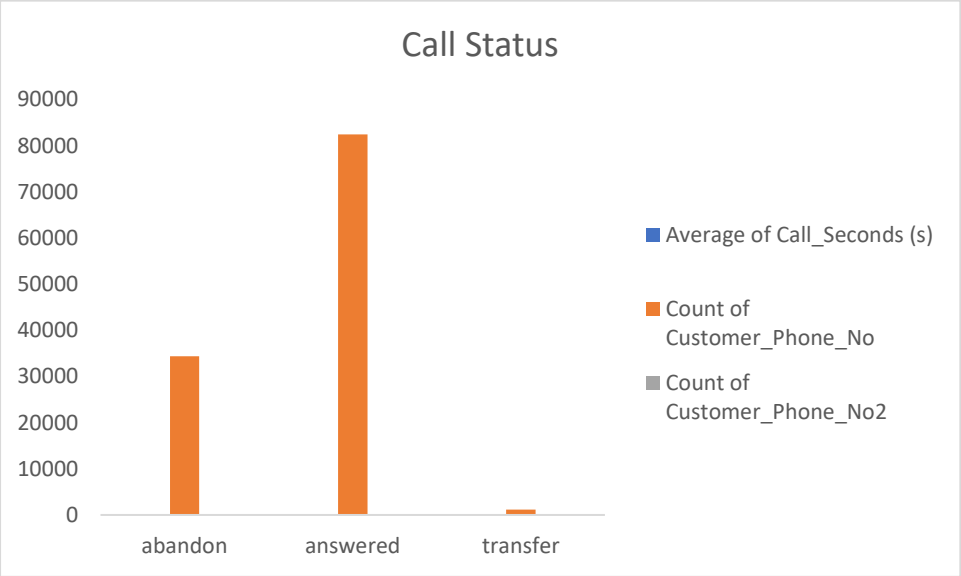
2. Can you create a chart or graph that shows the number of calls received in each time bucket?



3.What is the minimum number of agents required in each time bucket to reduce the abandon rate to 10%?

1-Jan sum of all call in second	**676664**
sum of hours	**187.96**
Total agent of 60%	**37.6**

Agent required for 90%	**57**



4.Propose a manpower plan for each time bucket throughout the day, keeping the maximum abandon rate at 10%.

Average No of calls in the Night is	1539
Increase call rate to 90% in Night	76
No of agent Needed in Night	17

No. of agents required in **9am-9pm** to attain 90%	57
No. of agents required in **9pm-9am** to attain 90%	17

No. of agents required in **9am-9pm** to attain 90%	57
No. of agents required in **9pm-9am** to attain 90%	17
Overall Total agents required is	74

Conclusion:

1. Throughout this project, I have gained valuable insights into the impact of an analyst in the customer service department. It is evident

that a company strives to ensure maximum customer satisfaction through effective customer handling strategies.

2. One of the notable tools used is the Interactive Voice Response (IVR) system, which employs Al technology to address customer queries by identifying their specific concerns and routing the calls to the appropriate agents for resolution.

3. The analysis of the provided data was made easier by the pre-calculated time buckets and call duration converted into seconds, saving time and effort in calculations.

4. Additionally, I have delved into the real in of behavioural analytics, which involves studying customer behaviour patterns to identify trends, preferences, and opportunities for enhancing the overall customer experience.

5. Overall, this project has provided me with valuable knowledge and insights into the dynamics of customer service and the role of an analyst in optimizing customer satisfaction

Appendix

1.Instagram User Analytics

https://drive.google.com/file/d/11PqtlAn2mk0F6s3NEf1bjkg_d7cbyy0f/view?usp=sharing

2. Operation & Metric Analytics

https://drive.google.com/file/d/1SSeQUeqFqSY1Jq7CijjHcgdt34KSq42D/view?usp=sharing

3. Hiring Process Analytics

https://docs.google.com/spreadsheets/d/1n9B8ftIISJ18QQr1aSosh3v9PZD1B9ha/edit?usp=sharing&ouid=106742878432922835995&rtpof=true&sd=true

4. IMDB Movie Analysis

https://docs.google.com/spreadsheets/d/1RXzfBN4BbikNtuTlBwNIMXphNobtFJ3u/edit?usp=sharing&ouid=106742878432922835995&rtpof=true&sd=true

5. Bank loan case study

https://docs.google.com/spreadsheets/d/1T1JAEfHxgDYRlPE3JI5EOZ3Z_wkKVVmZ/edit?usp=sharing&ouid=106742878432922835995&rtpof=true&sd=true

6. Analyzing the Impact of Car Features on Price and Profitability

https://docs.google.com/spreadsheets/d/1iOA7JfaExcfsCYlexZxgvYRHZ6otrJqk/edit?usp=sharing&ouid=106742878432922835995&rtpof=true&sd=true

7. ABC Call volume trend analysis

https://docs.google.com/spreadsheets/d/1HvX2RfXUgEY2XUfvSX_e9_B6gfTmF8B/edit?usp=sharing&ouid=106742878432922835995&rtpof=true&sd=true