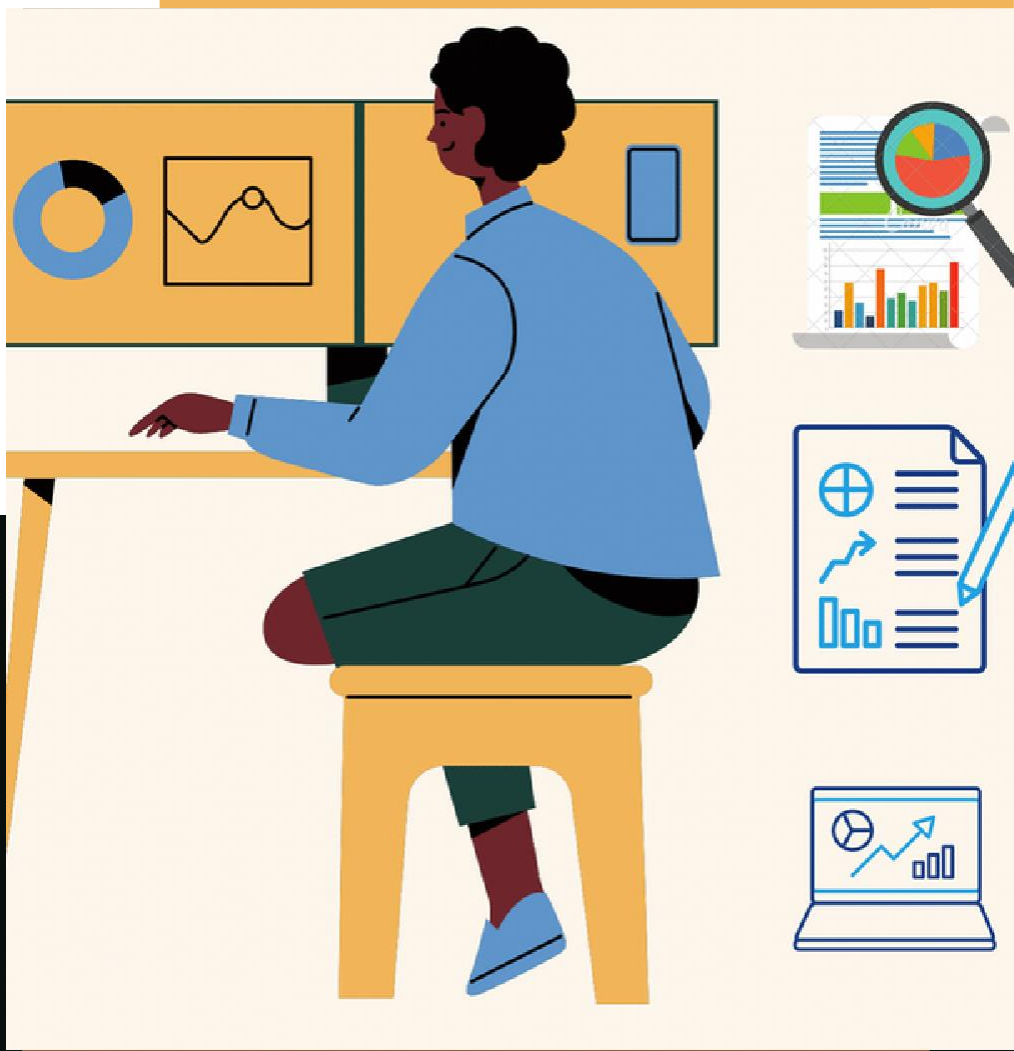
DATA ANALYTICS

PORTFOLIO PROJECT

PREPARED BY SAMI GAWANDI



CONTACT:

LinkedIn

GitHub

Email Id: samigawandi07@gmail.com

PROFESSIONAL BACKGROUND

ABOUT ME

Detail-oriented IT student who excels at maintaining large databases and performing data collection and data analysis and conducting detailed reports. Possess great analytical skills, excellent multitasking skills, strong attention to detail, and significant ability to work in a team.

EDUCATION

I am pursuing a Bachelor of Technology degree in Information Technology from Mumbai University with a 9.22 CGPA.

I have done a Data Analytics Trainee certification course from TRAINITY for 2 months and got familiar with different real-life projects.

HARD SKILLS

- Structured Query Language (SQL)
- MS Excel
- Basic analysis with Python

SOFT SKILLS

- Storytelling
- Communication Skills
- Critical Thinking
- Teamwork
- Research
- Presentation and Report

CERTIFICATIONS:

Data Analytics Training:
https://drive.google.com/file/d/1DYt1uhox_lwvSRaAruC8uYs1vWM7Rwoe/view?usp=share_link

Data Analytics Live Project:
https://drive.google.com/file/d/1kPRwDrPy66HHOxh2t2zEfWy3lwLJU2sN/view?usp=share_link

Data Analytics Virtual Internship:
https://drive.google.com/file/d/16XaUPl_YA0FiffdzLoeVDUO9i7zGYcxs/view?usp=share_link

TABLE OF CONTENT

Data Analytics Process

Description: We use Data Analysis in everyday life for e.g**Visiting mobile store for purchasing a mobile.**

- **Plan -**

Initially you need to decide which company model (Mobile) you are going to buy from that store may be it of Redmi, Samsung, Oneplus, Vivo, etc. as per your choice you will select one of the brand to go with.

- **Prepare –**

Later you will decide your budget range i.e. how much money you are going to give for that specific model. As per your budget range, you will select a mobile from that store.

- **Process –**

As in Mobile store there can be varieties of mobiles of different production companies(brand) if suppose you decided to purchase a Xiaomi mobile phone, there are many models manufactured by Xiaomi company specified my names of different model number according to their specifications such Redmi note 9, Redmi note 9 pro, Redmi note 9 pro max, Redmi note 10 pro, Redmi note 10 pro max and so on. So from this options you had to select one of them.

- **Analyse –**

If suppose you are using a Xiaomi Redmi brand phone since past 5 years and if it is convenient to you, then your attraction towards Redmi mobile phones will more as compared to others. Surely, you will purchase latest mobile phone according to purpose of use for e.g. If you had to buy mobile for gaming purpose, you will sure filter only those mobiles which have long lasting battery life, better display, frames per second(fps) above or equal to 60, good sound quality, etc. and check does this type of mobile fits my budget range or not. Some people may buy mobile phones with good camera and video quality for vlogging purpose. Overall decision making depends on specifications of device and purpose of use. The same process might be applied during purchasing any laptop. People feedbacks and reviews on that selected device might be taken into consideration by you which can help you to know the props and cons of that device. And you will ensure that the brand (selected mobile phone company) will provide satisfaction to their customers by their service.

- **Communication –**

Lastly, you will communicate with the shopkeeper about your selected device which fits in your budget as well as fulfil you purpose of buying and will also discuss about accessories getting along with device.

- **Act –** Then you finally buy it.

ProjectLink-

https://drive.google.com/file/d/1Z6nHLOa1hULfrYwTC4f55NMT_1b_YBYr/view?usp=share_link

Instagram User Analytics

Description:

In this project, we have performed Instagram User Analytics. User analytics is a process for evaluating data that depicts how users adopt, throw themselves into, and experience products and services. The project is all about tracking user activity like how a user interacts, manages their account, and interconnects with the posts updated by other users. Then this gathered data is used by the product team and product manager across the business to decide on several features for an app, track some success factors by estimating the data extracted by performing user analytics using SQL (Structured Query Language) and also improve the user experience for better user feedback which leads in the growth of a business.

Findings:

- **Rewarding Most Loyal Users:** The 5 oldest users of Instagram
- **Remind Inactive Users to Start Posting:** The users who have never posted a single photo on Instagram
- **Declaring Contest Winner:** The user who gets the most likes on a single photo
- **Hashtag Researching:** The top 5 most commonly used hashtags
- **Launch AD Campaign:** Day of the week do most users register on
- **User Engagement:** Average user posts on Instagram
- **Bots & Fake Accounts:** Users (bots) who have liked every single photo on the site

Approach: The report is all about user interaction with the Instagram app, the extracted data of the users sure going to help them as well as a management team to improve and add features to the app. By using a simple approach, with the help of SQL queries the data needed for the management team is smoothly drawn out and the output is shown in the result section below.

Insights:

- Initially, we rewarded the top 5 oldest users (Darby_Herzog, Emelio_Berneir52, Elenor88, Nicole71, Jordyn.Jacobson2) who were using Instagram for the longest time as the most loyal users. These users were constantly in touch with the app since over-long time.
- If a user had registered on the Instagram handle for a long time but never posted a single photo on the Instagram handle, I have enlisted some users from the dataset who have never posted a single photo on Instagram for sending them an email to remind them to post their 1st photo on the platform. This can help the marketing team to catch up the attention of the users towards the app after receiving mail from the Instagram team, he/she gets ready to post a photo on the Instagram platform.
- Then I have taken out the data of the user who gets the highest number of likes on a single photo and declared him/her as a winner of the contest. David.Osinski47 with the most likes (48) on a single photo is the winner of the contest. This narrates that the user (contest winner) has followers in excellent numbers and his/her interaction with the Instagram handle is often

good.

- After recognizing the top 5 most commonly used hashtags (#smile, #beach, #party, #fun, #concert) by users, I get to know users are using these tags in large numbers to get likes on their posts in huge numbers as well to increase their follower too.
- Identified most user registers on Thursday and Sunday as compared to other days of the week these measures helped the marketing team to promote advertisements of brands or to schedule any campaign on these days as most of the user remains active on these two days of the week.
- The as large number of users were able to discern the promoted ads on these days and the marketing team will focus more on these couple of days for scheduling the campaign and promoting advertisements to gain users' attraction in massive numbers. The above data is more useful to the marketing team to track how users engage and interact with their software, product, or application in an attempt to improve their product, bring more users in, improve user engagement with their product, and the general success of their application.
- Also provided data to Investors concerning the total number of users registered and how many times does average user posts on Instagram (2.57). In addition, the report also contains data about fake and dummy accounts whether an account is handled by a real person or a bot.
- These data might help the investors to judge whether their financing in Instagram is profitable or not and gave them a glimpse of users' collaboration on the platform

Project Link:
https://drive.google.com/file/d/1jB5hD9rSlnpGCEtb8lbAxg_a4ItETtH0/view?usp=share_link

Operation Analytics & Investigating Metric Spikes

Description:

In this project, we have performed operation analytics and investigated metric spike which is used to perform end-to-end operations for company growth and which areas to improve on. This kind of analysis is further used to predict the overall growth or decline of a company's fortune. It means better automation, better understanding between cross-functional teams, and more effective workflows. Investigating metric spikes is also an important part of operation analytics to understand or make other teams understand questions like- Why is there a dip in daily engagement? Why have sales taken a dip? Etc. Questions like these must be answered daily and for that, it is very important to investigate metric spikes. This analysis helps the company to look into the ongoing business's growth and downfall too.

Findings:

In this project, I have found out the –

1. The number of jobs reviewed per hour per day for November 2020.
2. 7-day rolling average of throughput.
3. The Percentage share of each language in the last 30 days.
4. Rows that have the same value present in them.
5. The weekly user engagement.
6. The user growth for the product.
7. The weekly retention of users-sign up cohort.
8. The weekly engagement per device.
9. The email engagement metrics. From the provided dataset.

Approach:

1. Initially I downloaded the provided datasets to the device. Both datasets were in the form of CSV (Comma Separate Value) format.
2. The datasets are imported into MySQL Workbench 8.0.31 for performing queries.
3. At times, I have used CMD (Command Prompt) for performing operations on the databases.
4. After execution of proper SQL queries the snapshot of the answers to the questions is given in the result section before being properly reviewed.

Insights:

First of all, I have studied the provided datasets and their columns before operations, which helped me to understand what exactly the dataset is about and what information it holds.

Case Study 1 (Job Data)

- From the tables, I discovered the number of jobs reviewed per hour per day for November 2020 i.e. **0.0083** jobs got reviewed in the month of November 2020 per hour per day.
- Then calculated 7-day rolling average of throughput (between 25-11-2020 and 30-11-2020) is displayed in the result section.
- And evaluated the percentage share of each language in the last 30 days in which the Persian language has the highest amount of contribution i.e. **50.00%** meanwhile English, Italian, and French have **16.67%** each.
- Lastly displayed duplicate rows (**job_ id 23**) from the table dataset.

Case Study 2 (Investigating Metrics Spikes)

- Subsequently, investigating metrics spike is carried out where at first calculated the weekly user engagement is about **2774** and the number of users growing over time for a product is evaluated (**9381**).
- Then calculated the weekly retention of users sign-up which results in **0** and the weekly engagement of users per device where MacBook pro sits at the top with **484** users.
- Lastly, weekly emails(**15688**), re-engagement emails(**0**), emails opens(**4692**) and email clickthrough (**1662**) have also been evaluated.

Project Link:

https://drive.google.com/file/d/141vQLpJAvhA5f-bjTzOtqleSMkHSVkZX/view?usp=share_link

Hiring Process Analysis

Description:

The hiring process refers to finding, selecting, and hiring new employees for a company. Simply, it is a procedure of placing employees according to posts available in the company in exchange for money(salary). Here, the HR leader aims to bring the best-skilled employees to the company to meet its business objectives. The hiring process is the fundamental and most important function of a company. Here, the MNCs get to know about the major underlying trends in the hiring process. Trends such as the number of rejections, number of interviews, types of jobs, vacancies, etc. are important for a company to analyze before hiring freshers or any other individual.

Findings:

From the given dataset I have answered some of the following questions using statistics and different formulas in MS Excel:

- **Hiring:** How many males and females are Hired?
- **Average Salary:** What is the average salary offered in this company?
- **Class Intervals:** Draw the class intervals for salary in the company?
- **Charts and Plots:** Draw Pie Chart / Bar Graph (or any other graph) to show proportion of people working different department?
- **Charts:** Represent different post tiers using chart/graph?

Approach:

I am working for a MNC such as Google as a lead Data Analyst and the company has provided with the data records of their previous hirings and has asked me to answer certain questions making sense of that data.
We will use EDA to generate different insights and to answer the questions asked by the company.
The dataset given by the company contains details about people who registered for a particular post in a department of this company. I used MS Excel to analyze the data with different tables and columns.

Insights:

For solving the asked questions I have used the EDA (Exploratory Data Analysis) approach to represent the summaries of data in a graphical manner. After resolving the asked questions I discovered:

1. Males working in the company i.e. 2563 are more as compared to Females i.e. 1856 in numbers meanwhile rejection rate of males is also 30% higher than females.
2. The average salary given in the company is Rs. 49,983 (approx. 50k) whereas the General Management department has the highest average salary in the company Rs. 58,722.
3. The maximum number of people (414 people) receive salaries ranging between 40000-49999 whereas only 3 people may be managers or someone in high posts receive salaries between 2Lakhs and 4Lakhs.
4. 39.24% of employees of the company are working in the Operations department and the Human Resource department has the least employees.
5. Maximum employees are hired on post c9 following post c5 and the company has shown interest in hiring applicants on both posts and none of the employees are hired on the n9, n10, and m7 posts.
6. Operation department has the highest rejection rate at 39% and the HR department with 1% of rejection rate.

Project Link:

https://drive.google.com/file/d/1OhTC_6uLM8AXMe4n4O9HGiDtDTA1PGZ8/view?usp=share_link

IMDB Movie Analysis

Description:

In this project, I have performed analysis on the IMDb movies dataset and displayed meaningful information from the dataset through MS Excel. The **Internet Movie Database (IMDb)** is a website that serves as an online database of world cinema containing a large number of public data on films such as the title of the film, the year of release of the film, the genre of the film, the audience, budget, revenue, the rating of critics, the duration of the film, the summary of the film, actors, directors and much more

Findings:

We are required to provide a detailed report for the below data record mentioning the answers to the questions that follow:

A. **Movies with the highest profit:** Find the movies with the highest profit.
B. **Top 250:** Find IMDB Top 250
C. **Best Directors:** Find the best directors
D. **Popular Genres:** Find popular genres
E. **Charts:** Find the critic-favorite and audience-favorite actors

Approach:

Firstly, after downloading the provided dataset I used MS Excel, understood each column's data and what exactly they mean, and used mostly pivot tables meanwhile some data can be extracted through statistics and different formulas to lay out required data to the company.

Insights:

- There South Korean film "*The Host*" directed by **Joon-ho Bong** in **2006** is the lowest-profit movie.
- "*The Avengers*" is the highest-grossing film with the highest profit of approx. 806 million whose director is Joss Whedon.
- The Avatar, Jurassic World, Titanic, Star Wars: Episode IV – A New Hope, The Lion King, etc. like English language films mostly dominated the film industry over the world.
- According to IMDb ratings, "*The Shawshank Redemption"* movie is the top-rated movie with an IMDb score of 9.3.
- Other than English films, the French movie "*Amelie*" with 8.4 IMDb ratings has the highest votes by users i.e. 534264 votes. And the Italian movie "*The Good, The Bad and The Ugly*" is the highest IMDb-rated movie with a rating of **8.9.**

- The director of the "*Modern Times*" movie named **Charles Champlin** and the director of the **"***American History X***"** movie **Tony Clave** have the highest mean IMDb score.
- **Crime|Drama|Fantasy|Mystery** and **Adventure|Animation|Drama|Family|Musical** types of genres are the most rated genres by users.
- People love such types of genres such as "*The Lion King"* movie which is an animated movie and "*Bahubali: The Beginning"* movie which is the **Action+Adventure** genre.
- Over the globe, audiences have shown lots of interest in Science-Fiction (Sci-Fi) movies where "*The Avengers", "The Avatar", and "Jurassic World"* are among them
- Throughout the entire career of Leonardo DiCaprio, Meryl Streep, and Brad Pitt,

Actors	Total Movies	Total User reviews	Total Critic reviews
Leonardo DiCaprio	21	19204	6934
Brad Pitt	17	12620	4165
Meryl Streep	11	3269	1996

- **Heather Donahue** has been named the audience's favorite actor and **Albert Finney** is the critic's favorite actor.
- From the past 10 decades, **2010**s film has the highest number of users' votes (approx. 108 million) whereas from 1921 users' votes seem to be increased over decades.

Project Link:
https://drive.google.com/file/d/1OhTC_6uLM8AXMe4n4O9HGiDtDTA1PGZ8/view?usp=share_link

Bank Loan Case Study

Description:

In this project, we have performed Exploratory Data Analysis (EDA) on Bank Loan Dataset. This case study aims to give you an idea of applying EDA in a real business scenario. The loan-providing companies find it hard to give loans to people due to their insufficient or non-existent credit history. Because of that, some consumers use it to their advantage by becoming defaulters.

Business Understanding:
The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it as their advantage by becoming a defaulter.
Suppose we work for a consumer finance company which specialises in lending various types of loans to urban customers. You have to use EDA to analyse the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected.

Findings:

Our aim is to identify the patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.
The driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.
Presenting the overall approach of the data analysis, cleaning the dataset, finding outliers, data imbalance, univariate, segmented univariate, bivariate analysis, etc.
The top 10 correlation for the Client with payment difficulties and all other cases (Target variable)

Approach:

1. Used MS Excel to understand what the dataset is about and its columns' meaning, which is important.
2. Then imported the required python libraries (Pandas, Numpy, Seaborn, etc.) into Jupyter notebook.
3. After libraries, datasets (Application_Data and Previous_Application) are imported.
4. Identified missing values and removed columns which are not useful for analysis work.
5. Identified outliers of different columns and their relations.
6. Identified data imbalance and its ratio and presented using the graphs.

7. Performed univariate, segmented univariate, and bivariate analysis of data.
8. Found out the top 10 correlations with respect to TARGET variables.
9. Lastly presented all the senseful data (analysis work) in the form of different graphs and charts.

Insights:

1. Females were found to be more defaulters as compared to males.
2. Out of 100 only 16% of clients have applied on weekend days(Saturday & Sunday) for current loans meanwhile on TUESDAY most of the clients about 17% have shown interest in applying.
3. Clients living in apartments and houses have the most defaulters and it's not safe side to approve such clients' loan applications.
4. 18% of clients have applied for the very first time whereas almost 73% of clients have applied for loans again.
5. After XAP, the common reason behind the rejection of a loan is HC.
6. In Previous applications, about 41% of clients have applied for POS followed by Cash with 22%.
7. Other than XAP and XNA, for repairing purposes most of the defaulters have taken advantage of the loan amount
8. Buying a used car is also a major reason for applying for a loan after urgent needs.
9. Company should think before approving loan applications for such reasons.
10. People living in office apartments is having higher credit for defaulting and people in co-op apartments have repaid the loan despite taking huge amounts as a loan.
11. Municipal apartments also have huge bars in not repaying the loan.
12. Defaulters use the unused offer for their benefit whereas only more than the half-approved loans are repaid.
13. New clients have returned their loan payments but there are still defaulters more than repayers.
14. Whereas clients applying for loans again has most counts in terms of not paying loan status.
15. Company should pay attention on client whose previuos application was for POS, Cash,etc beacuse they are also contributing to the defaulters counts.

Project Link:
https://drive.google.com/file/d/1IeSj2_8SHxE1B80zo5RmaSnaf9wBB9Lz/view?usp=share_link

XYZ Ads Airing Report Analysis

Description:

In this project, we have analyzed the meaningful data from the XYZ Ads dataset. Advertising is a way of marketing your business in order to increase sales or make your audience aware of your products or services. Until a customer deals with you directly and buys your products or services, your advertising may help form their first impressions of your business. The target audience for businesses could be local, regional, national, international, or a mixture. So they use different ways for advertisement. Some of the types of advertisements are Internet/online directories, Trade and technical press, Radio, Cinema, Outdoor advertising, National papers, magazines, and TV.

Findings:

1. What is Pod Position? Does the Pod position number affect the amount spent on Ads for a specific period of time by a company?
2. What is the share of various brands in TV airings and how has it changed from Q1 to Q4 in 2021?
3. Conduct a competitive analysis for the brands and define advertisement strategy of different brands and how it differs across the brands.
4. Mahindra and Mahindra wants to run a digital ad campaign to complement its existing TV ads in Q1 of 2022. Based on the data from 2021, suggest a media plan to the CMO of Mahindra and Mahindra. Which audience should they target?

Approach:

Different charts are used to represent the meaningful insights from the raw data. Almost for all analysis work, we have used bar charts and column charts and for summarization, a technique of Microsoft Excel known as Pivot table is also used.

Insights:

- The amount spent on Ads by Brands is directly proportional to pod position as pod position increases amount spent also increases.
- Brands spent more money to display their Ads at the top than others.
- Honda Cars, Maruti Suzuki, Toyota, and Mahindra and Mahindra almost follow the same trend (the amount spent decreases with a decrease in pod position number).
- Quarterly, Honda cars spent more money on Ads in Q1 and Q3 and 13% of the Ads for Honda products were shown in Q3.
- Hyundai Motors India has focused more on the start of the year i.e., in Q1.
- Mahindra and Mahindra have spent 25% in the last months of the year 2021 where their Ads are shown were only 13%.
- Maruti Suzuki has the highest percentage of shares and they have focused more on the Q4 months of the year 2021.

- The percentage of shares of Toyota decreases from Q1 to Q4 i.e., 8.64% to 4.89%.
- In the case of Day Parts, most of the brands have invested in Prime time but as per their shares, their Ads were not shown up to that margin.
- Whereas Ads of almost all brands were shown in the Daytime.
- Only Hyundai Motors India has a greater number of Ads in Primetime as compared to others.
- Brand Toyota's numbers look in profit as their spent money over the year and the numbers of Ads of their products are shown below:

	Percentage of shares ($)	Percentage of Ads shown
Q1	5.92%	10.12%
Q2	5.01%	7.62%
Q3	4.72%	8.03%
Q4	12.60%	20.93%

- Most of the Ads were shown through cable network type whereas money spent by brands for broadcast network type was also in high numbers.
- For the digital Ads campaign by Mahindra and Mahindra, Chief Marketing Officer (CMO) should strengthen their Ads strategies at Night time (between 8 pm to 2 am).
- Also, they should target the audience of daytime, as the rate of the number of Ads in the Daytime were often high than the rate of money spent.
- The investment for Primetime seems to be in loss for Mahindra and Mahindra because the rate of the number of Ads shown is half of the rate of the amount spent.

Project Link:

https://drive.google.com/file/d/1N1Kf_OgWMsPJk-anfmUqyCfCb_Inh4R9/view?usp=share_link

ABC Call Volume Trend Analysis

Description:

In this project, we have studied and analyzed the call volume data and drawn some meaningful insights from it as a Customer Experience (CX) team member. A customer experience (CX) team consists of professionals who analyze customer feedback and data and share insights with the rest of the organization. Typically, these teams fulfil various roles and responsibilities such as Customer experience programs (CX programs), Digital customer experience, Design and processes, Internal communications, Voice of the customer (VoC), User experiences, Customer experience management, Journey mapping, Nurturing customer interactions, Customer Success, Customer support, Handling customer data, Learning about the customer journey.

Inbound customer service is the methodology of attracting, engaging, and delighting your customers to turn them into your business' loyal advocates. By solving your customers' problems and helping them achieve success using your product or service, you can delight your customers and turn them into a growth engine for your business.

Findings:

The average call time duration for all incoming calls received by agents (in each Time_Bucket).

1. The total volume/ number of calls coming in via charts/ graphs [Number of calls v/s Time].
2. Propose a manpower plan required during each time bucket [between 9am to 9pm] to reduce the abandon rate to 10%. Propose a manpower plan required during each time bucket in a day[9 pm to 9 am]. Maximum Abandon rate assumption would be same 10%.

Approach:

Initially understanding the columns containing different data is the most important step before analyzing the data. Whereas the dataset is of a Customer Experience (CX) Inbound calling team for 23 days. Data includes **Agent_Name**, **Agent_ID**, **Queue_Time** [duration for which customers have to wait before they get connected to an agent], **Time** [time at which a call was made by a customer in a day], **Time_Bucket** [for easiness we have also provided you with the time bucket], **Duration** [duration for which a customer and executives are on call, **Call_Seconds** [for simplicity we have also converted those time into seconds], **call status** (Abandon, answered, transferred).

We assumed an agent works 20 days a month (6 days a week and 4 days are considered as unplanned leaves in a month) and 7.5 hrs per day whereas only 60% of working hours are given on calls by an agent therefore **(7.5*0.6) = 4.5** hours an agent gives to customer calls.

Insights:

1. The lowest average call answering time is recorded between 12 pm and 1 pm at midday i.e., **193** seconds.
2. Whereas **199** seconds is the average time required to answer a call per day.
3. About 30% of the calls got abandoned daily means those calls are not answered due to the high number of calls and a smaller number of headcounts to answer those calls.
4. The abandoned rate in the early morning between 9 am to 11 am is comparatively larger than the remaining hours approx. 52%.
5. It is identified that the greatest number of calls are made at 11 pm and the least number of calls at 8 pm.
6. On a daily average about **5130 calls** were done by the customers, out of which 3585 calls are **answered** daily, and the remaining them were **abandoned** (1496) and **transferred** (49) daily.
7. At the abandon rate of 10%, the average time required to answer 90% of daily average calls is 255 hours.
8. As an agent spent only 4.5 hours on calls, then to work 255 hours i.e., to answer 90% of daily average calls, about 57 headcounts should be required.
9. In order to reduce the abandon rate from 30% to 10%, about 57 headcounts need to manage or answer 90% of daily average calls.
10. For the night calls management, it is discovered that additionally, 17 headcounts seem to be required.
11. Especially between 9 pm to 12 am and 5 am to 9 am more concentration on incoming calls should be needed.
12. Considering the company has two shifts of working if some employees work in the early morning and if some work till late at night, then the 30% of calls coming at night daily can be managed easily.
13. In the case of the abandoned rate, about 20% (highest abandoned rate of calls as compared to other working hours) of calls were abandoned between 10 am to 11 am.
14. In terms of answering calls, most of the calls were answered between 12 pm and 1 pm.
15. The highest number of calls were transferred between 3 pm to 5 pm, transferred calls are those calls that are made for another purpose.

Project Link:

https://drive.google.com/file/d/1hzJ8dVDixdEM4zn9OWsX--HzFuujprrQ/view?usp=share_link

Conclusion

Learning points:

1. I got to know the data analysis process in day-to-day life.
2. Understood the SQL concept more clearly after solving the problem.
3. The Advance SQL queries are performed in Operation Analysis and Investigating Metrics spike.
4. Got an idea to use SQL in real life case scenario.
5. Also got familiar of python data analysis.
6. Cleaning the data, removing null values columns, imputing null values, correlations, visualization of insights through different graphs using python libraries, etc. also been studied.
7. How to present meaningful insights from data using different graphical representations.
8. Experienced process of HR analytics (Hiring process), Risk Analysis (Bank Loan Case study), and many more.
9. How to use Microsoft Excel tool in real-world case scenarios.
10. Hands-on practice with SQL and MS Excel on real-life problems made me much more knowledgeable of analyzing data.

Thank You