

# DATA ANALYTICS PORTFOLIO



**PREPARED BY:-SURAJ KUMAR**



# **INTRODUCTION**

**COMPLETED MY B.TECH IN ELECTRICAL AND ELECTRONICS ENGINEERING IN 2023. MY CGPA IS 8.24, AND I POSSESS A VARIETY OF ABILITIES SUCH AS WEB DEVELOPMENT, DATA ANALYSIS, PYTHON, AND C++.**

**AS AN INTERN, I WORKED AS A WEB DEVELOPER FOR A NUMBER OF FIRMS, INCLUDING COURSERA, FREECODECAMP AND 1STOP. ADDITIONALLY, I HAVE EXPERIENCE WORKING ON PROJECTS INVOLVING E-COMMERCE WEBSITES USING MY TALENTS TO IMPROVE THE WEBSITE.**

**SINCE I'M NEW, IT WOULD BE WONDERFUL TO ENCOUNTER THE ACTUAL DIFFICULTIES FACED BY THE CORPORATE SECTOR AND GET INSIGHT INTO ITS OPERATIONS. SINCE I AM FRESHER, I BELIEVE I AM HIGHLY ADAPTABLE AND FLEXIBLE WHEN IT COMES TO LEARNING NEW THINGS.**

**I KNOW THINGS THEORETICALLY. HOWEVER, I'M STILL WAITING TO PUT MY THEORETICAL KNOWLEDGE TO USE. AND I THINK THAT IF I WORK HARD ENOUGH, I WILL GET KNOWLEDGE.**

## **TABLE OF CONTENTS**

<b>PROFESSIONAL BACKGROUND -----</b>	<b>2</b>
<b>TABLE OF CONTENTS -----</b>	<b>3</b>
<b>DATA ANALYTICS PROCESS -----</b>	<b>4-7</b>
<b>INSTAGRAM USER ANALYTICS-----</b>	<b>8-12</b>
<b>OPERATION ANALYTIC-----</b>	<b>13-18</b>
<b>HIRING PROCESS ANALYTICS-----</b>	<b>19-22</b>
<b>IMDB MOVIE ANALYSIS-----</b>	<b>23-26</b>
<b>BANK LOAN CASE STUDY-----</b>	<b>27-30</b>
<b>IMPACT OF CAR FEATURES -----</b>	<b>31-37</b>
<b>ABC CALL VOLUME TREND ANALYSIS -----</b>	<b>38-41</b>
<b>APPENDIX -----</b>	<b>42</b>

# **DATA ANALYTICS PROCESS**

## **APPLICATION IN REAL LIFE SCENARIO CASE**

### **STUDY**

**TASK IS TO GIVE THE EXAMPLE(S) OF SUCH A REAL-LIFE SITUATION WHERE WE USE DATA ANALYTICS AND LINK IT WITH THE DATA ANALYTICS PROCESS. YOU CAN PREPARE A PPT/PDF ON A REAL-LIFE SCENARIO EXPLAINING IT WITH THE ABOVE PROCESS (PLAN, PREPARE, PROCESS, ANALYZE, SHARE, ACT) AND SUBMIT IT AS PART OF THIS TASK.**

**Real-life Example of Data Analytics Process in E-commerce**

NOTE : The primary reason I chose this example is that I'm already familiar with it from my two projects—the first is in e-commerce, and the other is advanced e-commerce.

cocsamcoc69@gmail.com <https://www.linkedin.com/in/suraj-kumar-950795222/>

**INTRODUCTION: WELCOME TO OUR EXPLORATION OF HOW DATA ANALYTICS REVOLUTIONIZES THE E-COMMERCE LANDSCAPE. IN THIS PRESENTATION, WE'LL DELVE INTO THE INTRICACIES OF THE DATA ANALYTICS PROCESS WITHIN THE REALM OF ONLINE RETAIL, SHOWCASING ITS PIVOTAL ROLE IN DRIVING SUCCESS AND GROWTH.**

**1. PLAN: UNDERSTANDING MARKET DYNAMICS**

- **MARKET RESEARCH:** ANALYZING MARKET TRENDS, CUSTOMER PREFERENCES, AND COMPETITOR STRATEGIES.
- **TARGET AUDIENCE IDENTIFICATION:** SEGMENTING CUSTOMERS BASED ON DEMOGRAPHICS, BEHAVIOR, AND PURCHASE HISTORY.

**2. PREPARE: INFRASTRUCTURE AND DATA COLLECTION**

- **SETTING UP ANALYTICS TOOLS:** IMPLEMENTING ROBUST ANALYTICS PLATFORMS LIKE GOOGLE ANALYTICS, ADOBE ANALYTICS, OR CUSTOM-BUILT SOLUTIONS.
- **DATA COLLECTION STRATEGY:** DEFINING KEY METRICS, SETTING UP TRACKING MECHANISMS, AND ENSURING DATA ACCURACY AND INTEGRITY.

**3. PROCESS: EXTRACTING INSIGHTS FROM DATA**

- **DATA AGGREGATION:** CONSOLIDATING DATA FROM VARIOUS SOURCES, INCLUDING WEBSITE TRAFFIC, SALES TRANSACTIONS, AND CUSTOMER INTERACTIONS.
- **DATA PROCESSING:** CLEANING, TRANSFORMING, AND STRUCTURING RAW DATA INTO ACTIONABLE INSIGHTS.
- **STATISTICAL ANALYSIS:** APPLYING STATISTICAL TECHNIQUES TO UNCOVER PATTERNS, CORRELATIONS, AND TRENDS WITHIN THE DATA.

#### **4. ANALYZE: DERIVING ACTIONABLE INSIGHTS**

- **CUSTOMER BEHAVIOR ANALYSIS: UNDERSTANDING BROWSING PATTERNS, PURCHASING TRIGGERS, AND PRODUCT PREFERENCES.**
- **PERFORMANCE EVALUATION: ASSESSING THE EFFECTIVENESS OF MARKETING CAMPAIGNS, PRODUCT OFFERINGS, AND PRICING STRATEGIES.**
- **PREDICTIVE ANALYTICS: FORECASTING FUTURE TRENDS, CUSTOMER BEHAVIOR, AND DEMAND FLUCTUATIONS.**

#### **5. SHARE: COMMUNICATING INSIGHTS ACROSS TEAMS**

- **DATA VISUALIZATION: CREATING INTUITIVE DASHBOARDS, CHARTS, AND GRAPHS TO PRESENT INSIGHTS IN A VISUALLY APPEALING MANNER.**
- **COLLABORATIVE ANALYSIS: FACILITATING CROSS-FUNCTIONAL COLLABORATION AMONG MARKETING, SALES, AND PRODUCT TEAMS.**
- **REGULAR REPORTING: SHARING ACTIONABLE INSIGHTS AND PERFORMANCE METRICS WITH KEY STAKEHOLDERS.**

#### **6. ACT: IMPLEMENTING DATA-DRIVEN STRATEGIES**

- **PERSONALIZED MARKETING: TAILORING PROMOTIONAL OFFERS, PRODUCT RECOMMENDATIONS, AND CONTENT BASED ON INDIVIDUAL CUSTOMER PREFERENCES.**
- **DYNAMIC PRICING: ADJUSTING PRODUCT PRICES IN REAL-TIME BASED ON DEMAND, COMPETITOR PRICING, AND CUSTOMER SEGMENTATION.**
- **INVENTORY OPTIMIZATION: OPTIMIZING INVENTORY LEVELS AND PRODUCT ASSORTMENTS TO MEET DEMAND WHILE MINIMIZING STOCKOUTS AND OVERSTOCK.**

**CONCLUSION: DATA ANALYTICS IS THE CORNERSTONE OF SUCCESS IN THE E-COMMERCE INDUSTRY, EMPOWERING BUSINESSES TO MAKE INFORMED DECISIONS, ENHANCE CUSTOMER EXPERIENCES, AND DRIVE SUSTAINABLE GROWTH. BY LEVERAGING THE DATA ANALYTICS PROCESS OUTLINED IN THIS PRESENTATION, E-COMMERCE COMPANIES CAN STAY AHEAD OF THE CURVE IN AN INCREASINGLY COMPETITIVE MARKETPLACE.**

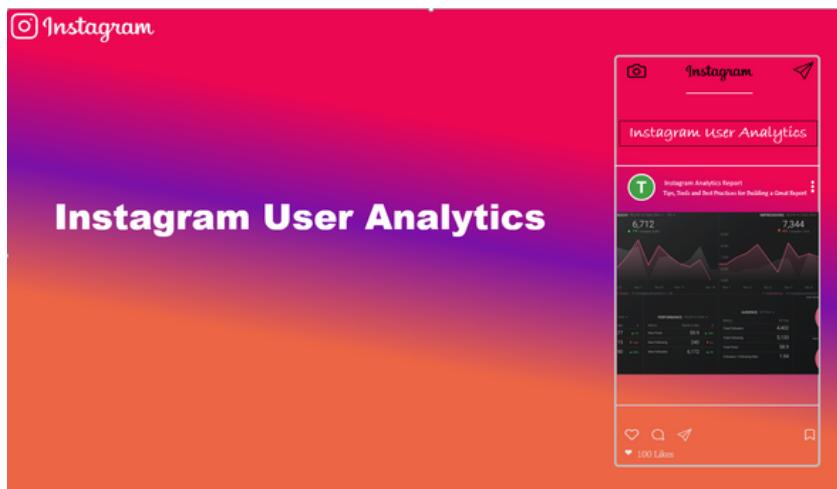
**REFERENCES LINK:**

**DATA ANALYTICS PROCESS IN E-COMMERCE**

The slide has a dark green background with a subtle texture. At the top left is a small circular icon with a white 'T'. At the top right is a URL: <https://trainity.in/data.html>. In the center is a white icon of a computer monitor with a pie chart on its screen. Below the icon, the word "Conclusion" is written in a large, bold, white sans-serif font. Underneath "Conclusion", there is a paragraph of white text: "Through the data analytics process, the e-commerce company can improve customer engagement, increase sales, and create a more personalized and optimized shopping experience for its customers." At the bottom left, there is an envelope icon followed by the email address [cocsamcoc69@gmail.com](mailto:cocsamcoc69@gmail.com). At the bottom right, there is a LinkedIn icon followed by the URL <https://www.linkedin.com/in/suraj-kumar-950795222/>.

# INSTAGRAM USER ANALYTICS

## SQL FUNDAMENTALS



### A) MARKETING ANALYSIS:

**LOYAL USER REWARD: THE MARKETING TEAM WANTS TO REWARD THE MOST LOYAL USERS, I.E., THOSE WHO HAVE BEEN USING THE PLATFORM FOR THE LONGEST TIME.**

A screenshot of an Instagram-themed dashboard titled "Loyal User Reward:". The dashboard features a header with the Instagram logo and a sub-header "Loyal User Reward:". Below the header is a section with the text "#Loyal User Reward: CODE" followed by a SQL query: "SELECT \* FROM users order by created\_at limit 5;". To the right of this text is a small box containing the same SQL code. Below the text is a table with columns "Id", "Username", and "Created at date and time". The table contains five rows of data. To the right of the table is a smaller box showing the same five rows of data in a different format. The data in both boxes is identical.

Id	Username	Created at date and time
38	Jordyn.Jacobson2	2016-05-14 07:56:26
63	Elenor88	2016-05-08 01:30:41
67	Emilio_Bernier52	2016-05-06 13:04:30
80	Darby_Herzog	2016-05-06 00:14:21
95	Nicole71	2016-05-09 17:30:22

id	username	created_at
80	Darby_Herzog	2016-05-06 00:14:21
67	Emilio_Bernier52	2016-05-06 13:04:30
63	Elenor88	2016-05-08 01:30:41
95	Nicole71	2016-05-09 17:30:22
38	Jordyn.Jacobson2	2016-05-14 07:56:26

**IN THIS GIVEN ABOVE FIG THE FIVE OLDEST USERS ON INSTAGRAM FROM THE PROVIDED DATABASE.**

# INACTIVE USER ENGAGEMENT: THE TEAM WANTS TO ENCOURAGE INACTIVE USERS TO START POSTING BY SENDING THEM PROMOTIONAL EMAILS.

**Instagram**

Inactive User Engagement

```
#Inactive User Engagement: CODE
SELECT username FROM users
left join photos
on users.id=photos.user_id
where photos.id is null;
```

username
Annye_Hackett
Kassandra_Hommerick
Jaylond3
Abigail_Orsi
Marcell_Holmstrom
Tanner_Bartolomeo
Pearl7
Ollie_Lindner32
Hildegard_McCullough
David_Olemaek47
Haley_Kassulke
Lorraine939
Duanew93
Lulu_Gronlund
Hilary_Auer39
Hannah_Marquardt44
Ree_Haag
Habiba_Haqqepovic
Leddie
Jeanette_Nikolic921
Charly_Peterson
Eduardo_Juliano81
BarTholomee_Bernard
Savanna
Emmanuelle_Meier57
Bethany20

USERNAME
Annye_Hackett
Kassandra_Hommerick
Jaylond3
Abigail_Orsi
Marcell_Holmstrom
Tanner_Bartolomeo
Pearl7
Ollie_Lindner32
Hildegard_McCullough
David_Olemaek47
Haley_Kassulke
Lorraine939
Duanew93
Lulu_Gronlund
Hilary_Auer39
Hannah_Marquardt44
Ree_Haag
Habiba_Haqqepovic
Leddie
Jeanette_Nikolic921
Bethany20

**IN THIS FIG USERS WHO HAVE NEVER POSTED A SINGLE PHOTO ON INSTAGRAM.**

**CONTEST WINNER DECLARATION:** THE TEAM HAS ORGANIZED A CONTEST WHERE THE USER WITH THE MOST LIKES ON A SINGLE PHOTO WINS.

**Instagram**

Contest Winner Declaration

```
#Contest Winner Declaration:CODE
SELECT username,photos.id,photos.image_url,
count(likes.user_id) as total FROM photos
inner join likes on likes.photo_id=photos.id
inner join onlikes.photos.photo_id=photos.id
inner join users on photos.user_id=users.id
group by photos.id
order by total desc
limit 1;
```

Username	Id	Image_url	Total
Zack_Kemmer93	145	https://jarret.name	48

username	id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

**IN THIS GIVEN FIG THE WINNER OF THE CONTEST AND PROVIDE THEIR DETAILS TO THE TEAM.**

## HASHTAG RESEARCH: A PARTNER BRAND WANTS TO KNOW THE MOST POPULAR HASHTAGS TO USE IN THEIR POSTS TO REACH THE MOST PEOPLE.

The screenshot shows the Instagram interface for hashtag research. At the top, there's a header with the Instagram logo and the text '# Hashtag Research'. Below the header, there's a code snippet for a MySQL query:

```
#Hashtag Research:CODE  
select tags.tag_name, count(*) as total  
from photo_tags join tags  
on photo_tags.tag_id=tags.id  
group by tags.id  
order by total desc  
limit 5;
```

On the right side, there's a preview window showing the results of the query:

```
Hashtag Research:  
select tags.tag_name, count(*) as total  
from photo_tags join tags  
on photo_tags.tag_id=tags.id  
group by tags.id  
order by total desc  
limit 5;
```

Tag_Name	Total
smile	59
beach	42
party	39
fun	38
concert	24

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

IN THIS FIG THE TOP FIVE MOST COMMONLY USED HASHTAGS ON THE PLATFORM.

## AD CAMPAIGN LAUNCH: THE TEAM WANTS TO KNOW THE BEST DAY OF THE WEEK TO LAUNCH ADS.

The screenshot shows the Instagram interface for ad campaign launch. At the top, there's a header with the Instagram logo and the text 'Ad Campaign Launch'. Below the header, there's a code snippet for a MySQL query:

```
#Ad Campaign Launch:  
select dayname(created_at) as day ,  
count(*) as total from users  
group by day  
order by total desc  
limit 2;
```

On the right side, there's a preview window showing the results of the query:

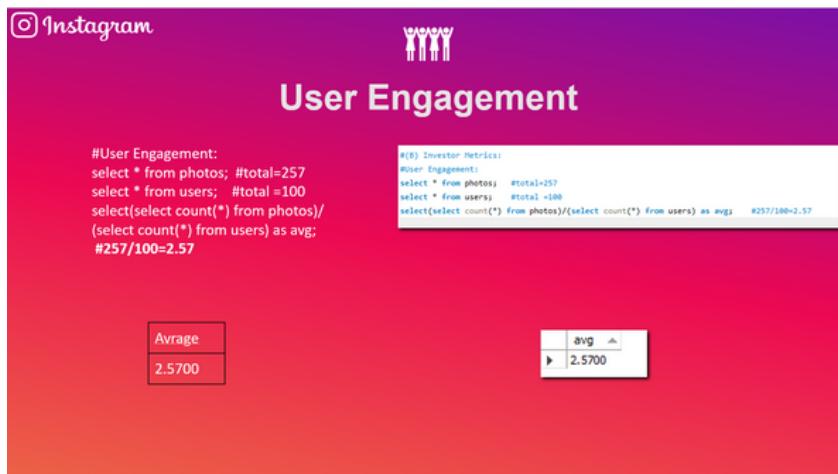
```
Ad Campaign Launch:  
select dayname(created_at) as day ,count(*) as total from users  
group by day  
order by total desc  
limit 2;
```

Day	Total
Thursday	16
Sunday	16

day	total
Thursday	16
Sunday	16

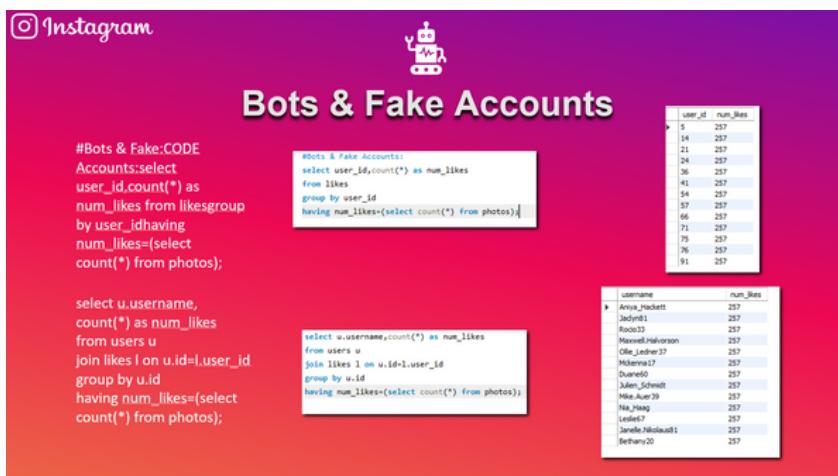
IN THIS FIG THE DAY OF THE WEEK WHEN MOST USERS REGISTER ON INSTAGRAM. PROVIDE INSIGHTS ON WHEN TO SCHEDULE AN AD CAMPAIGN.

## USER ENGAGEMENT: INVESTORS WANT TO KNOW IF USERS ARE STILL ACTIVE AND POSTING ON INSTAGRAM OR IF THEY ARE MAKING FEWER POSTS.



IN THIS FIG THE AVERAGE NUMBER OF POSTS PER USER ON INSTAGRAM. ALSO, PROVIDE THE TOTAL NUMBER OF PHOTOS ON INSTAGRAM DIVIDED BY THE TOTAL NUMBER OF USERS.

## BOTS & FAKE ACCOUNTS: INVESTORS WANT TO KNOW IF THE PLATFORM IS CROWDED WITH FAKE AND DUMMY ACCOUNTS.



IN THIS FIG IDENTIFY USERS (POTENTIAL BOTS) WHO HAVE LIKED EVERY SINGLE PHOTO ON THE SITE, AS THIS IS NOT TYPICALLY POSSIBLE FOR A NORMAL USER.

## **CONCULLISION**

- 1. LOYAL USER REWARD:** THE FIVE OLDEST USERS ON INSTAGRAM HAVE BEEN IDENTIFIED, ALLOWING THE MARKETING TEAM TO CONSIDER REWARDING THEM FOR THEIR LONG-TERM ENGAGEMENT WITH THE PLATFORM.
- 2. INACTIVE USER ENGAGEMENT:** USERS WHO HAVE NEVER POSTED A SINGLE PHOTO ON INSTAGRAM HAVE BEEN IDENTIFIED, ENABLING TARGETED EFFORTS TO ENCOURAGE THEM TO START POSTING THROUGH PROMOTIONAL EMAILS OR OTHER ENGAGEMENT STRATEGIES.
- 3. CONTEST WINNER DECLARATION:** THE WINNER OF THE CONTEST, BASED ON THE USER WITH THE MOST LIKES ON A SINGLE PHOTO, HAS BEEN DETERMINED, PROVIDING CLARITY FOR THE TEAM AND FACILITATING PRIZE DISTRIBUTION.
- 4. HASHTAG RESEARCH:** THE TOP FIVE MOST COMMONLY USED HASHTAGS ON THE PLATFORM HAVE BEEN IDENTIFIED, WHICH CAN BE SHARED WITH PARTNER BRANDS FOR THEIR POSTS TO REACH A WIDER AUDIENCE.
- 5. AD CAMPAIGN LAUNCH:** INSIGHTS ON THE BEST DAY OF THE WEEK TO LAUNCH ADS HAVE BEEN PROVIDED BASED ON THE DAY WHEN MOST USERS REGISTER ON INSTAGRAM, AIDING IN THE OPTIMIZATION OF AD CAMPAIGN SCHEDULING.
- 6. USER ENGAGEMENT METRICS:** THE AVERAGE NUMBER OF POSTS PER USER ON INSTAGRAM HAS BEEN CALCULATED, ALONG WITH THE TOTAL NUMBER OF PHOTOS ON INSTAGRAM DIVIDED BY THE TOTAL NUMBER OF USERS, PROVIDING INSIGHTS INTO USER ACTIVITY LEVELS.
- 7. BOTS & FAKE ACCOUNTS:** POTENTIAL BOTS OR FAKE ACCOUNTS HAVE BEEN IDENTIFIED BASED ON UNUSUAL BEHAVIOR SUCH AS LIKING EVERY SINGLE PHOTO ON THE SITE, WHICH CAN HELP IN MAINTAINING THE INTEGRITY AND AUTHENTICITY OF THE PLATFORM.

THESE CONCLUSIONS CAN GUIDE DECISION-MAKING PROCESSES WITHIN THE INSTAGRAM PRODUCT TEAM AND PROVIDE VALUABLE INSIGHTS FOR BUSINESS GROWTH AND USER ENGAGEMENT STRATEGIES.

**REFERENCES LINK:**

**INSTAGRAM USER ANALYTICS**

# OPERATION ANALYTICS AND INVESTIGATING METRIC SPIKE ADVANCED SQL

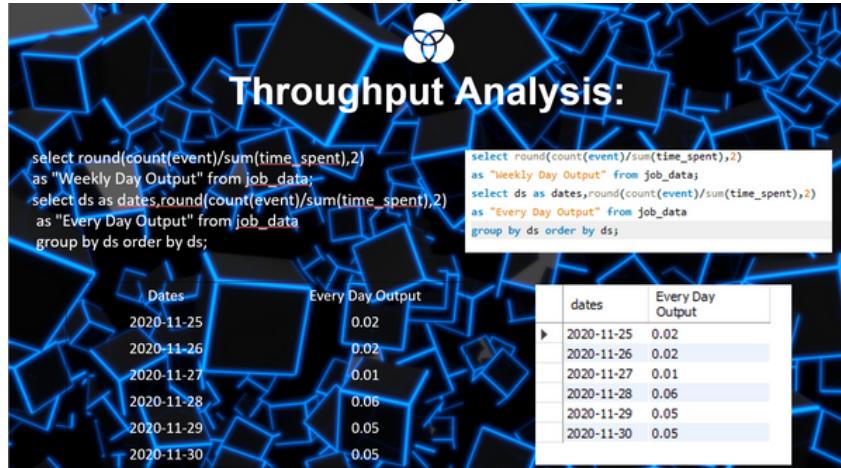


**JOB REVIEWED OVER TIME: OBJECTIVE: CALCULATE THE NUMBER OF JOBS REVIEWED PER HOUR FOR EACH DAY IN NOVEMBER 2022**



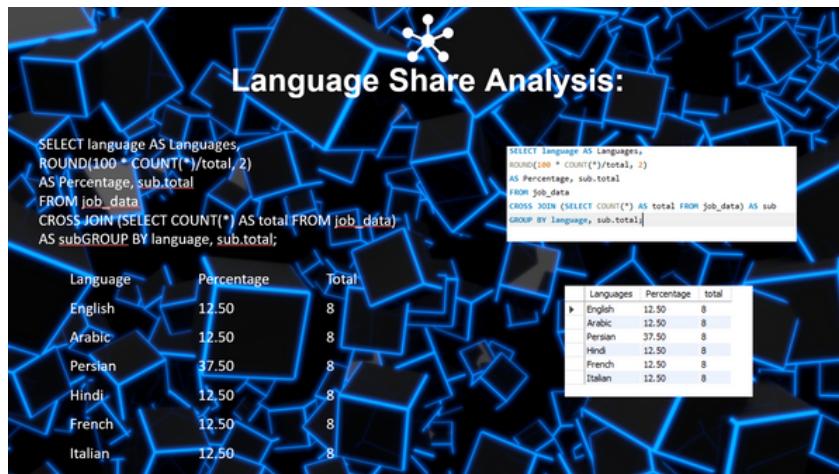
**IN THIS FIG AN SQL QUERY TO CALCULATE THE NUMBER OF JOBS REVIEWED PER HOUR FOR EACH DAY IN NOVEMBER 2020.**

## THROUGHPUT ANALYSIS: OBJECTIVE: CALCULATE THE 7-DAY ROLLING AVERAGE OF THROUGHPUT (NUMBER OF EVENTS PER SECOND).



IN THIS FIG SQL QUERY TO CALCULATE THE 7-DAY ROLLING AVERAGE OF THROUGHPUT. ADDITIONALLY, USING THE DAILY METRIC OR THE 7-DAY ROLLING

AVERAGE  
YOUR PARAGRAPH TEXT  
**LANGUAGE SHARE ANALYSIS: OBJECTIVE: CALCULATE THE PERCENTAGE SHARE OF EACH LANGUAGE IN THE LAST 30 DAYS.**



IN THIS FIG SQL QUERY TO CALCULATE THE PERCENTAGE SHARE OF EACH LANGUAGE OVER THE LAST 30 DAYS.

## DUPLICATE ROWS DETECTION: OBJECTIVE: IDENTIFY DUPLICATE ROWS IN THE DATA.

The background of the slide features a dark blue and black abstract pattern of geometric shapes. In the center, there is a white rectangular area containing a SQL query and its results. At the top left of this area is a small icon of three people. At the top right is a small icon of a puzzle piece. The title "Duplicate Rows Detection:" is centered above the code.

```
SELECT actor_id, COUNT(*) AS Duplicates FROM job_data GROUP BY actor_id HAVING COUNT(*) > 1;
```

```
SELECT actor_id, COUNT(*) AS Duplicates FROM job_data GROUP BY actor_id HAVING COUNT(*) > 1;
```

actor_id	Duplicates
1003	2

IN THIS FIG SQL QUERY TO DISPLAY DUPLICATE ROWS FROM THE JOB\_DATA TABLE.

## WEEKLY USER ENGAGEMENT: OBJECTIVE: MEASURE THE ACTIVENESS OF USERS ON A WEEKLY BASIS.

The background of the slide features a dark blue and black abstract pattern of geometric shapes. In the center, there is a white rectangular area containing a SQL query and its results. At the top left of this area is a small icon of three people. At the top right is a small icon of a puzzle piece. The title "Weekly User Engagement:" is centered above the code.

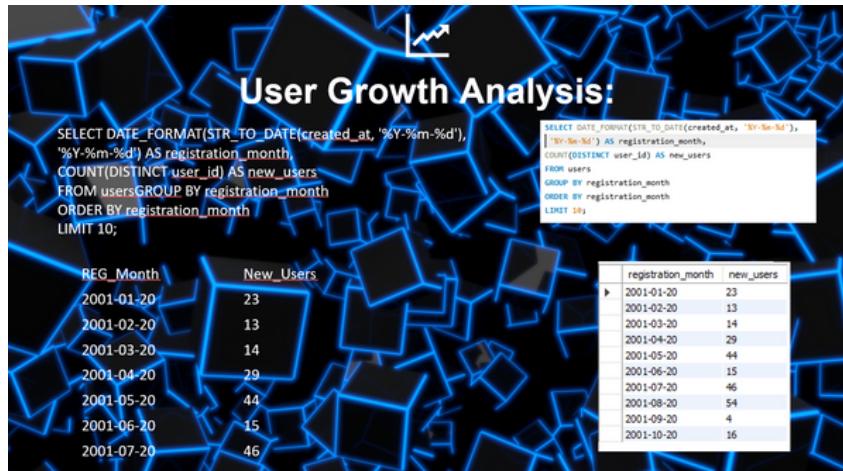
```
SELECT user_id, WEEK(STR_TO_DATE(occurred_at, '%Y-%m-%d'), 1) AS week_number, COUNT(*) AS weekly_engagement FROM events GROUP BY user_id, week_number ORDER BY user_id, week_number LIMIT 10;
```

```
SELECT user_id, WEEK(STR_TO_DATE(occurred_at, '%Y-%m-%d'), 1) AS week_number, COUNT(*) AS weekly_engagement FROM events GROUP BY user_id, week_number ORDER BY user_id, week_number LIMIT 10;
```

User_id	Week_num	Weekly_engag
4	20	16
4	21	25
4	24	15
4	25	23
4	29	14
8	20	14
8	21	15
8	29	7
11	25	64
11	29	37

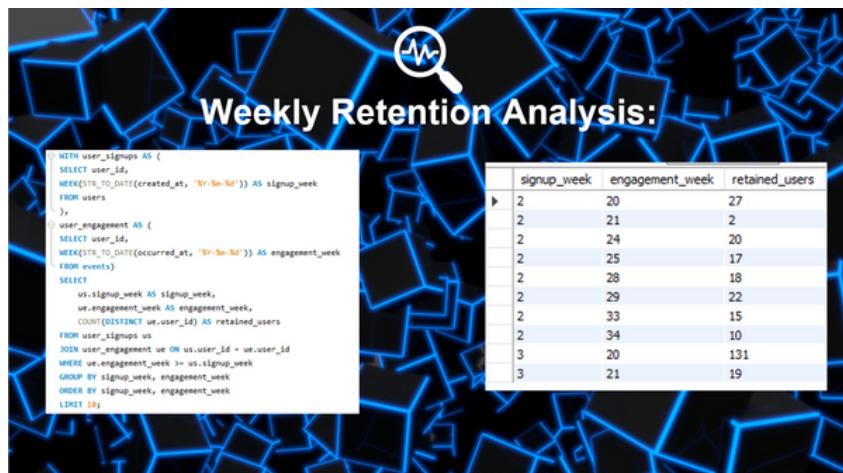
IN THIS FIG SQL QUERY TO CALCULATE THE WEEKLY USER ENGAGEMENT.

## USER GROWTH ANALYSIS: OBJECTIVE: ANALYZE THE GROWTH OF USERS OVER TIME FOR A PRODUCT.



IN THIS FIG SQL QUERY TO CALCULATE THE USER GROWTH FOR THE PRODUCT.

## WEEKLY RETENTION ANALYSIS: OBJECTIVE: ANALYZE THE RETENTION OF USERS ON A WEEKLY BASIS



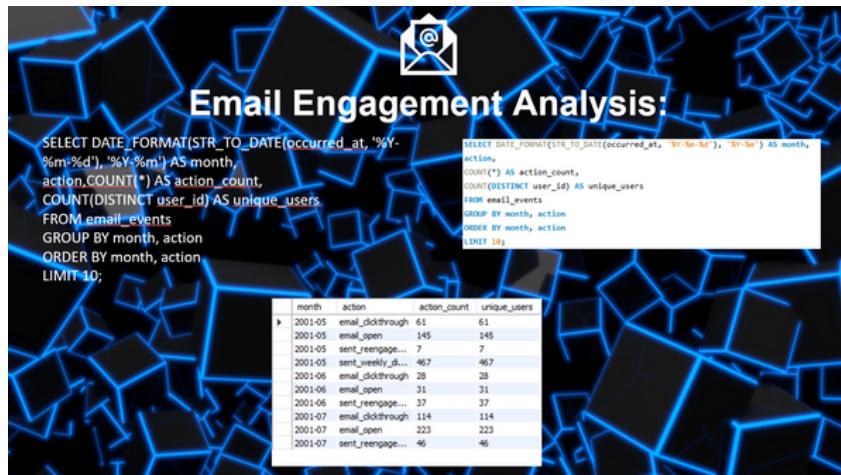
IN THIS FIG SQL QUERY TO CALCULATE THE WEEKLY RETENTION OF USERS BASED ON THEIR SIGN-UP COHORT.

## WEEKLY ENGAGEMENT PER DEVICE: OBJECTIVE: MEASURE THE ACTIVENESS OF USERS ON A WEEKLY BASIS PER DEVICE.



IN THIS FIG SQL QUERY TO CALCULATE THE WEEKLY ENGAGEMENT PER DEVICE

## EMAIL ENGAGEMENT ANALYSIS: OBJECTIVE: ANALYZE HOW USERS ARE ENGAGING WITH THE EMAIL SERVICE.



IN THIS FIG SQL QUERY TO CALCULATE THE EMAIL ENGAGEMENT METRICS.

## **CONCLUSION**

**BASED ON THE ADVANCED SQL ANALYSIS CONDUCTED FOR THE TWO CASE STUDIES, THE FOLLOWING CONCLUSIONS CAN BE DRAWN:**

**FOR CASE STUDY 1: JOB DATA ANALYSIS:**

- 1. JOBS REVIEWED OVER TIME: THE NUMBER OF JOBS REVIEWED PER HOUR FOR EACH DAY IN NOVEMBER 2020 HAS BEEN CALCULATED, PROVIDING INSIGHTS INTO THE DAILY ACTIVITY TRENDS.**
- 2. THROUGHPUT ANALYSIS: THE 7-DAY ROLLING AVERAGE OF THROUGHPUT HAS BEEN CALCULATED TO SMOOTH OUT FLUCTUATIONS AND PROVIDE A STABLE REPRESENTATION OF THROUGHPUT TRENDS OVER TIME.**
- 3. LANGUAGE SHARE ANALYSIS: THE PERCENTAGE SHARE OF EACH LANGUAGE IN THE LAST 30 DAYS HAS BEEN CALCULATED, OFFERING INSIGHTS INTO LANGUAGE PREFERENCES AMONG USERS.**
- 4. DUPLICATE ROWS DETECTION: DUPLICATE ROWS IN THE DATA HAVE BEEN IDENTIFIED, WHICH MAY REQUIRE FURTHER INVESTIGATION OR DATA CLEANING PROCESSES TO ENSURE DATA INTEGRITY.**

**FOR CASE STUDY 2: INVESTIGATING METRIC SPIKE:**

- 1. WEEKLY USER ENGAGEMENT: WEEKLY USER ENGAGEMENT HAS BEEN MEASURED, PROVIDING INSIGHTS INTO USER ACTIVITY PATTERNS OVER TIME.**
- 2. USER GROWTH ANALYSIS: USER GROWTH OVER TIME FOR THE PRODUCT HAS BEEN ANALYZED, HELPING TO UNDERSTAND THE PRODUCT'S ADOPTION AND EXPANSION.**
- 3. WEEKLY RETENTION ANALYSIS: WEEKLY RETENTION OF USERS BASED ON THEIR SIGN-UP COHORT HAS BEEN CALCULATED, OFFERING INSIGHTS INTO USER RETENTION EFFORTS.**
- 4. WEEKLY ENGAGEMENT PER DEVICE: WEEKLY ENGAGEMENT PER DEVICE HAS BEEN MEASURED, HELPING TO UNDERSTAND USER ENGAGEMENT ACROSS DIFFERENT DEVICES.**
- 5. EMAIL ENGAGEMENT ANALYSIS: EMAIL ENGAGEMENT METRICS HAVE BEEN ANALYZED, PROVIDING INSIGHTS INTO HOW USERS ARE INTERACTING WITH THE EMAIL SERVICE.**  
THESE CONCLUSIONS CAN GUIDE DECISION-MAKING PROCESSES WITHIN THE COMPANY, HELPING TO IDENTIFY AREAS FOR IMPROVEMENT, UNDERSTAND USER BEHAVIOR, AND OPTIMIZE OPERATIONS FOR BETTER PERFORMANCE AND USER EXPERIENCE. ADDITIONALLY, FURTHER EXPLORATION AND ANALYSIS MAY BE NEEDED BASED ON THE INSIGHTS OBTAINED TO DERIVE ACTIONABLE STRATEGIES FOR THE COMPANY'S GROWTH AND SUCCESS.

## **REFERENCES LINK:**

**OPERATION ANALYTICS AND INVESTIGATING METRIC SPIKE**

# HIRING PROCESS ANALYTICS

## STATISTICS

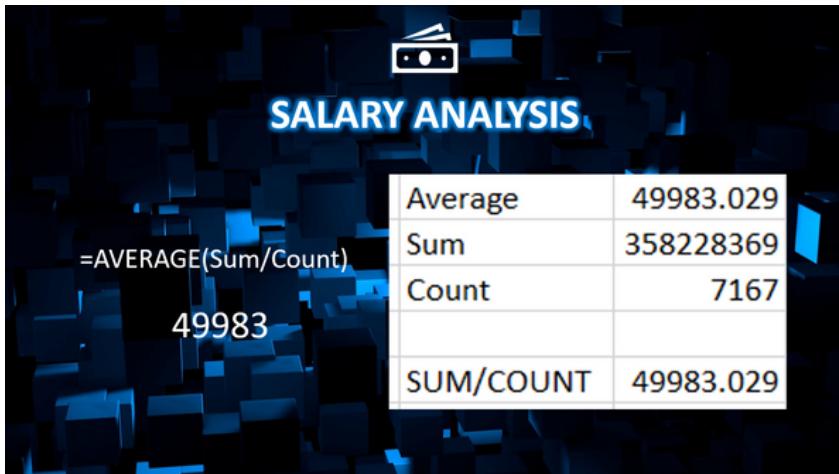


**HIRING ANALYSIS: THE HIRING PROCESS INVOLVES BRINGING NEW INDIVIDUALS INTO THE ORGANIZATION FOR VARIOUS ROLES.**



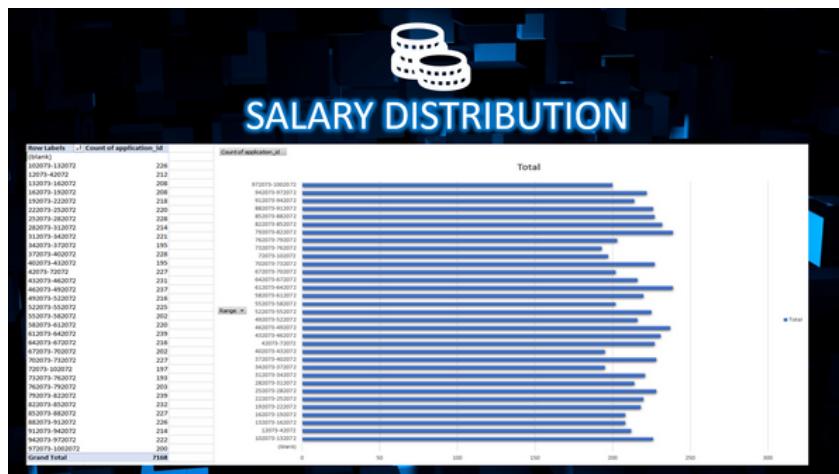
**IN THIS FIG THE GENDER DISTRIBUTION OF HIRES. & MALES AND FEMALES HAVE BEEN HIRED BY THE COMPANY.**

**SALARY ANALYSIS: THE AVERAGE SALARY IS CALCULATED BY ADDING UP THE SALARIES OF A GROUP OF EMPLOYEES AND THEN DIVIDING THE TOTAL BY THE NUMBER OF EMPLOYEES.**



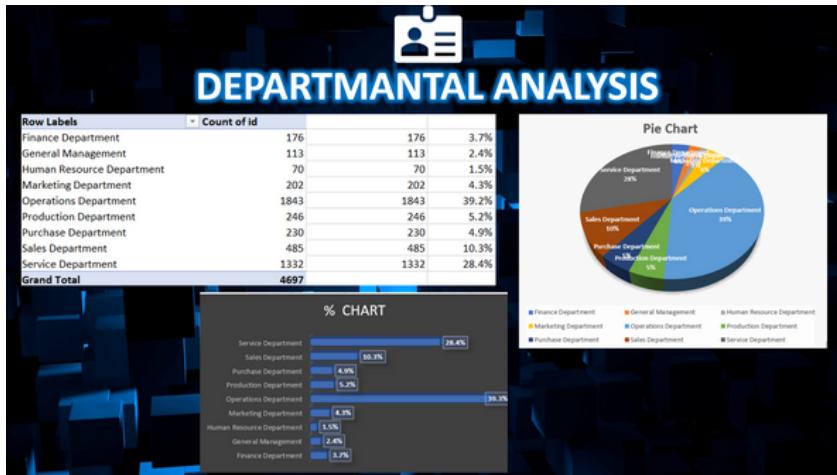
**IN THIS FIG THE AVERAGE SALARY OFFERED BY THIS COMPANY? USE EXCEL FUNCTIONS TO CALCULATE THIS**

**SALARY DISTRIBUTION: CLASS INTERVALS REPRESENT RANGES OF VALUES, IN THIS CASE, SALARY RANGES. THE CLASS INTERVAL IS THE DIFFERENCE BETWEEN THE UPPER AND LOWER LIMITS OF A CLASS.**



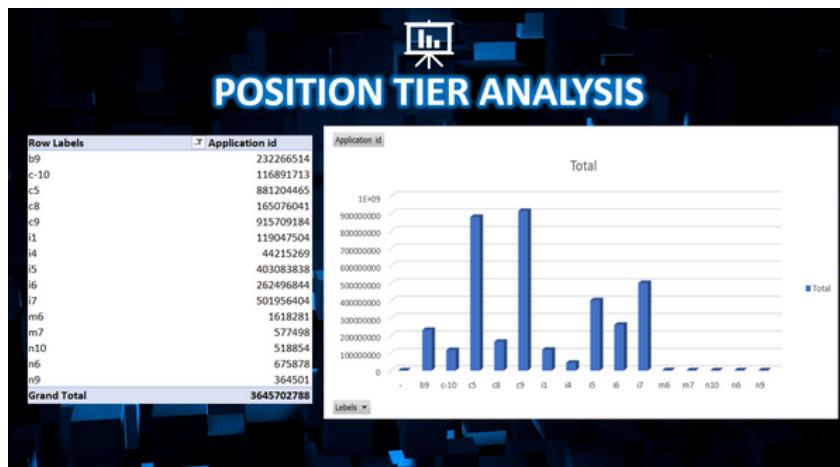
**IN THIS FIG CLASS INTERVALS FOR THE SALARIES IN THE COMPANY. THIS WILL HELP YOU UNDERSTAND THE SALARY DISTRIBUTION.**

# DEPARTMENTAL ANALYSIS: VISUALIZING DATA THROUGH CHARTS AND PLOTS IS A CRUCIAL PART OF DATA ANALYSIS.



IN THIS FIG USE A PIE CHART, BAR GRAPH, OR ANY OTHER SUITABLE VISUALIZATION TO SHOW THE PROPORTION OF PEOPLE WORKING IN DIFFERENT DEPARTMENTS.

# POSITION TIER ANALYSIS: DIFFERENT POSITIONS WITHIN A COMPANY OFTEN HAVE DIFFERENT TIERS OR LEVELS.



IN THIS FIG USE A CHART OR GRAPH TO REPRESENT THE DIFFERENT POSITION TIERS WITHIN THE COMPANY. THIS WILL HELP YOU UNDERSTAND THE DISTRIBUTION OF POSITIONS ACROSS DIFFERENT TIERS.

## **CONCLUSION**

- A. HIRING ANALYSIS: THE GENDER DISTRIBUTION OF HIRES SHOWS , INDICATING .**
- B. SALARY ANALYSIS: THE AVERAGE SALARY OFFERED BY THE COMPANY IS , SUGGESTING.**
- C. SALARY DISTRIBUTION: THE SALARY DISTRIBUTION IS CHARACTERIZED BY , HIGHLIGHTING.**
- D. DEPARTMENTAL ANALYSIS: THE PROPORTION OF PEOPLE WORKING IN DIFFERENT DEPARTMENTS IS VISUALIZED THROUGH , REVEALING .**
- E. POSITION TIER ANALYSIS: THE DISTRIBUTION OF POSITIONS ACROSS DIFFERENT TIERS IS REPRESENTED BY, DEMONSTRATING .**

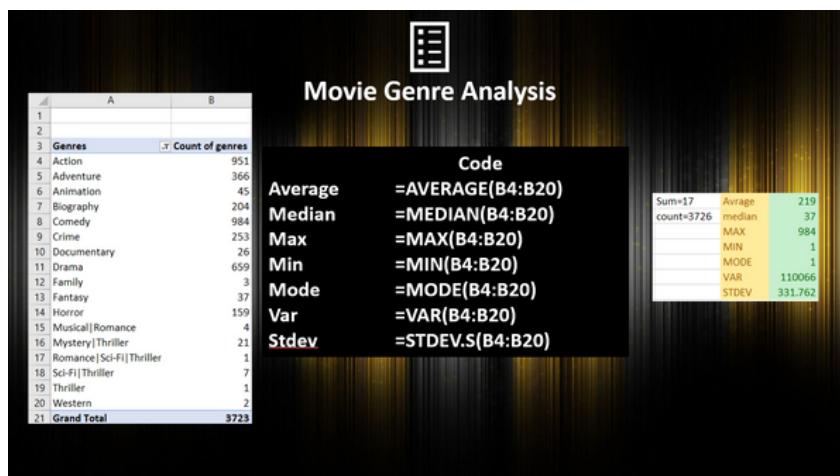
**OVERALL, THESE INSIGHTS PROVIDE VALUABLE INFORMATION FOR THE COMPANY TO IMPROVE ITS HIRING PROCESS, UNDERSTAND SALARY STRUCTURES, OPTIMIZE DEPARTMENTAL ALLOCATIONS, AND ANALYZE POSITION TIERS EFFECTIVELY. FURTHER ANALYSIS AND ACTIONS CAN BE UNDERTAKEN BASED ON THESE FINDINGS TO ENHANCE THE EFFICIENCY AND EFFECTIVENESS OF THE COMPANY'S HIRING PRACTICES.**

**REFERENCES LINK:  
HIRING PROCESS ANALYTICS**

## IMDB MOVIE ANALYSIS

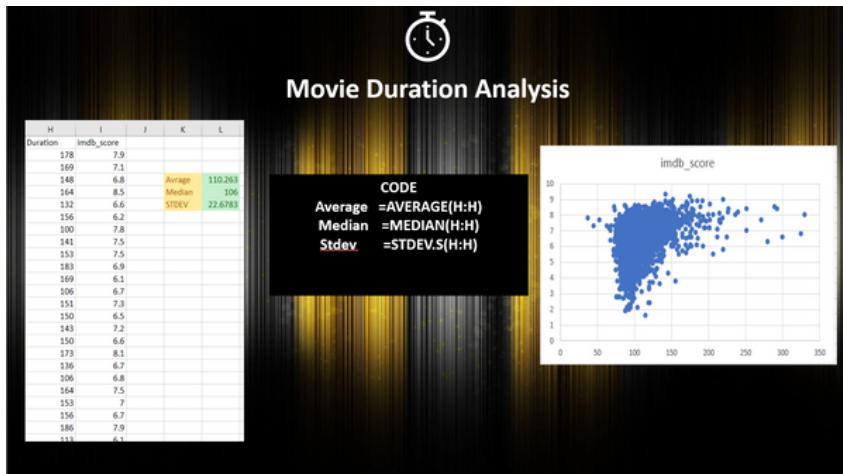


### MOVIE GENRE ANALYSIS: ANALYZE THE DISTRIBUTION OF MOVIE GENRES AND THEIR IMPACT ON THE IMDB SCORE.



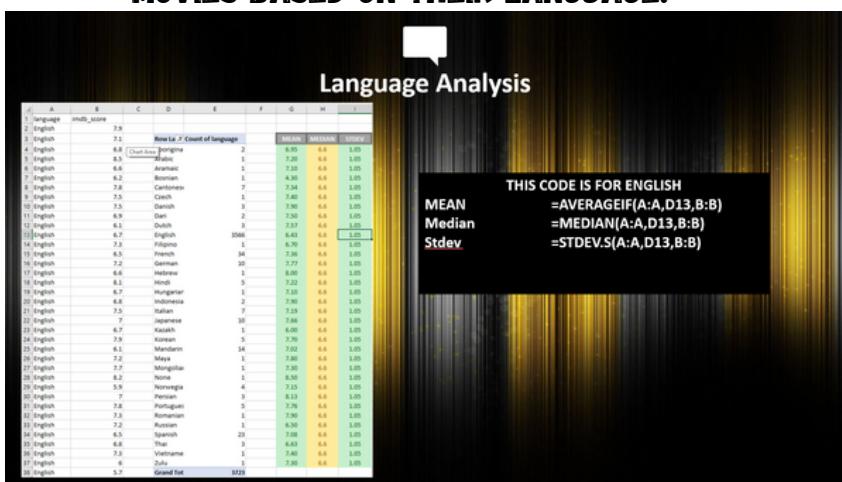
IN THIS FIG MOST COMMON GENRES OF MOVIES IN THE DATASET. THEN, FOR EACH GENRE, CALCULATE DESCRIPTIVE STATISTICS (MEAN, MEDIAN, MODE, RANGE, VARIANCE, STANDARD DEVIATION) OF THE IMDB SCORES.

# MOVIE DURATION ANALYSIS: ANALYZE THE DISTRIBUTION OF MOVIE DURATIONS AND ITS IMPACT ON THE IMDB SCORE.



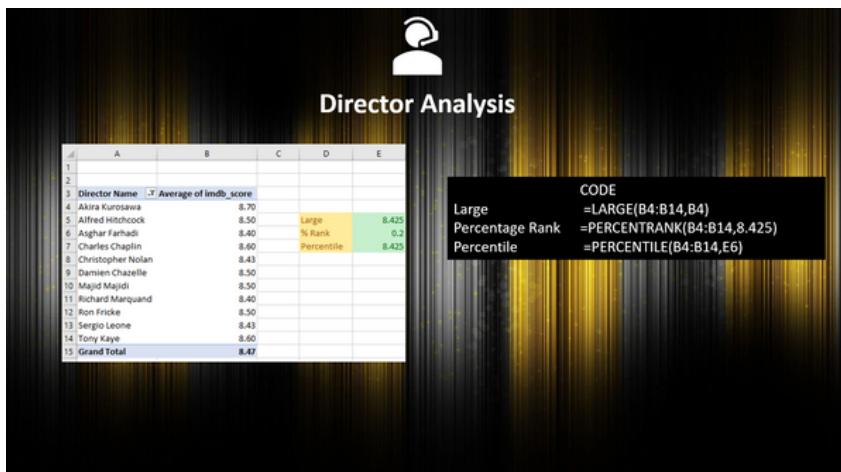
IN THIS FIG ANALYZE THE DISTRIBUTION OF MOVIE DURATIONS AND IDENTIFY THE RELATIONSHIP BETWEEN MOVIE DURATION AND IMDB SCORE.

LANGUAGE ANALYSIS: SITUATION: EXAMINE THE DISTRIBUTION OF MOVIES BASED ON THEIR LANGUAGE.



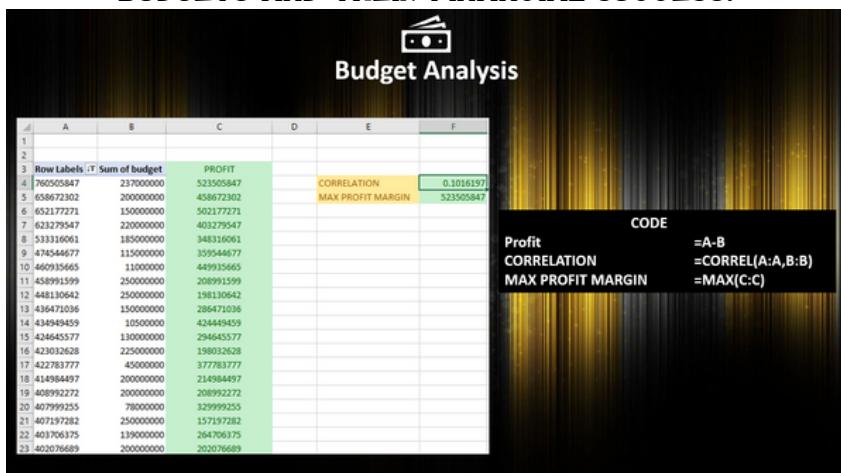
IN THIS FIG DETERMINE THE MOST COMMON LANGUAGES USED IN MOVIES AND ANALYZE THEIR IMPACT ON THE IMDB SCORE USING DESCRIPTIVE STATISTICS.

## DIRECTOR ANALYSIS: INFLUENCE OF DIRECTORS ON MOVIE RATINGS.



IN THIS FIG IDENTIFY THE TOP DIRECTORS BASED ON THEIR AVERAGE IMDB SCORE AND ANALYZE THEIR CONTRIBUTION TO THE SUCCESS OF MOVIES USING PERCENTILE CALCULATIONS.

## BUDGET ANALYSIS: EXPLORE THE RELATIONSHIP BETWEEN MOVIE BUDGETS AND THEIR FINANCIAL SUCCESS.



IN THIS FIG ANALYZE THE CORRELATION BETWEEN MOVIE BUDGETS AND GROSS EARNINGS, AND IDENTIFY THE MOVIES WITH THE HIGHEST PROFIT MARGIN.

## CONCLUSION

IN CONCLUSION, THE ANALYSIS OF THE IMDB MOVIES DATASET HAS PROVIDED VALUABLE INSIGHTS INTO THE FACTORS INFLUENCING THE SUCCESS OF A MOVIE ON IMDB. HERE ARE THE KEY FINDINGS:

1. **GENRE ANALYSIS:** THE MOST COMMON GENRES IN THE DATASET WERE IDENTIFIED, AND THEIR IMPACT ON IMDB SCORES WAS ANALYZED. CERTAIN GENRES SHOWED HIGHER AVERAGE RATINGS COMPARED TO OTHERS, SUGGESTING A CORRELATION BETWEEN GENRE AND AUDIENCE PREFERENCES.
2. **MOVIE DURATION ANALYSIS:** THE DISTRIBUTION OF MOVIE DURATIONS WAS EXAMINED, AND THE RELATIONSHIP BETWEEN MOVIE DURATION AND IMDB SCORES WAS EXPLORED. WHILE THERE MAY NOT BE A STRONG LINEAR RELATIONSHIP, CERTAIN TRENDS REGARDING OPTIMAL MOVIE DURATION FOR HIGHER RATINGS WERE OBSERVED.
3. **LANGUAGE ANALYSIS:** THE PREVALENCE OF DIFFERENT LANGUAGES IN MOVIES WAS INVESTIGATED, AND THEIR IMPACT ON IMDB SCORES WAS ASSESSED. SOME LANGUAGES SHOWED HIGHER AVERAGE RATINGS, INDICATING POTENTIAL PREFERENCES AMONG VIEWERS.
4. **DIRECTOR ANALYSIS:** TOP DIRECTORS WERE IDENTIFIED BASED ON THEIR AVERAGE IMDB SCORES, AND THEIR CONTRIBUTION TO THE SUCCESS OF MOVIES WAS EVALUATED. CERTAIN DIRECTORS CONSISTENTLY PRODUCED HIGHLY-RATED MOVIES, SUGGESTING THEIR INFLUENCE ON AUDIENCE RECEPTION.
5. **BUDGET ANALYSIS:** THE CORRELATION BETWEEN MOVIE BUDGETS AND GROSS EARNINGS WAS EXAMINED, AND MOVIES WITH THE HIGHEST PROFIT MARGINS WERE IDENTIFIED. WHILE A CORRELATION BETWEEN BUDGET AND GROSS EARNINGS WAS OBSERVED, OTHER FACTORS MAY ALSO INFLUENCE A MOVIE'S FINANCIAL SUCCESS.

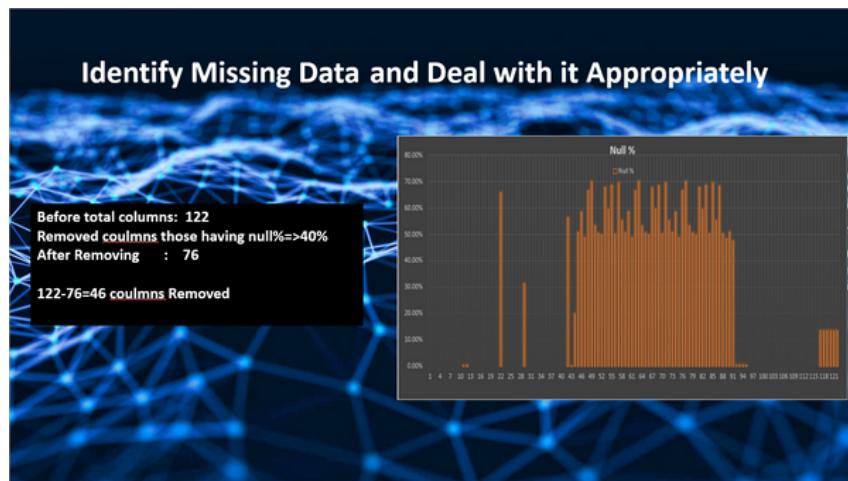
OVERALL, THESE INSIGHTS PROVIDE VALUABLE GUIDANCE FOR MOVIE PRODUCERS, DIRECTORS, AND INVESTORS SEEKING TO UNDERSTAND WHAT MAKES A MOVIE SUCCESSFUL ON IMDB. BY LEVERAGING THESE FINDINGS, STAKEHOLDERS CAN MAKE INFORMED DECISIONS IN THEIR FUTURE PROJECTS, ULTIMATELY IMPROVING THEIR CHANCES OF ACHIEVING HIGH RATINGS AND FINANCIAL SUCCESS IN THE COMPETITIVE MOVIE INDUSTRY. FURTHER ANALYSIS AND EXPLORATION MAY BE WARRANTED TO DELVE DEEPER INTO SPECIFIC ASPECTS OF THE DATASET AND UNCOVER ADDITIONAL INSIGHTS.

**REFERENCES LINK:**  
**IMDB MOVIE ANALYSIS**

## BANK LOAN CASE STUDY

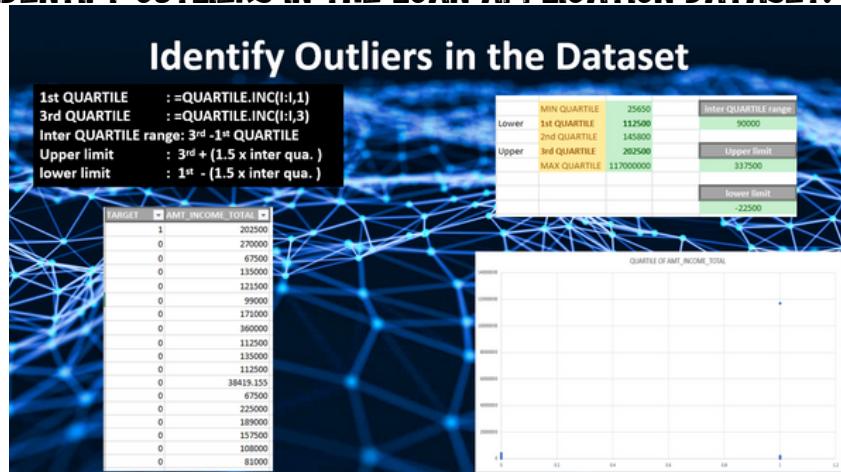


**IDENTIFY MISSING DATA AND DEAL WITH IT APPROPRIATELY: AS A DATA ANALYST, YOU COME ACROSS MISSING DATA IN THE LOAN APPLICATION DATASET. IT IS ESSENTIAL TO HANDLE MISSING DATA EFFECTIVELY TO ENSURE THE ACCURACY OF THE ANALYSIS.**



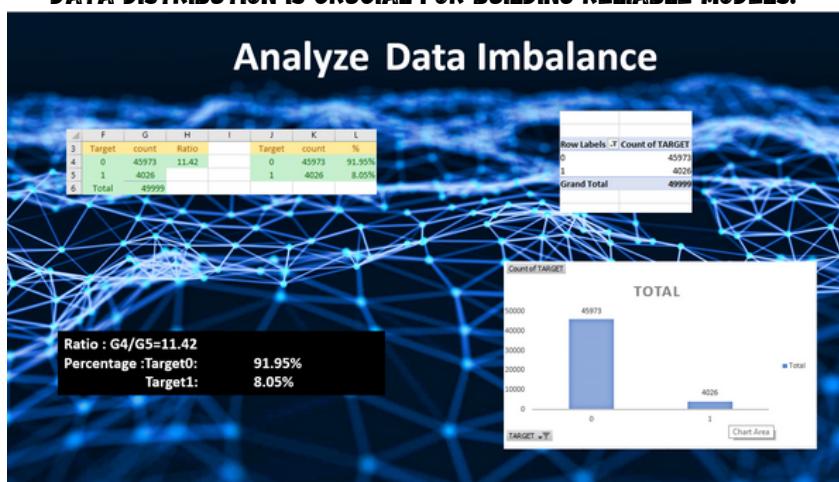
**IN THIS FIG IDENTIFY THE MISSING DATA IN THE DATASET AND DECIDE ON AN APPROPRIATE METHOD TO DEAL WITH IT USING EXCEL BUILT-IN FUNCTIONS AND FEATURES.**

# IDENTIFY OUTLIERS IN THE DATASET: OUTLIERS CAN SIGNIFICANTLY IMPACT THE ANALYSIS AND DISTORT THE RESULTS. YOU NEED TO IDENTIFY OUTLIERS IN THE LOAN APPLICATION DATASET.



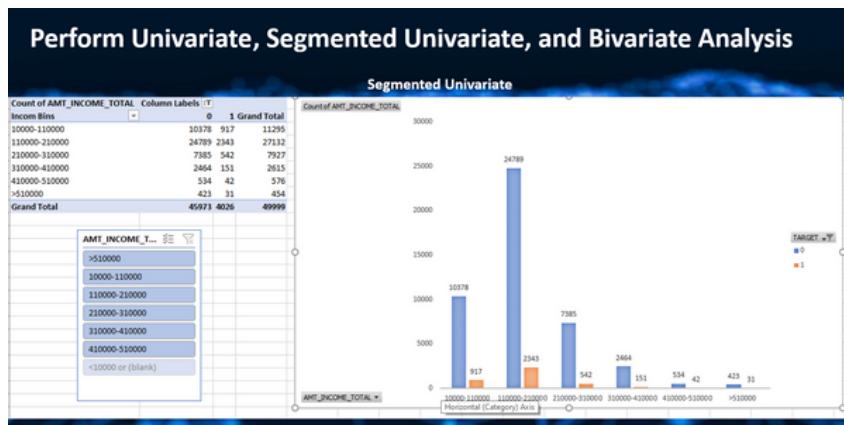
IN THIS FIG DETECT AND IDENTIFY OUTLIERS IN THE DATASET USING EXCEL STATISTICAL FUNCTIONS AND FEATURES, FOCUSING ON NUMERICAL VARIABLES.

**ANALYZE DATA IMBALANCE: DATA IMBALANCE CAN AFFECT THE ACCURACY OF THE ANALYSIS, ESPECIALLY FOR BINARY CLASSIFICATION PROBLEMS. UNDERSTANDING THE DATA DISTRIBUTION IS CRUCIAL FOR BUILDING RELIABLE MODELS.**



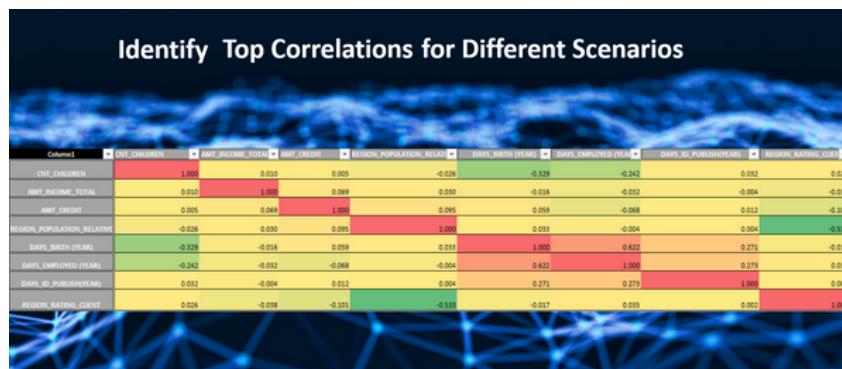
IN THIS FIG DETERMINE IF THERE IS DATA IMBALANCE IN THE LOAN APPLICATION DATASET AND CALCULATE THE RATIO OF DATA IMBALANCE USING EXCEL FUNCTIONS.

## PERFORM UNIVARIATE, SEGMENTED UNIVARIATE, AND BIVARIATE ANALYSIS: TO GAIN INSIGHTS INTO THE DRIVING FACTORS OF LOAN DEFAULT, IT IS IMPORTANT TO CONDUCT VARIOUS ANALYSES ON CONSUMER AND LOAN ATTRIBUTES.



IN THIS FIG PERFORM UNIVARIATE ANALYSIS TO UNDERSTAND THE DISTRIBUTION OF INDIVIDUAL VARIABLES, SEGMENTED UNIVARIATE ANALYSIS TO COMPARE VARIABLE DISTRIBUTIONS FOR DIFFERENT SCENARIOS, AND BIVARIATE ANALYSIS TO EXPLORE RELATIONSHIPS BETWEEN VARIABLES AND THE TARGET VARIABLE USING EXCEL FUNCTIONS AND FEATURES.

**IDENTIFY TOP CORRELATIONS FOR DIFFERENT SCENARIOS:** UNDERSTANDING THE CORRELATION BETWEEN VARIABLES AND THE TARGET VARIABLE CAN PROVIDE INSIGHTS INTO STRONG INDICATORS OF LOAN DEFAULT.



IN THIS FIG SEGMENT THE DATASET BASED ON DIFFERENT SCENARIOS (E.G., CLIENTS WITH PAYMENT DIFFICULTIES AND ALL OTHER CASES) AND IDENTIFY THE TOP CORRELATIONS FOR EACH SEGMENTED DATA USING EXCEL FUNCTIONS.

## **CONCLUSION**

IN CONCLUSION, THE EXPLORATORY DATA ANALYSIS (EDA) CONDUCTED ON THE LOAN APPLICATION DATASET HAS PROVIDED VALUABLE INSIGHTS INTO FACTORS INFLUENCING THE LIKELIHOOD OF LOAN DEFAULT FOR THE FINANCE COMPANY SPECIALIZING IN LENDING VARIOUS TYPES OF LOANS TO URBAN CUSTOMERS. HERE ARE THE KEY CONCLUSIONS DRAWN FROM THE ANALYSIS:

**1. HANDLING MISSING DATA:**

- MISSING DATA WERE IDENTIFIED AND APPROPRIATELY HANDLED USING METHODS SUCH AS IMPUTATION WITH MEAN, MEDIAN, OR MODE.
- VISUALIZATIONS WERE USED TO SHOWCASE THE PROPORTION OF MISSING VALUES FOR EACH VARIABLE, ENSURING TRANSPARENCY IN DATA HANDLING.

**2. IDENTIFYING OUTLIERS:**

- OUTLIERS IN THE DATASET WERE DETECTED AND IDENTIFIED, ENABLING FURTHER INVESTIGATION INTO POTENTIALLY ERRONEOUS OR UNUSUAL DATA POINTS.
- VISUAL REPRESENTATIONS SUCH AS BOX PLOTS OR SCATTER PLOTS AIDED IN HIGHLIGHTING OUTLIERS WITHIN THE DISTRIBUTION OF NUMERICAL VARIABLES.

**3. ANALYZING DATA IMBALANCE:**

- DATA IMBALANCE IN THE LOAN APPLICATION DATASET WAS ASSESSED, PROVIDING INSIGHTS INTO THE DISTRIBUTION OF LOAN APPROVAL OUTCOMES.
- VISUALIZATIONS SUCH AS PIE CHARTS OR BAR CHARTS EFFECTIVELY ILLUSTRATED THE CLASS DISTRIBUTION IMBALANCE, FACILITATING UNDERSTANDING OF DATA DISTRIBUTION.

**4. PERFORMING UNIVARIATE, SEGMENTED UNIVARIATE, AND BIVARIATE ANALYSIS:**

- UNIVARIATE, SEGMENTED UNIVARIATE, AND BIVARIATE ANALYSES WERE CONDUCTED TO EXPLORE THE DISTRIBUTIONS OF INDIVIDUAL VARIABLES, COMPARE DISTRIBUTIONS ACROSS DIFFERENT SCENARIOS, AND INVESTIGATE RELATIONSHIPS WITH THE TARGET VARIABLE.
- EXCEL FUNCTIONS AND FEATURES WERE UTILIZED TO PERFORM DESCRIPTIVE AND COMPARATIVE ANALYSES, UNCOVERING INSIGHTS INTO FACTORS INFLUENCING LOAN DEFAULT.

**5. IDENTIFYING TOP CORRELATIONS FOR DIFFERENT SCENARIOS:**

- CORRELATION ANALYSIS WAS PERFORMED ON SEGMENTED DATASETS TO IDENTIFY TOP CORRELATIONS BETWEEN VARIABLES AND THE TARGET VARIABLE WITHIN EACH SCENARIO.
- CORRELATION MATRICES OR HEATMAPS VISUALIZED THE RELATIONSHIPS BETWEEN VARIABLES, HIGHLIGHTING STRONG INDICATORS OF LOAN DEFAULT FOR DIFFERENT SCENARIOS.

OVERALL, THE EDA PROCESS HAS EQUIPPED THE FINANCE COMPANY WITH ACTIONABLE INSIGHTS TO MAKE INFORMED DECISIONS REGARDING LOAN APPROVAL, RISK ASSESSMENT, AND MITIGATION STRATEGIES. BY UNDERSTANDING THE PATTERNS AND FACTORS INFLUENCING LOAN DEFAULT, THE COMPANY CAN ENHANCE ITS LENDING PRACTICES, MINIMIZE FINANCIAL RISKS, AND IMPROVE OVERALL BUSINESS OUTCOMES. FURTHER ANALYSIS AND REFINEMENT OF STRATEGIES BASED ON THESE INSIGHTS CAN CONTRIBUTE TO THE COMPANY'S SUCCESS IN EFFECTIVELY MANAGING LOAN PORTFOLIOS AND SERVING ITS URBAN CUSTOMER BASE.

**REFERENCES LINK:  
BANK LOAN CASE STUDY**

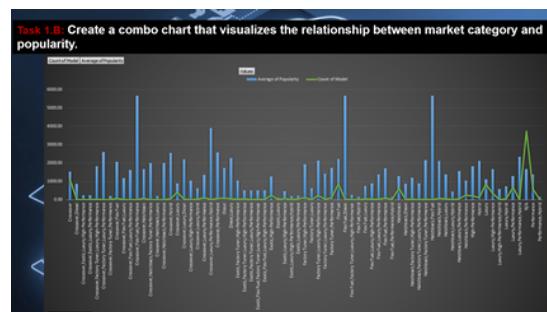
# **ANALYZING THE IMPACT OF CAR FEATURES ON PRICE AND PROFITABILITY**



## INSIGHT REQUIRED: HOW DOES THE POPULARITY OF A CAR MODEL VARY ACROSS DIFFERENT MARKET CATEGORIES?

**TASK 1.A: CREATE A PIVOT TABLE THAT SHOWS THE NUMBER OF CAR MODELS IN EACH MARKET CATEGORY AND THEIR CORRESPONDING POPULARITY SCORES.**

## **TASK 1.B: CREATE A COMBO CHART THAT VISUALIZES THE RELATIONSHIP BETWEEN MARKET CATEGORY AND POPULARITY.**



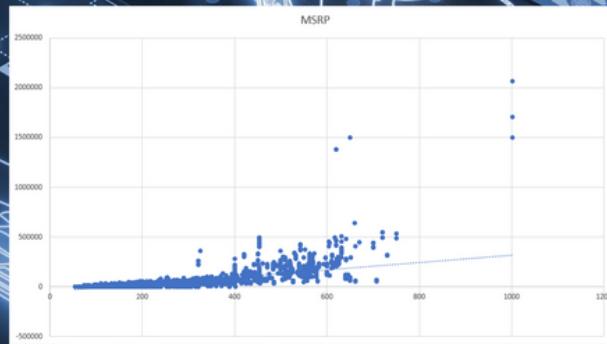
**Business and Data Analytics Skills:**  
How does the popularity of a car model vary across different market categories?  
**Task 1.A:** Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.

## INSIGHT REQUIRED: WHAT IS THE RELATIONSHIP BETWEEN A CAR'S ENGINE POWER AND ITS PRICE?

**TASK 2: CREATE A SCATTER CHART THAT PLOTS ENGINE POWER ON THE X-AXIS AND PRICE ON THE Y-AXIS. ADD A TRENDLINE TO THE CHART TO VISUALIZE THE RELATIONSHIP BETWEEN THESE VARIABLES.**

Insight Required: What is the relationship between a car's engine power and its price?

Task 2: Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.

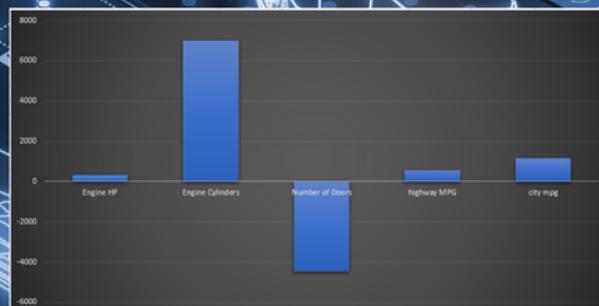


## INSIGHT REQUIRED: WHICH CAR FEATURES ARE MOST IMPORTANT IN DETERMINING A CAR'S PRICE?

**TASK 3: USE REGRESSION ANALYSIS TO IDENTIFY THE VARIABLES THAT HAVE THE STRONGEST RELATIONSHIP WITH A CAR'S PRICE. THEN CREATE A BAR CHART THAT SHOWS THE COEFFICIENT VALUES FOR EACH VARIABLE TO VISUALIZE THEIR RELATIVE IMPORTANCE.**

Insight Required: Which car features are most important in determining a car's price?

Task 3: Use regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.



# INSIGHT REQUIRED: HOW DOES THE AVERAGE PRICE OF A CAR VARY ACROSS DIFFERENT MANUFACTURERS?

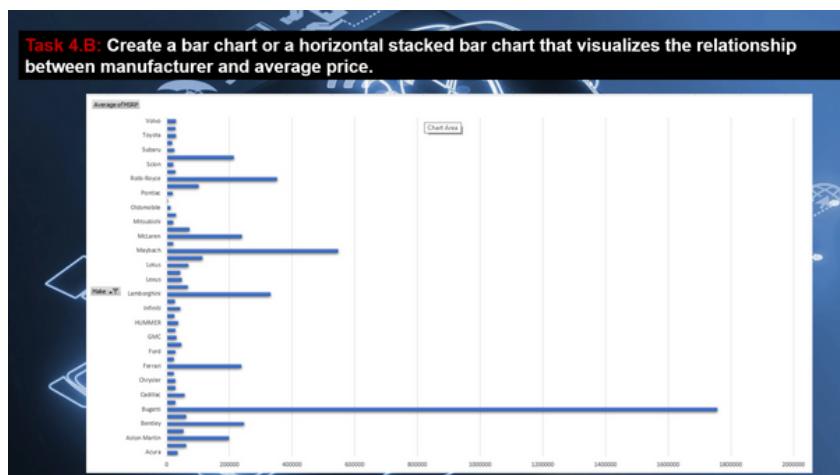
**TASK 4.A: CREATE A PIVOT TABLE THAT SHOWS THE AVERAGE PRICE OF CARS FOR EACH MANUFACTURER.**

**TASK 4.B: CREATE A BAR CHART OR A HORIZONTAL STACKED BAR CHART THAT VISUALIZES THE RELATIONSHIP BETWEEN MANUFACTURER AND AVERAGE PRICE.**

Insight Required: How does the average price of a car vary across different manufacturers?

Task 4.A: Create a pivot table that shows the average price of cars for each manufacturer.

manufacturers	Average of MSRP
Acura	34887.5873
Alfa Romeo	61600
Aston Martin	197910.3763
Audi	53452.1128
Bentley	247169.3243
BMW	61546.76347
Bugatti	1757223.667
Buick	28206.61224
Cadillac	56231.31738
Chevrolet	28273.35695
Chrysler	26722.96257
Dodge	22390.05911
Ferrari	237383.8235
FIAT	22206.01695
Ford	27393.42051
Genesis	46616.66667
GMC	30493.29903
Honda	26629.81879
HUMMER	36464.41176
Hyundai	24597.0363
Infiniti	42394.21212
Kia	25112.38938
Lamborghini	331567.3077



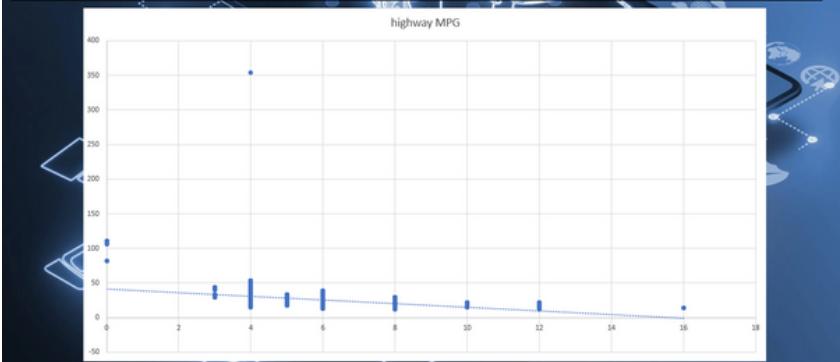
## INSIGHT REQUIRED: WHAT IS THE RELATIONSHIP BETWEEN FUEL EFFICIENCY AND THE NUMBER OF CYLINDERS IN A CAR'S ENGINE?

**TASK 5.A: CREATE A SCATTER PLOT WITH THE NUMBER OF CYLINDERS ON THE X-AXIS AND HIGHWAY MPG ON THE Y-AXIS. THEN CREATE A TRENDLINE ON THE SCATTER PLOT TO VISUALLY ESTIMATE THE SLOPE OF THE RELATIONSHIP AND ASSESS ITS SIGNIFICANCE.**

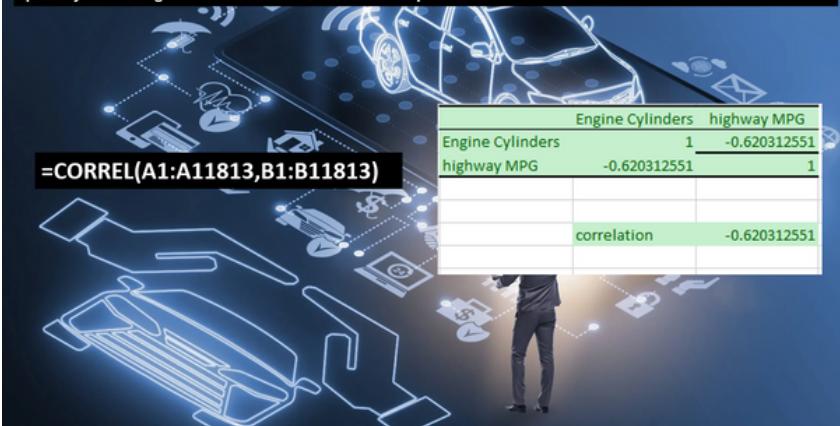
**TASK 5.B: CALCULATE THE CORRELATION COEFFICIENT BETWEEN THE NUMBER OF CYLINDERS AND HIGHWAY MPG TO QUANTIFY THE STRENGTH AND DIRECTION OF THE RELATIONSHIP.**

Insight Required: What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

**Task 5.A:** Create a scatter plot with the number of cylinders on the x-axis and highway MPG on the y-axis. Then create a trendline on the scatter plot to visually



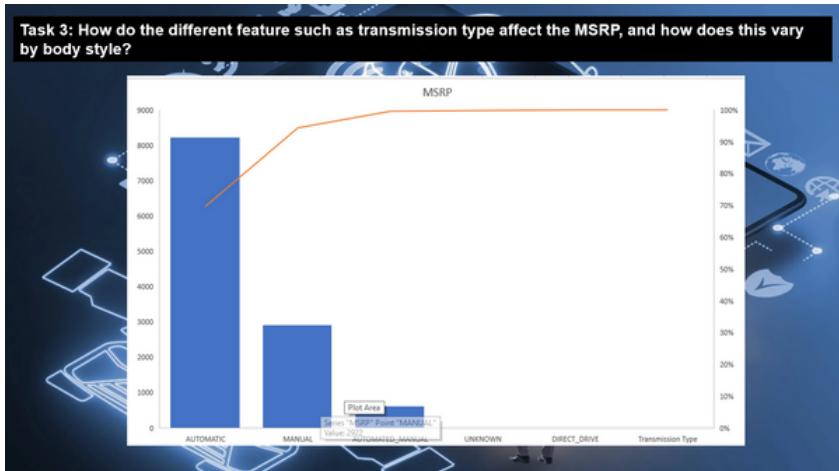
**Task 5.B:** Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.



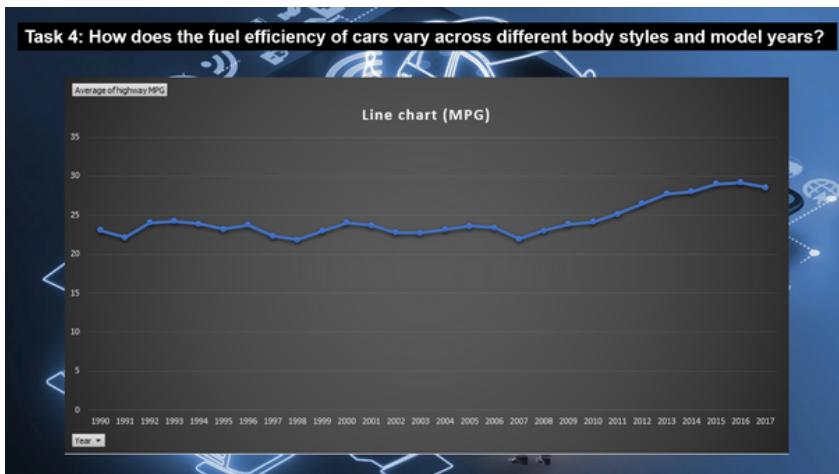
# TASK 1: HOW DOES THE DISTRIBUTION OF CAR PRICES VARY BY BRAND AND BODY STYLE?

Building the Dashboard																	
Task 1: How does the distribution of car prices vary by brand and body style?																	
<b>Summary of MRP Data</b>																	
Brand	Body Style	Adm SUV	Adm HB	Adm Hatchback	Adm MPV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Station Wagon	Unlisted
Acura	Integra	4529117	251440	2000000						129000	715200	175200	175200	175200	175200	175200	300000
Audi	Q5	40000		2074900					5291400	48000	175200	175200	175200	175200	175200	175200	300000
Audi	Q7	40000		2074900					5291400	48000	175200	175200	175200	175200	175200	175200	300000
Audi	Q8	60097	1144000	1600000					4952671	36000	175200	175200	175200	175200	175200	175200	300000
BMW	3 Series	2141770		179325					18534	330005						3214783	
BMW	5 Series	2141770		179325					18534	330005						3214783	
Buick	Regal	36000	16000	200000	20000	20000	20000	20000	20000	20000	20000	20000	20000	20000	20000	20000	
Cadillac	Sedan	80000	2E-005	1209738	8000000	4261180	79868	2083245	1003000	46100	599100	2285564	12711300	2431988	2299240	2285564	281570
Cadillac	SUV	80000	2E-005	1209738	8000000	4261180	79868	2083245	1003000	46100	599100	2285564	12711300	2431988	2299240	2285564	281570
Cadillac	XT5	325315		369365					472841	16000	864172	867426	70708	719408	2417685	793058	1401777
Cadillac	XT6	360000	5E-005	40155	437071	793007				16000	861253						131165
Cadillac	XTS	360000	5E-005	40155	437071	793007				16000	861253						131165
Chevrolet	Colorado	112200	2016270	5841919	142790	460085	282158	28100	4002424	2183800	18000	863270	1308328	2340498	1670449	2340498	
Chevrolet	Equinox	1030600	826880	277490					140000	724079							1670449
Chevrolet	Impala	1030600	826880	277490					140000	724079							1670449
Chevrolet	Malibu	1030600	826880	277490					140000	724079							1670449
Chevrolet	Trax	140000	826880	277490					140000	724079							1670449
Chevrolet	Volt	140000	826880	277490					140000	724079							1670449
Chevrolet	Z71	140000	826880	277490					140000	724079							1670449
Chevrolet	ZL1	140000	826880	277490					140000	724079							1670449
Chevrolet	ZR2	140000	826880	277490					140000	724079							1670449
Chevrolet	ZSh	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT1	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT2	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT3	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT4	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT5	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT6	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT7	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT8	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT9	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT10	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT11	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT12	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT13	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT14	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT15	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT16	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT17	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT18	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT19	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT20	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT21	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT22	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT23	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT24	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT25	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT26	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT27	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT28	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT29	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT30	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT31	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT32	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT33	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT34	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT35	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT36	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT37	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT38	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT39	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT40	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT41	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT42	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT43	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT44	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT45	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT46	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT47	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT48	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT49	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT50	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT51	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT52	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT53	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT54	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT55	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT56	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT57	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT58	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT59	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT60	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT61	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT62	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT63	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT64	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT65	140000	826880	277490					140000	724079							1670449
Chevrolet	ZT66	140000	826880	277490													

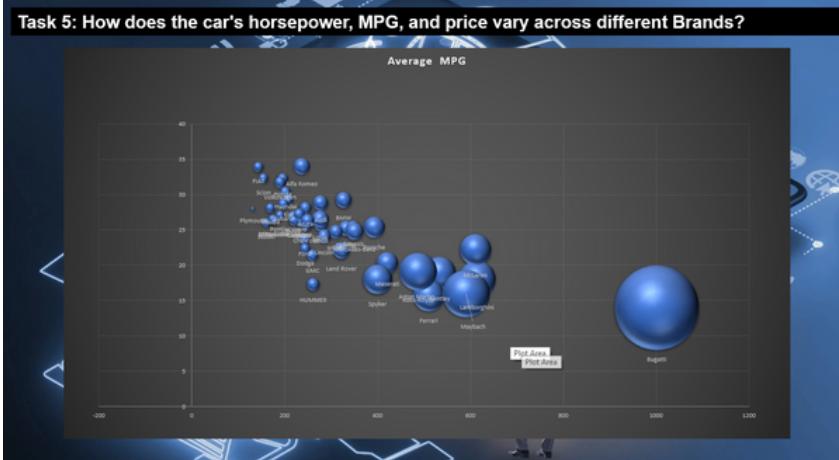
## **TASK 3: HOW DO THE DIFFERENT FEATURES SUCH AS TRANSMISSION TYPE AFFECT THE MSRP, AND HOW DOES THIS VARY BY BODY STYLE?**



## **TASK 4: HOW DOES THE FUEL EFFICIENCY OF CARS VARY ACROSS DIFFERENT BODY STYLES AND MODEL YEARS?**



## TASK 5: HOW DOES THE CAR'S HORSEPOWER, MPG, AND PRICE VARY ACROSS DIFFERENT BRANDS?

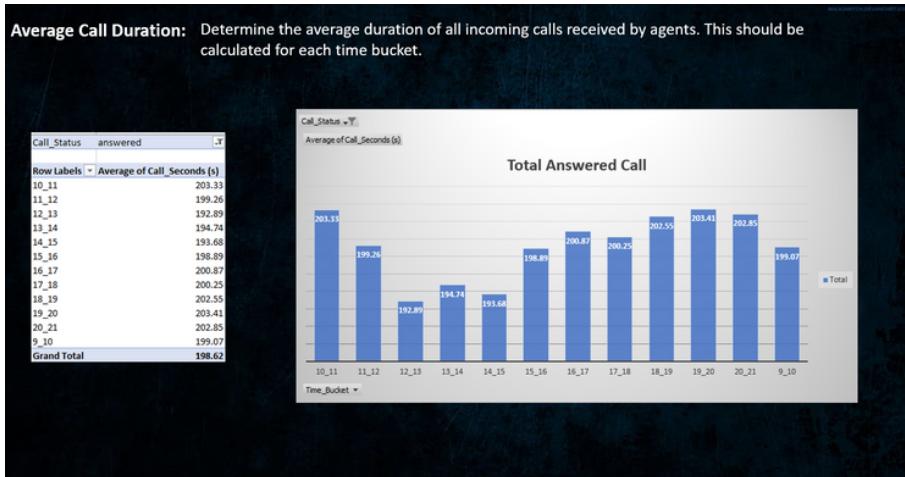


**REFERENCES LINK:  
IMPACT OF CAR FEATURES**

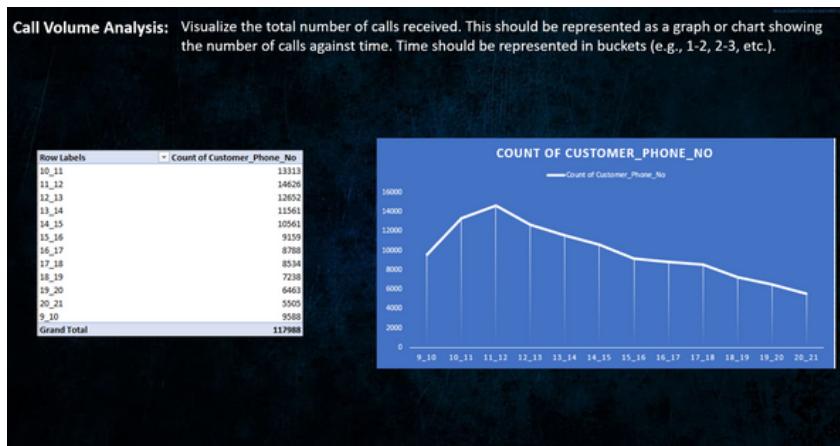
## ABC CALL VOLUME TREND ANALYSIS



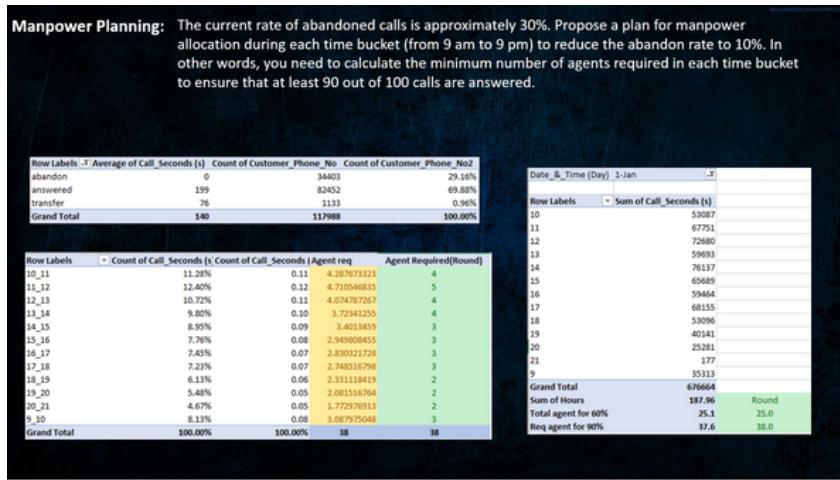
**AVERAGE CALL DURATION: DETERMINE THE AVERAGE DURATION OF ALL INCOMING CALLS RECEIVED BY AGENTS. THIS SHOULD BE CALCULATED FOR EACH TIME BUCKET.**



# CALL VOLUME ANALYSIS: VISUALIZE THE TOTAL NUMBER OF CALLS RECEIVED. THIS SHOULD BE REPRESENTED AS A GRAPH OR CHART SHOWING THE NUMBER OF CALLS AGAINST TIME. TIME SHOULD BE REPRESENTED IN BUCKETS (E.G., 1-2, 2-3, ETC.).



**MANPOWER PLANNING: THE CURRENT RATE OF ABANDONED CALLS IS APPROXIMATELY 30%. PROPOSE A PLAN FOR MANPOWER ALLOCATION DURING EACH TIME BUCKET (FROM 9 AM TO 9 PM) TO REDUCE THE ABANDON RATE TO 10%. IN OTHER WORDS, YOU NEED TO CALCULATE THE MINIMUM NUMBER OF AGENTS REQUIRED IN EACH TIME BUCKET TO ENSURE THAT AT LEAST 90 OUT OF 100 CALLS ARE ANSWERED.**



- 1. NIGHT SHIFT MANPOWER PLANNING: CUSTOMERS ALSO CALL ABC INSURANCE COMPANY AT NIGHT BUT DON'T GET AN ANSWER BECAUSE THERE ARE NO AGENTS AVAILABLE. THIS CREATES A POOR CUSTOMER EXPERIENCE. ASSUME THAT FOR EVERY 100 CALLS THAT CUSTOMERS MAKE BETWEEN 9 AM AND 9 PM, THEY ALSO MAKE 30 CALLS AT NIGHT BETWEEN 9 PM AND 9 AM. THE DISTRIBUTION OF THESE 30 CALLS IS AS FOLLOWS:**

- ## **2. YOUR TASK: PROPOSE A MANPOWER PLAN FOR EACH TIME BUCKET THROUGHOUT THE DAY, KEEPING THE MAXIMUM ABANDON RATE AT 10%**

**ASSUMPTIONS: AN AGENT WORKS FOR 6 DAYS A WEEK; ON AVERAGE, EACH AGENT TAKES 4 UNPLANNED LEAVES PER MONTH; AN AGENT'S TOTAL WORKING HOURS ARE 9 HOURS, OUT OF WHICH 1.5 HOURS ARE SPENT ON LUNCH AND SNACKS IN THE OFFICE. ON AVERAGE, AN AGENT SPENDS 60% OF THEIR TOTAL ACTUAL WORKING HOURS (I.E., 60% OF 7.5 HOURS) ON CALLS WITH CUSTOMERS/USERS. THE TOTAL NUMBER OF DAYS IN A MONTH IS 30.**

Distribution of 30 calls coming in night for every 100 calls coming in between 9am - 9pm (i.e. 12 hrs slot)												
10pm	10pm - 11pm	11pm- 12am	12am- 1am	1am - 2am	2am - 3am	3am - 4am	4am - 5am	5am - 6am	6am - 7am	7am - 8am	8am - 9am	
3	3	2	2	1	1	1	1	3	4	4	5	

**Night Shift Manpower Planning:** Customers also call ABC Insurance Company at night but don't get an answer because there are no agents available. This creates a poor customer experience. Assume that for every 100 calls that customers make between 9 am and 9 pm, they also make 30 calls at night between 9 pm and 9 am. The distribution of these 30 calls is as follows:

Call Center Performance Metrics - Q4 2023									
Time Period	Call Type	Call Volume		Agent Utilization		Customer Satisfaction		Operational Efficiency	
		Inbound	Outbound	Avg Call Duration	Agents Available	Wait Time Avg	Success Rate	Abandon Rate	Transfer Rate
Q4_2023	Total	1000	1000	120m	100	15m	95%	5%	2%
Q4_2023	Day	500	500	120m	50	15m	95%	5%	2%
Q4_2023	Night	500	500	120m	50	15m	95%	5%	2%
Q4_2023	Weekend	500	500	120m	50	15m	95%	5%	2%
Q4_2023	Month	1000	1000	120m	100	15m	95%	5%	2%
Time		Call		Time		Agent		Round	
9_10		3	10.00	1.00	1.00				
10_11		3	10.00	1.00	1.00				
11_12		2	15.00	0.67	1.00				
12_13		2	15.00	0.67	1.00				
13_14		1	30.00	0.33	0.00				
14_15		1	30.00	0.33	0.00				
15_16		1	30.00	0.33	0.00				
16_17		1	30.00	0.33	0.00				
17_18		3	10.00	1.00	1.00				
18_19		4	7.50	1.33	1.00				
19_20		4	7.50	1.33	1.00				
20_21		5	6.00	1.67	2.00				
Total		30	1.00	10	9				
Grand Total									
Aug call		Round							
Call 30%		5129.913		5130					
Call 30%		1518.9739		1539					
Hour Req for 90%									
Hour Req for 90%		76.420112		76					
Agent Req		10.133333		10					
Count of Call_Status									
Row Labels		abandon		answered		transfer		Grand Total	
1-Jan		684		3883		77		464	
2-Jan		356		2935		60		335	
3-Jan		599		4079		111		476	
4-Jan		595		4404		114		511	
5-Jan		536		4140		114		484	
6-Jan		991		3875		85		495	
7-Jan		1218		3875		42		487	
8-Jan		1103		3519		50		461	
9-Jan		962		2628		62		360	
10-Jan		1212		3699		72		491	
11-Jan		856		3695		86		466	
12-Jan		1299		3297		47		466	
13-Jan		738		3326		59		411	
14-Jan		291		2832		32		311	
15-Jan		304		2770		24		307	
16-Jan		1191		3910		41		514	
17-Jan		16636		5706		5		2234	
18-Jan		1738		4024		12		57	
19-Jan		974		3717		12		474	
20-Jan		833		3485		43		505	
21-Jan		566		3104		5		367	
22-Jan		239		3045		7		325	
23-Jan		381		2832		12		322	
Grand Total		34403		82452		11133		11796	
Average		1496		3885		49		511	

## **CONCLUSION**

**IN CONCLUSION, OUR ANALYSIS OF THE INBOUND CALL DATA FOR ABC INSURANCE COMPANY HAS PROVIDED VALUABLE INSIGHTS INTO CALL VOLUME PATTERNS AND MANPOWER PLANNING STRATEGIES TO ENHANCE CUSTOMER EXPERIENCE AND OPERATIONAL EFFICIENCY. HERE'S A SUMMARY OF OUR KEY FINDINGS AND PROPOSED SOLUTIONS:**

**1. AVERAGE CALL DURATION:**

- WE DETERMINED THE AVERAGE DURATION OF CALLS FOR EACH TIME BUCKET, ENABLING US TO IDENTIFY PERIODS OF HIGH CALL VOLUME AND ASSESS AGENT WORKLOAD.

**2. CALL VOLUME ANALYSIS:**

- VISUALIZATIONS WERE CREATED TO ILLUSTRATE THE TOTAL NUMBER OF CALLS RECEIVED IN EACH TIME BUCKET, REVEALING PEAK CALL HOURS AND HELPING IN RESOURCE ALLOCATION.

**3. MANPOWER PLANNING:**

- PROPOSED PLANS FOR MANPOWER ALLOCATION DURING EACH TIME BUCKET WERE DEVISED TO REDUCE THE ABANDON RATE TO 10%, ENSURING THAT A SUFFICIENT NUMBER OF AGENTS ARE AVAILABLE TO HANDLE INCOMING CALLS EFFICIENTLY.

**4. NIGHT SHIFT MANPOWER PLANNING:**

- MANPOWER PLANS FOR EACH TIME BUCKET, INCLUDING THE NIGHT SHIFT, WERE DEVELOPED TO MAINTAIN A MAXIMUM ABANDON RATE OF 10%, CONSIDERING THE DISTRIBUTION OF CALLS AT NIGHT AND AGENT AVAILABILITY ASSUMPTIONS.

BY IMPLEMENTING THESE STRATEGIES, ABC INSURANCE COMPANY CAN OPTIMIZE ITS WORKFORCE, IMPROVE CALL HANDLING EFFICIENCY, AND ENHANCE OVERALL CUSTOMER SATISFACTION. THESE INSIGHTS AND RECOMMENDATIONS SERVE AS VALUABLE TOOLS FOR THE COMPANY'S CX TEAM TO MAKE INFORMED DECISIONS AND DRIVE IMPROVEMENTS IN THEIR INBOUND CALLING OPERATIONS.

**REFERENCES LINK:**

**ABC CALL VOLUME TREND ANALYSIS**

# **APPENDIX**

**DATA ANALYTICS PROCESS:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1P00UDJD\_BHX3DQVB5UWF0DWZUSKWYCAW?USP=DRIVE\_LINK**

**INSTAGRAM USER ANALYTICS:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1CVQUM7DNNVLRFOXUWXCGH\_D7RLBHHSEW?USP=DRIVE\_LINK**

**OPERATION & METRIC ANALYTICS:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1W60MP4HBT7KP\_NF8FE4G1X5HHHHBDI0G?USP=DRIVE\_LINK**

**HIRING PROCESS ANALYTICS:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1UT92B38NVRCF8URUJTBRV5X4UTKMQEQQ?USP=DRIVE\_LINK**

**IMDB MOVIE ANALYSIS:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/10TXQC6IODHAMJ5DV623WS-Y-3TRKJVOS?USP=DRIVE\_LINK**

**BANK LOAN CASE STUDY:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1UPKUFV3GVLJY5YSQ-0GA8QEUV6XAXKUE?USP=DRIVE\_LINK**

**IMPACT OF CAR FEATURES:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/127RZVVH4PNUKYGGFYLAC9FHCMDRLJOMV?USP=DRIVE\_LINK**

**ABC CALL VOLUME TREND:-**

**HTTPS://DRIVE.GOOGLE.COM/DRIVE/FOLDERS/1TMUREZKHVBGW-2WMYJLB08BSNOD\_GFCW?USP=DRIVE\_LINK**

**THANK YOU**  
**THANK YOU**  
**THANK YOU**  
**THANK YOU**  
**THANK YOU**

**SURAJ  
KUMAR**