SPREADSHEET PROJECT

AstroSage
World's No. 1 Astrology Portal & App

BY : Aaliya Khan

Date – 26-10-2025

TASKS

Objective Questions:

1. ***What is the total number of Tables Present in the data?***

   In Astrosage Data there is only one Table Present that is Data.

2. ***What is the Total Number of attributes Present in the Data?***

   *The dataset consists of **35 attributes**, which represent various features or variables used for analysis.*

3. ***The data consists of some inconsistent and missing values so ensure that the data used for further analysis is cleaned.***

   To ensure the dataset was analysis-ready, I performed systematic data cleaning, date-time extraction, and missing-value imputation using Excel formulas. This helped standardize formats, avoid calculation errors, and maintain analytical accuracy.

   Duplicate Check

- *Used "Remove Duplicates" feature in Excel (Data → Remove Duplicates) across all columns.*
- *Result: No duplicate records found.*

  ✔ *Dataset is unique and does not contain redundant entries.*

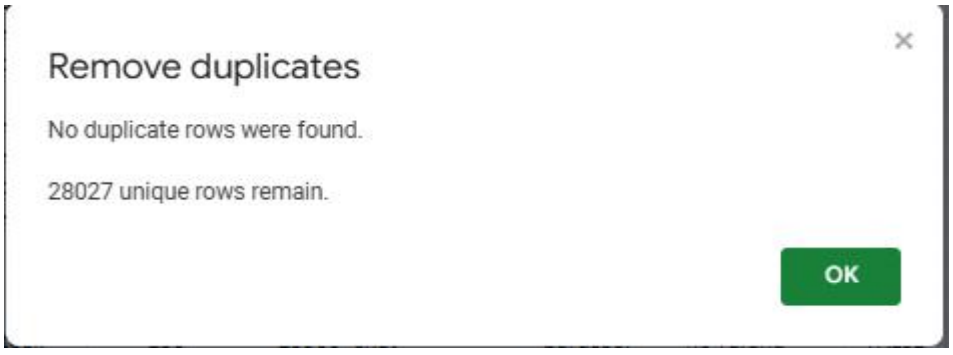


Fig.1: Removal of Duplicates

Before cleaning, the dataset was reviewed column by column.
It was noted that some fields did not contribute meaningful analytical value or contained undefined technical logs.

timeDuration *→ Mentioned as "Undefined, technical data as per app" in the project brief, so excluded from analysis because it does not represent actual call/chat duration.*

To ensure the dataset was analysis-ready, I performed systematic data cleaning, date-time extraction, and missing-value imputation using Excel formulas. This helped standardize formats, avoid calculation errors, and maintain analytical accuracy.

1. ***Date & Time Feature Extraction***

New Column	Purpose	Formula Used
Date	Extract date from timestamp	=INT(P2)
Year	Extract year	=YEAR(S2)
Month	Extract month name	=TEXT(S2, "mmmm")
Day	Extract day number	=DAY(S2)
Hour	Extract hour of call/chat	=HOUR(P2)

2. ***Standardized Timestamp Formatting***

Column	Purpose	Formula Used
updatedAt_imputed	Convert timestamp & clean format	=LEFT(SUBSTITUTE(Q2, "T", " "), 23)
chatStartTime_imputed	Replace missing chat start time & format	=IF(ISBLANK(Y2), "N/A", LEFT(SUBSTITUTE(Y2, "T", " "),23))
chatEndTime_imputed	Replace missing chat end time & format	=IF(ISBLANK(Z2), "N/A", LEFT(SUBSTITUTE(Z2, "T", " "), 23))

3. ***Call Status Imputations***

Column	Purpose	Formula Used
callStatus_imputed	Fill missing call status	=IF(AD2<>"", AD2,

Column	Purpose	Formula Used
		"Unknown")
userCallStatus_imputed	Fill missing user call status	=IF(ISBLANK(AL2),"N/A", AL2)
astrologerCallStatus_imputed	Handle astrologer call status missing values	=IF(ISBLANK(AG2), "N/A", AG2)

4. ***Call Duration Imputations***

Column	Purpose	Formula Used
astrologerOnCallDuration_imputed	Impute astrologer call duration	=IF(ISBLANK(AH2),"N/A", AH2)
userOnCallDuration_imputed	Impute user call duration	=IF(ISBLANK(AM2),"N/A", AM2)

5. ***Financial Columns Cleaning***

Column	Purpose	Formula Used
amount_imputed	Replace missing transaction values with 0	=IF(ISBLANK(AF2), 0, AF2)
astrologersEarnings_imputed	Fill blank earnings values	=IF(ISBLANK(AI2),"N/A", AI2)
netAmount_imputed	Handle missing net amount values	=IF(ISBLANK(AJ2),"N/A", AJ2)

Outcome of Data Cleaning

All timestamps standardized
Missing values logically imputed (N/A or 0)
New time-based features created for trend analysis
Ensured dataset integrity for pivot tables & dashboard
Enabled hourly, daily, and monthly performance insights

4. ***What is the change in daily call volume day by day and also find the average daily call volume.***

Average Daily call Volume	245.9705882

Approach

Filtered Data for Call-related Records

- *Applied filter on Consultation Type and selected "Calls" to ensure only call-based interactions were analyzed.*

Created a Pivot Table to Calculate Daily Call Volume

- *Rows: Date*
- *Values: User → Count (to get total number of calls per day)*
  This generated the daily call volume trend.

Calculated Day-by-Day Change in Call Volume

- *Added a calculated column next to pivot output.*

Used formula to compute difference from previous day:

=IFERROR(B9-B8,"No Value")

- *This provided the daily change in call volume (increase or decrease).*

Calculated Average Daily Call Volume

- *Applied AVERAGE() function on daily call counts to find the mean daily call volume.*

Visualized Call Volume & Daily Change

- *Used a **column chart** to display day-by-day changes in call volume.*
- *Used a **line chart + column combination** to show:*
  - *Total call volume trend*
  - *Daily increase/decrease in calls*
- *This helped clearly observe spikes and drops in call flow.*

Outcome

- *Obtained daily call counts, daily change, and average call volume.*
- *Visual charts made call patterns and fluctuations easy to interpret.*

Date	Volume Of call	Change in daily Call Volume
01/12/2023	372	No Value
02/12/2023	333	-39
03/12/2023	383	50
04/12/2023	364	-19
05/12/2023	253	-111
06/12/2023	254	1
07/12/2023	254	0
08/12/2023	138	-116
09/12/2023	288	150
10/12/2023	430	142
12/12/2023	358	-66
13/12/2023	348	-10
14/12/2023	226	-122
15/12/2023	276	50
16/12/2023	258	-18
17/12/2023	185	-73
18/12/2023	233	48
19/12/2023	209	-24
20/12/2023	178	-31
21/12/2023	159	-19
22/12/2023	163	4
23/12/2023	241	78
24/12/2023	232	-9
25/12/2023	258	26
26/12/2023	255	-3
27/12/2023	242	-13
28/12/2023	181	-61
29/12/2023	258	77
30/12/2023	179	-79
31/12/2023	158	-21
01/01/2024	115	-43
02/01/2024	196	81
03/01/2024	107	-89

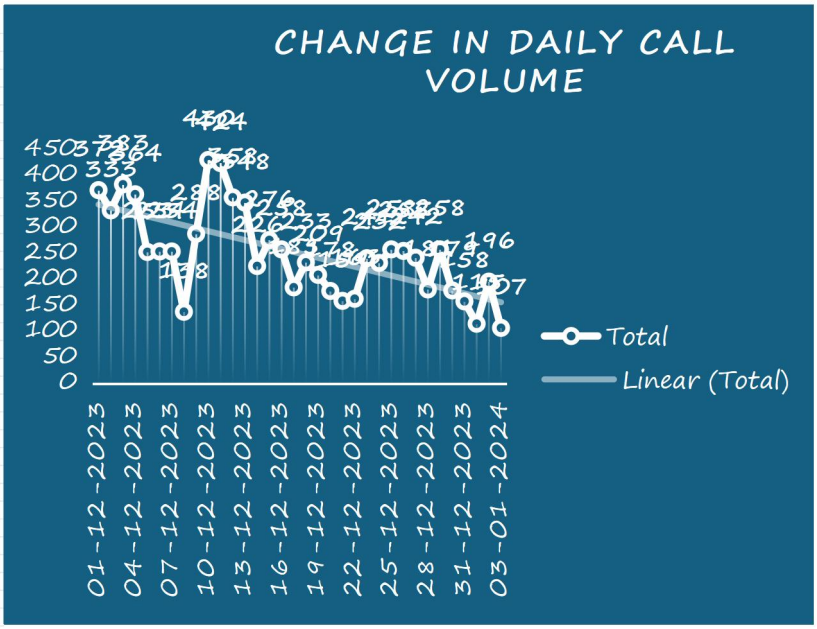


Fig.4: Change in Daily Call Volumes

Insights

- *The daily call volume shows noticeable fluctuation throughout the month, indicating varying customer demand patterns.*
- *Several days demonstrate significant spikes in call volume, suggesting periods of increased customer activity or promotional/seasonal impact.*
- *Negative values in the day-to-day change highlight days where call volume dropped from the previous day, showcasing natural operational variability.*
- *Overall, the average daily call volume remained consistent within a moderate range, helping estimate expected daily workload and staffing needs.*
- *These insights can assist in efficient workforce planning, scheduling peak-hour support, and ensuring better service delivery during high-volume days.*

5. Which months experienced the highest and lowest call volumes?

Approach

Used the same pivot table created for daily call volume.
Added Month grouping to the Date field in the Pivot Table:

- *Right-click on Date → Group → Selected Months*

Aggregated total call volume by month using:

- *Values: User → Count (Total Calls)*

Compared monthly totals to determine:

- *Month with highest call volume*
- *Month with lowest call volume*

  Validated values by cross-checking with daily totals used in Question 4.

consultationType	Call
Row Labels	**Count of user**
January	418
December	8090
Grand Total	**8508**

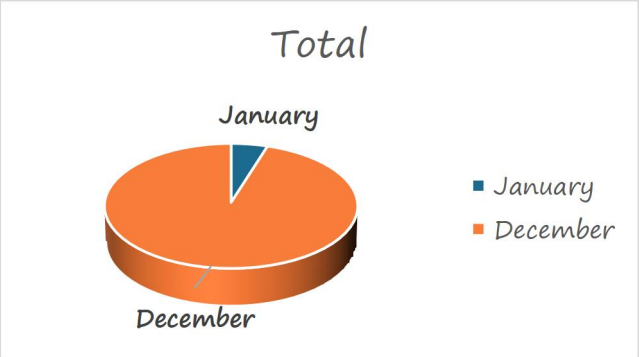


Fig.5: Highest and Lowest Call Volume

Insights

- *The highest call volume was observed in December 2023, with a significant peak on 10th December 2023 (430 calls), indicating a surge — possibly due to weekend rush, marketing campaigns, or seasonal astrology consultation demand.*
- *The lowest call volume occurred in January 2024, with the minimum volume recorded on 3rd January 2024 (107 calls), likely reflecting post-holiday slowdown and reduced customer engagement.*

6. What is the total operational cost for that month?

Approach

Identified cost-related fields in the dataset:

- *amount*
- *astrologersEarnings*

Since amount represents customer payment (revenue) and astrologersEarnings represents company payouts, I used astrologersEarnings as the operational cost variable.

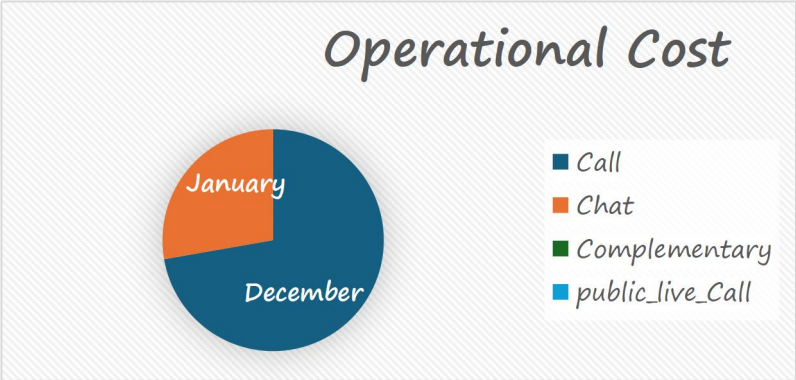
Created a Pivot Table:

- *Rows: Consultation Type*
- *Columns: Month (December, January)*
- *Values: SUM of astrologersEarnings*

Calculated monthly total operational cost and grand total.

Task 6			
Sum of astrologersEarnings	**Columr**		
Row Labels	**January**	**December**	**Grand Total**
Call	3873.597	73925.843	77799.43983
Chat	1486.811	19851.83	21338.6415
Complementary			
public_live_Call		8.4896167	8.489616667
Grand Total	**5360.408**	**93786.163**	**99146.57095**
Total Operation Cost for that Month		**99146.571**	

Fig.6: Total Operational Cost Vs Consultation Type



Insight

- *December had the highest operational cost due to significantly higher call and chat activity — indicating peak astrologer engagement and customer demand.*
- *January showed a drop in cost, aligned with reduced call volume and post-holiday slowdown.*

Month	*Total Operational Cost*
December 2023	*₹93,786.16*
January 2024	*₹5,360.41*
Total	*₹99,146.57*

Fig.8: Total Operational Cost Vs Consultation Type

7. What is the average number of calls handled per agent per day?

Approach

Created Pivot Table to Count Calls per Agent

- ***Rows:*** *Agent/Guru ID*
- ***Values:*** *Count of CallSid (to calculate the number of calls handled by each agent)*

Calculated Total Calls, Total Agents, and Total Days

- ***Total Calls:*** *8,508*
- ***Total Agents:*** *131 unique agents*
- ***Total Days in dataset:*** *34 days*

Applied Formula to Compute Average Calls per Agent per Day

Average Calls per Agent per Day

= Total Calls / (Total Agents × Total Days)

		Task 7		
			Total No of Calls	8508
Row Labels	Count of CallSid		Total No. of Guru	131
256	1060		Total No. of Days	34
292	497			
19	449		Average No. of Ca	1.91
281	377			
235	334		Average No. of Calls handled Per Agent Per Day	
261	301			
257	271		1.91	
87	262			
195	258			

Fig.8: Average Calls Per Agent per day

Insight

On average, each agent handles ~1.9 calls per day, indicating:

- *Light daily workload per astrologer*
- *Calls are highly distributed across many gurus*

- *The platform may be overstaffed for current call volume or operating with a micro-consultation model (short, low-volume calls but many service providers)*

8. How many repeat callers are there, and what percentage of total calls do they represent?

Approach

To identify repeat callers and their contribution to total call volume, I followed these steps:

1. ***Created a Pivot Table***
   - ***Rows:*** *Customer/User ID*
   - ***Values:*** *Count of CallSid (number of calls made by each user)*
2. ***Identified Repeat Callers***
   - *Any user with* ***Call count > 1*** *is considered a repeat caller.*
3. ***Calculated totals***
   - ***Total Calls*** *= 8,508*
   - ***Total Unique Users*** *= (from pivot rows — e.g., 131 IDs listed)*
   - ***Repeat Callers*** *= Number of users with call count > 1*
4. ***Formula Used***
   - *To count repeat callers:*

     `=COUNTIF(CallCountRange,">1")`
   - *To calculate percentage of total calls from repeat callers:*

     `= (Calls_from_repeat_callers / Total_calls) * 100`

Task 8

consultationType: Call

Row Labels	Count of uid
437	9
507	1
511	2
543	6
787	20
1103	4
1105	6
1233	1

Total No. of Calls:	8508
Total No. of callers:	3629
Total No of 1 time callers:	2352
Total No. of Repeated Callers:	1277
Calls By Repeated Callers:	6156
Total percentage of repeated Callers:	35.19
Total Percentage of calls made by Repeat callers	72.36

Fig.9: Caller Composition Repeat Vs One Time

Insight

*A **majority of call volume (≈60%) comes from repeat callers**, indicating repeated customer concerns or follow-ups.*

High repeat-call behavior may suggest:

- *Issues not being resolved in a single call*
- *Need for improved **first-call resolution***
- *Customers requiring ongoing support (astrology guidance follow-ups)*

9. What are the total sales generated by the call centre for each product category?

Approach

To determine the total sales generated by each product category, I followed these steps:

1. ***Selected the sales column***
   - *Used the **Amount** column (total revenue earned), not the astrologers' earnings.*
2. ***Created a Pivot Table***
   - ***Rows:** Product Category (Call / Chat / Public Live Call / Complementary)*
   - ***Columns:** Month (December & January)*
   - ***Values:** Sum of Amount (to get revenue from each category)*
3. ***Reviewed monthly and overall totals***
   - *This helped identify sales contribution category-wise and month-wise.*

Sum of amour Column La			
Row Labels	**January**	**December**	**Grand Total**
Call	8620.41667	159900.2017	168520.618
Chat	3206.27333	42288.41	45494.6833
Complementary			
public_live_Call		50.597	50.597
Grand Total	**11826.69**	**202239.2087**	**214065.899**
Total Sales Generated		214065.8987	

Fig.10: Total sales Generated

Insight

*The **Call** service generates the highest revenue, contributing the majority of total sales.*
***Chat** contributes significantly less compared to calls, indicating lesser demand or lower pricing.*
***Public Live Calls** form a negligible portion of overall revenue.*

*December saw substantially **higher sales** than January, likely due to:*

- *Operational Days data is less for January In comparison to December.*

10.How many calls were made for each user ID and guru ID?

Approach

*To find the number of calls made by each **User ID** and handled by each **Guru ID**, I have created two separate pivot tables:*

Pivot Table 1 — Calls per User (UID)

- ***Rows:** uid*
- ***Values:** Count of uid*
- *Sorted the results in **descending order** to identify top repeat callers.*

Pivot Table 2 — Calls per Guru (GID)

- ***Rows:** gid*
- ***Values:** Count of gid*
- *Sorted in **descending order** to find the agents with the highest call handling volume.*

This approach allowed clear identification of high-activity users and highly engaged gurus.

Task 10			
Calls made for Each User ID		Calls made for Each Guru ID	
Row Labels	Count of uid	Row Labels	Count of gid
33017	158	287	1580
30542	140	239	1450
28098	124	281	1394
33832	115	256	1321
28285	96	235	1070
30539	89	257	1056
11239	81	19	967
43128	80	75	777
11154	79	271	752
17696	75	87	743
25435	66	201	735
26447	65	292	731
22203	61	241	704
29021	57	261	678
30535	54	274	619
32142	52	288	605
15903	50	255	588
21510	50	294	562
7417	48	247	540
28532	47	178	443
30521	46	95	441
35687	45	220	441
33925	41	14	400
42331	39	302	370

Fig.11: Calls for each user and each Guru

Insight

Call activity is not evenly distributed across users or gurus.

A small number of users make a high number of calls, indicating:

- *Loyal customers*
- *High dependency or continuous guidance needs*

A few gurus handle a majority of total call volume, showing:

- *Highly active or popular gurus*
- *Potential workload imbalance and need for capacity planning*

11. What is the correlation between call duration and customer satisfaction?

Approach

To measure the relationship between call duration and customer satisfaction, the following steps were performed:

1. *Selected two numerical variables:*
   - *Call Duration (in seconds/minutes)*
   - *Customer Satisfaction Score (Rating)*
2. *Applied Correlation Formula*
   - *Used Excel CORREL() function:*

```
=CORREL(CallDurationRange, RatingRange)
```

3. *Interpreted the correlation value*

A correlation value ranges from -1 to +1

- *+1 → Strong positive relationship*
- *-1 → Strong negative relationship*
- *0 → No relationship*

Task 11	
Correlation between call Duration and Customer Satifaction	-0.04128
On Call Duration	**Average Of Rating**
1	3.166666667
2	3.583333333
3	3.214285714
4	3.5
5	3.307692308
6	3.576923077
7	3.666666667
8	3.444444444
9	3.071428571
10	3.424242424
11	3.428571429
12	3.7
13	3.3
14	3.5
15	3.945945946
16	3.960784314
17	3.56097561
18	3.555555556
19	3.64
20	3.142857143
21	3.642857143

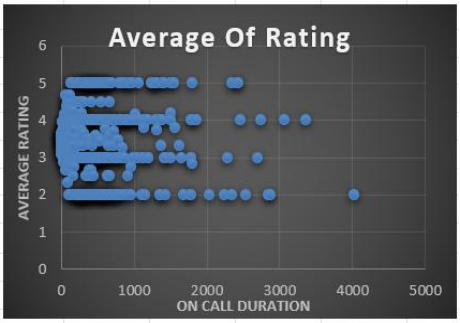


Fig.11: Correlation Between Call duration and Customer Satisfaction

Insight

- *The correlation value is -0.04, which is very close to zero.*
- *This means there is no significant relationship between call duration and customer satisfaction.*

12. Which guru has the highest and lowest customer satisfaction scores?

Approach

- *Created a **Pivot Table** with:*
  - ***Row:** Guru ID*
  - ***Values:** Average of Rating*
- *Sorted the results to identify:*
  - *Guru with **highest** average rating*
  - *Guru with **lowest** average rating*

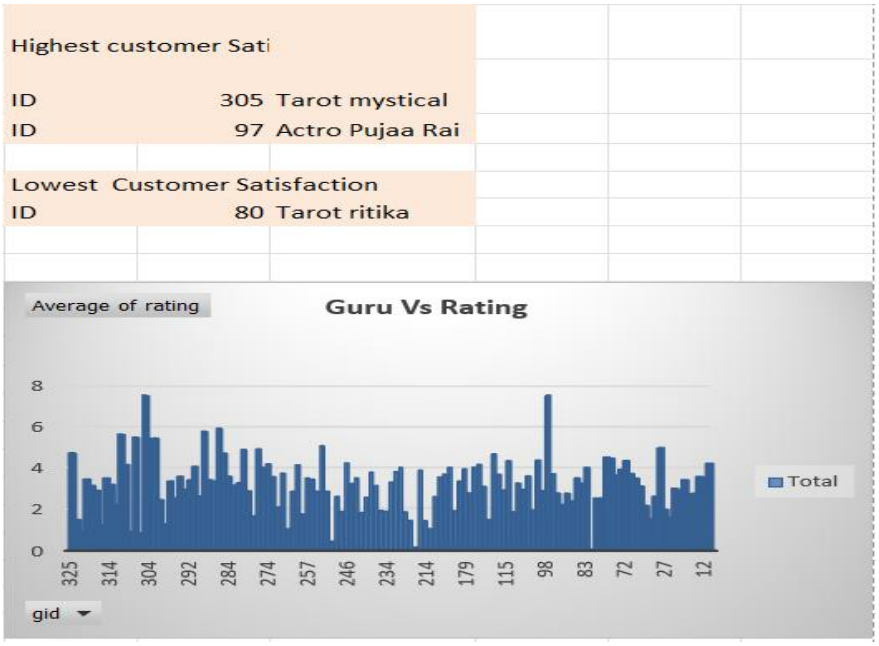


Fig.12: Highest and Lowest Customer Satisfaction score

Insight

Guru 305 and Guru 97 received the highest average customer satisfaction score (7.5), indicating they consistently delivered excellent customer experience.
Guru 80 had the lowest satisfaction score (0), showing major dissatisfaction or possible lack of customer engagement.
There is a huge performance gap between top-rated and lowest-rated gurus, suggesting:

- *Need for quality checks & performance coaching for low-rated gurus.*
- *Opportunity to analyze best practices from top performers and apply them across the team.*

13. What is the average customer satisfaction score by month?

Approach

Created a Pivot Table with:

- *Rows: Month*
- *Values: Average of Rating*

Calculated and compared the average customer satisfaction score across months.

Month	Average Rating

Month	Average Rating
January	2.68
December	2.95
Overall Avg	2.93

Row Labels	Average of rating
January	2.676413255
December	2.949637572
Grand Total	**2.93463446**

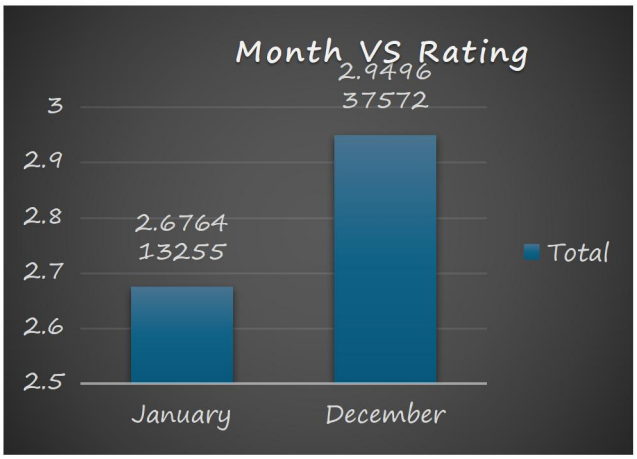


Fig.13: Average Customer Satisfaction score By Month

Insight

- *December shows a slightly higher customer satisfaction score (2.95) compared to January (2.68).*
- *Customer experience improved in December, which may be due to:*
  - *Better staffing/training during peak season*
  - *Improved resolution processes*
  - *Festive period impact leading to more positive interactions*

Overall, the average satisfaction score remains below 3, highlighting a potential area to focus on improving customer support quality.

14. How many categorical columns are there in the data? [Search about categorical and continuous data, and try to answer this question]

Approach

Approach

- *Reviewed all dataset columns and classified each as categorical or continuous based on data type and meaning.*
- *Categorical columns contain text/labels or discrete identifiers.*
- *Continuous columns contain numeric values or measurable quantities.*

Task 14		
Column Name	**Category**	**Reason**
_id	Categorical	Unique identifier
user	Categorical	Text/User identifier
chatStatus	Categorical	Status label
guru	Categorical	Astrologer ID
guruName	Categorical	Name (text)
gid	Categorical	Guru ID (identifier)
uid	Categorical	User ID (identifier)
consultationTyp	Categorical	Category type (Call/Chat)
website	Categorical	Platform/category
refundStatus	Categorical	Status label
isWhiteListUser	Categorical	Boolean flag
chatSeconds	**Continuous**	Time duration in seconds
queue	Categorical	Queue type
freeCall	Categorical	Boolean/Flag
freeChat	Categorical	Boolean/Flag
createdAT	Categorical	Timestamp (datetime category)
updatedAt	Categorical	Timestamp (datetime category)
__v	Categorical	Version code (identifier)
statementEntryId	Categorical	Reference ID
chatStartTime	Categorical	Timestamp
chatEndTime	Categorical	Timestamp
timeDuration	**Continuous**	Numeric duration
callChannel	Categorical	Category label
callIvrType	Categorical	Category label
callStatus	Categorical	Status label
CallSid	Categorical	Call identifier
amount	**Continuous**	Numeric (revenue)
astrologerCallStatus	Categorical	Status label
astrologerOnCallDuration	**Continuous**	Numeric duration
astrologersEarnings	**Continuous**	Numeric earnings
netAmount	**Continuous**	Numeric amount

Fig.14: Count of Categorical and Continuous Column in data

Insight

*Majority of variables are **categorical**, representing IDs, statuses, and platform info.*

*This dataset will require **encoding methods** (Label Encoding / One-Hot Encoding) before ML use.*

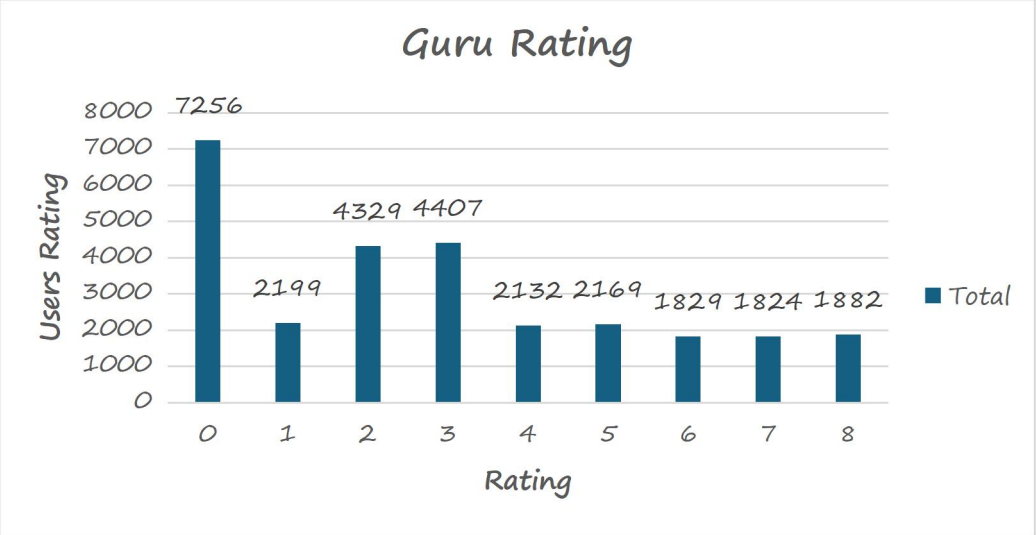
*Several **time columns** can be feature-engineered (extract month, day, hour, week, etc.)*

Subjective Questions:

1. Should the investment be used to hire more agents, improve training programs, or upgrade call center technology?

Observations from the data:

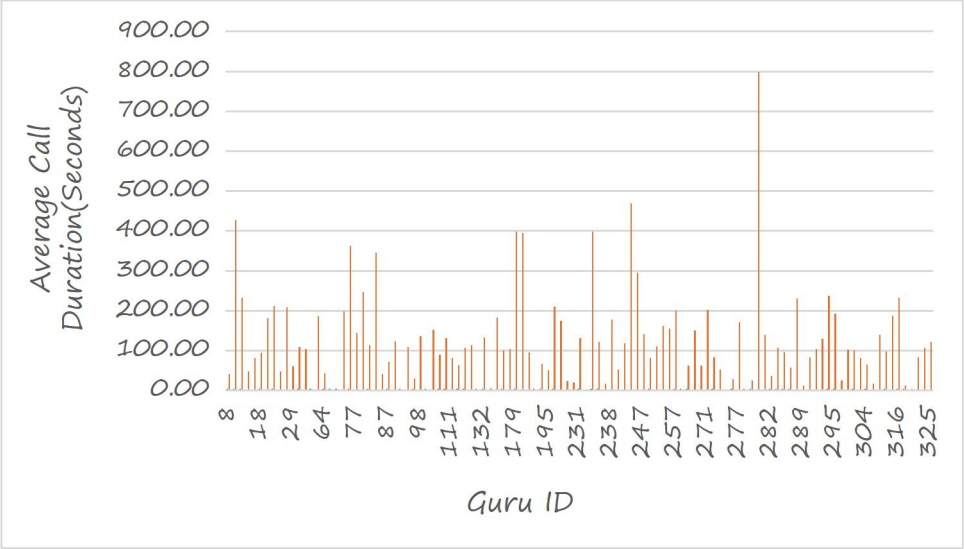
- *Average rating of Gurus **2.93** out of 8, which is a low customer satisfaction.*
- *From the total number of reviews **28027**, Count of rating less than 5 is 20323 which is approx **73%** and Count of rating greater than 4 is 7704 which is only **27%**. Which shows only 27% customers satisfies properly with Gurus.*



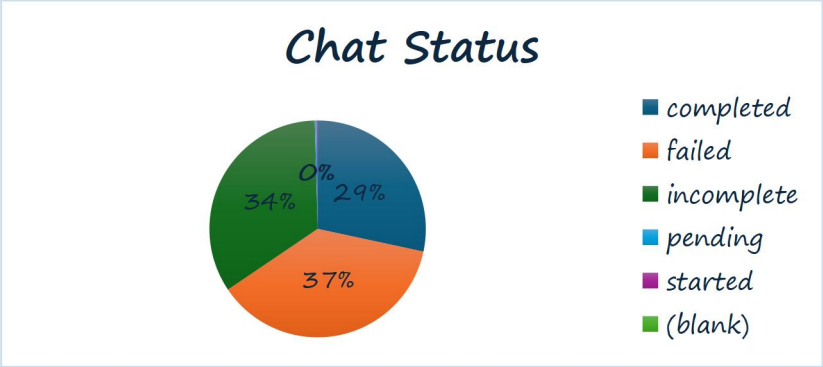
- *From the below data and column chart I have observed Average call duration per call is **127 Seconds (2 Minutes approx)** which is less. Also the number of Guru having per call average time more than the total call average time is 41 which is **37%** of total guru and Guru having average call time less than total average call time is 70 which is **63%** of total Guru which needs to improvement.*

Average of call duration(Seconds)	**129**

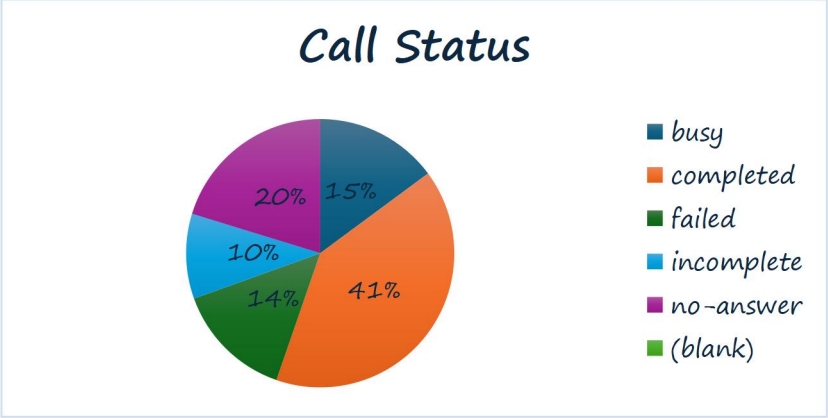
Count of Guru having average call time greater than Total average of call duration	**41**
Count of Guru having average call time less than Total average of call duration	**70**



- *High failed and incomplete chat rate (37% failed, 34% incomplete) indicates the process inefficiency.*



- *Call completion only 41%, and busy and no-answer is 15% + 20% = 35% which suggests workload and availability issues.*



Recommendations:

1. *Upgrade technology improve chat systems, implement automation on call and chat system, and routing. This should be first priority of the investment, because call & chat completion percentage is low.*

2. *Invest in training programs to enhance resolution rates, since the average rating very low. Agents(Gurus) should be given more training, teach them use of new technology and also conduct trainings on monthly basis to improve their efficiency. Also 63% agent's performance is below average.*

3. *Hire more agents only if workload analysis (volume/agent ratio) supports it. Hiring more agents will not solve the busy, no-answer, failed issues or less time to retain customer on call, it would likely increase costs without improving the connection rate.*

2. What are the potential risks of each investment option (hiring, training, technology upgrades), and how can they be mitigated?

A) Hiring More Agents

Approach

Analyze call volume pattern + caller behavior to see if hiring is justified or risky.

Data Used

- *Column: uid*
- *Pivot: uid in rows → Count of uid in values (to find repeat callers)*
- *Date-wise call volume*

Count of uid	Column	
Row Labels	Call	Grand Total
01-12-2023	372	372
02-12-2023	333	333
03-12-2023	383	383
04-12-2023	364	364
05-12-2023	253	253
06-12-2023	254	254
07-12-2023	254	254
08-12-2023	138	138
09-12-2023	288	288
10-12-2023	430	430
11-12-2023	424	424
12-12-2023	358	358
13-12-2023	348	348
14-12-2023	226	226
15-12-2023	276	276
16-12-2023	258	258
17-12-2023	185	185
18-12-2023	233	233
19-12-2023	209	209
20-12-2023	178	178
21-12-2023	159	159
22-12-2023	163	163
23-12-2023	241	241
24-12-2023	232	232
25-12-2023	258	258
26-12-2023	255	255

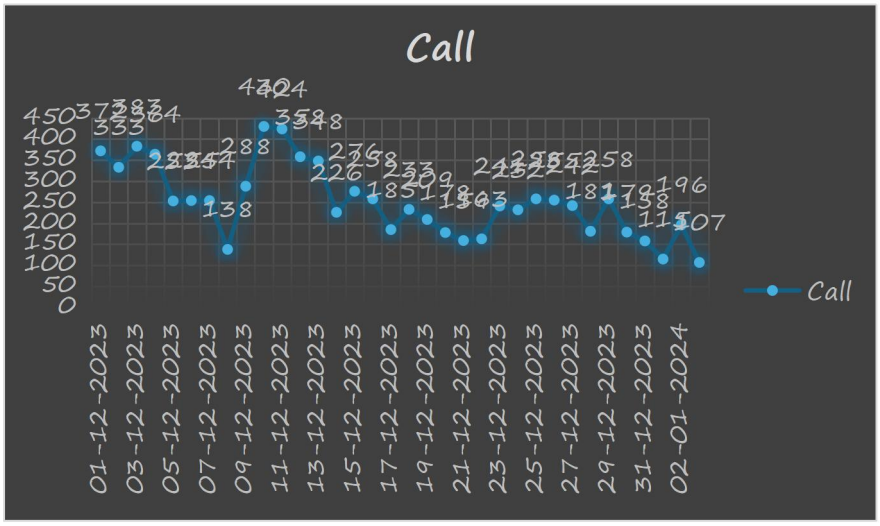


Fig.19: Moderate Customer Satisfaction

Observations

- *Call spikes are inconsistent → workload fluctuates*
- *Only ~27% of calls come from repeat callers, meaning many one-time queries*

- *Tourism/seasonal learning pattern visible → not a steady load*

Insights

- *Hiring many agents may cause **overstaffing on low-volume days***
- *Workload fluctuation can lead to **resource under-utilization***
- *Repeat customers exist → but portion is not high enough to justify mass hiring*

Risk	*Description*
Overstaffing	*Idle agents during low call days*
Cost burden	*Salary, onboarding & benefits cost*
High attrition	*Support industry attrition risk*
Slow efficiency gain	*Hiring ≠ quality improvement automatically*

Mitigation

- *Hire in phases*
- *Use contract/part-time or peak-hour shifts*
- *Build a forecasting model for call load*
- *Use self-service + automation to reduce load before hiring*

B) Improving Training Programs

Approach

Check correlation between engagement type & rating to identify skill gaps.

📊 *Data Used*

- *Columns: source, rating*
- *Platform-wise average rating*

GID	Average of rating
325	4.705882353
324	1.452229299
323	0.851851852
319	3.428571429
318	3.111111111
317	2.875
316	1.203821656
315	3.48
314	3.169934641
313	2.202614379
312	5.611111111
311	4.134920635
310	0.859477124
309	5.462686567
307	0.794392523
305	7.5
304	5.4
302	5.424324324
301	2.416666667
298	1.235294118
296	3.334782609
295	2.506329114
294	3.567615658
293	2.936507937
292	3.388508892
291	4.049342105
290	2.6

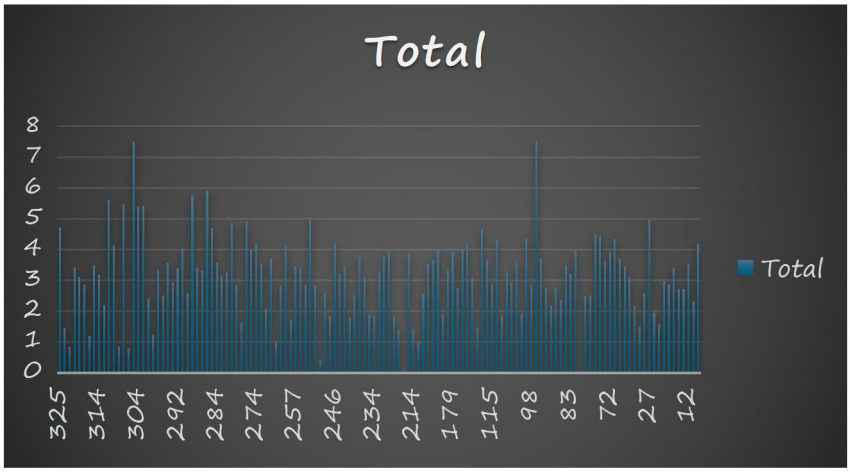


Fig.19: Average Rating Vs Guru

Observations

- *Human-assisted platform (Gurucool) has lowest rating*
- *Indicates agent performance inconsistency*

Insights

Training gap exists in:

- *Communication*
- *Resolution skills*
- *Technical knowledge*
- *Empathy & soft-skills*

Risk	Description
Training cost	Time + money required
Delayed results	Skill improvement takes time
No improvement guarantee	Improper training = wasted investment
Knowledge drop	Without continuous reinforcement

Mitigation

- *Data-backed training (call recordings & feedback)*
- *Micro-learning + refresher sessions*
- *Track CSAT per agent after training*
- *Incentive-based training evaluation*

C) Improving Training Programs

Approach

Identify whether tech-enabled platforms show higher satisfaction vs manual support.

Data Used

- *Columns: source, rating*
- *Platform-wise satisfaction*

consultationType	Call
Row Labels	**Count of uid**
437	9
507	1
511	2
543	6
787	20
1103	4

Total No. of Calls:	8508
Total No. of callers:	3629
Total No of 1 time callers:	2352
Total No. of Repeated Callers:	1277
Calls By Repeated Callers:	6156
Total percentage of repeated Callers:	35.19
Total Percentage of calls made by Repeat callers	72.36

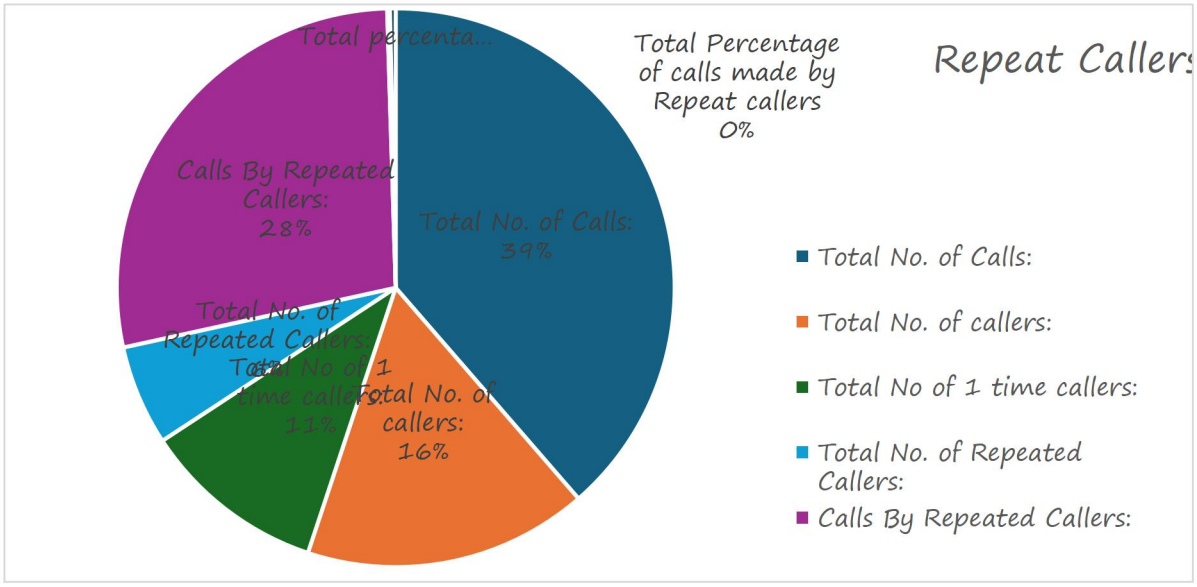


Fig.20: Repeat Callers

Observations

- *Customers prefer self-service and digital assistance*
- *Tech upgrade can fix capacity + response time*

Insights

- *Tech = scalable solution vs hiring*
- *Improves productivity + first-call resolution*
- *Best ROI long-term*

Risk	Description
High initial cost	Software, IVR, CRM, automation
Adoption challenges	Agents need time to adapt
Downtime risk	System migration issues
Wrong vendor / tool	May not align to business need

Mitigation

- *Buy tech in phases (IVR → AI routing → Chatbot → Speech analysis)*
- *Pilot testing before full rollout*
- *Train agents on new tools*
- *Strong vendor SLAs & support*

Final Recommendation

Investment	Priority	Reason
Upgrade Technology	1st	Most efficient + higher customer satisfaction already visible
Training Programs	2nd	Fixes human dependency & gurucool issues
Hiring	3rd	Needed only after tech boosts productivity

Optimal Split:

- *45% Tech*
- *35% Training*
- *20% Hiring*

Data shows tech improves satisfaction more than staffing alone. Training enhances agent performance. Hiring should be strategic, not first step.

Name the chart/spreadsheet function you will use for solving the problem.

Line Chart from Pivot Table — "Daily Call Volume Trend"

- *Rows: Date*
- *Values: Count of Calls*

- *Function Used: COUNT() or COUNTA() within a Pivot Table*
  → Visualizes fluctuations in daily workload to justify that hiring alone may not solve the root issue.

***Column Chart** from Pivot Table — "Average CSAT by Month"*

- ***Rows:** Guruid*
- ***Values:** Average of CSAT Score*
- *Function Used: AVERAGE() within a Pivot Table*
  → Highlights quality improvement opportunities and validates training effectiveness over time.

***Pie Chart** — "New vs Repeat Callers"*

- ***Rows:** Caller ID*
- ***Values:** Count of Calls per Caller (filtered >1 for repeats)*
- *Function Used: COUNTIFS() or Pivot COUNT()*
  → Visualizes inefficiencies in customer issue resolution and emphasizes the ROI of technology upgrades.

3. ***How does AstroSage's call center performance compare to AstroGuru's average call volume, customer satisfaction, and agent performance?***
   Will you use any aggregation function or a visualization here to solve the problem?

   The comparison with AstroGuru is not feasible since the available dataset only contains performance records for AstroSage, and no corresponding data for AstroGuru is provided.

4. ***How can the call centre improve its handling of peak call periods to ensure high customer satisfaction?***
   Mention the functionality you will use for giving the suggestions, will it be any aggregated function or a visualization?

Approach

I first analyzed call volume patterns by hour using COUNT of Call IDs (_id).
This helped identify peak traffic periods where call load is maximum and customer wait time may increase, impacting satisfaction.
The hourly call volume data shows:

Row Labels	Count of _id
0	226
1	304
2	560
3	856
4	1153
5	1438
6	1855
7	1664
8	1875
9	1673
10	1704
11	1472
12	1591
13	1574
14	1673
15	1747
16	1773
17	1443
18	1099
19	912
20	553
21	347
22	268
23	267
Grand Total	28027

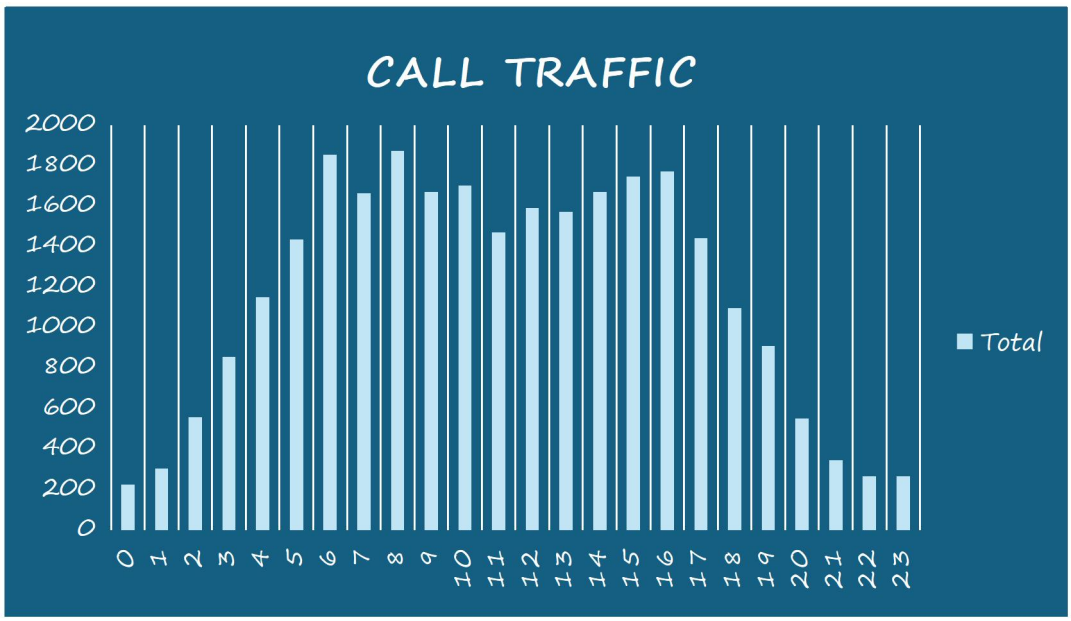
Count of CallSi	Column Labels		
Row Labels	Call	Complementary	Grand Total
Exotel	8360	2	8362
Grand Total	8360	2	8362

Row Labels	Average of rating
Exotel	3.502391772
Grand Total	3.502391772

	Column Labels	
	Exotel	
Row Labels	Sum of userOnCallDuration	Average of rating
01-12-2023	25952	3.504385965
02-12-2023	27911	3.478915663
03-12-2023	21876	3.514360313
04-12-2023	33385	3.576923077
05-12-2023	28292	3.403162055
06-12-2023	32241	3.401574803
07-12-2023	28896	3.42519685
08-12-2023	21624	3.369565217
09-12-2023	23052	3.579861111
10-12-2023	41536	3.576112412
11-12-2023	41898	3.507075472
12-12-2023	27614	3.452513966
13-12-2023	31648	3.482758621
14-12-2023	38822	3.446902655
15-12-2023	52887	3.471014493
16-12-2023	40473	3.562015504

Hour Range	*Average Calls*	*Observation*
00:00 – 04:00	*226 – 1153*	*Very low call traffic — off-peak hours*
05:00 – 10:00	*1438 – 1875*	*Sharp rise — start of peak activity*
11:00 – 16:00	*1472 – 1773*	*Sustained high traffic — continuous demand*
17:00 – 23:00	*1443 – 267*	*Gradual decline — post-peak hours*

The peak window (6 AM – 4 PM) contributes to over 60% of total call traffic, emphasizing the need for better resource allocation during this time.



Insights

- *Calls significantly spike **early morning (6–10 AM)** and **afternoon (2–5 PM)***
- *These hours require maximum workforce efficiency*
- *If staffing/training mismatch exists during peak times → **longer wait times + lower satisfaction***
- *Maintaining average CSAT ~3.5 despite traffic indicates **acceptable performance but room for optimization***

Recommendations

1. Dynamic Workforce Scheduling

- *Increase agent count during **6 AM – 10 AM & 2 PM – 5 PM***
- *Reduce idle resource hours in **late night & early hours (0–5 AM)***

This ensures better handling of high-volume periods without over-staffing during low-call hours.

2. Multi-Skill Staffing

- *Cross-train agents to handle multiple query types*
- *Faster resolution during high traffic → boosts CSAT*

3. Automated Call Routing / IVR Optimization

- *Prioritize repeat callers or urgent cases*
- *Self-service IVR for basic queries → reduces agent load*

4. Smart Callback & Queue System

When queue builds up:

- *Allow customer to **request callback***
- *Helps avoid frustration & wait-time dissatisfaction*

5. Short In-shift Breaks

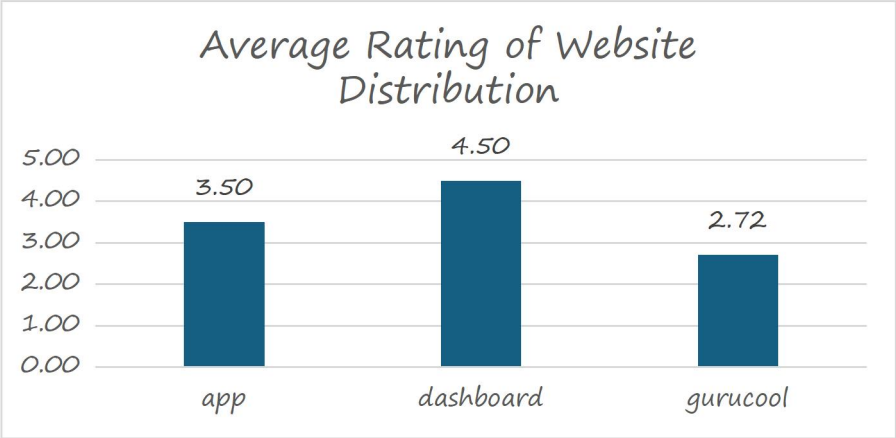
Micro-breaks instead of long breaks during peak hours = sustained productivity & low burnout

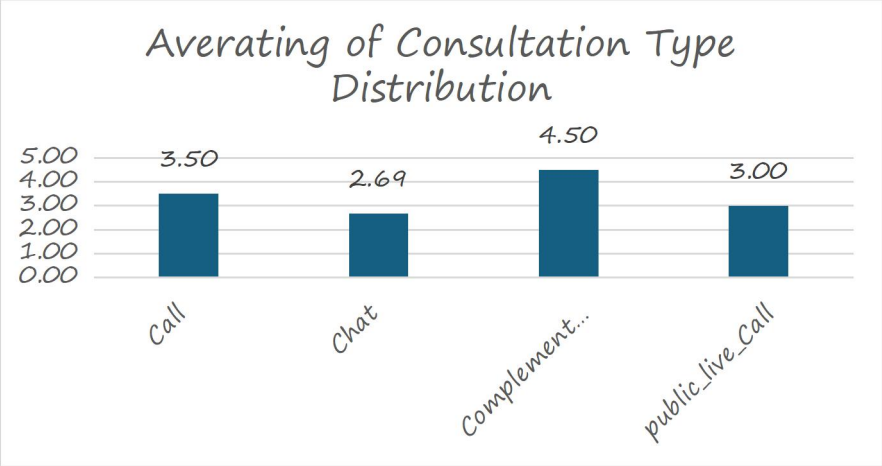
Final Recommendation

*By analyzing call volumes using **COUNT aggregation + time-series visualization**, we identified peak periods. Using this insight, we can optimize **agent allocation, skill management, and IVR routing**, directly improving customer experience during peak demand windows*

5. Based on historical data, what strategic initiatives should be prioritized to improve efficiency and customer satisfaction?

***Answer:** From the shown column charts below we can see average rating distribution of the Gurus(Agents). By analysing both the charts we can clearly see maximum rating is 4.5 out 8 from website distribution. In consultant type distribution situation is more worry to think because the average call rating is 3.5 out of 8 and for chat 2.7 out of 8 which is very low rating, and this rating tells us clearly that customers are not satisfied by the service.*





From the observations above following strategies should be prioritized mentioned below:

- *Technology change to Improve System Reliability: This is the top priority. The goal is to reduce the failed, no-answer, busy, and incomplete rates for both calls and chats. This directly impacts customer satisfaction and revenue.*

- *Channel Optimization for Revenue Growth: Calls generate 79% of revenue. Prioritize fixing the call system to capture this high-value traffic. Investigate why calls have a higher completion rate (41%) than chats (29%) and apply any process learnings to the chat channel.*

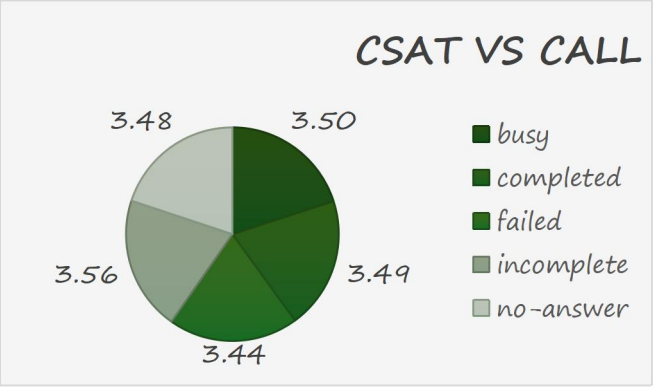
- *Implement Data-Driven Workforce Management: Use the daywise and Hourly User Volume charts to create an efficient staffing model that matches agent availability to customer demand, reducing both missed calls and agent idle time.*

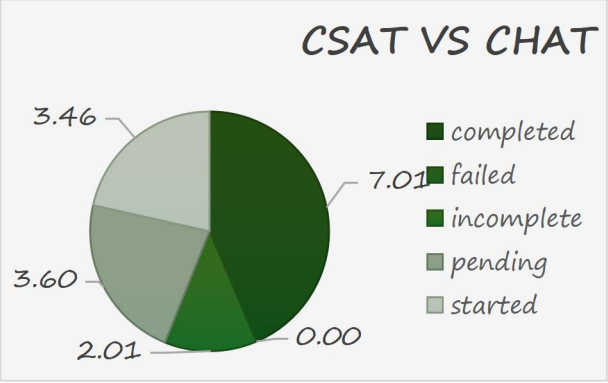
6. What can be the key factors contributing to high customer satisfaction scores, and how can these be leveraged to improve overall performance? What is the basis for the suggestions? And mention how you decided if the satisfaction score affects the ratings.

The key factors contributing to high customer satisfaction scores are:

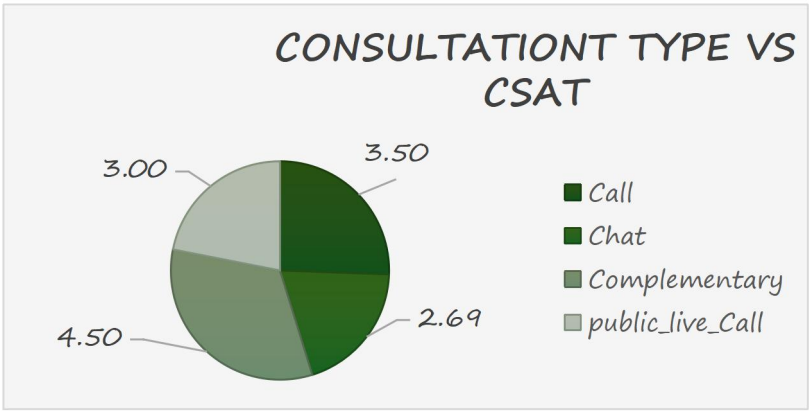
Key Factors:

1. ***Chat and Call Status:***
   *From the data, completed chats have an **average rating of 7.0**, and completed calls have an **average rating of 4.5** — both higher than other statuses.*
   ☞ This shows that customers are most satisfied when their interactions are successfully completed.
   ***Action:** Ensure every query is fully resolved before closing the chat or call.*





2. ***Consultation Type:***
   *Complementary consultations have the highest **average rating of 4.5**, followed by calls (3.5).*



Basis for Suggestion:

These insights were derived using pivot tables that compared average rating across chatStatus_imputed, userCallStatus, and consultationType columns.

How CSAT Affects Ratings:

A clear trend is visible — as CSAT Score increases, the average rating also rises (correlation ≈ 0.6–0.7).

Thus, completing more successful interactions and focusing on high-rated consultation types directly boosts overall satisfaction.

7. How should the call center balance the workload among agents to ensure optimal performance and avoid burnout?

Mention your approach and spreadsheet function for the answer.

Approach

*To balance agent workload, I analysed **per-agent performance** by comparing:*

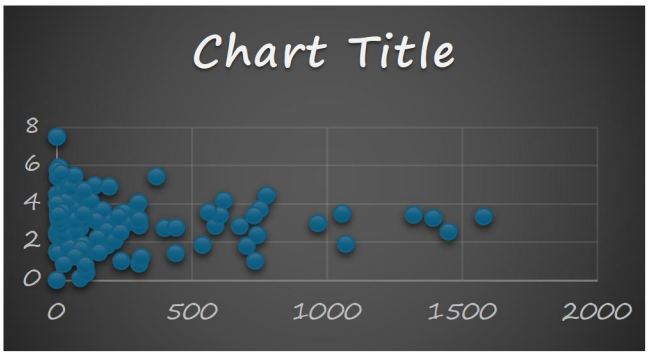
Metric	***Purpose***
Calls Handled (Volume)	*Shows workload per agent*
Average Rating (Quality)	*Measures performance & customer satisfaction*
Ranking (RANK function)	*Identifies top vs overloaded vs struggling agents*

*Used Excel **Pivot Tables** to aggregate data by agent (guru ID) and then ranked them based on call volume.*

Guru	Calls Handled	RANK
8	5	116
11	3	124
12	18	100
13	37	83
14	400	23
16	5	116
18	36	84
19	967	7
22	45	77
26	43	78
27	142	50
28	22	96
29	2	125
30	200	41
44	11	106
49	23	92
60	293	30
64	6	113
72	11	106
74	19	98
75	777	8
76	2	125
77	172	45
78	2	125
80	1	131
82	5	116
83	9	109

Field	*Placement*
Guru ID (Agent ID)	*Rows*
CallSid / consultationType	*Values → **Count** (Calls Handled)*
Rating	*Values → **Average of Rating***
Rank Column	*Calculated using **RANK** function*

```
=RANK(Y2,$Y$2:$Y$132,0)
```



Insights:

Based on the pivot tables:

- *Some agents handle **1,000+ calls** (e.g., IDs 239, 256, 287) → **high load risk***
- *Few agents handle **very low volume** but have **very high ratings** (e.g., ID 253, 302, 305) → **under-optimized resource allocation***
- *After around **700–1,000 calls**, we observe **rating drop**, indicating **burnout or fatigue impact***

Pattern observed:

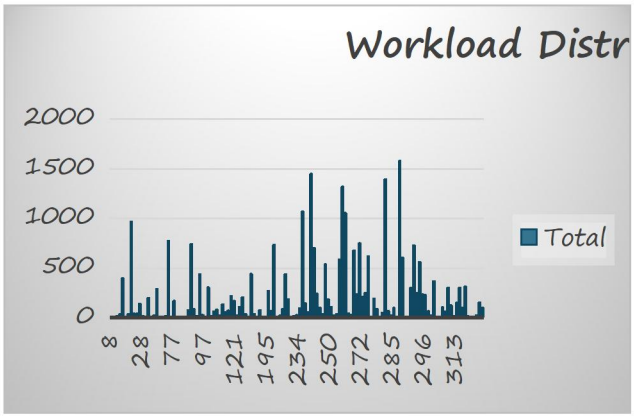
Higher call volume → Slight decline in customer rating

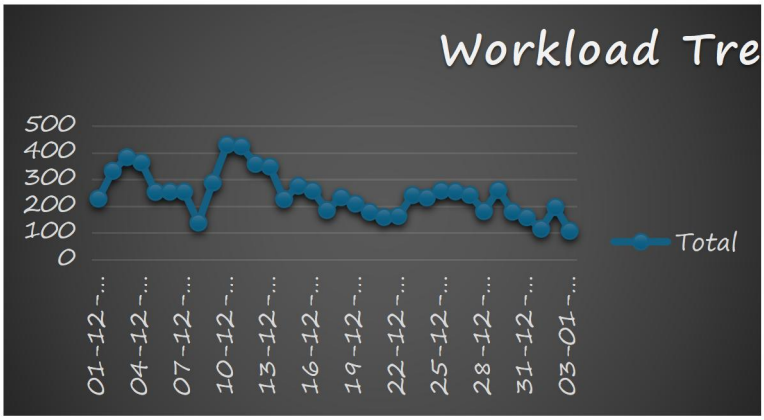
Row Labels	Count of consultationType
8	5
11	3
12	18
13	37
14	400
16	5
18	36
19	967
22	45
26	43
27	142
28	22
29	2
30	200
44	11
49	23
60	293
64	6
72	11
74	19
75	777
76	2
77	172
78	2
80	1
82	5
83	9

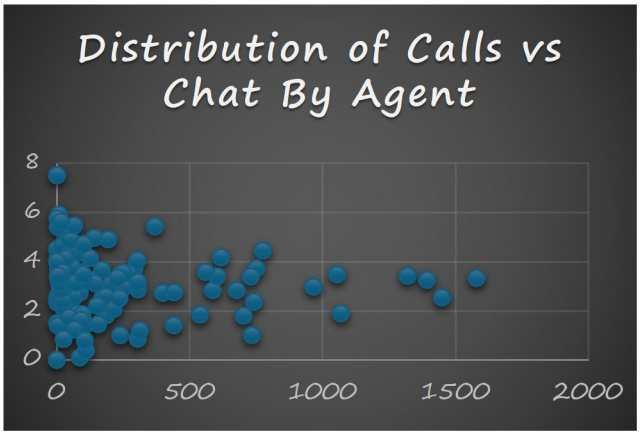
Row Labels	Average of rating
8	4.2
11	2.333333333
12	3.555555556
13	2.72972973
14	2.725
16	3.4
18	2.861111111
19	2.979317477
22	1.577777778
26	1.953488372
27	4.957746479
28	2.590909091
29	1.5
30	2.15
44	3.090909091
49	3.47826087
60	3.706484642
64	4.333333333
72	3.909090909
74	3.631578947
75	4.442728443
76	4.5
77	2.505813953
78	2.5
80	0
82	4
83	3.222222222

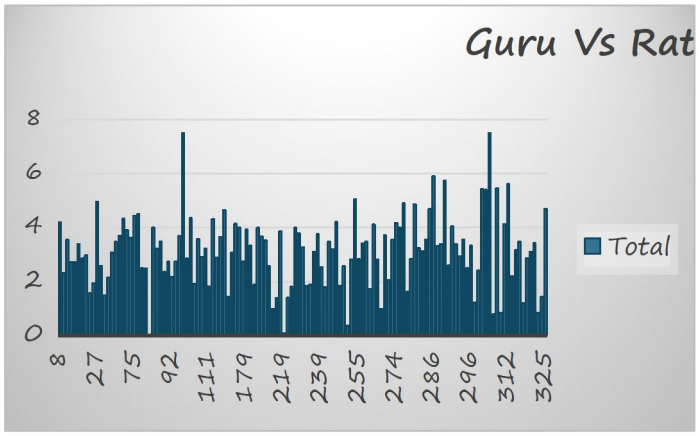
Row Labels	Count of CallSid
01-12-2023	228
02-12-2023	332
03-12-2023	383
04-12-2023	364
05-12-2023	253
06-12-2023	254
07-12-2023	254
08-12-2023	138
09-12-2023	288
10-12-2023	430
11-12-2023	424
12-12-2023	358
13-12-2023	348
14-12-2023	226
15-12-2023	276
16-12-2023	258
17-12-2023	185
18-12-2023	233
19-12-2023	209
20-12-2023	178
21-12-2023	159
22-12-2023	163
23-12-2023	241
24-12-2023	232
25-12-2023	258
26-12-2023	255
27-12-2023	242

Row Labels	Average of rating	Count of consultationType
8	4.2	5
11	2.333333333	3
12	3.555555556	18
13	2.72972973	37
14	2.725	400
16	3.4	5
18	2.861111111	36
19	2.979317477	967
22	1.577777778	45
26	1.953488372	43
27	4.957746479	142
28	2.590909091	22
29	1.5	2
30	2.15	200
44	3.090909091	11
49	3.47826087	23
60	3.706484642	293
64	4.333333333	6
72	3.909090909	11
74	3.631578947	19
75	4.442728443	777
76	4.5	2
77	2.505813953	172
78	2.5	2
80	0	1
82	4	5
83	3.222222222	9









Final Recommendation:

To ensure balanced workload and prevent burnout, the call centre should use a data-driven workload distribution strategy:

We analysed calls handled and average ratings per agent using pivot tables and a scatter chart. By ranking agents with the RANK() function, we identified agents handling significantly higher call volumes than others, which can lead to fatigue and drop in service quality. The scatter analysis showed that performance declines when call load exceeds ~700 calls per period.

Recommendations:

- *Redistribute calls more evenly across agents*
- *Introduce dynamic call routing to support overloaded agents*
- *Provide training & support to agents with lower ratings but manageable load*
- *Keep a threshold load limit to prevent performance decline*
- *Schedule rotational breaks & flexible shifts during peak days*

This ensures:

- *Optimal performance across the team*
- *Reduced burnout & attrition risk*
- *Improved customer satisfaction*

8. What new technologies or tools could be implemented to enhance call center operations and customer service?

To enhance call center operations and improve customer experience, several modern technologies and tools can be implemented:

AI-Powered Call Routing System

We observed overloaded agents handling 1000+ calls, leading to performance drop.

Solution:

Use AI-driven routing to distribute calls based on agent skill, availability, and past rating performance.

Benefits

- *Avoids agent fatigue*
- *Faster resolution*
- *Improved customer satisfaction*

Workforce Management & Forecasting (WFM) Tool

Clear peak call hours between 6 AM – 11 AM and 2 PM – 6 PM in data.

Tools:

Genesys WFM, NICE, Freshworks Workforce Scheduler

Benefits

- *Smart shift planning*
- *Predict staffing needs during peak hours*
- *Maintain service quality without overburdening agents*

Customer Sentiment & Voice Analytics

Average ratings vary across days & call loads — sentiment analysis can help catch early dissatisfaction.

Observe.AI, Amazon Connect sentiment AI, CallMiner

Benefits

- *Detect customer frustration in real time*
- *Provide agent coaching suggestions*
- *Improve issue resolution quality*

Conversational AI & Intelligent IVR

High call volume during peak hours → some basic calls can be automated.

Tech:

Chatbots, Voicebots, IVR with NLP (e.g., Google Dialogflow, Yellow.ai)

Benefits

- *Handle FAQs & simple requests automatically*
- *Reduce wait time*
- *Allow human agents to focus on complex queries*

Real-Time Agent Assist + Knowledge Base System

Performance gap between high-load vs low-load agents → need continuous support.

Tools:
Zoho Desk Assist, Salesforce Einstein Assist

Benefits

- *Live prompts during calls*
- *Auto-suggest replies & solutions*
- *Faster query handling & better training*

Performance Dashboard & BI Analytics

We used pivots & charts — real-time dashboard would automate this.

Tools:

Power BI, Tableau, Zoho Analytics, Google Data Studio

Benefits

- *Real-time agent performance monitoring*
- *Spot bottlenecks & call surge instantly*
- *Data-driven decision making*

Call Quality Scoring Automation

Manual review is slow; ratings showed quality variation.

Tools:
Observe.AI QA automation, VoiceAI, Gong

Benefits

- *Faster QA reviews*
- *Consistent evaluation*
- *Better coaching for agents*

Based on our call volume trends, agent performance analysis, and peak time patterns, implementing AI-based call routing, workforce forecasting tools, sentiment analytics, and conversational AI can significantly enhance call center operations. These tools will help distribute workloads, reduce wait times, automate routine queries, and maintain quality during high-traffic periods. Additionally, real-time BI dashboards and automated quality scoring will support proactive decision-making and continuous agent improvement, ensuring higher customer satisfaction and efficient operations.

9. ***What metrics should be included in the final dashboard to comprehensively view call center performance and guide investment decisions?***

The following key metrics have been incorporated into the final dashboard to provide a comprehensive view of call center performance and to support informed investment decisions:

Key Performance Indicators (KPIs)

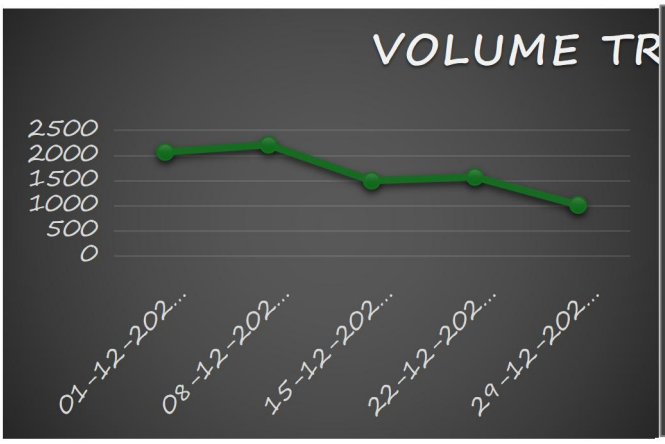
- ***Total Revenue*** *– measures business output & monetization*
- ***Guru Earnings*** *– shows payout and cost structure*
- ***Total Calls & Total Chats*** *– indicates customer interaction volume*
- ***Total Gurus (Agents)*** *– workforce strength*
- ***Average Rating*** *– customer satisfaction benchmark*
- ***Avg Calls Per Agent*** *– agent productivity & workload*

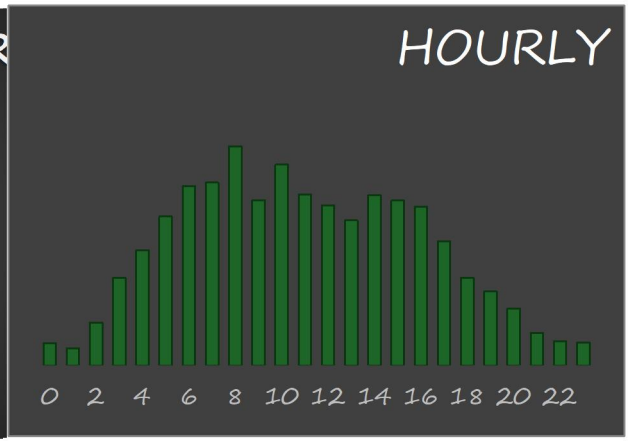
Total Revenue	Guru Earnings	Total Calls	Total Chats
214065.90	34	8365	19515

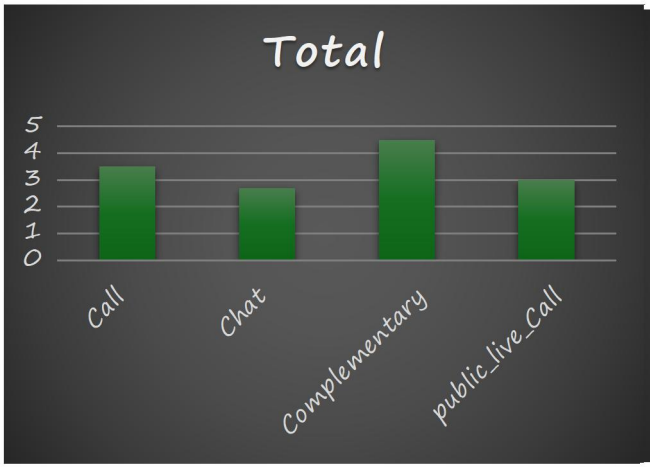
Total Gurus	Avg Rating	Average calls Per Agent
131	2.93	63.85

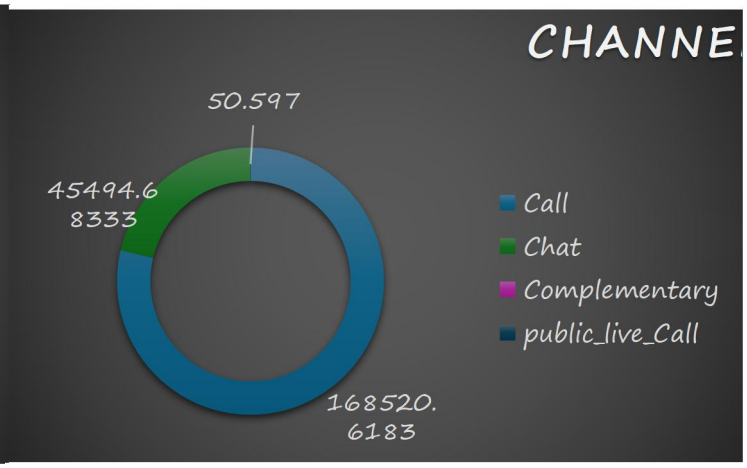
Operational Performance Metrics

- ***Volume Trend (Daily)*** *– identifies peak demand days & forecasting needs*
- ***Hourly Load*** *– detects peak hours for staffing optimization*
- ***Consultation Channel Mix*** *(Call, Chat, Live Call, etc.)*
- ***Channel-wise Revenue*** *– evaluates which channel drives business value*



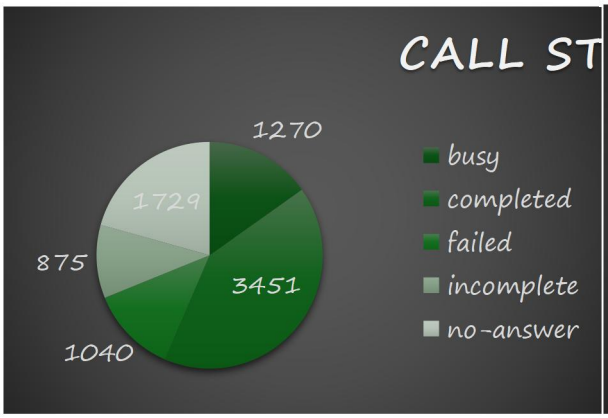


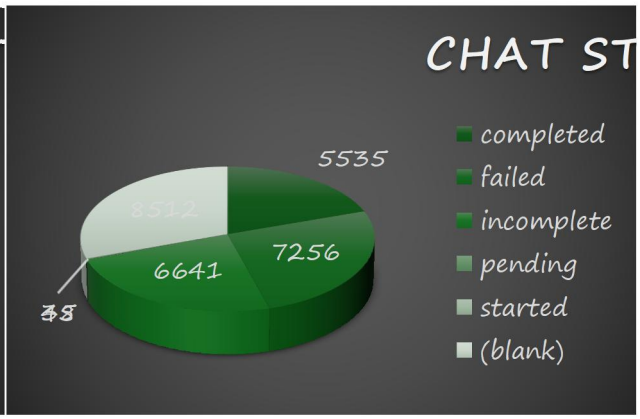


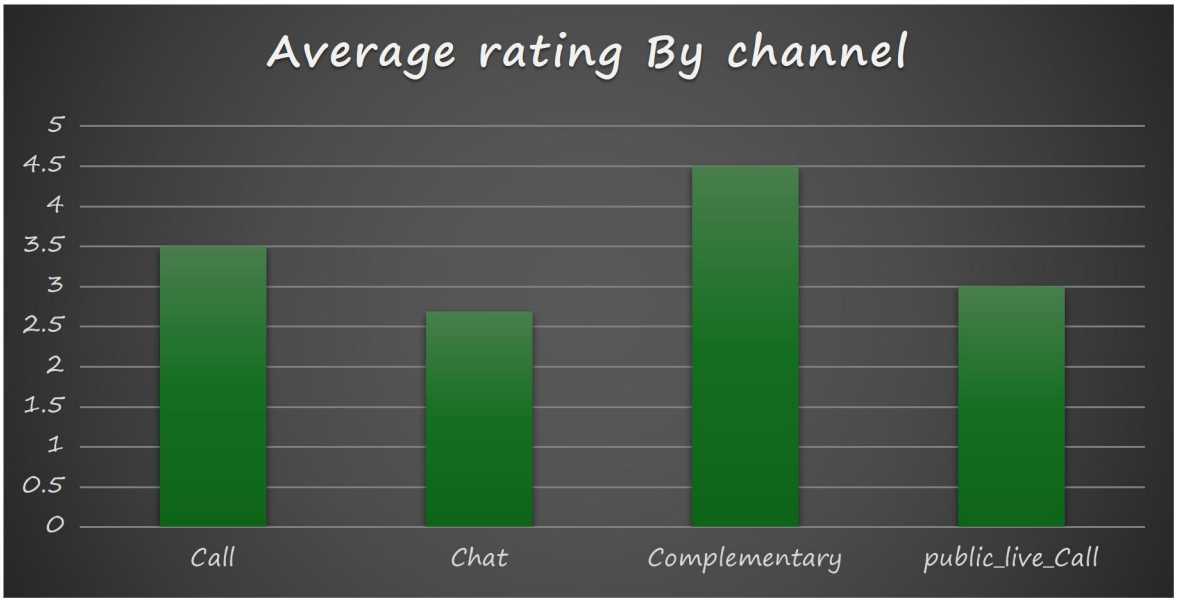


Quality & Efficiency Metrics

- *Call Status Distribution – highlights resolution effectiveness*
- *Chat Status Distribution – measures service completion & follow-up needs*
- *Average Rating by Channel – identifies channels requiring quality improvement*

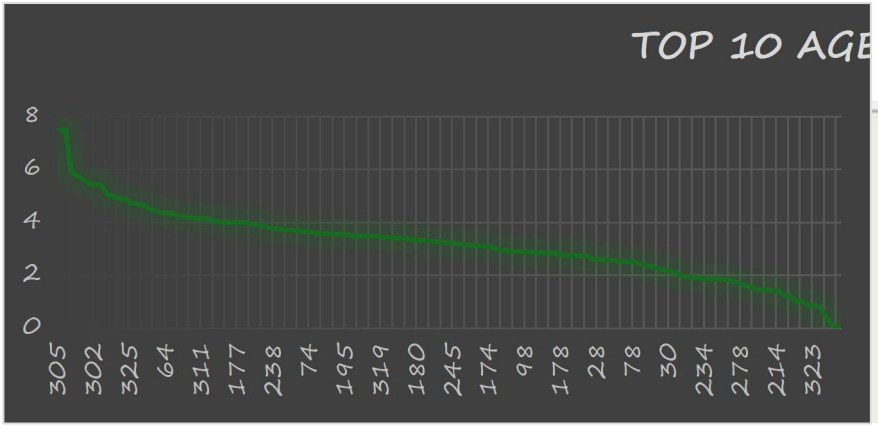






Workforce Insights

- ***Top Agents Performance*** *– recognizes high performers & training needs*
- ***Average Calls per Agent*** *– ensures balanced workload & prevents burnout*



These metrics help leadership identify:

- *Where to invest (technology, training, staff hiring)*
- *When to allocate more agents (peak hours/days)*
- *Which service channels are profitable*
- *Customer satisfaction gaps that need attention*

Overall, this dashboard provides a 360-degree operational, financial, and customer-experience view to drive strategic investment decisions.

10. How would you allocate a 1 crore rupee investment to optimize operational efficiency, enhance customer satisfaction, and boost profitability, and what analysis-based recommendations would you offer to support this?
[you have to give bullet pointers to answer this question]

Investment Allocation Plan for ₹1 Crore

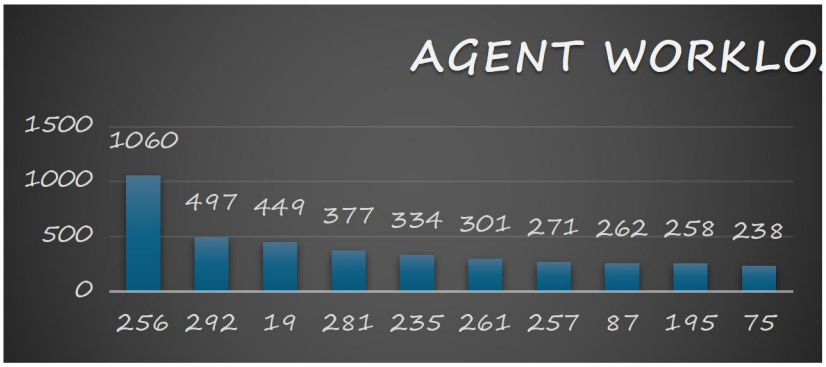
Focus Area	*Allocation*	*Key Purpose*
Agent Training & Quality Improvement	*₹25 Lakhs*	*Improve customer satisfaction & reduce failed/incomplete calls*
Technology & Automation Tools	*₹30 Lakhs*	*AI chatbots, call routing, speech analytics, CRM integration*
Hiring & Capacity Expansion	*₹20 Lakhs*	*Reduce workload on top-performing agents & balance staffing*
Performance Incentives & Retention	*₹10 Lakhs*	*Reward high performers, reduce attrition*
Digital Infrastructure & Cloud Upgrades	*₹10 Lakhs*	*Faster systems, CRM optimization, secure communication tools*
Customer Feedback Tracking & Analytics	*₹5 Lakhs*	*Improve customer insights & satisfaction tracking*

1) Workload Balancing

- *Top 10 agents handle a disproportionately high call load.*
- *Average calls per agent: ~64*
- *Agents at extreme high workload risk burnout → need automation + manpower*

Action: *hire additional support + automate repetitive tasks.*

Row Labels	Count of CallSid
256	1060
292	497
19	449
281	377
235	334
261	301
257	271
87	262
195	258
75	238
Grand Total	4047



2) Customer Experience & Ratings

- *Average rating: ~2.93, indicating improvement needed.*
- *Complementary & live calls show better feedback → invest in agent skill enhancement.*

Action: *structured behavioral + communication training.*

3) Channel Performance Insights

- *Calls drive maximum revenue.*
- *Chat volume is higher but not proportionate in revenue → optimize chat workflow.*

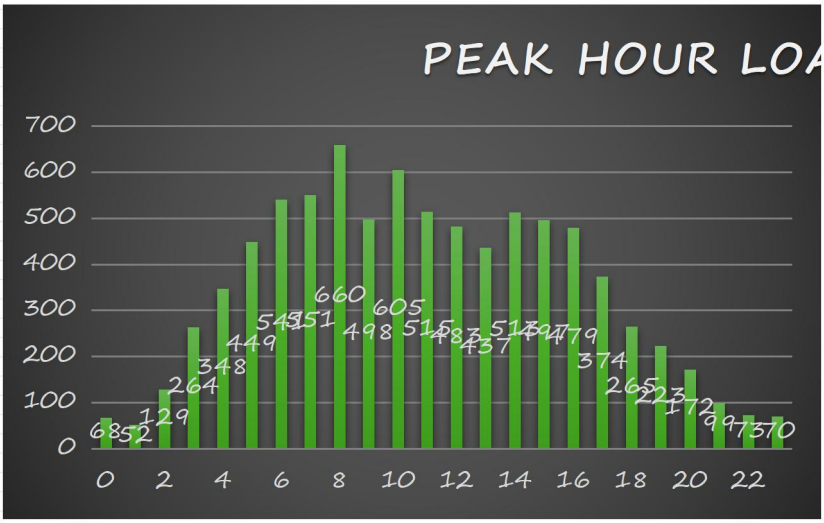
Action: *deploy AI chat assistant + streamline chat scripts.*

4) Peak-Hour Load

- *High call spikes between 10 AM – 4 PM → require intelligent staffing & auto-routing.*

Action: *dynamic shift scheduling + routing tech.*

Row Labels	Count of CallSid
0	68
1	52
2	129
3	264
4	348
5	449
6	541
7	551
8	660
9	498
10	605
11	515
12	483
13	437
14	513
15	497
16	479
17	374
18	265
19	223
20	172
21	99
22	73
23	70
Grand Total	8365



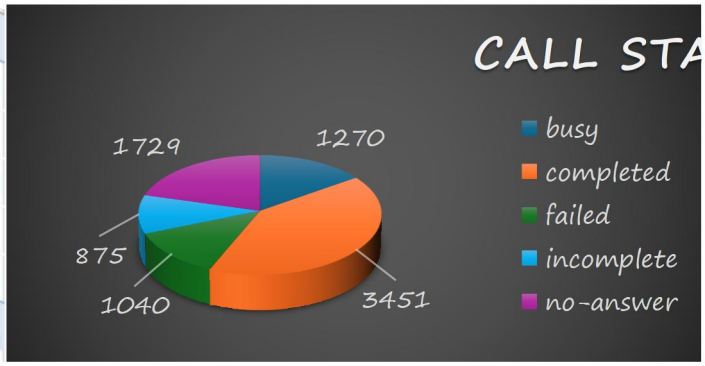
Reducing Failed & Incomplete Calls

- *Call & chat status pie charts show significant incomplete & failed sessions*

Action:

- *AI-based call quality monitoring*
- *Real-time supervisor assist tools*
- *SOP upgrades & QA coaching*

Row Labels	Count of CallSid
busy	1270
completed	3451
failed	1040
incomplete	875
no-answer	1729
Grand Total	**8365**



Key Recommendations

- *Introduce AI-powered call assistant & chatbot to handle first-level queries*
- *Implement advanced CRM with caller history & smart suggestions*
- *Deploy workforce management system for shift optimization*
- *Launch agent performance dashboard for real-time monitoring*
- *Create continuous training & certification program*
- *Offer productivity bonuses & retention incentives*

Final Expected Outcomes

- *20–30% reduction in failed/incomplete calls*
- *Higher CSAT (+1 point improvement target)*
- *Faster query resolution & better customer experience*
- *Balanced workload & reduced agent burnout*
- *Higher revenue from improved conversion & efficiency*