**Overview:**

This project focuses on extracting valuable insights from key performance metrics such as customer’s satisfaction scores, customer sentiment, call timestamps, reasons for contact, response times, and call durations, the project seeks to uncover actionable insights that can drive improvements in customer service and operational performance.

Utilizing MySQL for thorough data pre-processing and analysis, along with Tableau for clear and impactful data visualization, the project will provide a comprehensive understanding of both customer interactions and agent productivity.

**Aim:**

The ultimate goal of this project is to support informed decision-making, optimize resource allocation, and elevate the overall customer experience within the call center.

**Tools Utilized:**

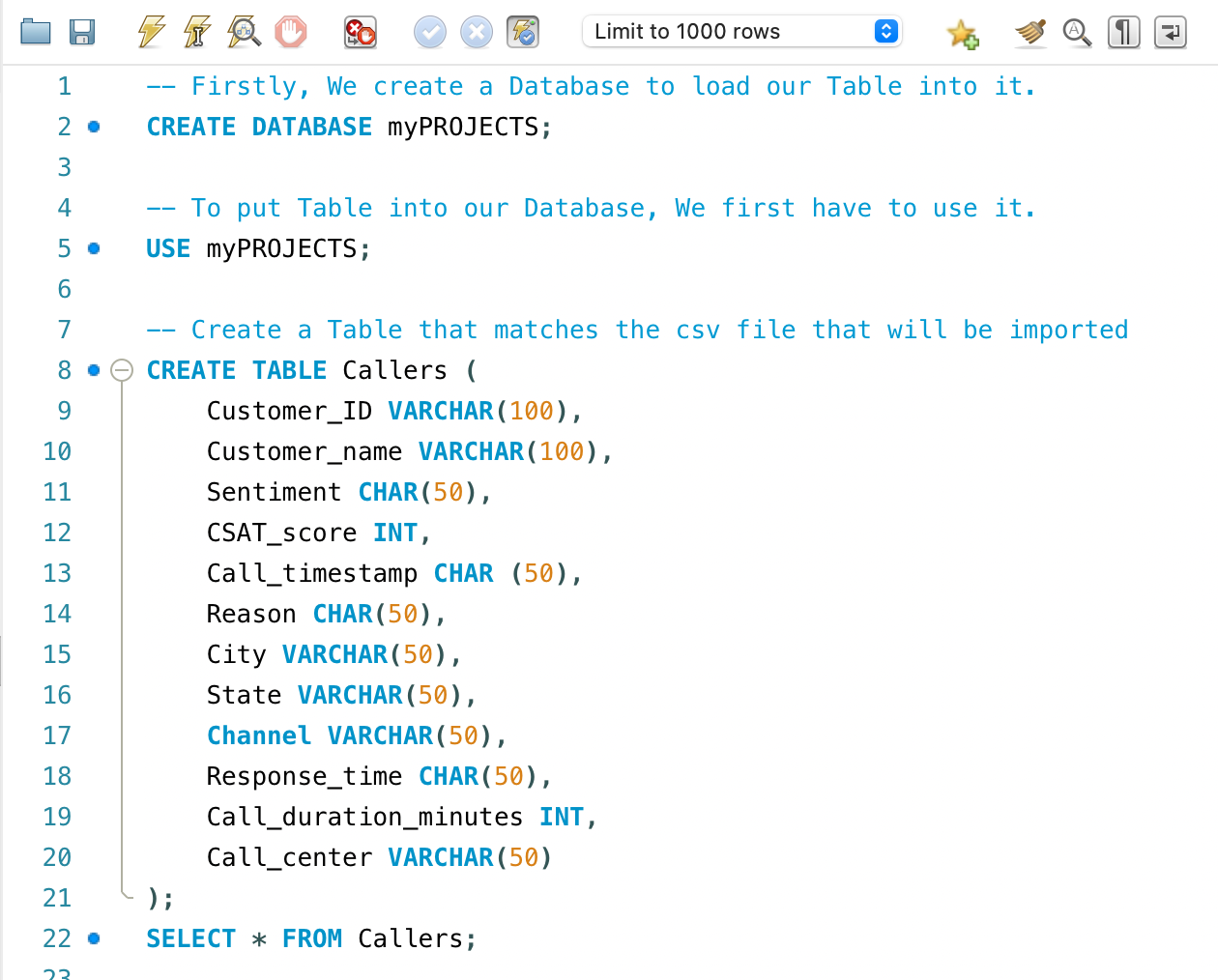
* MySQL Server
* MySQL Workbench
* Tableau

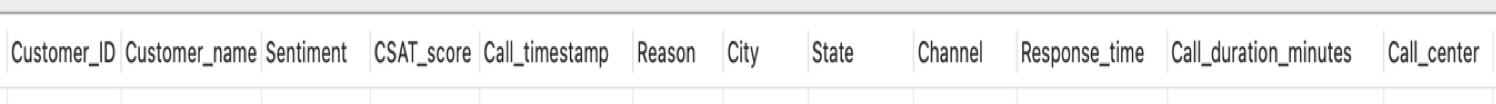
**Customers Information:**

* **id:** Unique identifier for each record
* **customer\_name:** Name of the customer
* **sentiment:** Sentiment associated with the interaction
* **csat\_score:** Customer Satisfaction (CSAT) score
* **call\_timestamp:** Timestamp of the call
* **reason:** Reason for the call
* **city:** City where the customer is located
* **state:** State where the customer is located
* **channel:** Channel used for the interaction
* **response\_time:** Whether the response time was within SLA (Service Level

Agreement) or not

* **call\_duration:** Duration of the call in minutes
* **call\_center:** Call center handling the call





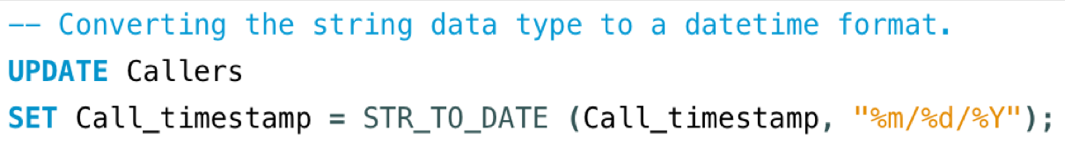
**Note:** Statement line 13 is in a String format file, but should be converted to DATETIME as we proceed.

Load data into the table using Data Import Wizard.

**Data Exploration and Processing**



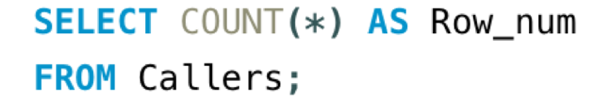


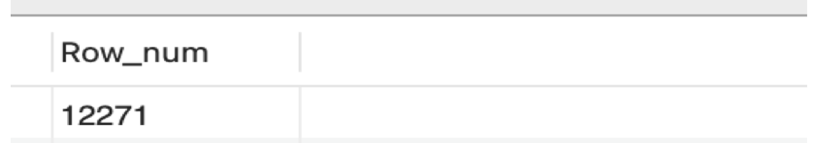




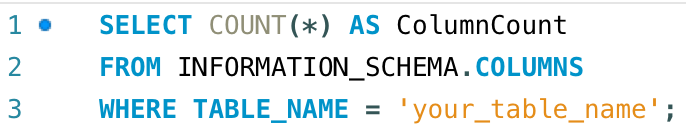


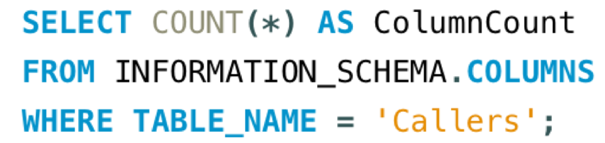
**Exploring the shape of the table**

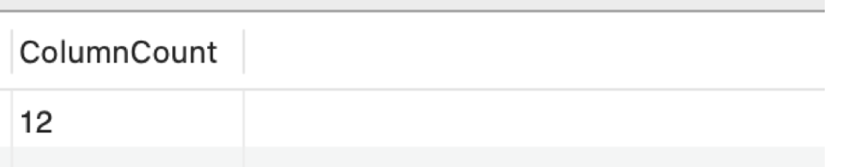
****

****

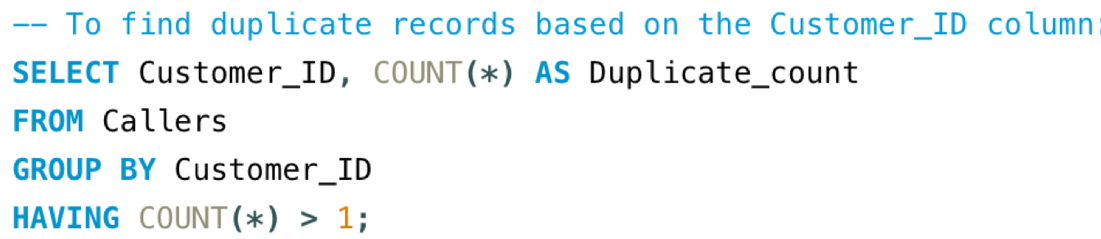
**For numbers of Columns, use the SQL Syntax below**

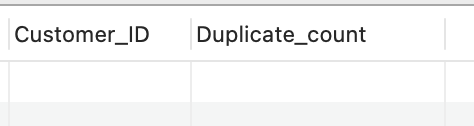
****

****

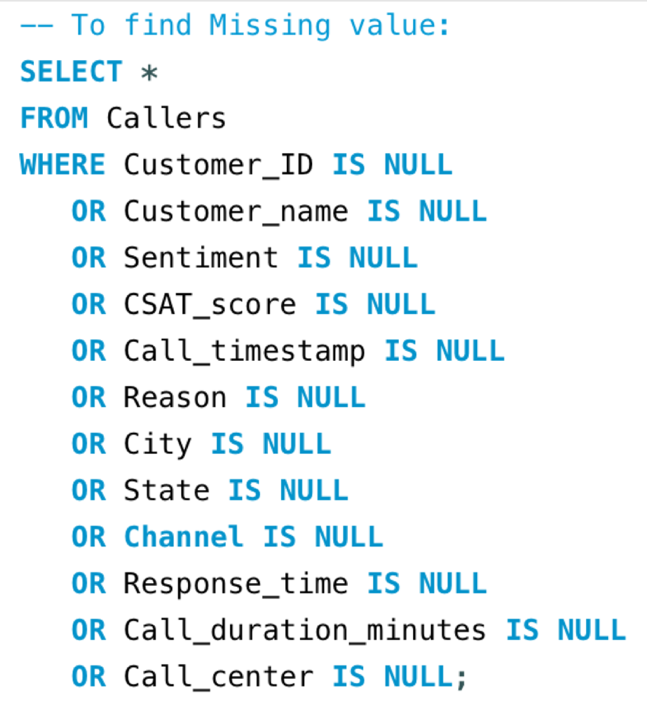
****

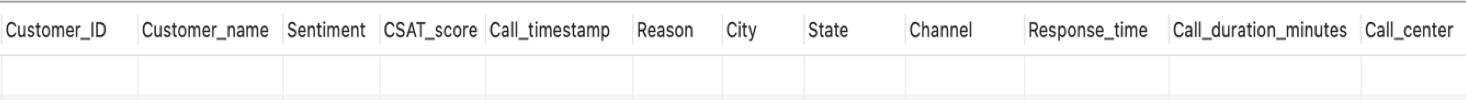
**Finding Duplicate Record**

****

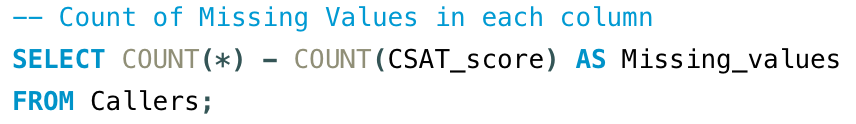
****

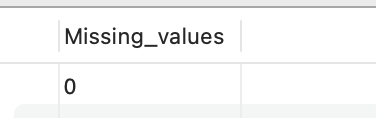
**Finding Missing Value**

****

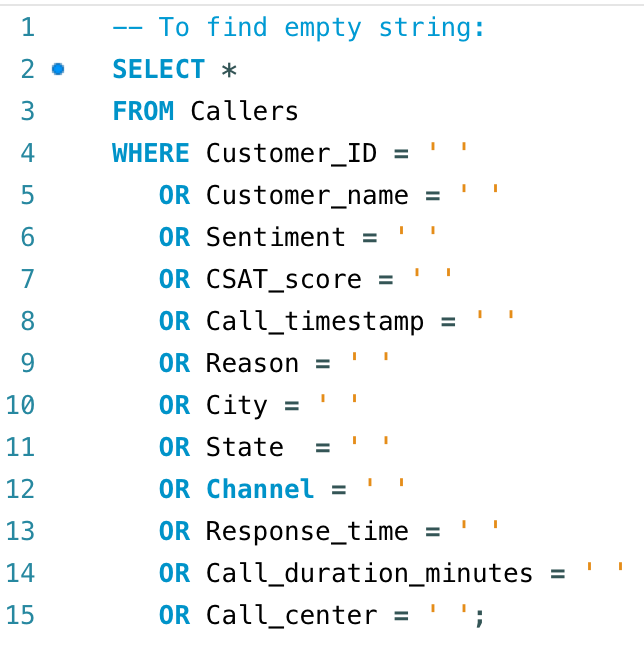
****

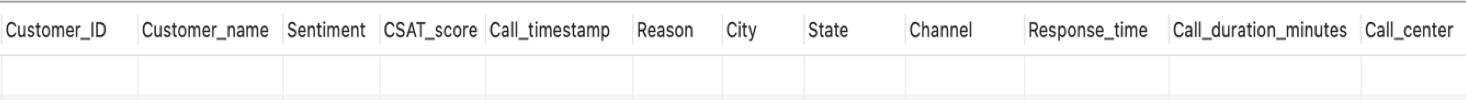
**Alternatively:**





It is obvious that our dataset has no null value

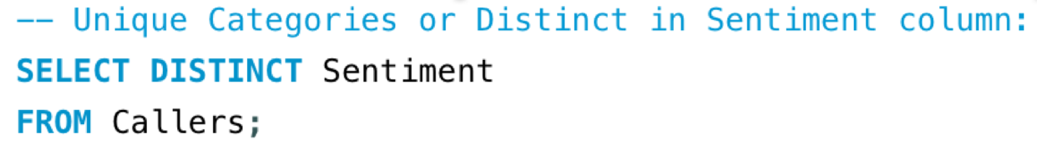




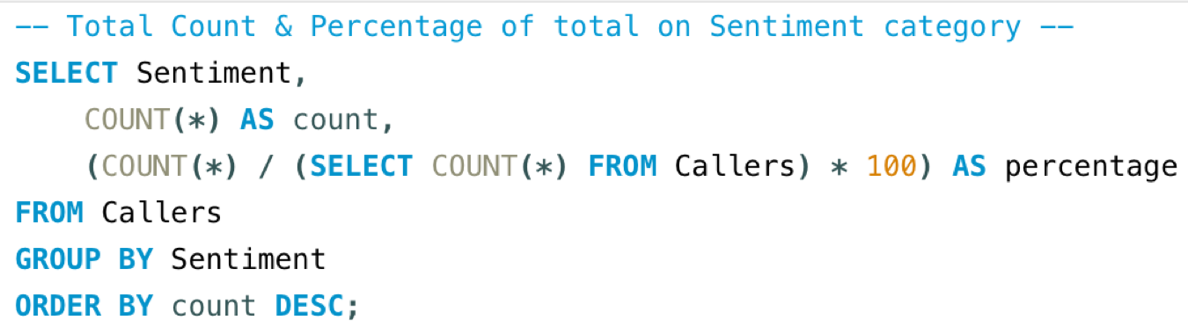
The query returns no empty value.

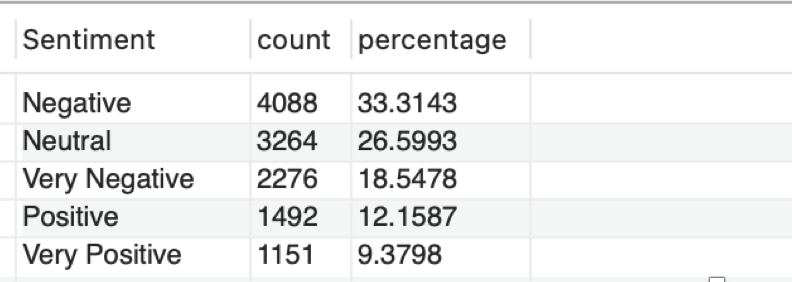
**Frequency Distribution of Categorical Variables**

**Sentiment:**

****

****

****

****

The query shows customer satisfaction by Sentiment. It calculates the Count and Percentage of occurrences for each unique value on Sentiment column.

**Sentiment Distribution:** This column lists the different sentiment categories found in the dataset i.e. "Very Positive," "Positive," "Neutral," "Negative," "Very Negative".

**Count**: This column shows the number of occurrences for each sentiment categories.

* **Very Positive:** appear 1151 times with a Percentage total of 9.38%
* **Positive:** appeared 1492 times with (12.15%) Percentage total
* **Neutral:** appeared 3264 times with (26.60%) Percentage total
* **Negative:** appeared 4088 times with (33.31%) Percentage total
* **Very Negative:** appeared 2276 times with (18.55%) Percentage total

**Negative and Very negative Sentiments**: Negative and Very negative Sentiments categories are be the most common, making up the majority (51.86%) of the dataset. This suggests that most customer interactions lead to dissatisfaction and it result in negative experiences.

Negative Sentiments can have a significant impact on customer loyalty and brand reputation. Understanding the root causes of these negative experiences and addressing them can help improve overall satisfaction.

**Neutral Sentiment**: This portion (26.60%) of the interactions may result in neutral feelings, indicating that these customers might not have strong positive or negative emotions about their experience.

**Very Positive and Positive Sentiments**: These categories are the least common, comprising only (21.54%) of the dataset. This indicate that a smaller fraction of customers interactions leads to satisfaction with positive experience. Indicating a lower proportion of positive feedback.

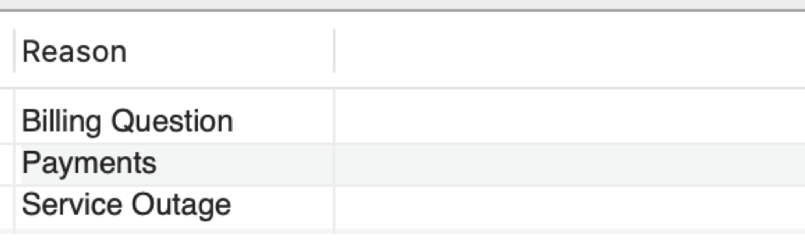
This needs to be addressed to improve the overall customer satisfaction.

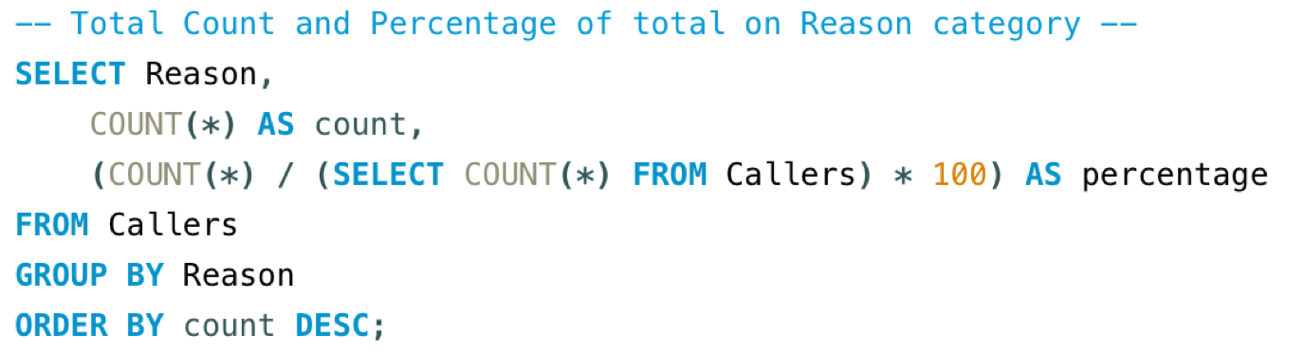
This analysis can be used to prioritize improvements in customer service, focusing on reducing negative sentiments and turning the negatives and neutral experiences into positive ones.

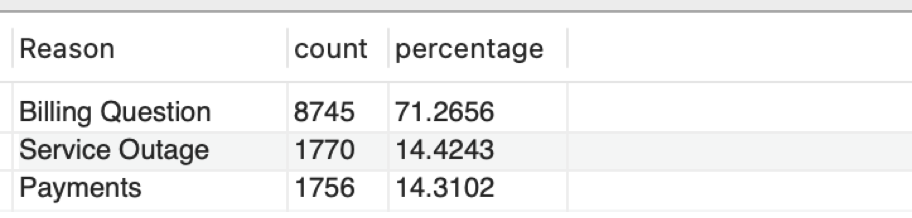
However, having a larger dataset might provide a more accurate representation of the overall sentiment.

**Reason distribution:**



****

****

****

The SQL query calculate the Count and Percentage of occurrences for each unique value in the Reason column to shows the customer satisfaction by Reason of call.

***Interpretation of the Output***

**Billing Question**:

* **Count:** 8,745
* **Percentage:** 71.27%

Billing Question is the most common reason for calls, representing 71.27% of the total calls. This suggests that the majority of customer interactions are related to billing inquiries.

**Service Outage**:

* **Count**: 1,770
* **Percentage**: 14.42%

Service Outage Is the second most common reason, accounting for 14.42% of the total calls. This indicates that a significant portion of customers are calling about service disruptions.

**Payments**:

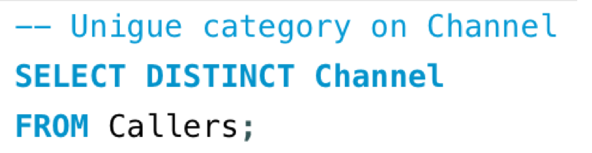
* **Count**: 1,756
* **Percentage:** 14.31%

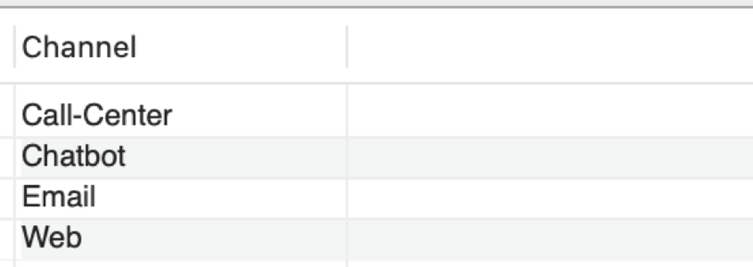
Payments is the third most common reason for calls, with 14.31% of the total. This suggests that payment-related inquiries are also a frequent reason for customer interactions.

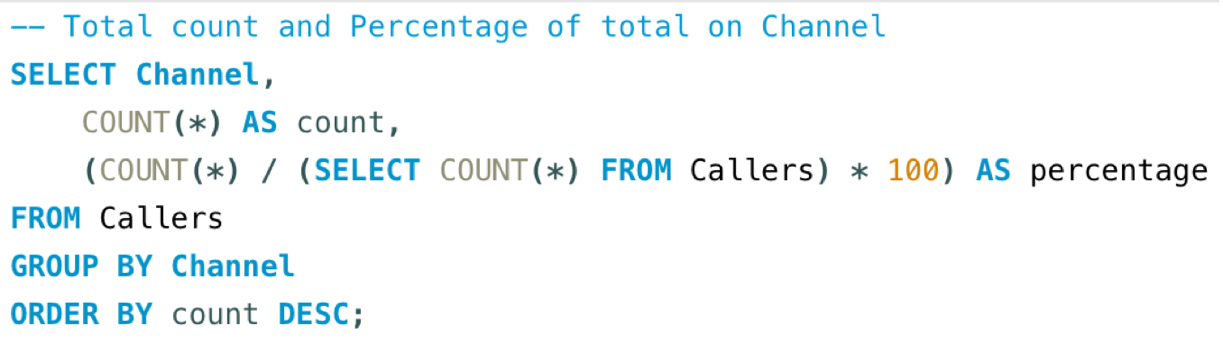
In summary, the output indicates that the majority of calls are related to billing issues, followed by service outages and payment inquiries.

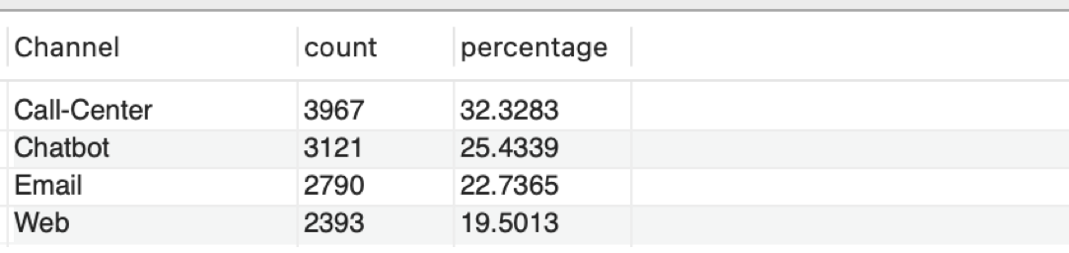
The high percentage of billing-related calls (71.27%) may suggest a need for better billing processes, clearer communication, or self-service options to reduce the volume of calls. Addressing these key areas could potentially improve customer satisfaction and reduce the overall call volume.

**Channel distribution:**



****

****

****

The query returned total Count and Percentage of customer interactions for each communication channel e.g., Call-Center, Chatbot, Email and Web.

Indicating the customer satisfaction by Channel of communication.

***Interpretation of Output***

**Call-Center**

* **Count:** 3,967
* **Percentage**: 32.33%

The Call-Center is the most frequently used channel, accounting for 32.33% of all customer interactions. This suggests that a significant portion of customers prefer direct human interaction to resolve their issues or inquiries.

**Chatbot**

* **Count**: 3,121
* **Percentage**: 25.43%

The Chatbot is the second most used channel, making up 25.43% of interactions. This indicates that a large number of customers are comfortable using automated systems for support.

**Email**

* **Count**: 2,790
* **Percentage**: 22.74%

Email is used in 22.74% of interactions, showing that a notable portion of customers prefer written communication, likely for issues that don't require immediate responses.

**Web**

* **Count**: 2,393
* **Percentage**: 19.50%

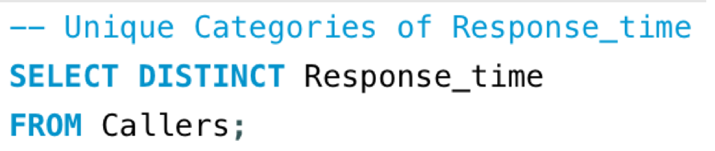
The Web channel, accounting for 19.50% of interactions, is the least used among the listed options. This could involve online forms, FAQs, or other web-based customer service tools.

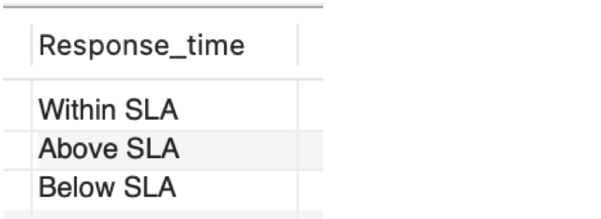
Summary, the data suggests that the Call-Center is the most preferred channel for customer interactions, followed by the Chatbot, Email, and Web. The high usage of the Call-Center indicates a strong demand for direct, human interaction (assistance), while the significant usage of Chatbot and Email highlights the growing acceptance of automated and asynchronous communication methods.

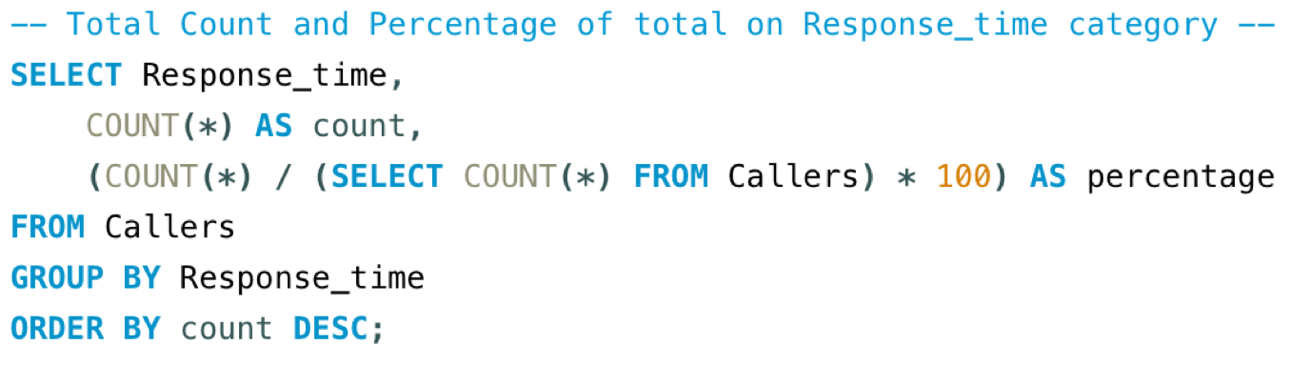
The lower percentage for Web-based interactions suggests that while web tools are useful, they may not be the first choice for many customers seeking support.

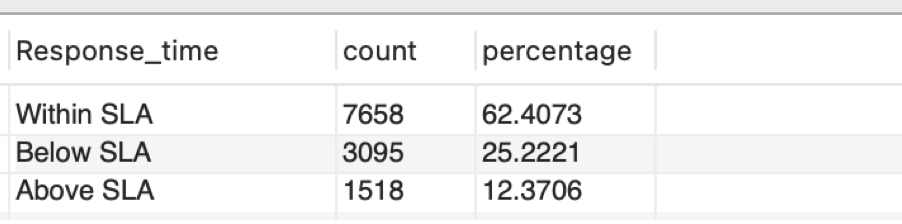
This information could be valuable in allocating resources and improving the efficiency of customer support services, particularly in balancing between human-operated and automated channels.

**Response\_time distribution:**

****

****

****

****

The query displays the Count and Percentage of customer interactions categorized by "Response\_time"

This data helps to understand how often response times fall within, below, or above the Service Level Agreement (SLA).

***Output Interpretation***

**Within SLA**

* **Count**: 7,658
* **Percentage**: 62.41%

The majority of interactions, 62.41%, are handled within the agreed-upon Service Level Agreement (SLA). This indicates that most responses are timely and meet the expected standards.

**Below SLA**

* **Count**: 3,095
* **Percentage**: 25.22%

25.22% of the interactions fall below the SLA, meaning they were handled more quickly than required. This might suggest efficient handling of some cases, potentially exceeding customer expectations.

**Above SLA**

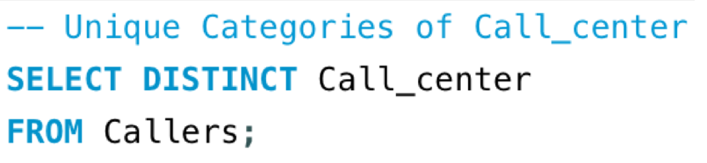
* **Count**: 1,518
* **Percentage**: 12.37%

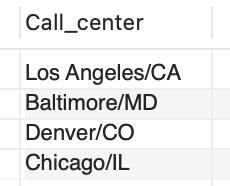
12.37% of the interactions exceed the SLA, meaning they were not handled within the expected timeframe. This is a smaller portion, but it indicates an area where improvement is needed to ensure all interactions are timely.

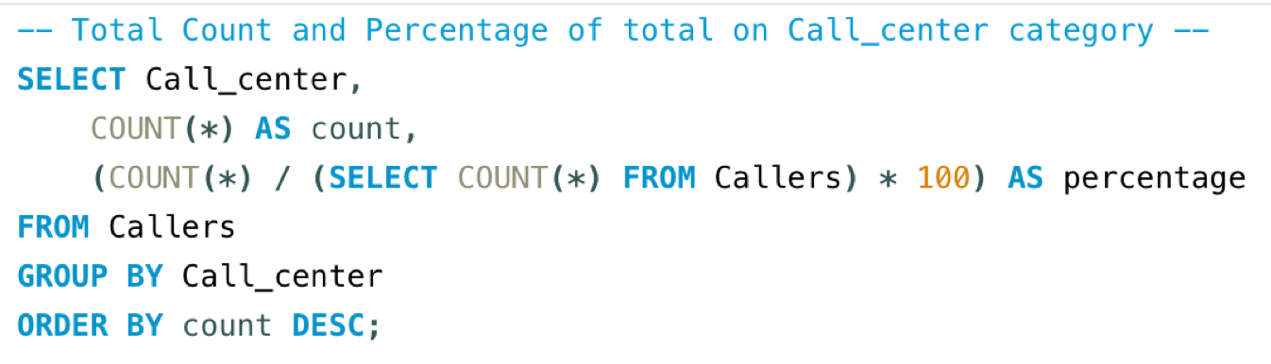
In summary, the output shows that most customer interactions (62.41%) are resolved within the agreed SLA, demonstrating strong adherence to response time commitments. A notable portion (25.22%) of cases are resolved faster than required, reflecting efficiency in handling these interactions. However, 12.37% of interactions exceed the SLA, suggesting there is room for improvement to ensure that all inquiries are addressed in a timely manner.

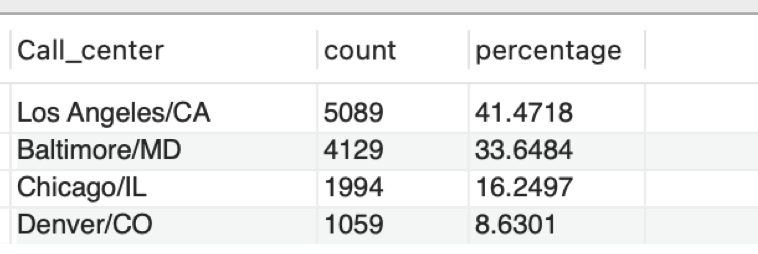
This analysis can help pinpoint areas where service efficiency can be enhanced to better meet customer expectations.

**Call Center distribution:**

****

****

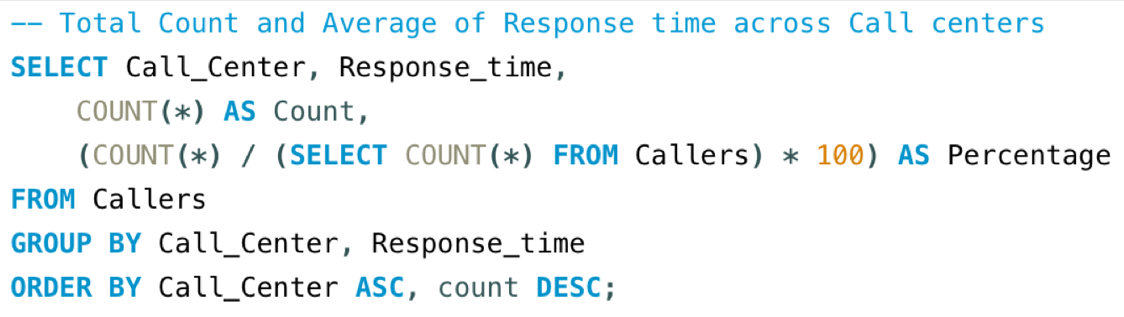
****

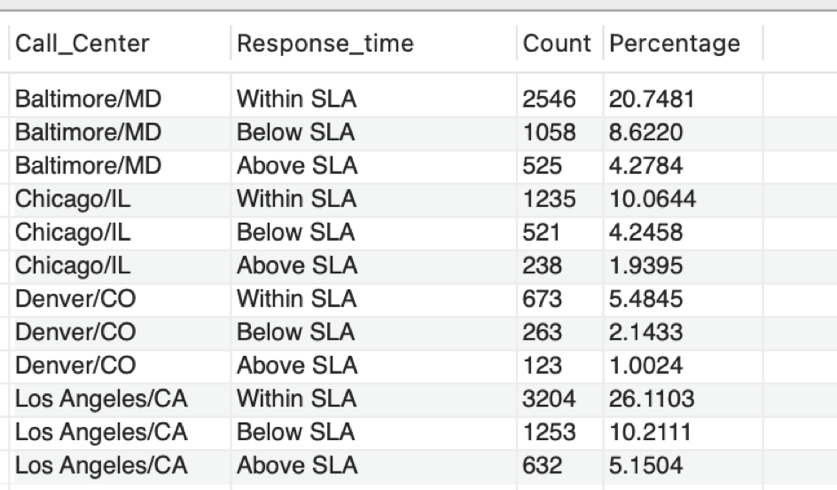
****

***Interpretation of Output***

The output provides a summary of call volume distribution across different call centers.

* **Los Angeles/CA** is the busiest call center, handling 41.47% of all calls.
* **Baltimore/MD** is the second busiest, responsible for 33.65% of calls.
* **Chicago/IL** and **Denver/CO** handle a smaller portion of calls, with 16.25% and 8.63%, respectively.

****

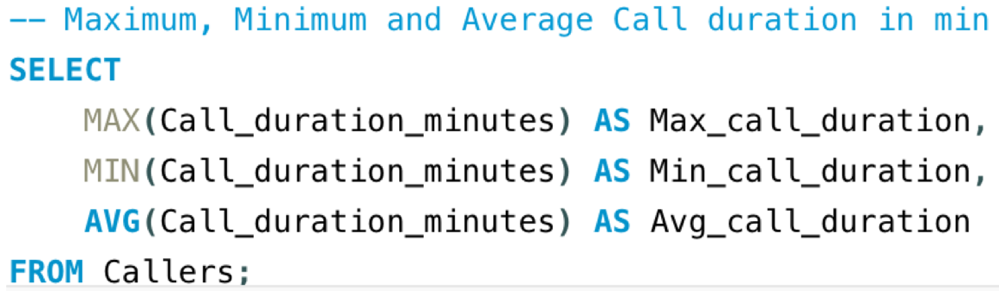
****

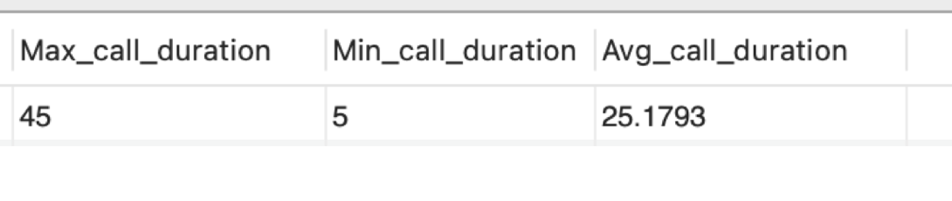
To gain more insights into each call center performance, the query returns the Count and Average of calls resolved within, below and above Service Level Agreement (SLA) for each Call\_center.

We can now determine which call center consistently exceed or fall short of performance targets. Then investigate or explore factors contributing to longer response time e.g. high call volume, agent skills level, system issues.

We can also analyse centers with high SLA adherence to identify successful strategies.

**Minimum, Maximum and Average call duration**



****