Telecom Market Analysis

Javier Lopez

D205, Data Acquisition

Table of Contents

A. Research Questions

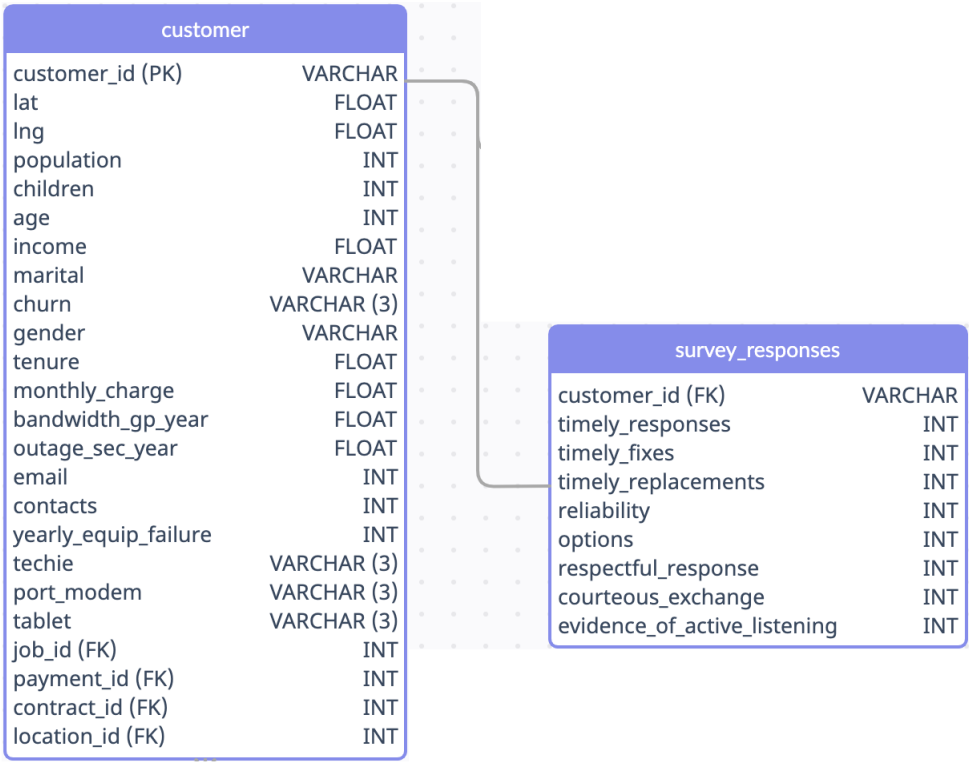
How did active customers rate the importance of each survey category compared to churned customers?

A1. Identify Required Data

We will use two tables from our churn database to answer our research question, the customer table and the survey responses table. From the customer table, we will use the customer id column and the churn column. From the survey responses table, we will use the customer_id, timely_response, timely_fixes, timely_replacements, reliability, options, respectful_response, courteous_exchange, and evidence_of_active_listening columns.

B. Logical Data Model

The logical data model below shows the one-to-one relationship between the two tables needed to answer the research question, customer, and survey responses. The customer table and survey responses table have a one-to-one relationship as one customer can only have one respond two the survey once. The survey responses table has a foreign key constraint that references the customer table’s primary key, the customer id field.



B1. Create Table SQL Code

The code snippet below creates the survey responses table if it does not already exist within the churn database. We then list the column name, data type, and any constraints within the query.

```
CREATE TABLE IF NOT EXISTS public.survey_responses (
    customer_id VARCHAR NOT NULL,
    timely_responses INT,
    timely_fixes INT,
    timely_replacements INT,
    reliability INT,
    options INT,
    respectful_response INT,
    courteous_exchange INT,
    evidence_of_active_listening INT,
    CONSTRAINT survey_responses_pkey PRIMARY KEY (customer_id),
    CONSTRAINT survey_responses_fkey FOREIGN KEY (customer_id)
    REFERENCES customer (customer_id)
);
```

B2. Load Data SQL Code

The code snippet below loads data from the CSV data file into the survey responses table.

```
\\copy public.survey_responses (customer_id, timely_responses, timely_fixes, timely_replacements, reliability, options, respectful_response, courteous_exchange, evidence_of_active_listening) FROM '/Users/javilopezcastillo/Downloads/churnfiles/Survey_Responses.csv' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '"' FORCE NOT NULL customer_id ;
```

C. Queries SQL Code

The code snippet below joins the customer and survey responses tables on the customer id column. We then group the data using the Yes and No values on the churn columns to identify active and churned customers. We continue to count the number of customers within each category. Lastly, we calculated the average importance score that each customer answered for each survey question.

```sql
SELECT DISTINCT churn,
       COUNT(churn) AS customers,
       ROUND(AVG(timely_responses), 2) AS avg_timely_responses,
       ROUND(AVG(timely_fixes), 2) AS avg_timely_fixes,
       ROUND(AVG(timely_replacements), 2) AS avg_timely_replacements,
       ROUND(AVG(reliability), 2) AS avg_reliability,
       ROUND(AVG(options), 2) AS avg_options,
       ROUND(AVG(respectful_response), 2) AS avg_respectful_response,
       ROUND(AVG(courteous_exchange), 2) AS avg_courteous_exchange,
       ROUND(AVG(evidence_of_active_listening), 2) AS avg_evidence_of_active_listening
FROM customer AS c
INNER JOIN survey_responses AS s
ON c.customer_id = s.customer_id
GROUP BY churn;
```

D. Data Refresh Rate

To gain an accurate answer to our research question, the data should be refreshed monthly as an understanding of customer sentiment is needed to make business decisions regarding the effects of actions taken to shift overall sentiment.

E. SQL Script to Load Data

Using the method listed on the Postgres Tutorial Website, we created a script using the copy command on a local Mac device that will load the data from the *services* and *survey_responses* CSV files into their corresponding tables (Postgres Tutorial, 2022).

```
COPY survey_responses (
    customer_id,
    timely_responses,
    timely_fixes,
    timely_replacements,
    reliability,
    options,
    respectful_response,
    courteous_exchange,
    evidence_of_active_listening
)
FROM '/Users/javilopezcastillo/Downloads/churnfiles/Survey_Responses.csv'
DELIMITER ','
CSV HEADER;
```

G. Web Sources

The source used to acquire the data supporting our application was Lab on Demand. The virtual machine on Labs on Demand hosts the churn database and the survey responses CSV file that holds the data used to answer our research question.

1. Labs on Demand, https://labclient.labondemand.com/LabClient/a71055f3-a9fa-4487-9b32-5832f98b02ae?rc=10.

H. Sources

1. “Import CSV File into Postgresql Table.” *PostgreSQL Tutorial*, https://www.postgresqltutorial.com/postgresql-tutorial/import-csv-file-into-posgresql-table/.