WGU MSDA Program

SQL Exploration Analysis on Survey Database

Write Up

Christopher Kamper

August 31st, 2022

1. Research Question: What is the average reliability rating given by each marital grouping from the survey?

A1. To answer the research question, data from both the original customer table and the newly created survey\_responses table was needed. The customer\_id and marital columns were pulled from the customer table. After the CSV file provided was loaded to the survey\_responses table, the reliability column was pulled, matching the values based on the unique customer\_id from both tables.

1. Table creation and addition to the database.

B1. The survey\_responses table was created to load the provided CSV file. The columns were set up to match the information in the CSV along with the data type. Below is the code used to create the survey\_responses table:

CREATE TABLE survey\_responses (

customer\_id text,

timely\_responses numeric,

timely\_fixes numeric,

timely\_replacements numeric,

reliability numeric,

options numeric,

respectful\_reponse numeric,

courteous\_exchange numeric,

evidence\_of\_active\_listening numeric,

CONSTRAINT PK\_customer PRIMARY KEY (customer\_id));

The survey\_responses table was set up with a primary key based on the customer\_id as that column is full of unique values that can be used to identify data. Below is the Entity Relationship Diagram (ERD) for the database after the survey\_responses table was added:

Diagram

Description automatically generated with medium confidence

B2. The survey\_responses table was filled with data by importing the provided CSV file. Below is the code used to import this data (Kartones, 2019):

--command " "\\copy public.survey\_responses (customer\_id, timely\_responses, timely\_fixes, timely\_replacements, reliability, options, respectful\_responses, courteous\_exchange, evidence\_of\_active\_listening) FROM 'C:/Users/LabUser/Desktop/SURVEY~1.CSV' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '''';""

1. To gather the data needed to answer the research question, the customer and survey\_responses tables were joined using a left join. This combined the customer\_id, marital, and reliability columns to show just the information used to calculate the necessary information in the next step. Below is the code used to join these two tables:

SELECT c.customer\_id, c.marital, s.reliability

FROM customer AS c

LEFT JOIN survey\_responses AS s

ON c.customer\_id=s.customer\_id;

This code was then modified to calculate the necessary information to answer the research question. This added an average calculation of the reliability column where the data was grouped by the marital column. The result shows the average rating given by each grouping in order of lowest rating to the highest rating. Below is the code used to calculate this information:

SELECT c.marital, ROUND(AVG(s.reliability),2) AS avg\_reliability

FROM customer AS c

LEFT JOIN survey\_responses AS s

ON c.customer\_id=s.customer\_id

GROUP BY c.marital

ORDER BY avg\_reliability;

C1. CSV files of the output from the previous two code sets are included in this submission.

1. If this data is collected as a survey conducted over a specified time that does not recur often, then the data should be added as the add-on file after each round.

If this data is collected continuously through a website or email that may occur anytime a service request is completed, the data should be added as the add-on file on a regular cadence such as weekly or monthly.

1. The code is the same as the code mentioned in B2. Below is the code used to import this data (Kartones, 2019):

--command " "\\copy public.survey\_responses (customer\_id, timely\_responses, timely\_fixes, timely\_replacements, reliability, options, respectful\_responses, courteous\_exchange, evidence\_of\_active\_listening) FROM 'C:/Users/LabUser/Desktop/SURVEY~1.CSV' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '''';""

G. Sources:

Kartones, D. (n.d.). PostgreSQL command line Cheatsheet. Gist. Retrieved August 28, 2022, from https://gist.github.com/Kartones/dd3ff5ec5ea238d4c546