D205 PA

Jon Isak

D205: Data Acquisition

William Sewell

D205 PA

1. This paper will research the question, "What is the average outage time for customers with high reliability ratings?" This question seeks to identify the average outage time for customers who rated reliability as high. Identifying this could help the business determine the average downtime it can have without significantly affecting customers' perception of reliability.

A1.  The data required includes the numeric "outage\_sec\_week" column and the text "customer\_id" column from the "customer" table. The numeric "reliability" column and the text "customer\_id" column from the "survey" table will also be used. The "customer\_id" column will serve as the primary key in the "customer" table and as a foreign key in the "survey" table.

1. The ERD shows the relationship between the "customer" and "survey" tables. It also illustrates that the "survey" table's "customer\_id" has a many-to-one relationship with the "customer\_id" in the "customer" table because a customer can complete multiple surveys, but each survey can only belong to one customer.

A screenshot of a computer

Description automatically generated

B1.  This SQL code creates two tables, "customer" and "survey." "Customer\_id" is set as the primary key in the "customer" table and as a foreign key in the "survey" table, establishing a relationship between the two tables.

CREATE TABLE public.survey (

customer\_id text NOT NULL,

timely\_responses integer,

timely\_fixes integer,

timely\_replacements integer,

reliability integer,

options integer,

respectful\_response integer,

courteous\_exchange integer,

evidence\_of\_active\_listening integer,

CONSTRAINT fk\_survey\_customer FOREIGN KEY (customer\_id) REFERENCES public.customer (customer\_id)

);

CREATE TABLE public.customer (

customer\_id text NOT NULL,

lat numeric,

lng numeric,

population integer,

children integer,

age integer,

income numeric,

marital text,

churn text,

gender text,

tenure numeric,

monthly\_charge numeric,

bandwidth\_gp\_year numeric,

outage\_sec\_week numeric,

email integer,

contacts integer,

yearly\_equip\_faiure integer,

techie text,

port\_modem text,

tablet text,

PRIMARY KEY (customer\_id)

);

B2.  --command " "\\copy public.survey (customer\_id, timely\_responses, timely\_fixes, timely\_replacements, reliability, options, respectful\_response, courteous\_exchange, evidence\_of\_active\_listening) FROM 'C:/Users/LabUser/DOWNLO~1/survey.csv' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '''';""

This code, generated by pgAdmin using the import tool, imports "survey.csv" into the "survey" table. Import settings were adjusted to include a header and set a comma as the delimiter to match the file structure.

C.   SELECT AVG(c.outage\_sec\_week) AS average\_outage\_time

FROM public.customer AS c

INNER JOIN public.survey AS s

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

WHERE s.reliability in (4,5);

This SQL query returns one column with the average of the data in the "outage\_sec\_week" column for customers who rated reliability as high (a rating of 4 or 5). The SQL statement joins the "customer" and "survey" tables using an inner join on "customer\_id." The query shows that the average downtime per week for customers who rated reliability as high was 10.02 seconds.

C1.  A CSV file is attached to show the query result. The file shows one column.

D. The question in part A requires survey data acquired through collecting data from customers. The data should be refreshed quarterly.

D1. A quarterly refresh rate is relevant to the business's needs because it allows the business to collect customer satisfaction data often enough to reflect changes in the customer base and their needs, without inconveniencing them with frequent survey requests.

E.  <https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=2707579e-7967-4e67-bdc2-b08a007a8d9d>

F.  No web sources were used.