Alma Munoz

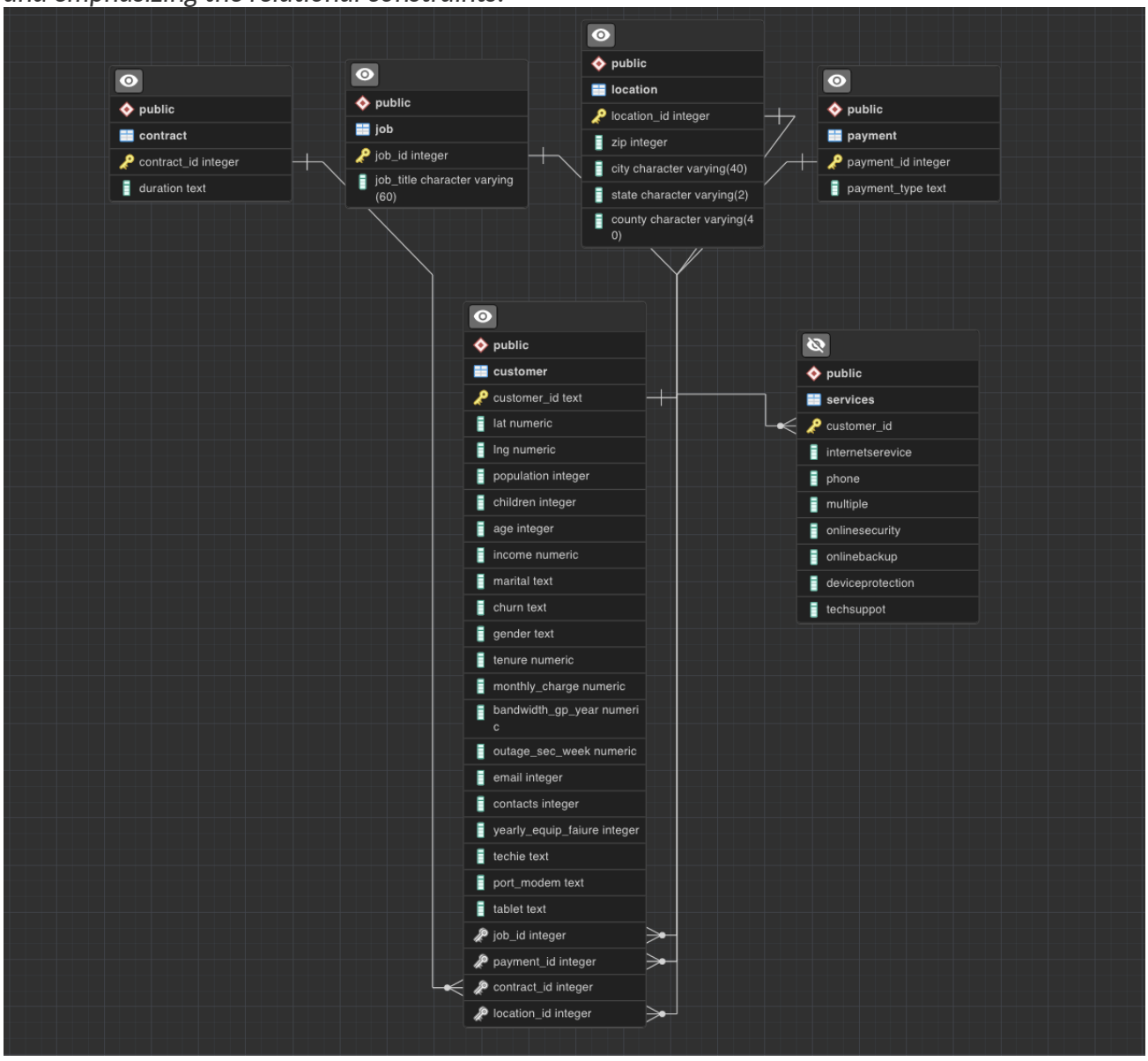
D205 DATA ACQUISITION

Performance Assesment

A. Summarize a research question that can be answered using both *the original database and the add-on CSV data. The question should require data from* both *these data sources.*

Do customers with children tend to use protection on their device more due to incidents that could happen with their device?
The data needed for this is customer table from the original given data and the services table from the add-on data. These two will give us a count of the customers with children and in connection with those who have device protection.

B. Create a logical data model for the add-on CSV file by evaluating the data contained in the file and emphasizing the relational constraints.



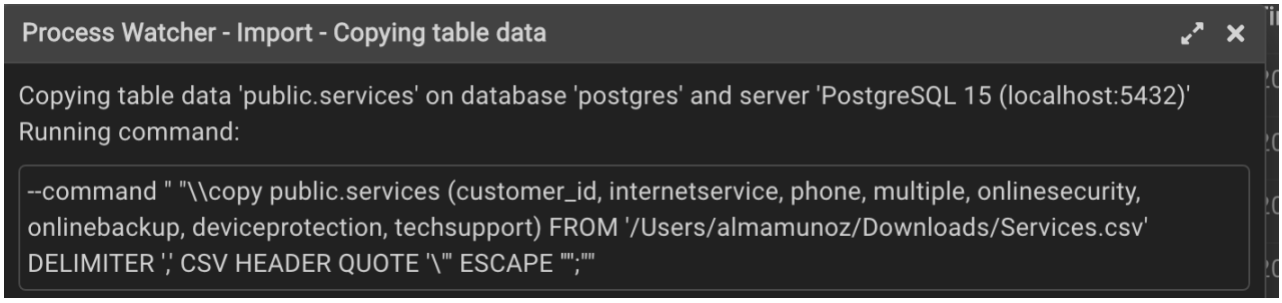
1. Write SQL code that creates a table that accommodates the extension of the logical data model to a physical data model by specifying the field types and relevant keys.

```
CREATE TABLE IF NOT EXISTS public.services
(
    customer_id character varying(30) COLLATE pg_catalog."default" NOT NULL,
    internetservice character varying(30) COLLATE pg_catalog."default",
    phone character varying(10) COLLATE pg_catalog."default",
    multiple character varying(3) COLLATE pg_catalog."default",
    onlinesecurity character varying(3) COLLATE pg_catalog."default",
    onlinebackup character varying(3) COLLATE pg_catalog."default",
    deviceprotection character varying(3) COLLATE pg_catalog."default",
    techsupport character varying(3) COLLATE pg_catalog."default",
    CONSTRAINT service_pkey PRIMARY KEY (customer_id)
)
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public.services
    OWNER to postgres;
```

```sql
-- Table: public.services

-- DROP TABLE IF EXISTS public.services;

CREATE TABLE IF NOT EXISTS public.services
(
    customer_id character varying(30) COLLATE pg_catalog."default" NOT NULL,
    internetservice character varying(30) COLLATE pg_catalog."default",
    phone character varying(10) COLLATE pg_catalog."default",
    multiple character varying(3) COLLATE pg_catalog."default",
    onlinesecurity character varying(3) COLLATE pg_catalog."default",
    onlinebackup character varying(3) COLLATE pg_catalog."default",
    deviceprotection character varying(3) COLLATE pg_catalog."default",
    techsupport character varying(3) COLLATE pg_catalog."default",
    CONSTRAINT service_pkey PRIMARY KEY (customer_id)
)

TABLESPACE pg_default;

ALTER TABLE IF EXISTS public.services
    OWNER to postgres;
```

2. Write SQL code that loads the data from the add-on CSV file into the table created in part B1.

```
-command " "\\copy public.services (customer_id, internetservice, phone, multiple, onlinesecurity,
    onlinebackup, deviceprotection, techsupport) FROM '/Users/almamunoz/Downloads/Services.csv'
    DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '"';""
```

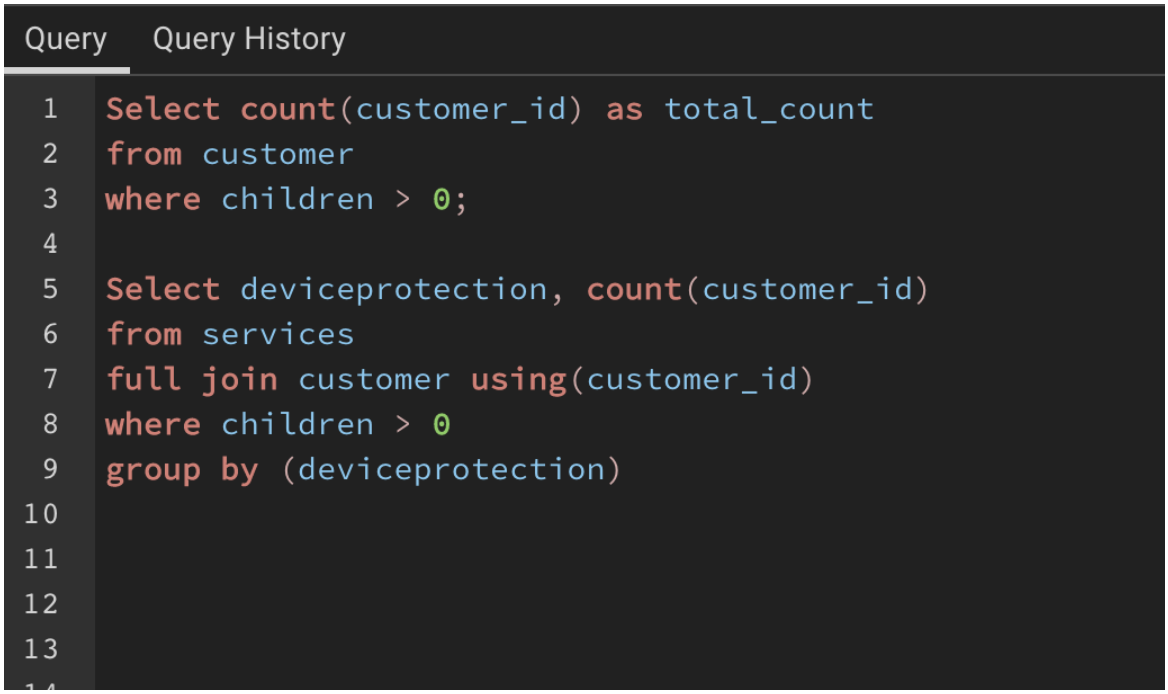


(Agarwal et al., 2023)

C. Write SQL statement(s) for a query or queries that inform the research question summarized in part A.

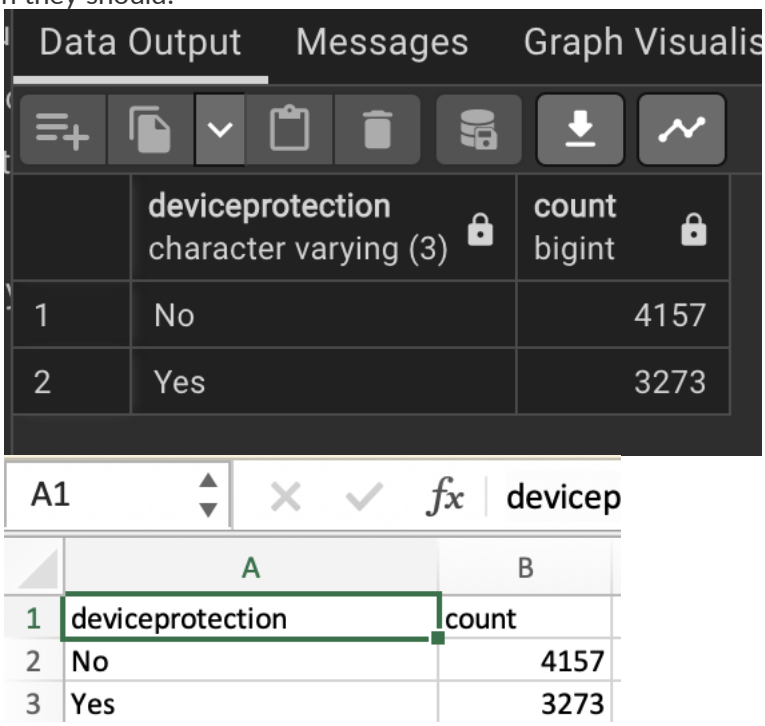
This SQL statement counts the number of customers with children and full joins with the table that have device protection.

```
Select count(customer_id) as total_count
from customer
where children > 0;
Select deviceprotection, count(customer_id)
from services
full join customer using(customer_id)
where children > 0
group by (deviceprotection)
```



1. *Provide a CSV file or files that capture the results from the query or queries.*

Here we can see that those with children, there are less with device protection. The results revealed the although precaution is always warranted, those while children still do not have more device protection than they should.



D205results <- CSV file

D. Determine how often the add-on file should be acquired and refreshed in the database for the data to remain relevant to the business and the research question.

The add-on file should be acquired and refreshed in the database every quarter. This way every quarter the company can review and prepare a case for making sure those with children get device protection. With every idea coming out quarterly, the company can double check if the marketing project to get more adults with children to inquire about device protection is working or not.

E. Create an SQL script that performs the process of loading the add-on data.

```
Copy public."Services"
From 'Users/almamunoz/Downloads/Services.csv'
Delimiter ',' CSV Header;
```

(Agarwal et al., 2023)

F. Provide a Panopto video recording that includes a demonstration of the functionality of the code used for the analysis and a summary of the programming environment.

Note: For instructions on how to access and use Panopto, use the "Panopto How-To Videos" web link provided below. To access Panopto's website, navigate to the web link titled "Panopto Access" and then choose to log in using the "WGU" option. If prompted, log in using your WGU student portal credentials, and then it will forward you to Panopto's website.

To submit your recording, upload it to the Panopto drop box titled "XXXX." Once the recording has been uploaded and processed in Panopto's system, retrieve the URL of the recording from Panopto and copy and paste it into the Links option. Upload the remaining task requirements using the Attachments option.

G. Record the web sources used to acquire data or segments of third-party code to support the application. Be sure the web sources are reliable.

Agarwal, S., Dwivedi, P. and Talha PostgreSQL import CSV: 3 easy methods - learn https://hevodata.com/learn/postgresql-import-csv/ (accessed Feb, 2023).

H. Acknowledge sources, using in-text citations and references, for content that is quoted, paraphrased, or summarized.

I. Demonstrate professional communication in the content and presentation of your submission.