# Data Acquisition -D205

Student ID – **009412995**

Student Name – Rishika Gottepally

**Performance Assessment**

1. **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.**

🡪How many customers used 'Electronic Check' and 'Mailed Check' as their payment method for each internet service.

1. **Identify which data from the original data set and the add-on CSV file are needed to answer the research question.**

🡪From the original data set customer table – customer\_id and payment\_id data are required.

Payment table- Payment\_type data is required.

Contract table- Contract duration data is required to understand how often the data needs to be acquired and refreshed.

Add-on csv file internet service data is required from the services.csv file.

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

Graphical user interface

Description automatically generated

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 public.Internet**

**(customer\_id varchar(20),**

**internet\_service varchar(20) not null,**

**phone varchar(5) not null,**

**multiple varchar(5) not null,**

**online\_security varchar(5) not null,**

**online\_backup varchar(5) not null,**

**device\_protection varchar(5) not null,**

**tech\_support varchar(5) not null,**

**primary key (customer\_id) not null);**

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

**“Copy public.internet**

**From ‘c:\Users\Labusers\Desktop\Services.csv’**

**Delimiter ‘,’**

**CSV Header;”**

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

**create table public.Internet**

**(customer\_id varchar(20) not null,**

**internet\_service varchar(20) not null,**

**phone varchar(5) not null,**

**multiple varchar(5) not null,**

**online\_security varchar(5) not null,**

**online\_backup varchar(5) not null,**

**device\_protection varchar(5) not null,**

**tech\_support varchar(5) not null,**

**primary key (customer\_id));**

**copy public.internet**

**from ‘c:\Users\Labusers\Desktop\Services.csv’**

**delimiter ‘,’**

**csv Header;**

**select count(c.customer\_id),i.internet\_service,p.payment\_type,cont.duration**

**from public.customer as c**

**join public.Internet as i using(customer\_id)**

**join public.payment as p using(payment\_id)**

**join public.contract as cont using (contract\_id)**

**where payment\_type in('Electronic Check','Mailed Check')**

**group by i.internet\_service,p.payment\_type,cont.duration**

**order by i.internet\_service;**

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

Table

Description automatically generated

**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 Csv file should be acquired and refreshed every month as there are a total of 5456 customers on month-to-month contract\_id.

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

“**Copy public.internet**

**From ‘c:\Users\Labusers\Desktop\Services.csv’**

**Delimiter ‘,’**

**CSV Header;”**

**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.**

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=02667def-6790-4746-bfd3-adff0163bbc6>

**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.**

🡪 I have not used any third-party code or sources to acquire data.

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

🡪 I have not used any external references.