Koraun Wilson

000669871

Advanced Data Management D191

VDM1 Performance Assessment

1. **Summarize one real-world business report that can be created from the attached Data Sets and Associated Dictionaries.**
   1. Our company would like to create a business report to show the top 50 customers ordered by the number of rentals they’ve had. This will allow us to form marketing campaigns specifically geared toward our most loyal customers that help generate the most revenue for the company. We can also form a rewards program for our top customers to further encourage repeated business.
2. **Describe the data used for the report.**
   1. We will need to use the customer’s info to identify them for the report. This info will come from the customer table, while we will use the rental table to help order number of rentals the data by the customer id. We will also use the address, city, and country tables to pull in customer data. The customer\_id will link connect the data from the customer and rental tables to grab the address information from the address, city and country tables.
3. **Identify two or more specific tables from the given dataset that will provide the data necessary for the detailed and the summary sections of the report.**
   1. The tables that will be used to provide data are **customer, rental, address, city and country**
4. **Identify the specific fields that will be included in the detailed and the summary sections of the report.**
   1. The specific fields that will be used in the summary report are:  
      **full\_name, total\_rentals, email, city and country as full\_address**  
      The specific fields that will be used in the detailed report are:  
      **customer\_id, first\_name and last\_name as full\_name, total\_rentals, email, address, city and country**
5. **Identify**one**field in the detailed section that will require a custom transformation and explain why it should be transformed. For example, you might translate a field with a value of ‘N’ to ‘No’ and ‘Y’ to ‘Yes’.**
   1. One field in the detailed section that will require custom transformation is the first name and last name fields, which will be joined using concatenation into a field called full\_name, to make the report easier to read. I will also join the city and country together in the summary report in a field that will be named full\_address, to make it easier to read and to shorten the amount of columns in the summary.
6. **Explain the different business uses of the detailed and the summary sections of the report.**
   1. The detailed report will show our top 50 customers by the number of rentals, along with customer information such as the email address, address, and country. This will allow the marketing team to specifically target our top customers with offers and rewards that can be tailored to the customer based on their country of residence as well as the localized address.
   2. The summarized report includes the customer’s full name and the number of rentals and the email address, this allows our company to have a quick view of the total amounts of rentals of the top 50 customers, which can be used to help target marketing and ad campaigns.
7. **Explain how frequently your report should be refreshed to remain relevant to stakeholder**
   1. This report should be refreshed weekly to verify that the report is up to date with current rental numbers, so the report data can be used to contact the customers that are currently within our top 50 renters.
8. **Write a SQL code that creates the tables to hold your report sections.**

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

1. **Write a SQL query that will extract the raw data needed for the Detailed section of your report from the source database and verify the data’s accuracy.**Graphical user interface, text, application, email

   Description automatically generated

I verified the accuracy of the report by running the select \* from detailed code and was able to pull back the detailed table that has the customers ordered by the number of total rentals they have. I can see that it only includes 50 records and it lists them in descending order by the number of total rentals. I don’t see any missing values and all of the data types in each column are what was expected. Graphical user interface, text

Description automatically generated

1. **Write code for function(s) that perform the transformation(s) you identified in part A4.**

The code to transform the first name and last name are already present in the code to populate the detailed table.

Below I have the code to enter data into the summary table, this also has the code to use concatenation on the city and country to create the full address.

Graphical user interface, text, application, email

Description automatically generated

1. **Write a SQL code that creates a trigger on the detailed table of the report that will continually update the summary table as data is added to the detailed table.**

Graphical user interface, text, application, email

Description automatically generated **Graphical user interface, text, application, email

Description automatically generated**

1. **Create a stored procedure that can be used to refresh the data in *both* your detailed and summary tables. The procedure should clear the contents of the detailed and summary tables and perform the ETL load process from part C and include comments that identify how often the stored procedure should be executed.**

Graphical user interface, text, email

Description automatically generated

Graphical user interface, text, email

Description automatically generated

1. **Explain how the stored procedure can be run on a schedule to ensure data freshness**

To ensure data freshness, the procedure should be run at least on a monthly basis, but I recommend weekly. It can also be run on demand if needed as well. The admin should use PgAgent, the job scheduling agent for PostgreSQL. It does need to be installed as it doesn’t come with PgAdmin 4, but once it’s installed, a new job can be set up in PgAgent jobs, and the schedule will be set up under the Schedules tab of the job.

**G.   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.**

**H.   Record the web sources you used to acquire data or segments of third-party code to support the application if applicable. Be sure the web sources are reliable.**

I did not use web sources.

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

No in text citations or references.