CSN1 — CSN1 TASK 1: DATA ANALYSIS

ADVANCED DATA MANAGEMENT — D326

PRFA — CSN1

TASK OVERVIEW   SUBMISSIONS   EVALUATION REPORT

COMPETENCIES

4037.5.1 : **Writes Structured Query Language (SQL) Statements**
The learner writes complex Structured Query Language (SQL) statements for data analysis and manipulation

4037.5.2 : **Configures Automated Tasks**
The learner configures data extraction, transformation, and loading tasks for automated data integration

INTRODUCTION

Data analysts frequently transform data within a database so that it can be used for analysis and so that the data will be easier for nontechnical users to use and understand. You will emulate such a process in this task by choosing your own business question to analyze, creating tables and queries to use as a business report, and streamlining your analysis by writing your own SQL functions, triggers, and stored procedures.

This task defines a report as a collection of data that answers a real-world business question. Your report will have two distinct sections (SQL tables that you will create) that differ in the granularity of the data they present and how directly they support the answering of the business question you choose. The detailed table should contain all data that informs the answer to the question at a very granular level, and the summary table should contain aggregated data that provide a direct answer to the business question.

REQUIREMENTS

Your submission must be your original work. No more than a combined total of 30% of the submission and no more than a 10% match to any one individual source can be directly quoted or closely paraphrased from sources, even if cited correctly. The similarity report that is provided when you submit your task can be used as a guide.

You must use the rubric to direct the creation of your submission because it provides detailed criteria that will be used to evaluate your work. Each requirement below may be evaluated by more than one rubric aspect. The rubric aspect titles may contain hyperlinks to relevant portions of the course.

*Tasks may **not** be submitted as cloud links, such as links to Google Docs, Google Slides, OneDrive, etc., unless specified in the task requirements. All other*

submissions must be file types that are uploaded and submitted as attachments (e.g., .docx, .pdf, .ppt).

To work on this task, use the "Labs on Demand Assessment Environment and DVD Database" found in the Web Links section. In this environment, you will be able to write and test your PostgreSQL code and access the database to complete this task.

Plan for and compose the sections of a real-world business report that can be created from the "Labs on Demand Assessment Environment and DVD Database" web link, and demonstrate the functionality of the supporting SQL code by doing the following:

A. Summarize **one** real-world written business report that can be created from the DVD Dataset from the "Labs on Demand Assessment Environment and DVD Database" attachment.
    1. Identify the specific fields that will be included in the detailed table and the summary table of the report.
    2. Describe the types of data fields used for the report.
    3. Identify *at least* **two** specific tables from the given dataset that will provide the data necessary for the detailed table section and the summary table section of the report.
    4. Identify *at least* **one** field in the detailed table section that will require a custom transformation with a user-defined function and explain why it should be transformed (e.g., you might translate a field with a value of N to No and Y to Yes).
    5. Explain the different business uses of the detailed table section and the summary table section of the report.
    6. Explain how frequently your report should be refreshed to remain relevant to stakeholders.

B. Provide original code for function(s) in text format that perform the transformation(s) you identified in part A4.

C. Provide original SQL code in a text format that creates the detailed and summary tables to hold your report table sections.

D. Provide an original SQL query in a text format that will extract the raw data needed for the detailed section of your report from the source database.

E. Provide original SQL code in a text format 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.

F. Provide an original stored procedure in a text format that can be used to refresh the data in *both* the detailed table and summary table. The procedure should clear the contents of the detailed table and summary table and perform the raw data extraction from part D.
    1. Identify a relevant job scheduling tool that can be used to automate the stored procedure.

G. Provide a Panopto video recording that includes the presenter and a vocalized demonstration of the functionality of the code used for the analysis.

   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 "Advanced Data Management D191 | D326 (Student Creators) [assignments]." 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.

H. Acknowledge all utilized sources, including any sources of third-party code, using in-text citations and references. If no sources are used, clearly declare that no sources were used to support your submission.

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

File Restrictions

File name may contain only letters, numbers, spaces, and these symbols: ! - _ . * ' ( )
File size limit: 400 MB
File types allowed: doc, docx, rtf, xls, xlsx, ppt, pptx, odt, pdf, txt, qt, mov, mpg, avi, mp3, wav, mp4, wma, flv, asf, mpeg, wmv, m4v, svg, tif, tiff, jpeg, jpg, gif, png, zip, rar, tar, 7z

RUBRIC

A:BUSINESS REPORT

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
A summary of 1 real-world business report created from the dataset is not provided.	The submission summarizes the 1 real-world written business report created from the dataset, but it is either incomplete or inaccurate.	The submission accurately and completely summarizes 1 real-world written business report created from the dataset.

A1:SPECIFIC FIELDS

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
A submission identifying 1 or more specific fields that will be included in the detailed table and summary table of the report is not provided.	The submission does not identify 1 or more of the specific fields that will be included in the detailed table and summary table of the report.	The submission identifies *all* of the specific fields that will be included in the detailed table and summary table of the report.

A2:TYPES OF DATA

NOT EVIDENT

A description of the data fields to be used in the report is not provided.

APPROACHING COMPETENCE

The submission describing the types of data fields to be used in the report is either inaccurate or incomplete.

COMPETENT

The submission accurately and completely describes the types of data fields that will be used in the report.

A3:SPECIFIC TABLES

NOT EVIDENT

A submission identifying *at least* 2 specific tables from the given dataset is not provided.

APPROACHING COMPETENCE

The submission identifies *at least* 2 of the tables from the given data set, but 1 or both of the identified tables will not provide the data necessary for the detailed table section and summary table section of the report.

COMPETENT

The submission identifies *at least* 2 tables from the given dataset that provide the data necessary for the detailed table section and summary table section of the report.

A4:FIELD TRANSFORMATION

NOT EVIDENT

A submission identifying *at least* 1 field in the detailed table section of the report that will require a custom transformation is not provided.

APPROACHING COMPETENCE

The submission identifies 1 field in the detailed table section of the report but does not explain why the field should be transformed with a user-defined function.

COMPETENT

The submission identifies at least 1 field in the detailed table section of the report that will require a custom transformation and explains why it should be transformed with a user-defined function.

A5:BUSINESS USES

NOT EVIDENT

An explanation of the business uses of the summary and detailed table sections of the report is not provided.

APPROACHING COMPETENCE

The submission is missing either the explanation of the detailed table section or the explanation of the summary table section. Or the explanation did not accurately and logically explain how the sections can be used to improve business.

COMPETENT

The business uses of the summary and detailed table sections of the report accurately and logically explains how they can be used to improve business.

A6:REPORT FRESHNESS

NOT EVIDENT

An explanation of how frequently the report should be refreshed to remain relevant to stakeholders is not provided.

APPROACHING COMPETENCE

The explanation of how frequently the report should be refreshed has an unreasonable frequency, or the explanation is inadequate or illogical.

COMPETENT

The explanation of how frequently the report should be refreshed to remain relevant to stakeholders has a reasonable frequency. The explanation is adequate and logical.

B:FUNCTIONS

NOT EVIDENT

The code for the function(s) that perform the transformation(s) identified in part A4 is not provided.

APPROACHING COMPETENCE

The code for the function(s) that perform the transformation(s) identified in part A4 *either* did not result in the expected transformations, or 1 or more of the previously identified functions is not included in the submission, or the code is not in text format.

COMPETENT

The text format code for the function(s) that perform the transformation(s) identified in part A4 produced the expected transformations and *all* previously identified functions are included in the submission.

C:CREATING TABLES

NOT EVIDENT

The SQL code that creates the tables to hold the detailed and summary tables report sections is not provided.

APPROACHING COMPETENCE

The SQL code that creates the detailed and summary tables to hold the report sections is incorrect, incomplete, or not in a text format.

COMPETENT

The SQL code that creates the detailed and summary tables is correct, complete, and in a text format.

D:SQL QUERY

NOT EVIDENT

The SQL query to extract raw data is not provided.

APPROACHING COMPETENCE

The SQL query provided does not correctly query the raw data needed for the detailed section of the report from the source database or is not in a text format.

COMPETENT

The text-format SQL query provided correctly queries the raw data needed for the detailed section of the report from the source database.

E:TRIGGERS

NOT EVIDENT

The SQL code that creates a trigger on the detailed table of the report is not provided.

APPROACHING COMPETENCE

The provided SQL code does not create a trigger or the trigger does not continually update the summary table as data is added to the detailed table, or the code is not in text format.

COMPETENT

The SQL code in text format successfully 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.

F:STORED PROCEDURES

NOT EVIDENT

A stored procedure to refresh the data is not provided.

APPROACHING COMPETENCE

The stored procedure provided does not refresh the data in 1 or *both* of the detailed table and summary table, or it does not clear the contents of the detailed table and summary table, or it does not perform the raw data extraction from part D. Or the submission is not in text format.

COMPETENT

The stored procedure in a text format provided refreshes the data in *both* the detailed table and summary table, and it can clear the contents of the detailed table and summary table and perform the raw data extraction from part D.

F1:JOB SCHEDULING TOOL

NOT EVIDENT

A tool is not identified.

APPROACHING COMPETENCE

A tool is identified but cannot be used to automate the stored procedure.

COMPETENT

A relevant job scheduling tool that can be used to automate the stored procedure is identified.

G: PANOPTO VIDEO:

NOT EVIDENT

The submission does not provide a Panopto video recording.

APPROACHING COMPETENCE

A Panopto video recording is provided but does not include a demonstration of the functionality of the code used for the analysis or the video does not capture both the presenter and the vocalized

COMPETENT

A Panopto video recording is provided that includes a demonstration of the functionality of the code used for the analysis. For the duration of the presentation, the video captures both the presen-

presentation of the functioning code in a Panopto video recording.

ter and the vocalized presentation of the functioning code in a Panopto video recording.

H:SOURCES

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The submission does not include *both* in-text citations and a reference list for any sources that are quoted, paraphrased, or summarized.	Sources are utilized in the submission but are either not acknowledged in-text or not included in a reference list, or the sources cited are not reliable.	All utilized sources are reliable and acknowledged using in-text citations and references, or the submission clearly identifies that no sources were used.

I:PROFESSIONAL COMMUNICATION

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
Content is unstructured, is disjointed, or contains pervasive errors in mechanics, usage, or grammar. Vocabulary or tone is unprofessional or distracts from the topic.	Content is poorly organized, is difficult to follow, or contains errors in mechanics, usage, or grammar that cause confusion. Terminology is misused or ineffective.	Content reflects attention to detail, is organized, and focuses on the main ideas as prescribed in the task or chosen by the candidate. Terminology is pertinent, is used correctly, and effectively conveys the intended meaning. Mechanics, usage, and grammar promote accurate interpretation and understanding.

WEB LINKS

Labs on Demand Assessment Environment and DVD Database

Panopto Access
Sign in using the "WGU" option. If prompted, log in with your WGU student portal credentials, which should forward you to Panopto's website. If you have any problems accessing Panopto, please contact Assessment Services at assessmentservices@wgu.edu. It may take up to two business days to receive your WGU Panopto recording permissions once you have begun the course.

Panopto How-To Videos

Skillable Labs Knowledge Base Article
Please consult this WGU Knowledge Base article for general FAQs regarding your Skillable lab environment.