TGM1 — TGM1 TASK 1: DATA ACQUISITION

DATA ACQUISITON — D205
PRFA — TGM1

TASK OVERVIEW | SUBMISSIONS | EVALUATION REPORT

COMPETENCIES

4034.3.1 : Dimension, Quality Relations, and Limitations
The graduate examines the data available for analysis to determine their dimension, quality, relations, and limitations.

4034.3.2 : Physical Data Models
The graduate implements physical data models.

4034.3.3 : Table Operations
The graduate performs table operations and queries within the context of data acquisition for analysis.

INTRODUCTION

A common undertaking for a data analyst is to connect an external data collection to an existing data set. This process involves the exploration of the source and target data sets to merge the new data in a meaningful and logical way.

You will replicate this common process in this task. You will take external data from a CSV file and design conceptual, logical, and physical models that describe the data. You will load the new data into an existing database alongside the existing data and run queries across it.

The data sets, associated data dictionaries, and the CSV file are located at the Labs on Demand site where you will work on this performance assessment. Please follow the Labs on Demand link in the Web Links section.

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).*

Perform the following tasks to combine an existing data set with some external data:

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

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

C. Write SQL statement(s) for a query or queries that inform the research question summarized in part A.
   1. Provide a CSV file or files that capture the results from the query or queries.

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.

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

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.

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.

File Restrictions

File name may contain only letters, numbers, spaces, and these symbols: ! - _ . * ' ( )
File size limit: 200 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:RESEARCH QUESTION

NOT EVIDENT	APPROACHING	COMPETENT

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
A summary of the research question is not provided.	The summary of the research question is illogical or inaccurate, or it does not require data from the original database or the add-on CSV file data.	The summary of the research question is logical and requires data from the original database and the add-on CSV file data.

A1:IDENTIFYING DATA

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The submission does not identify which data from the original data set and the add-on CSV file are needed to answer the research question.	The submission identifies which data from the original data set and the add-on CSV file are needed, but the identified data do not provide sufficient information to answer the research question.	The submission identifies which data from the original data set and the add-on CSV file are needed and provides sufficient information to answer the research question.

B:LOGICAL DATA MODEL

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The logical data model for the add-on CSV file is not provided.	The logical data model for the add-on CSV is inaccurate or does not identify relational constraints.	The logical data model for the add-on CSV file that was created is accurate and identifies relational constraints.

B1:CODE FOR THE PHYSICAL DATA MODEL

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The SQL code is not provided.	The SQL is incorrect, or it does not create a table that extends the logical data model to the physical data model, or it does not correctly specify field types or relevant keys.	The SQL code correctly 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.

B2:LOADING CSV DATA

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
SQL code that loads the data from the add-on CSV file into the table created in part B1 is not provided.	The SQL code does not correctly load the data from the add-on CSV file into the table	The SQL code correctly loads the data from the add-on CSV file into the table created in part B1.

	created in part B1.	

C:SQL QUERY

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
SQL statement(s) for a query or queries that inform the research question summarized in part A is not provided.	SQL statement(s) for a query or queries does not inform the research question summarized in part A or is incorrect.	SQL statement(s) for a query or queries correctly informs the research question summarized in part A.

C1:CSV FILE(S)

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The CSV files(s) were not provided.	The CSV file(s) did not accurately capture the results of the query or queries.	The CSV file(s) accurately capture the results of the query or queries.

D:ADD-ON FILE

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The submission does not 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 submission does not accurately or logically 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 submission accurately and logically determines 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.

E:SQL SCRIPT

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
An SQL script that performs the process of loading the add-on data is not provided.	The SQL script does not perform the process of loading the add-on data.	The SQL script successfully performs the process of loading the add-on data.

F:PANOPTO VIDEO

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
A Panopto video recording is not provided.	The Panopto video recording does not include a full demonstration of the functionality of the code used for the analysis or an accurate or complete summary of the programming environment.	The Panopto video recording includes a full demonstration of the functionality of the code used for the analysis and an accurate and complete summary of the programming environment.

G:WEB SOURCES

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
A record of the web sources used to acquire data or segments of third-party code to support the application is not provided.	The record the web sources used to acquire data or segments of third-party code to support the application is incomplete or inaccurate. Or the web sources cited are not reliable.	The record the web sources used to acquire data or segments of third-party code to support the application is both complete and accurate, and the web sources cited are reliable. Or the candidate did not use any web sources to acquire data or segments of third-party code and stated this in their submission.

H:SOURCES

NOT EVIDENT	APPROACHING COMPETENCE	COMPETENT
The submission does not include both in-text citations and a reference list for sources that are quoted, paraphrased, or summarized.	The submission includes in-text citations for sources that are quoted, paraphrased, or summarized and a reference list; however, the citations or reference list is incomplete or inaccurate.	The submission includes in-text citations for sources that are properly quoted, paraphrased, or summarized and a reference list that accurately identifies the author, date, title, and source location as available.

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	Content is poorly organized, is difficult to follow, or contains errors in mechanics, usage, or grammar that cause confusion. Terminology is misused or	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,

from the topic.

terminology is misused or ineffective.

is used correctly, and effectively conveys the intended meaning. Mechanics, usage, and grammar promote accurate interpretation and understanding.

WEB LINKS

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

Labs On Demand

Skillable Labs Knowledge Base Article

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