Competencies

**4157.1.1** : **Recommends an Appropriate Data Architecture**

The learner recommends an appropriate data architecture.

**4157.1.2** : **Examines the Data Available for Analysis**

The learner examines the data available for analysis to determine their dimension, quality, relations, and limitations.

**4157.1.4** : **Implements Physical Data Models**

The learner implements physical data models by incorporating required data.

**4157.1.5** : **Performs Queries to Answer a Business Question**

The learner performs database queries to answer a business question.

**4157.1.6** : **Uses Appropriate Data Manipulation Language Statements**

The learner uses appropriate data manipulation language statements to manage data.

Introduction

A common undertaking for a data analyst is to create databases using existing datasets. This process involves the exploration of the source and target datasets to merge the data in a meaningful and logical way.

You will replicate this common process as you complete this task. You will take external data from either a document dataset or a graphical dataset and design a non-relational database solution. You will load the data into a database solution and retrieve or modify the data using JSON datasets and MongoDB in the virtual lab environment.

Choose one of the scenarios provided in the Supporting Documents. Each scenario has associated JSON files that will be located on the desktop of the WGU Virtual Lab environment. You will complete the following:

1.  Design a **non-relational** database solution.

2.  Implement the solution.

3.  Present the solution using Panopto to an audience that consists of your peers within a project team.

*Note: It is recommended to use the other scenario, different from the one you used in Task 1.*

You will complete this performance assessment in the provided WGU Virtual Lab environment. Your submission will be a design document that includes screenshots of your work in the lab environment and a Panopto video.

Requirements

Your submission must represent your original work and understanding of the course material. Most performance assessment submissions are automatically scanned through the WGU similarity checker. Students are strongly encouraged to wait for the similarity report to generate after uploading their work and then review it to ensure Academic Authenticity guidelines are met before submitting the file for evaluation. See [Understanding Similarity Reports](https://cm.wgu.edu/t5/Frequently-Asked-Questions/Understanding-Similarity-Reports/ta-p/252) for more information.    
  
**Grammarly Note:**   
Professional Communication will be automatically assessed through Grammarly for Education in most performance assessments before a student submits work for evaluation. Students are strongly encouraged to review the Grammarly for Education feedback prior to submitting work for evaluation, as the overall submission will not pass without this aspect passing. See [Use Grammarly for Education Effectively](https://cm.wgu.edu/t5/Academic-Coaching-Center/Use-Grammarly-for-Education-Effectively/ta-p/52276) for more information.    
  
**Microsoft Files Note:**  
Write your paper in Microsoft Word (.doc or .docx) unless another Microsoft product, or pdf, is specified in the task directions. Tasks may not be submitted as cloud links, such as links to Google Docs, Google Slides, OneDrive, etc.  All supporting documentation, such as screenshots and proof of experience, should be collected in a pdf file and submitted separately from the main file. For more information, please see [Computer System and Technology Requirements.](https://cm.wgu.edu/t5/WGU-Student-Policy-Handbook/Computer-System-and-Technology-Requirements/ta-p/78)    
 *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.*

**Part 1: Design Document**

A.  Select one of the provided scenarios and complete the following:

1.  Describe a business problem that can be solved with a database solution and is in alignment with the chosen scenario.

2.  Justify why a NoSQL database solution will solve the identified business problem.

3.  Identify a NoSQL database type to solve the identified business problem.

4.  Explain how the business data will be used within the database solution.

B.  Discuss how the proposed database design addresses scalability concerns, including strategies that align with the chosen scenario.

C.  Outline the privacy and security measures that should be implemented in the proposed database design.

**Part 2: Implementation**

*Note: The data files for each scenario are located in a folder titled “D597 Datasets” on the desktop of the WGU Virtual Lab environment. Be sure to pull the files from “Task 2” that relate to your chosen scenario.*

*Note: Submit your screenshots from the WGU Virtual Lab for each prompt with your design document.*

D.  Implement the proposed database design in the WGU Virtual Lab environment by completing the following:

1.  Write script to create a database instance named “D597 Task 2” using the appropriate query language, based on your design in Part 1. Provide a screenshot showing the script and the database instance in the platform.

2.  Write script to insert or map the data records from the chosen scenario JSON files into the database instance. Provide a screenshot showing the script and the data correctly inserted or mapped into the database.

3.  Write script for **three** queries to retrieve specific information from the database that will help to solve the identified business problem. Provide a screenshot showing the script for each query and each query successfully executed.

4.  Apply optimization techniques to improve the run time of your queries from part D3, providing output results via a screenshot.