Competencies

**4158.1.2** : **Integrates Syntax, Control Elements, and Modular Components**

The learner integrates programming language syntax, control elements, and modular components from common data analytics programming languages to create simple programs.

**4158.1.3** : **Performs Data Acquisition and Organization Tasks**

The learner performs data acquisition and organization tasks using a data analytics programming language.

Introduction

In this task you will demonstrate your ability to convert pseudocode into a functional program that can carry out set tasks for data analysis.

Scenario

You have been hired by a small investment company that manages an equity fund comprised of 150 U.S. companies across multiple industries. The fund managers are looking to rebalance the fund’s holdings and would like you to provide an analysis of the companies’ performance based on data from the most recent quarter. The data can be found in the supporting documents section as “D598 Data Set”.

To aid in your analysis you must write programs in Python or R to perform the following tasks:

•   Import the data file into a data frame.

•   Identify any duplicate rows in the data set.

•   Group all IDs by state, then run descriptive statistics (mean, median, min, & max) for all numeric variables by state and store this result as a new data frame. (Code should be modified from “D598 Task 2 Original Code” in the supporting documents section)

•   Filter the data frame to identify all businesses with debt-to-equity ratios that are negative.

•   Create a new data frame that provides the debt-to-income ratio for every business in the data set. Debt-to-income ratio is defined as long-term debt divided by revenue.

•   Concatenate the debt-to-income ratio data frame you created with the original data frame.

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.*

In this task you will create the following program in Gitlab:

A.  Create a program in Gitlab using Python or R to perform the data analysis described in Task 1.

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

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