
GLOBAL ACADEMY OF MATHEMATICAL AND ECONOMIC SCIENCES

*October 2021*

**SQL PORTFOLIO PROJECT**

Exploratory Data Analysis on International Debt Data Using SQL

**COURSE: DATA AND BA FOUNDATIONS**

Supervisor: Zion Pibowei

## PART A: ANALYSIS

Set up a new database in **SQLite Studio** with the name **sqlproject.db**. From the "Databases" panel, double-click on your newly created database. The World Bank international debt data has been provided for this project. Import this dataset as a table into your newly created database (ensure to tick the field "first line represents CSV column names").

Upon successful loading of the data, open a new SQL editor file to start writing your queries. For each question below, use a separate SQL editor and save each script to your computer afterwards. Each script **must** be saved as Question*X*.sql where *X* represents the question number.

Write all your queries for each question within one script and use comments to separate your queries (Check out the lesson on **Adding Comments in SQL**).

In your queries, write column names as: "ColumnName" or [ColumnName]

### Question 1

Write a query to compute how many different countries were recorded in the data.

### Question 2

There is a column called **indicator_name** that briefly specifies the purpose of taking the debt. Just beside that column, there is another column called **indicator_code** which symbolises the category of these debts.

Write (i) a query to compute the total number of unique indicator codes, (ii) a query to retrieve the unique indicator codes.

### Question 3

Write queries to compute the total debt owed by all countries in each of the following years (i) 2000 (ii) 2001 (iii) 2002 (iv) 2007 (v) 2008 (vi) 2009 (vii) 2013 (viii) 2014 (ix) 2015 (x) 1988

Use the **AS** keyword to create an alias for each (e.g., **AS** *total_debt_2000*)

### Question 4

Write queries to compute the total debt owed by Nigeria in (i) 1997 (ii) 1999 (iii) 2006 (iv) 2007 (v) 2015 (vi) 2016. Your result set should each include country name, country code and the value computed.

**Question 5**

Write queries to compute the maximum debt owed by Nigeria in (i) 2007 (ii) 2008 (iii) 2009. Your result set should each include country name, country code and the value computed.

**Question 6**

Write queries to compute the minimum debt owed by Nigeria in (i) 2007 (ii) 2008 (iii) 2009. Your result set should each include country name, country code and the value computed.

**Question 7**

(a) In what year did Nigeria have the largest interest arrears from private creditors and what is the value? Write a query to compute this.

(b) In what year did Nigeria have the largest multilateral debt, and what is the value of this debt in percentage?

**Question 8**

Write queries to return a table of the maximum debt owed by each country in descending order for each of the following years (i) 1998 (ii) 1999 (iii) 2001 (iv) 2010 (v) 2012 (vi) 2015

Use the clause **GROUP BY** [Country Name] before the **ORDER BY** clause.

**Question 9**

Write queriesx to compute which country had the largest debt in (i) 2003 (ii) 2004 (iii) 2005 (iv) 2008 (v) 2009 (vi) 2012

Use the clause **GROUP BY** [Country Name] before the **ORDER BY** clause.

**Question 10**

Find the country that had the highest minimum debt and the value of this debt in each of the following years: (i) 2003 (ii) 2004 (iii) 2005 (iv) 2008 (v) 2009 (vi) 2012

### Question 11

Write a query to return a table of the average debt owed by each country in descending order for each of the following years, limiting your result set to the first 20 rows. (i) 2003 (ii) 2004 (iii) 2005 (iv) 2008 (v) 2009 (vi) 2012

Your result set should include the country code, country name, and the average debt. Use the clause **GROUP BY** [Country Code], [Country Name] before the **ORDER BY** clause.

### Question 12

Write a query to return a table of the country with the largest average debt in each of the following years (i) 1998 (ii) 1999 (iii) 2001 (iv) 2010 (v) 2012 (vi) 2015

This is similar to the previous question, except you are getting only the record with the highest average debt value.

## PART B: REPORTING

Write a report detailing each of the steps you took and the results you obtained in Part A. Your report should be structured according to the following headings:

- **Introduction:** In this section, briefly describe the purpose of the project and the motivation for the analysis of international debt data. Give a quick summary of the dataset, the tools you used for the analysis, the dataset,
- **About the Data:** The dataset provided was obtained from Kaggle (see link in Classroom). Use the details on Kaggle via the link provided to provide firsthand insights about the data in this section. Be creative – no points will be awarded for a copy-and-paste job!
- **Loading the Data** (or, alternatively, **Data Ingestion**): Describe the steps you took to import your data into SQLite studio, from setting up the database to loading the data. How do you check the data in the table after importing it? How many rows per page is loaded? Why does SQLite Studio load this data page by page, and not everything at once?

- **Analysis and Results:** This section should be subdivided into subsections each representing the questions in this project. Your responses in each subsection should clearly answer the question(s), and include the query(ies) you ran and the results or table you obtained. You can use descriptive titles relevant to the question for your subsections. For example, you could title your subsection for Question 1 as "Computing the number of distinct countries", etc. You can also run additional queries of your choice and create subsections to present your code and results.
- **Discussion:** In this section, you are to present any observations from your analysis in the previous section. Discuss any interesting insight or patterns you gleaned from your analysis.
- **References (if any)**
- **Appendix (if any)**

## SUBMISSION

For Part A, you are to write a separate script for each of the 12 questions and save each script as **Question*X*.sql** where *X* represents the question number.

You are to save your report in pdf or docx format. A cover page will be provided for you which captures the course title and project title, and will give room for your name and track. You are expected to prepare your report with this template.

Upon completion of the project, you are to create a zip folder containing all the files you are to submit, and name the zip folder **SQL Project – Your Name**, where {Your Name} includes your first name and last name.

You are to submit this project latest by **Wednesday, October 20, 2021.**

## GRADING

Total number of points allotted to this project is **500 points**, divided equally between Part A and Part B. Good luck!