Database Fundamentals and Applications

Assignment 3: SQL with Tableau and Python

Instructors:
Ronald de Vlaming
Ignas Krikštaponis

March 26, 2021

Contents

1 Database installation

2 Tableau assignment (X points in total)

2.1 Introduction

In this assignment you will create a number of visualisations on Tableau to effectively communicate the impact COVID-19 had on the EU. CoronaImpactEU database has a number of tables that posses information related to employment and vaccination roll-out. You can find the documentation of the database at the end of this document.

The main goal of this assignment is to allow you to gain some experience with the the Tableau software. Therefore, the main challenge will lay in figuring out how to perform certain things yourself. You are encouraged to use any educational Tableau resources you find online to create the visualisations described in the assignment.

2.2 Connecting to the CoronaImpactEU database

- Upon starting Tableau you will be greeted with the Connect page. Here you can connect to various data sources such as Excel spreadsheets, PDF files and SQL databases.
- Select MySQL under the "To a Server" section. Note: if you do not have a MySQL connector installed on your system you will be asked to install it - click on Download Driver and follow the instructions.
- Use details below to connect to the database you will use for the Tableau assignment:
  - Server: localhost
  - Port: 3306
  - Database: CoronaImpactEU
  - Username: root
  - Password: enter your root user password here

You should now be redirected to the Data Sources tab where you should see the CoronaImpactEU tables on the left.

2.3 Problems

2.3.1 Data preparation (X points in total)

- Go over all tables and their columns - make sure that the data types have been parsed correctly and fix any inconsistencies. (X points)
- Use Tableau's drag and drop interface to establish the necessary links between the tables - use the provided documentation. (X points)

2.3.2 Dashboard 1: Youth unemployment (X points in total)

The main table you will use for creating this dashboard is called "YouthUnemployment". Please read all bullet points before starting.

- Create an extra column in "YouthUnemployment" table named "MonthStart" by concatenating "-01" to the column "Period". Make sure to set the newly created column's data type to date. (X points)
- Map visualisation - create a map visualisation that colours countries based on their monthly youth unemployment rate. Choose Blue-Green sequential palette. (X points)
- Horizontal bar chart - put country names in rows and their monthly unemployment rates in columns. Order countries in descending order based on their monthly youth unemployment rate. (X points)
- Filter 1 - create a filter for MonthStart. It must be of single value slider type and apply to both Map and Horizontal bar chart visualisations. (X points)
- Line chart - create a line chart that shows monthly youth unemployment rate (set aggregation function to average). Show only year and month. Narrow the y-axis range to make the trend clearly visible. (X points)
- Filter 2 - create a filter (that apply to line chart visualisations) that allows the user to select which country is being displayed. It should be of drop-down (single value) type. Also, there should not be an option to display aggregated data of all countries. (X points)
- Combine all the above created items into one dashboard named "Youth unemployment". (X points)
- Allow the Vertical Bar Chart and the Map visualisations to act as filters. (X points)

2.3.3 Dashboard 2: Vaccination roll-out (X points in total)

The main tables you will use for creating this dashboard are called "vaccinationroll-out" and "vaccines". Please read all bullet points before starting.

- Stacked bars visualisation - create a stacked bar visualisation to display how many first doses were administrated every week. Color the bars based on the vaccine type. (X points)
- Parameter - create a parameter that would allow users to change what they see in the stacked bar visualisation you have created in the step above. The parameter should allow the user to change the row axis to either First dose, Second dose or Number of doses received. (X points)
- New data - make a new connection to CoronaImpactEU database. In this connection include a SQL statement that returns a rolling percentage of vaccine targeted population that have received first and second dose. This should be done on a country level (do not differentiate based on the vaccine), also, the statement should return separate percentages for first and second dose - do not sum them up. (X points)
- Area chart (continues) - use the data you have extracted in the step above to create an area chart (continues) visualisation that displays the percentage of target population vaccinated by first and second dose. (X points)

3 Python assignment

4 CoronaImpactEU database documentation

5 Yelp database documentation

References