**Instructions**

You must carefully prepare the data so that you feed it into your favorite visualization program and create the plots.

You can use MySQL to retrieve and join information from multiple tables, creating a sophisticated analysis. You are also welcome to perform this pre-processing step in Python or any other programming language you feel most comfortable with.  You can also manipulate the relevant data directly in Tableau, Power BI, or another visualization tool.

**Optional Instructions**

Let’s think about the information we need to retrieve for each of the visualizations, following the numbering from the Dashboard Structure section and the dashboard skeleton. The following tables can be constructed in, for example, MySQL.

* Construct a data source(**Source 1**), which will later form **Table** **1**in the dashboard. The source should contain the following columns:
  + Course name (Varchar)
  + Total minutes watched for each course (Decimal)
  + Average minutes for each course (Decimal)
  + Number of ratings for each source (Integer)
  + Average rating for each course (Decimal)
* Choose the User Type definition you want to work with from the **Data Dictionary** Construct a data source (**Source 2)**, which will later form **Chart** **1.4**, **KPI** **3.1**, and **KPI** **3.4**in the dashboard. The source should contain the following columns:
  + User ID (Integer)
  + Date of the user’s registration (Date)
  + Which country the user comes from (Varchar)
  + Whether the user is onboarded or not (Boolean)
  + Whether the user is paid or not (Boolean)
  + The subscription type of the user (Integer or NULL if the user is free)
* Construct a data source (**Source 3)**, which will later form **Chart** **1**, **Chart** **1.2**, **Chart** **1.3**, **KPI** **3.2**, and **KPI** **3.3**in the dashboard. The source should contain the following columns:
  + User ID (Integer)
  + When a video was watched (Date)
  + Which course the video is from (Integer)
  + The number of minutes watched (Decimal)
  + Whether the user is paid or not (Boolean)
  + The subscription type of the user (Integer or NULL if the user is free)
  + Date of the user’s registration (Date)
  + Which country the user comes from (Varchar)
  + Whether the user is onboarded or not (Boolean)
* The following columns from **Source 2** and **Source 3** would be used when filtering with **Parameters** **1-4.4**:
  + Date of registration
  + Date in which a video is watched
  + Whether the user is paid or not
  + The subscription type of the user (null if the user is free)
  + Which country the user comes from