Slido Analytics Assignment

Slido is a subscription software with customers purchasing year-long licences but customers can also use free Basic plans or purchase individual packages with rather low entry costs. However, mainly because of data sensitivity, large customers are keen to consolidate all accounts under one big deal.

To hit the target for this quarter, our sales team needs to close at least 5 company-wide deals. The sales director asked you to create a report and give recommendations on clients to approach about a company-wide deal offer.

Below you can find a data request from Pali, our Sales Director:

Hi data team! We need to hit our goal this quarter and the way to achieve it is GOING BIG and rolling out company-wide deals. However, we need your help in targeting those firms. We would appreciate the following report for a list of email domains. The struggle is real! Thanks for your help.

- email domain name

- # of all accounts

- # of users under the accounts

- # of paid annual subscriptions

- the total monetary value in EUR

- # events

- # of joined participants

- # of active participants

- AVG joined participants per event

- AVG active participants per event

- Activation 30 days

- Activation 90 days

# Data and definitions

* Go to [SQL Online Compiler](https://sqliteonline.com/) and – through *File > Open DB* – import [this database](https://drive.google.com/file/d/14uTV1bAcq7uaBUbx6H6XZl6_61cCfT9G/view?usp=sharing).
* You should see a dataset with 3 tables: accounts, subscriptions and events.
  + accounts: all accounts
  + subscriptions: all paid annual subscriptions
  + events: all used Slido events of accounts
* Information about users is in subscriptions, column users.
  + users: the number of all users under the account’s subscription including account owner
* Even if there is no subscription purchased under an account, it still has one user, i.e. account owner.
* Activation is the percentage of accounts with some Slido events within n days after signup.
* Prices are in EUR.

# What to do

* Process the provided three tables in SQL to create the list of email domains with the metrics as requested by Pali. *(Tip: use common table expressions.)*
* Export the created list and use the data to put together a final analysis of which customers are ripe for company-wide deals in Python\* using Jupyter Notebook or Google Colab.
* Feel free to analyse data out of Pali’s scope and show us anything helpful about our customers. Can you spot any red flags?
* Share with us:
  + SQL code for processing raw tables, incl. comments on your thinking process behind
  + Python notebook\* with your analysis, incl. comments on your thinking process behind
  + Summary of your conclusions and recommendations from your analysis, as a part of the Python notebook\* or separately

\* If you don’t know Python, you can use MS Excel or Google Sheets to make a spreadsheet, instead.