# Technical Interview Tasks

Interview has three tasks. Feel free to work on them in any order. You will be asked to present the solution during the 2nd round of the interview.

Feel free to use any open source library to accomplish the tasks if you think it is appropriate.

## Task 1 – Python

Finish the Python test here: ./Test Files/01\_Python\_test.py

## Task 2 – SQL

Let’s have a following table in PostgreSQL or Snowflake that contains the data about # of clicks on a websites per user and client

|  |  |  |
| --- | --- | --- |
| Name | Day | Clicks |
| Oakley Sosa | 2022-01-31 | 10 |
| Francesca Dotson | 2022-02-15 | 321 |
| Abel Valdez | 2022-01-15 | 56 |
| Yasin Rowland |  |  |
| … | … |  |

1. Show a number of clicks by User and calendar month
2. Show a share of clicks of each user by month

Test data are stored in ./Test Files/Website Statistics Data.csv

Please use the service <https://sqliteonline.com/> and use any database there.

## Task 3 – Architecture

The users mentioned in Task 2 also have records in the CRM database. At the end of each month, the sales department needs to match the usage data from website and with the CRM database for each user for further analysis. The user names in CRM and Website database are inconsistent, CRM has names in the following format:

|  |  |  |
| --- | --- | --- |
| User Name | Start Date | Subscription Revenue in Past 3 Months |
| O. Sosa | 1998-03-01 | $1.500 |
| Mrs. Francesca Dotson | 2005-04-01 | $200 |
| Abel Valdez, Phd. | 2000-12-01 | $350 |
| Yasin Rowland, MBA | 2010-01-01 | $100 |
| … | … | … |

1. Propose a Python solution that will solve the matching problem above – reads the data from source databases > matches both inputs by client name > save the result back to the database
2. Client names in both sources may not be consistent so exact matching/comparison may not work for 100% cases. The solution should include multiple matching strategies that can be applied to maximize the match rate.
3. Implement the solution using Python. The implementation should address key areas, although a production-ready implementation is not mandatory.
4. Describe how you would validate the result of the matching result during initial implementation and also during possible future enhancements of the matching algorithm by another developer

Please send us your completed task either as a zipped file attached to an email or as a link to your online drive.