

**Aim of this test task:** to assess your skills in data analytics and visualisation; note that there is no correct answer, and we'd like to hear your thoughts about the data set in this task.

**Task.** Briefly, the task can be formulated as follows:

1. Merge the data using sql;
2. Provide a description of the data and some quick insights.

**Output format:** report in pdf format or link to Google Docs or Jupyter Notebook (in English).

## Before you start

### Keep the goal in mind

The goal of the exercise is to properly merge and manipulate the data to show that there is potential for a further analysis of the problem. To the people involved in this problem, this analysis is only a small puzzle piece. We are not going to roll out a new model, or make major changes to our processes, based on your analysis alone. This means that you only need to touch on the particularities of the data and make them visible to other people.

Similarly, the real goal of the exercise is to get an interview. Making us want to meet you in person is more important than using the right jargon or coming up with the right numbers.

### There's no wrong answer

We're not looking for right answers. We're more interested in seeing the approach you take and your ability to make it easy to follow. We've tried to formulate the questions to go in that direction, but it might be worth making it even clearer. Don't hesitate to use screen shots, tell us what to look at, etc.

### Don't spend too much time

The exercise has the potential of being very time consuming. We don't want to waste your time. We'd recommend you timebox yourself to 3 hours with a view of having a rough document you could send us at the end of this time period.

## Step 1: Dataset and tools

For this task, we prepared a small database, where you can perform some queries.

You can use both Postgres (with file db\_postgres.sql) and MySQL (db\_mysql.sql).

The easiest way is to use pgAdmin for the Postgres database and MySQL Workbench for the MySQL database. Both allow you to easily import these databases.

The following links include an approachable description of how to import databases:

- <https://dev.mysql.com/doc/workbench/en/wb-admin-export-import-management.html>
- [https://www.pgadmin.org/docs/pgadmin4/6.18/import\\_export\\_data.html](https://www.pgadmin.org/docs/pgadmin4/6.18/import_export_data.html)

If you have any problems, we also included CSV files for each table.

**This task includes such tables:**

- **accounts** - Facebook public profiles with some data on them
- **posts** - posts of these public profiles, you can use the profile\_id column to connect to the accounts table
- **sources\_for\_followers** - number of followers for these public profiles

## Step 2: Merge the data

Create queries to merge data in SQL and provide short explanations for your queries.

## Step 3: Exploratory data analysis

We need to know if we have the complete data, can you please do exploratory analysis on merged data and provide short descriptions of the EDA steps and final summary on the data completeness. Hint: take a look at the posts' creation dates.

You do not need to write full sentences, but your answer should cover the following questions:

- Is the data complete?
- Does it all look fine?
- Are there obvious patterns in the data?

## Step 4: Look for insights and provide data sample of overperforming posts

Great! We have a better understanding of data! Let's finally do some analysis! We need to understand how the audience interacts with the post.

Can you please analyse engagement on provided posts and provide short summary of your findings?

We need to prepare a table with the overperforming posts. The next structure is great for this task (but you can add more information if needed):

- post id
- account username
- account id
- created time of post
- total engagement count
- comments count

### P.S. Some useful tips:

You can build some links using our data. Use the id of the posts to go to this post - [www.facebook.com/{id}](https://www.facebook.com/{id})

In the same way, you can use the id of the account to go to the page.