

Position	Data Scientist
Company	Botprise, Inc.
Test task	Text clustering

Problem setup:

- The Client has thousands of events produced by his monitoring system every day. In our case, events are parsed and represented as strings.
- The Client wants to have a daily report with events summary. Tickets data for 1 day is provided in the form of a csv file (tickets.csv) in the test task archive. There is 1 column: "ticket", without an index.

Task description:

Candidate should deliver Jupyter Notebook with the following functionality:

- Basic data analysis, visualisation.
- Clustering. Results evaluation and analysis.
- Summary of tickets in compact form (say, you need to include it in the report).

Required language:

- Python

Deadline:

- No deadline

Task task presentation:

- 15 minutes for presentation of conducted work and why candidate has chosen particular approach(es)
- 15 minutes for questions and discussion