# Purpose of the Task

This exercise should be seen as an opportunity to collaborative work in a real use case and help the candidate to experience the work style and the type of challenges she will be engaged with. At first, the candidate will have some time to review the problem and the data on his own, and get prepared for the first meeting. In this meeting, the necessary clarifications will be given, we will plan the next steps and a follow-up meeting. Meanwhile, the candidate will be able to send us email either for questions, or to communicate intermediate results. Finally, the candidate will prepare and deliver a brief report describing the outcomes of his analysis, accompanied with the respective code or the notebook. The candidate is encouraged to at least describe a proposed methodology if the available time does not allow for experimentation.

# Task description

The data in the provided csv file come from the process of users launching remote applications over the period of one month. All users are employees of the same organization. Each row in the provided file corresponds to an individual instance of a remote application being launched by a user. The meaning of the information recorded in each column of the file is given below:

* user\_name: The user id of the person who launched the application.
* machine\_name: The id of the machine where the application is hosted (running).
* logon\_time: The date and time when the application was launched.
* total\_sec: The total duration in seconds of the launch process.
* stage\_1 – stage\_7: These are durations (in seconds) of some of the sub-stages of the launch process.

We are interested in:

1. Detecting launch instances, or groups of launch instances, with an unusually high launch duration time.
2. Detecting any *temporal* changes in the magnitude of launch duration that the average user may have experienced. For such detections it would be nice to have a measure of how big the change is.
3. Identifying, if possible, potential root causes for the detections of Questions 1 and 2 (e.g. identify whether the host machine, or any of the sub-stages of the launch process are the root cause of the detections).