Data platform engineer assessment Yolt 2021

In this assessment we ask you to create a basic data application in Python that will transform a parquet file (this can be anything you choose; from removing a column, to heavy transformations in lambdas or UDF). It needs to write the dataframe to S3. It would be nice if you have a test workflow present.

We ask you to add the infrastructure as code (in Terraform) and provide all of this code in Github. If you have time left it would be nice to have a scheduler present as well. But it is also ok if you share your schedule approach in the presentation.

You will be asked to present your work and we will discuss the approach together in a virtual session.

If you have any questions please do not hesitate to contact us at: marcel.jepma@yolt.eu