



Data Engineer

Moody's

careers.moody's.com/

About the assignment:

Location

Vilnius, Lithuania

Rate (after tax)

€2000 - 3600/Month

Duration

Full time position

Extension (project)

No

Remotely (optionally)

Yes

Expire On

2022-01-31 (3 days ago)

This assignment expired :when

Description:

The data engineer is a key member of the Analytic and Technology Solutions (ATS) group in MIS. ATS is responsible for developing the quantitative models and analytical tools used in the rating process and across the rating agency, as well as MIS technology innovation activities, including advanced capabilities in data processing and machine learning. The data engineer will be part of a team of individuals responsible for applying the latest techniques in Data Processing and Distributed Computing to drive business value. A successful data engineer will not only be technically competent but will be able to work collaboratively with other data engineers and data scientists to build robust data processing pipelines in an automated and testable fashion. The role also includes advocating for operational and process changes to move towards a more data driven organizational paradigm.

The duties of the Data Engineer include:

- Building ingestion pipelines to bring data into the data lake.
- Transforming data into useful representations for use cases.
- Designing and solutioning methods to make data engineering tasks easier for data scientist.
- Applying sound software and architectural

development practices in development and deployment of models as software products.

- Using cloud and distributed computing platforms for model development and deployment.
- Communication of results to business stakeholders and decision makers.

Qualifications

- Practical experience in Data Engineering or statistics and/or distributed computing.
- Highly proficient in Python and/or SQL, additional languages like R, Java, Scala, C#, etc are preferred.
- Expert in Relational and NoSQL data modeling and databases, file formats such as Parquet, Avro, Json, CSV and how to apply different data modalities for the appropriate use cases.
- Experience using Cloud Service, particularly AWS, using managed services such as SageMaker, EMR, S3, ECS, EKS and DynamoDb.
- Experience using the latest DevOps techniques including Git, Github Actions (or other similar CI platform), and Cloudformation (AWS).
- Experience building ML modeling pipelines to ingest, process, train a model, and deploying to a serving platform especially around NLP.
- Experience with the following distributed compute and stream analytics platforms including Kafka, Apache Spark, Koalas, or Dask is preferred.

Required Skills

DATA BASE

MySQL 2-3 years

Data Warehouse 0-1 year

MORE

Cloud 1-2 years