Reasons why we chose Microsoft Azure over AWS for cloud deployment

The main resources used in the scope of the project for cloud architecture are:

Data lake

Server for MySQL database

Data Factory

Integration Runtime

These resources were compared and evaluated on AWS and Azure to choose the most suitable cloud platform for the project

Flexibility of Data Lake:

In AWS, S3 buckets needed to be created to store each file extracted from the API. But Azure Data lake lets us store the file in different folders in a view similar to the file explorer. Also, the option to save and view the file in different formats online is an added functionality in Azure.

Security of Integration Runtime

Integration Runtime is a Microsoft product that creates a secure channel each time the data is copied from a local file storage on to one of the cloud services (in our scope – Data Lake). This channel creates a private network every time data is moved from local storage to the cloud, making data secure during storage and in transit.

Easy to use and monitor interface of Azure Data Factory

The data factory in Azure helps create pipelines using an interface along with the options of a CLI. They also have an automated trigger and logging system that would be easier to run pipelines everyday and debug in case of failed runs

Pay per use of Azure

The Azure resources give the flexibility of pay per use instead of a fixed monthly amount by AWS. This makes it easier to host costly resources like a server on Azure and also makes it cost efficient.