1. Azure Data Factory is a managed data integration service that lives in the Azure cloud. So what does it mean? Azure provides a lot of technologies to help you derive value from your data.

For example, there is Azure Blob Storage, that provides cheaper cloud storage to store your data.

there is Data Lake Analytics, for massaging and transforming your data and

& there is Azure SQL Data Warehouse, that provides scalable relational data warehousing that you can use to understand your business, and there are many other Azure services available for different purposes.

You can mix-and-match these Azure services as needed to analyze both structured and unstructured data. But, hold on! there’s an important aspect of data analytics that none of them address, **which is data integration**. So, data integration means:

extracting data from an OLTP database,

then doing some transformation,

and then loading it into a data warehouse for data analysis.

And while each of these tasks can be done manually, it makes more sense to automate them.

Welcome to Azure Data Factory (ADF), a service that is designed to help you automate these tasks.

So using ADF, you can perform data integration across all your data sources, whether they’re on Azure, on-premises, or on another public cloud, such as AWS.

2)

ADF is an ELT or an Extract, Load & Transform tool. So, how it is different from ETL?

With SSIS, you extract the data, transform the data with built-in SSIS transformations and load the data in the target.

Whereas with Data Factory, you just extract the data from the source and load it in the target, and the transformation is done in a target data store.

3)

Think about Data Factory as a conductor in an orchestra;

In an orchestra, the conductor does not play the music, but he leads a group of musicians, who produce music at various stages of a symphony. The conductor has a big picture of the entire symphony and is aware of the music being performed, but the actual music is performed by individual musicians.

Similarly, the ADF will not perform the actual work required to transform the data but will instruct other services, like Data bricks or data lake analytics to, prepare and transform the data. So, it would be Data lake Analytics / Databricks that perform the actual work and not ADF. ADF merely orchestrates/oversees the execution of work.