European
SharePoint
Office 365 & Azure
Conference





Azure Data Factory Your data pipeline in the cloud Dr. Nico Jacobs



SQL Server and Azure data trainer at U2U @SQLWaldorf





"Azure Data factory helps populating and cleaning the data lake"

- What's a data lake?
- Data Factory terminology
- Demo: ELT in Data Factory
- Dataflows in Data Factory
- Demo: ETL in Dataflows





What is the most important part of the human body?



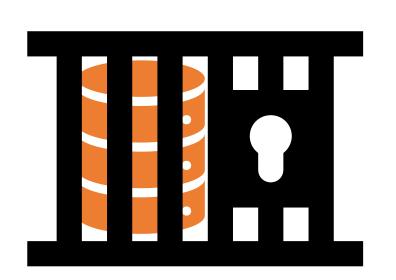


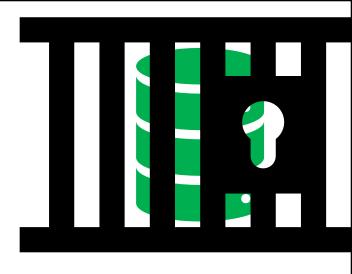




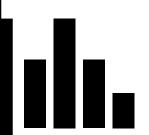
FaxTransmiss



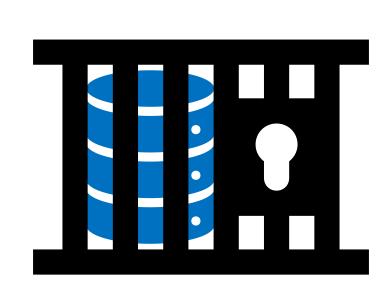






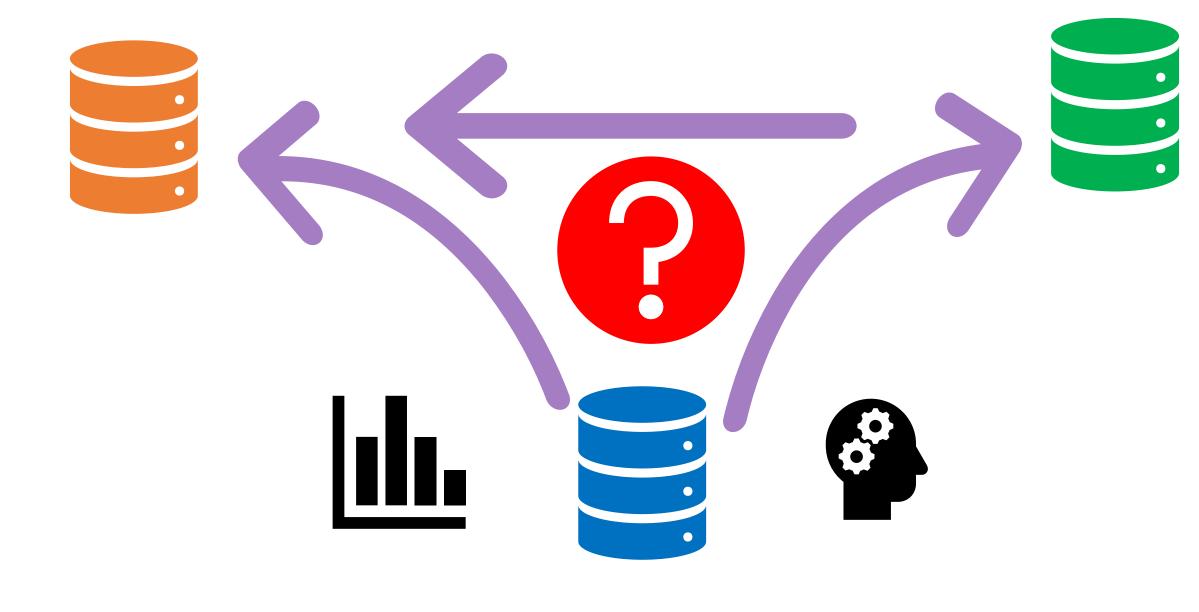




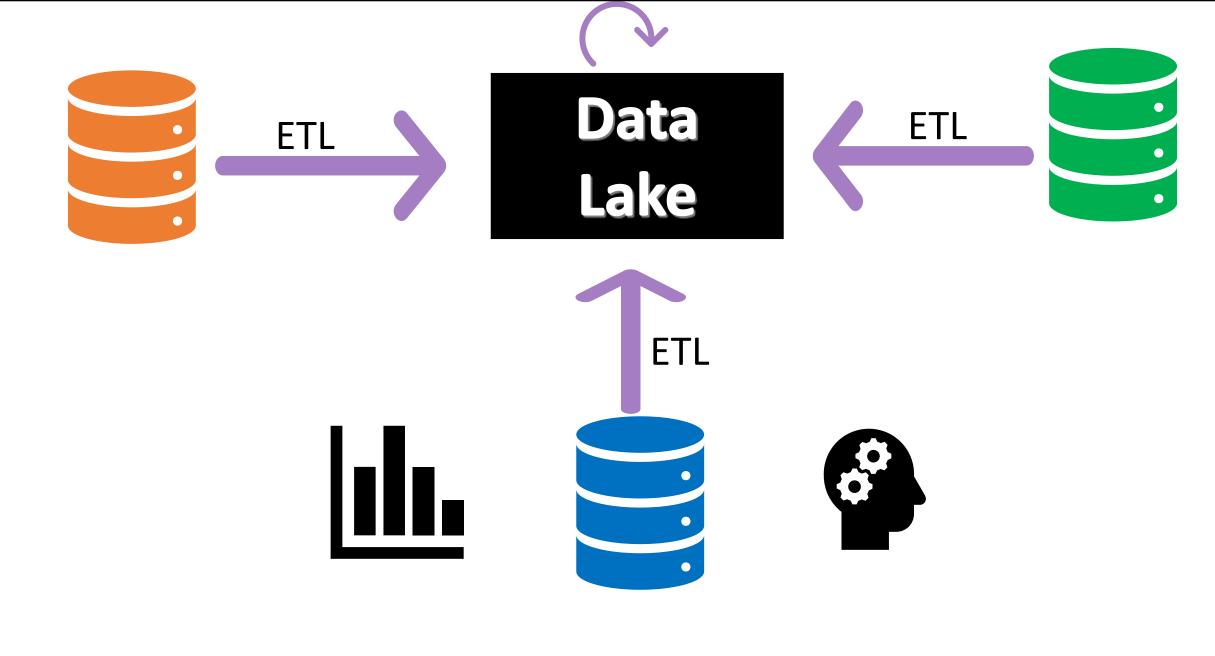












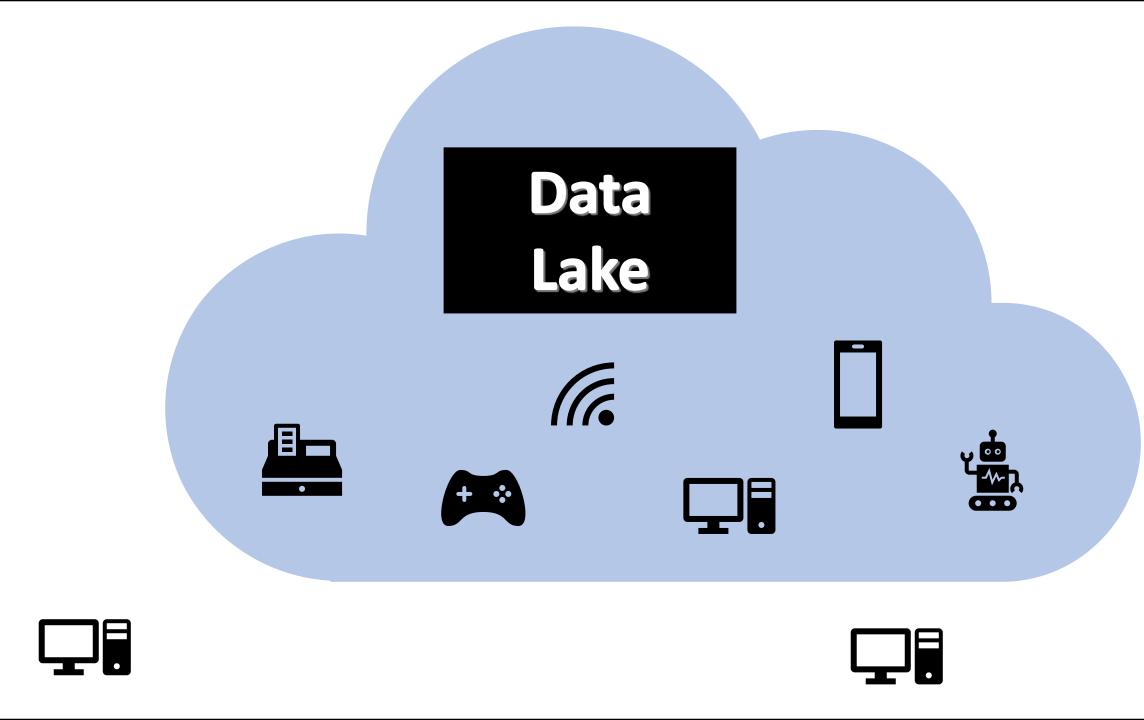
Data Lake

Raw/HR/.../2019/12/05/absenteeism.csv

Staging/...

Clean/...

Data Science/...



Azure Data Factory



What does the ADF do?

Pipeline with actions



Where does ADF run?

Integration runtimes



When does ADF run?

Trigger





Data Factory terminology

Action

Pipeline

Integration runtime

Trigger

Linked Service

Dataset

Data Factory





Demo 1: building an ELT pipeline







From ELT to ETL

Why do I need an external technology to transform my data?

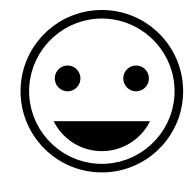






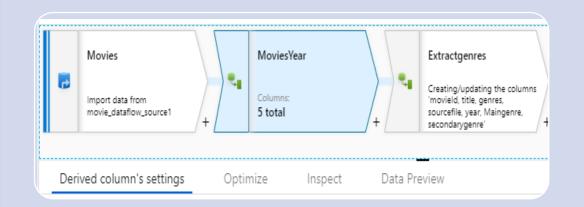
From ELT to ETL

Data flows allow me to transform the data within ADF!





Two flavors of data flows



$f_{\mathcal{X}}$ = Table.Group(#"Removed \checkmark			Name
landmark 🔻	1.2 Stations		UserQuery
Jose	16		Applied steps
wood City	7		Source
untain View	7		Changed colu
Alto	5		Removed col
Francisco	35		➤ Grouped row

Mapping data flow

- Visual flow
- Schema drift support

Wrangling data flow

- Power Query M script
- Live preview





Demo 2: building an ETL pipeline







Conclusions

- Azure Data factory helps populating and cleaning the data lake
- Both Business Intelligence and Artificial Intelligence require data from different parts
- Azure Data Factory can copy data between different cloud and onprem locations
- Transformations can either be defined in separate services or in Data Flows





Attribution

By I, Sailko, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=25095506







PLEASE RATE THIS SESSION ON THE APP













Azure Data Factory Your data pipeline in the cloud Dr. Nico Jacobs



SQL Server and Azure data trainer at U2U @SQLWaldorf





