MODERN DATA WAREHOUSE

Microsoft Azure

MISTER SPEX

MISTER SPEX

SO MANY DATA SOURCES

GERMAN SCREAM

MISTER SPEX

EVANDRO MUCHINSKI

- **Data Engineer at Mister Spex**
- **7 years as Microsoft Certified Trainer**
- **MCSE, MCSA…**
- **Brazilian living in Berlin for almost 3 years**
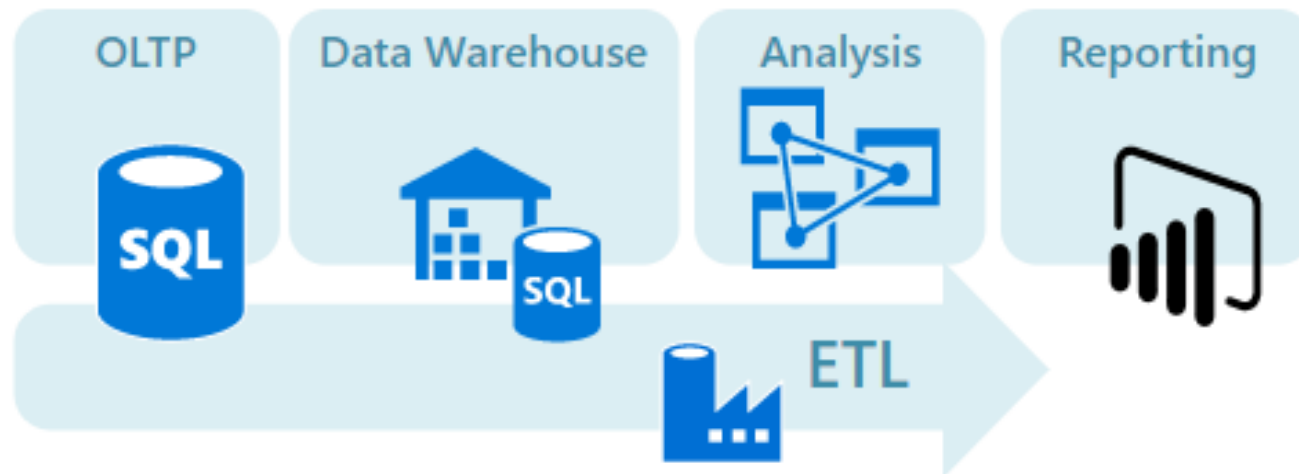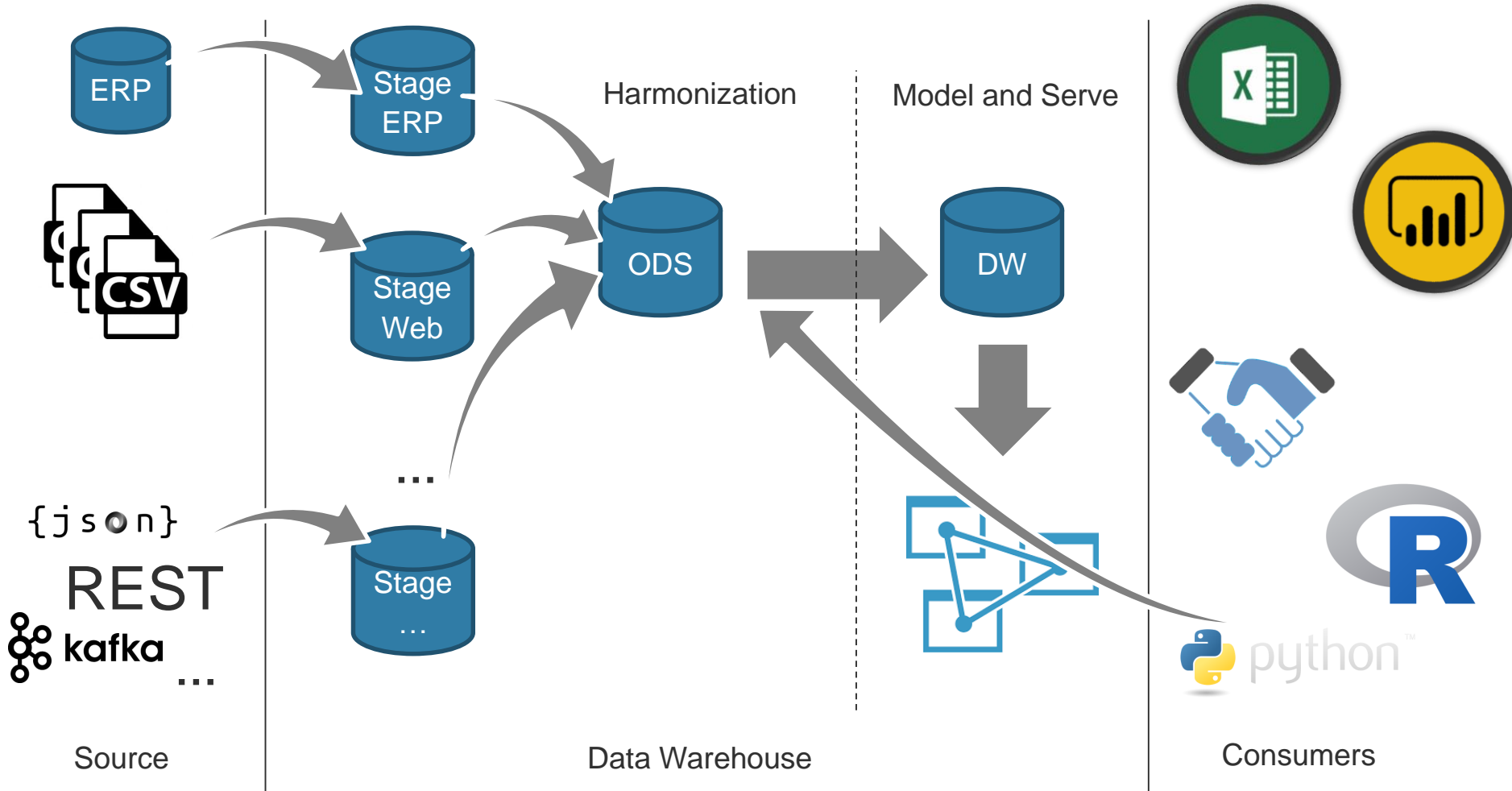- **Working with data driven projects since 2006**

LinkedIn https://www.linkedin.com/in/evandro-muchinski/

AGENDA

1. **PROBLEM**

2. **TRADITIONAL DATA WAREHOUSE ARCHITECTURE**

3. **MODERN DATA WAREHOUSE**

4. **TOOLSET OVERVIEW**

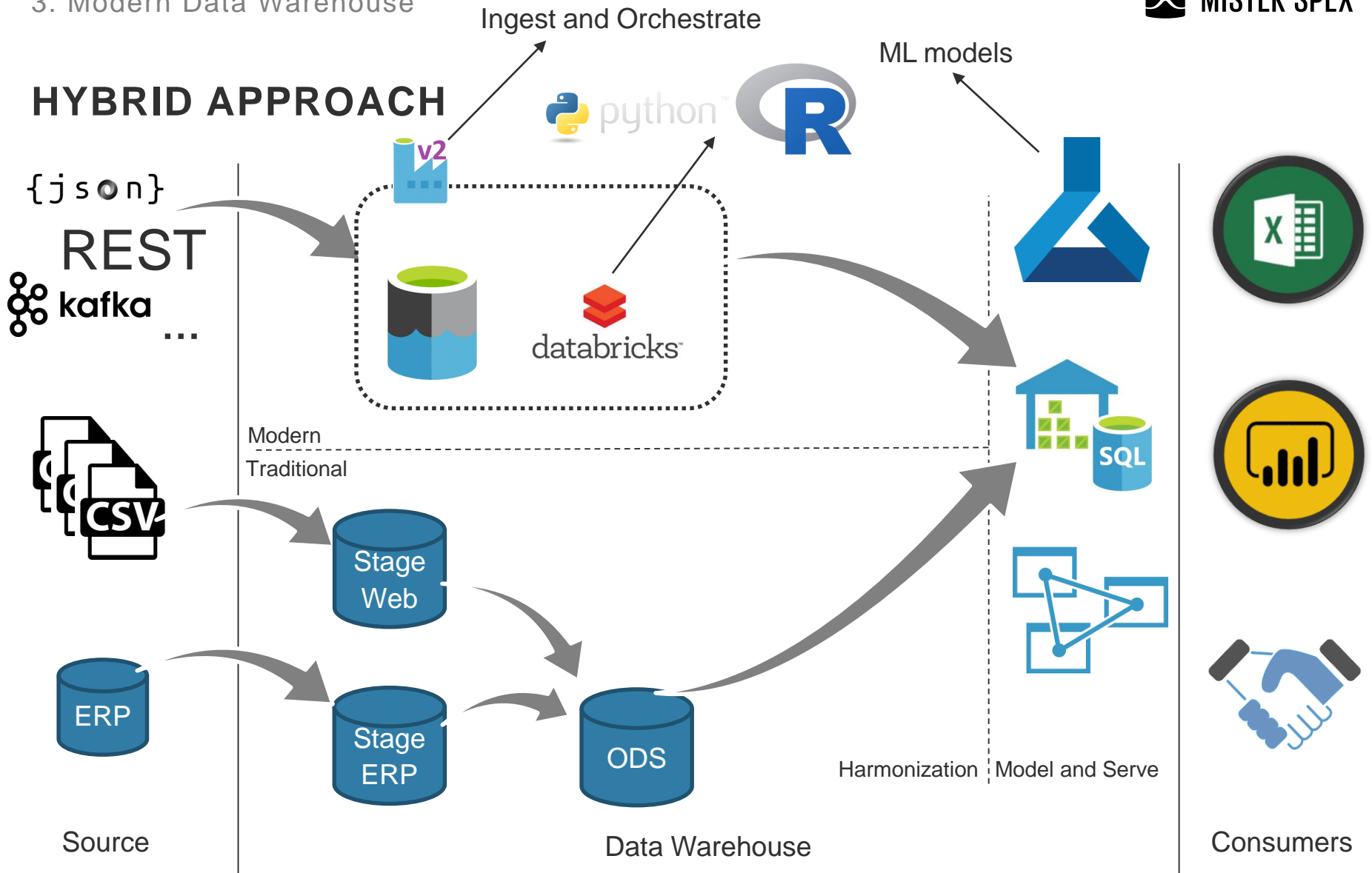5. **REFERENCES**

GENERAL ARCHITECTURE

MISTER SPEX

SOMETHING CLOSE TO WHAT WE HAVE

ERP

CSV

{json}

REST

kafka
...

Stage
ERP

Stage
Web

...

Stage
...

Harmonization

ODS

Model and Serve

DW

Source

Data Warehouse

Consumers

CONFIDENTIAL

GENERAL ARCHITECTURE

MISTER SPEX

Ingest and Orchestrate

ML models

HYBRID APPROACH

{json}

REST

kafka

...

Modern

Traditional

Stage
Web

CSV

ERP

Stage
ERP

ODS

Harmonization | Model and Serve
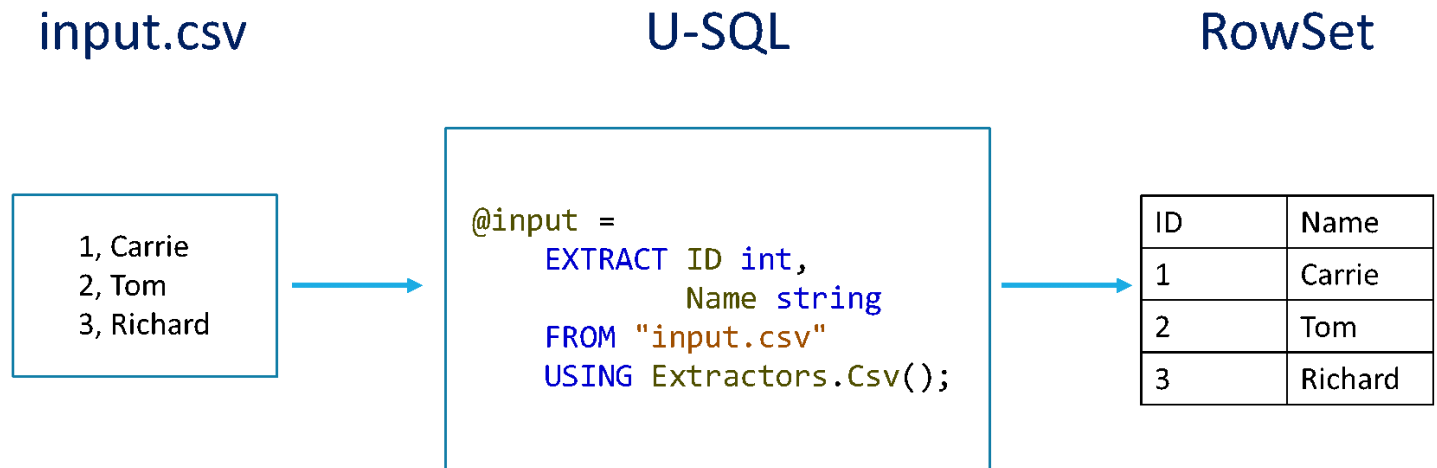
Source

Data Warehouse

Consumers

MISTER SPEX
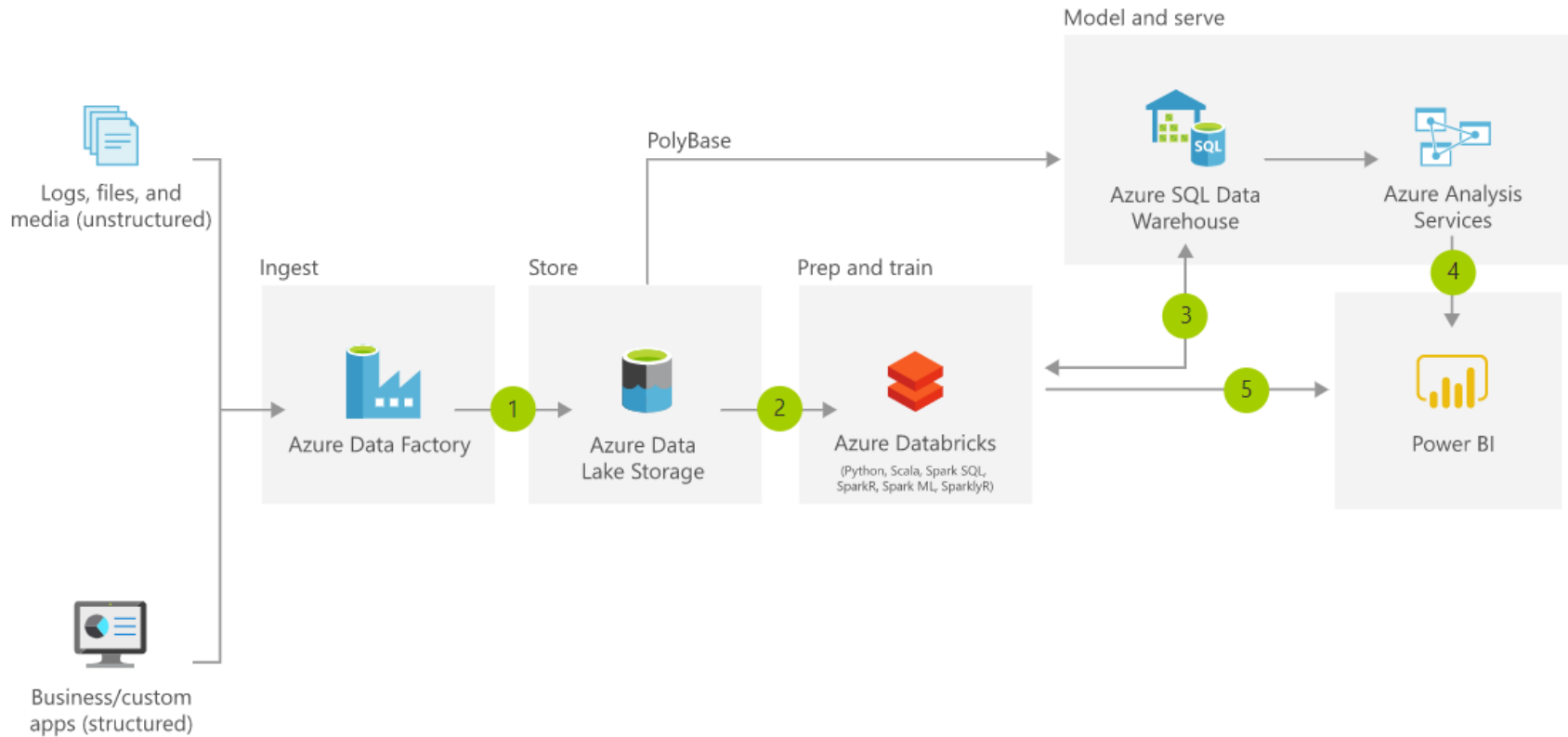
NEW TOOLS AVAILABLE ON AZURE

- **Data Lake**
- **Data Factory V2**
- **Data Bricks**
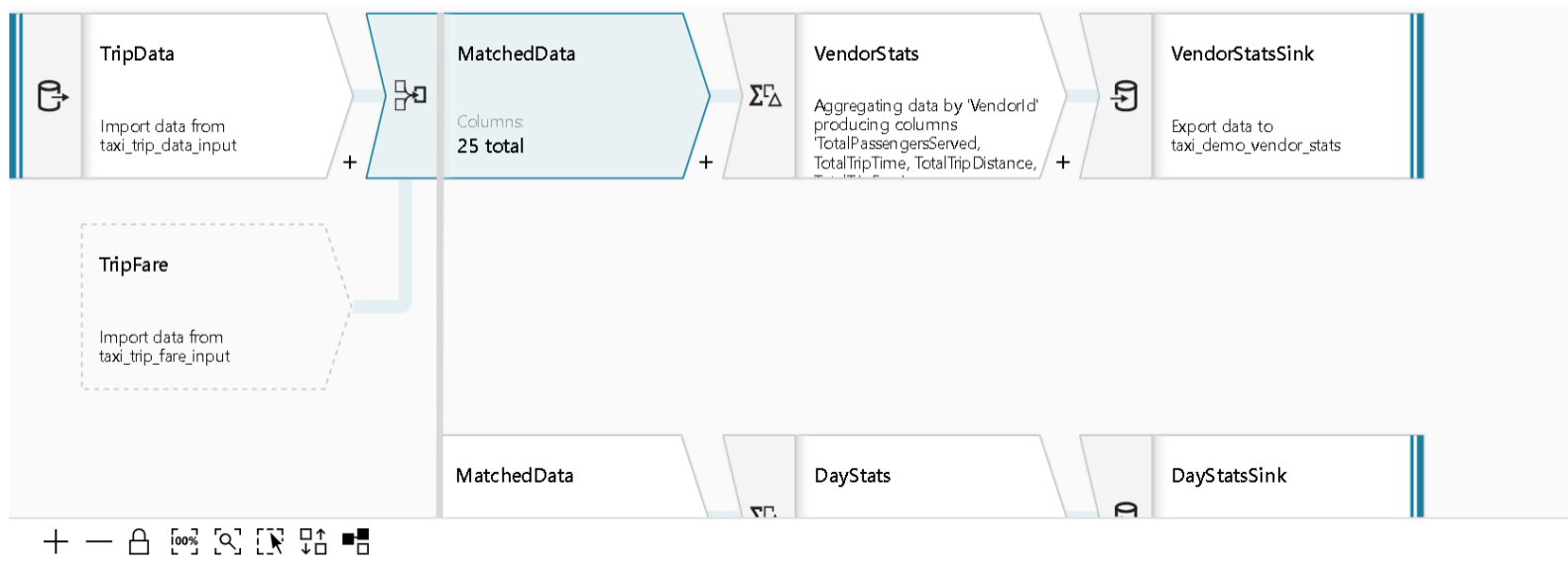- **Data Warehouse**
- **Machine Learning Services**

MISTER SPEX

DATA LAKE

- **High scalable storage**
- **Hadoop compatible access**
- **Files managed in directories and subdirectories**
- **Permission can de granted to directories or files**
- **U-SQL compatible**

input.csv   U-SQL   RowSet

1, Carrie
2, Tom
3, Richard

```
@input =
    EXTRACT ID int,
            Name string
    FROM "input.csv"
    USING Extractors.Csv();
```

ID	Name
1	Carrie
2	Tom
3	Richard

MISTER SPEX

DATA FACTORY V2

- **Ingest and Orchestrate tool**
- **80+ connectors**
- **Integrated with Git**
- **Develop with Code or Interface**
- **Compatible with Azure Data Bricks**
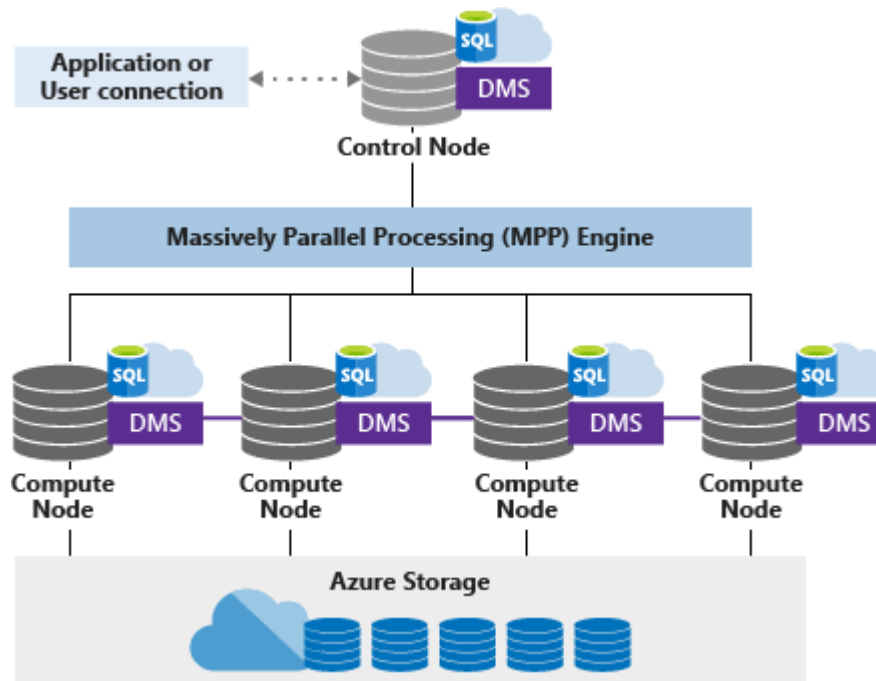- **Dataflows = Spark Executables**

CONFIDENTIAL 12

DATA BRICKS

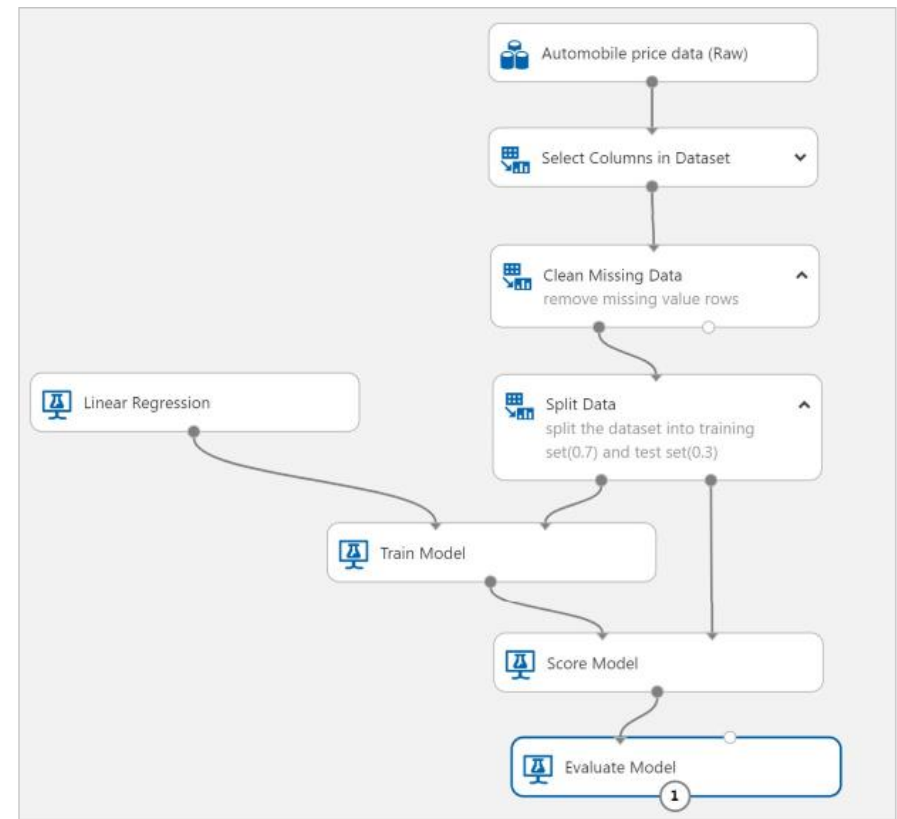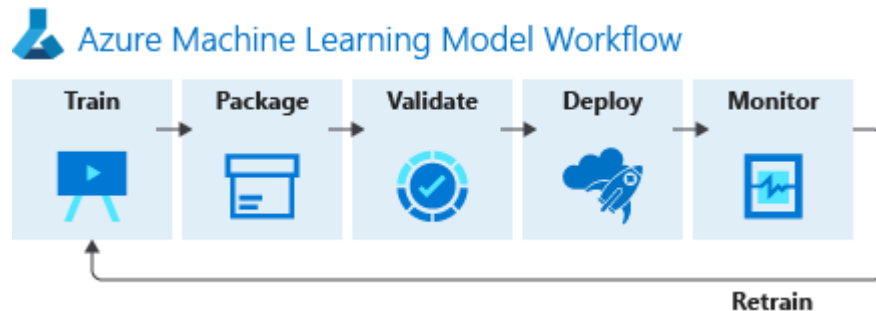- **Runs Python, R, Scala, SQL**

MISTER SPEX
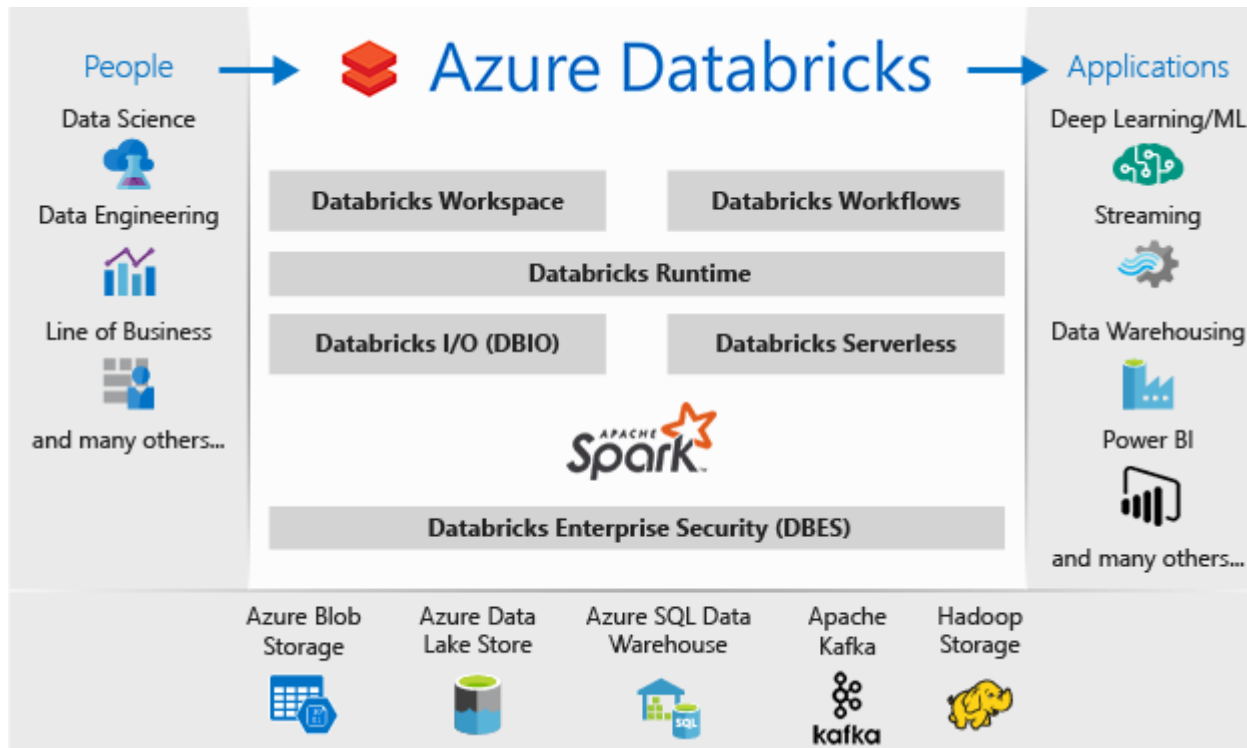
DATA WAREHOUSE

- **Massive Parallel Processing column-store database cluster**

MISTER SPEX

MACHINE LEARNING SERVICES

- **Automate and scale Data Science projects**
- **Work with Python only**
- **More on the next session** ☺

CONFIDENTIAL   15

MISTER SPEX

LEARN MORE AT:

Azure Data Architecture Guide

Microsoft Build 2019

Learning path for Azure Data Engineer

CONFIDENTIAL

THANK YOU!