Spanish Point technologies ltd

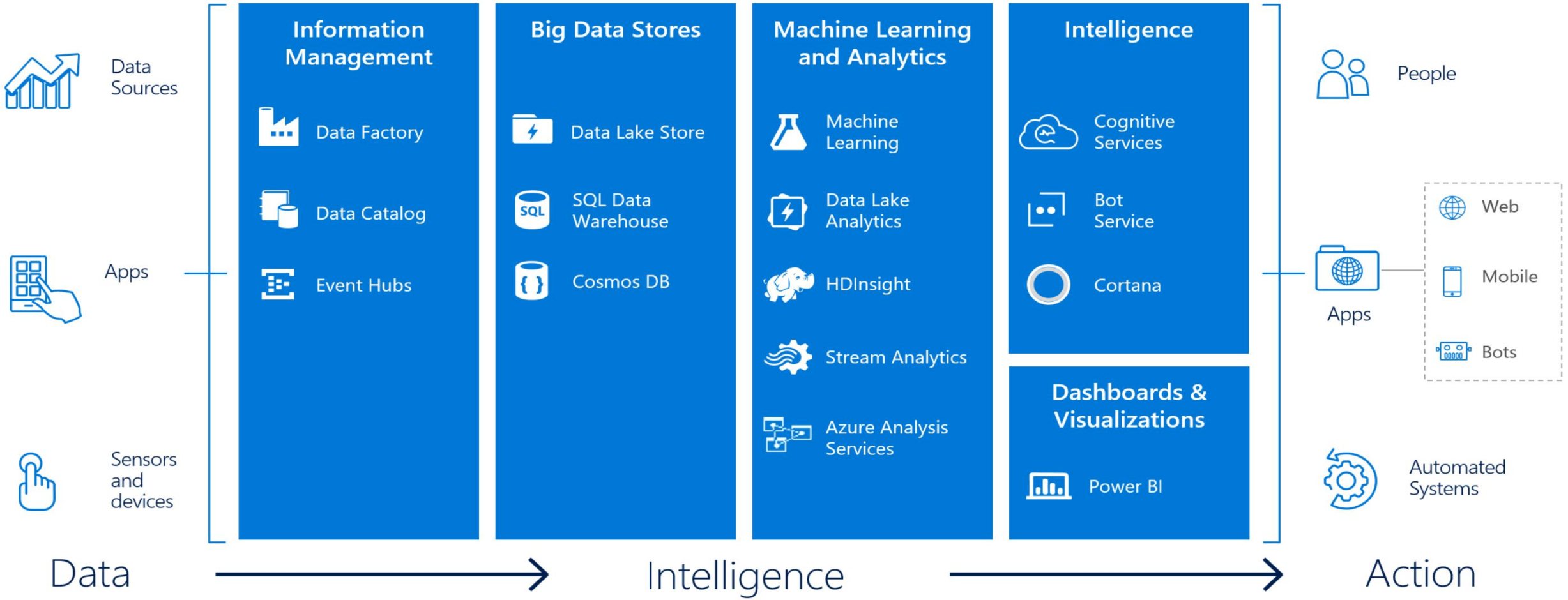
Azure Data Analytics & Machine Learning Seminar

Daire Cunningham:
BI Practice Area Manager

AGENDA

09:00 AM – 09:30 AM	Registration & Refreshments
09.30AM – 10:00 AM	Welcome & Keynote, Ger Perdisatt, Director – Enterprise Commercial Sales & Donal Cullen, CEO Spanish Point Technologies
10:00 AM – 10:30 AM	The Azure Data and Analytics Platform – Data Storage, Ingestion and Transformation
10:30 AM – 10:50 AM	Coffee Break
10:50 AM – 11:20 AM	Machine Learning demo
11:20 AM – 11:50 PM	Customer Story – National Treatment Purchase Fund
12:00 PM – 12.45 PM	Microsoft Tour
12:45 PM – 1:30 PM	Lunch & Networking

AZURE DATA ANALYTICS SUITE



MOVING DATA

AZURE DATA FACTORY - OVERVIEW

Creates, orchestrates, & automates the movement, transformation and/or analysis of data through the cloud

- **Data** can mean structured or unstructured, small or big, online or on-premise
- **Transform & Analyse** can mean Hadoop, Data Lake Analytics, Databricks, Machine Learning, custom code etc. etc.
- ADF ***orchestrates*** these processes

AZURE DATA FACTORY

- Create workflows called pipelines for data loading and transformation
- Managed and scalable cloud-based orchestrations
- Usual pattern is “EL & TL” i.e. Extract, Load & Transform, Load
- Can leverage PolyBase for large data to/from supported data sources
- Traditional SSIS packages can also be deployed to ADF

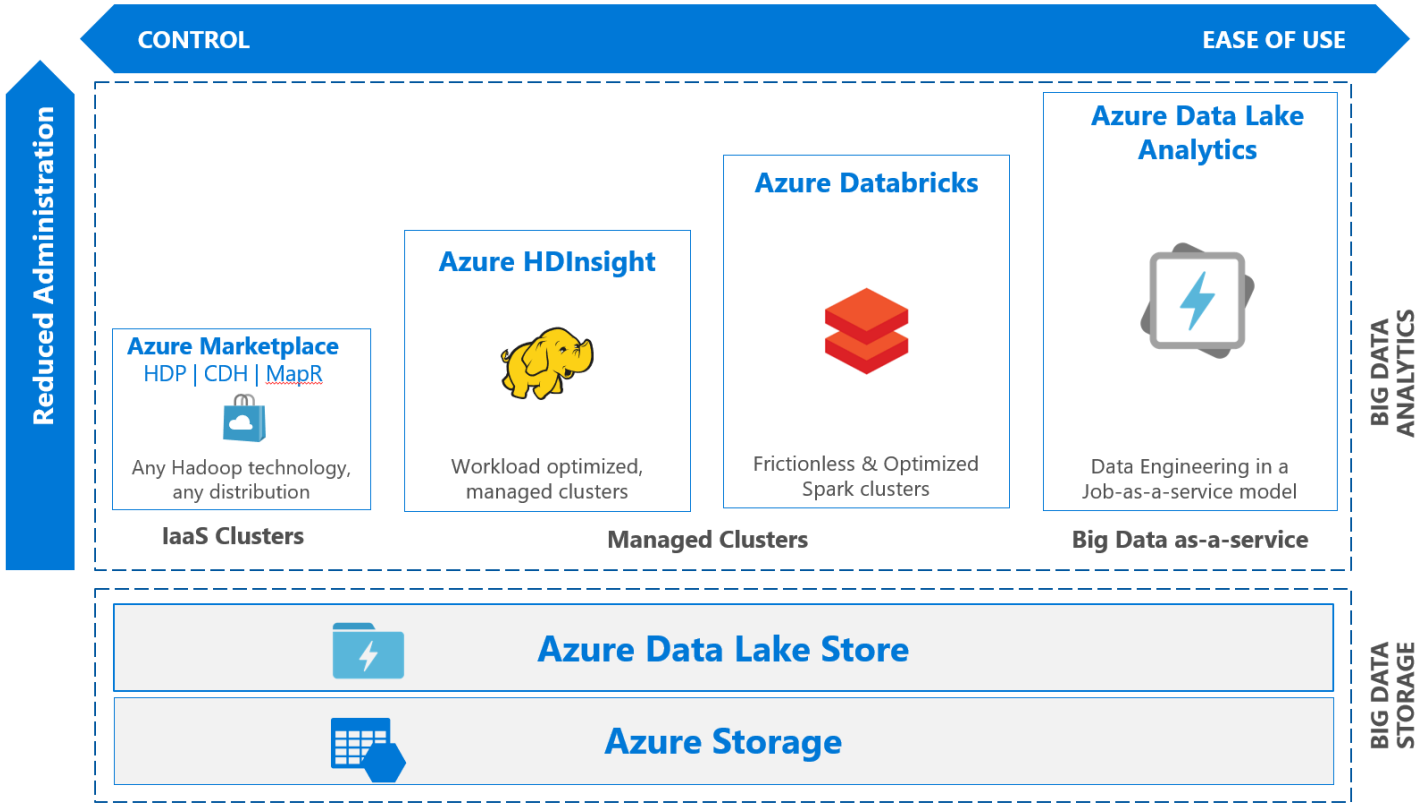
TRANSFORM & ANALYSE DATA

TRANSFORM & ANALYSE DATA

- There are several tools available to transform and processes your raw data into predictions and insights

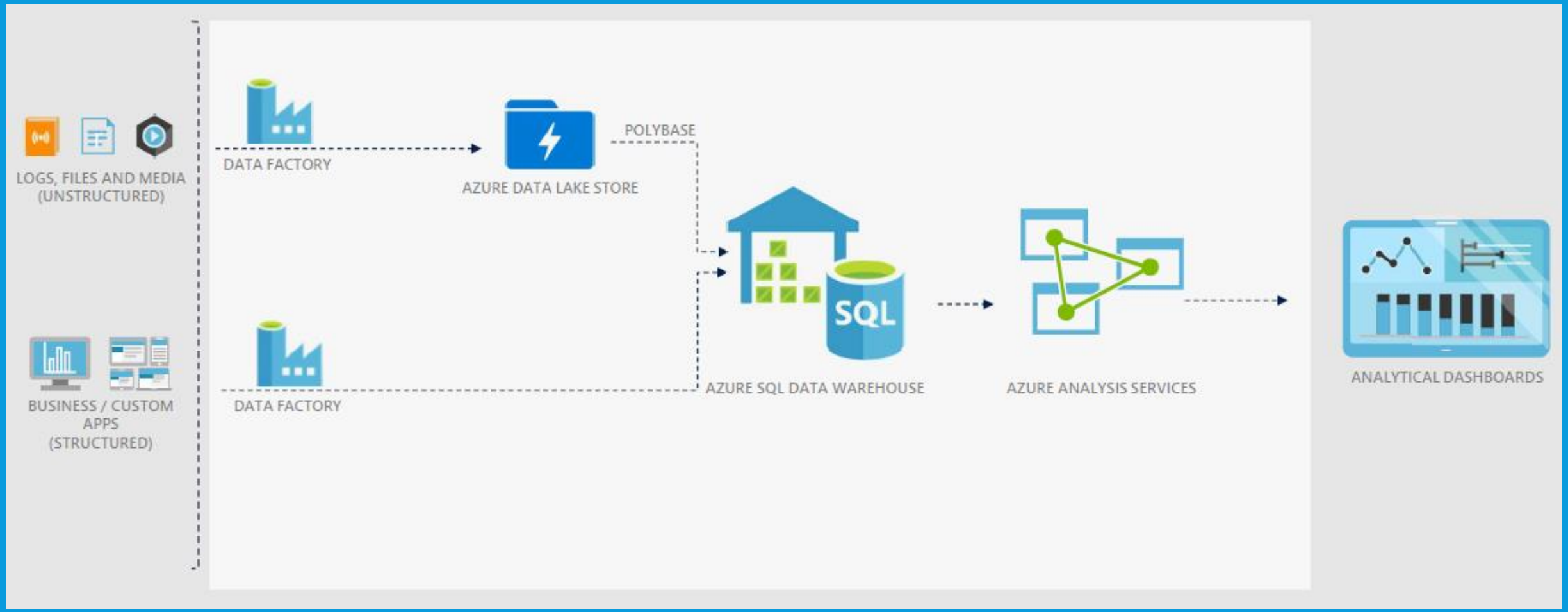
Compute environment	Activities
On-demand HDInsight cluster or your own HDInsight cluster	Hive, Pig, Spark, MapReduce, Hadoop Streaming
Azure Databricks	Notebook, Jar, Python
Azure Data Lake Analytics	Data Lake Analytics U-SQL
Azure Machine Learning	Machine Learning activities: Batch Execution and Update Resource
Azure SQL, Azure SQL Data Warehouse, SQL Server	Stored Procedure
Azure Batch	Custom

CHOOSE AN ANALYTICS SOLUTIONS

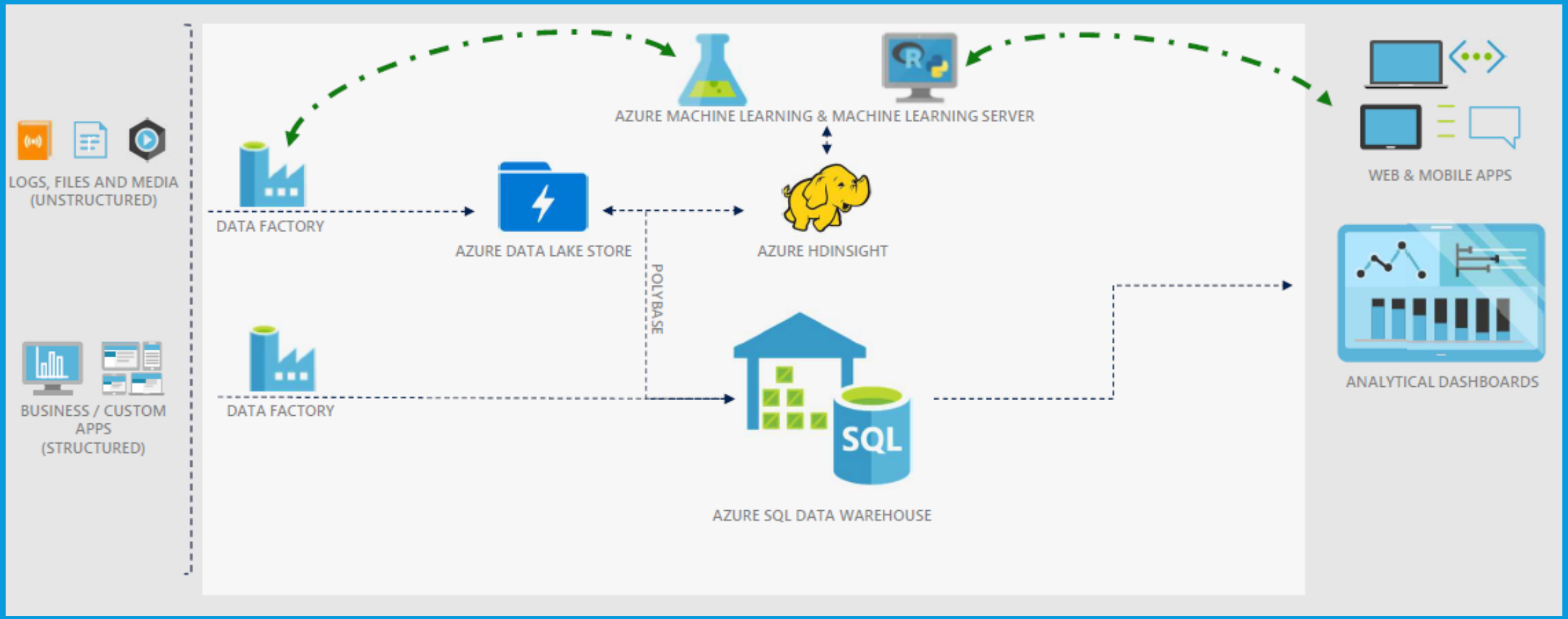


EXAMPLE ARCHITECTURES

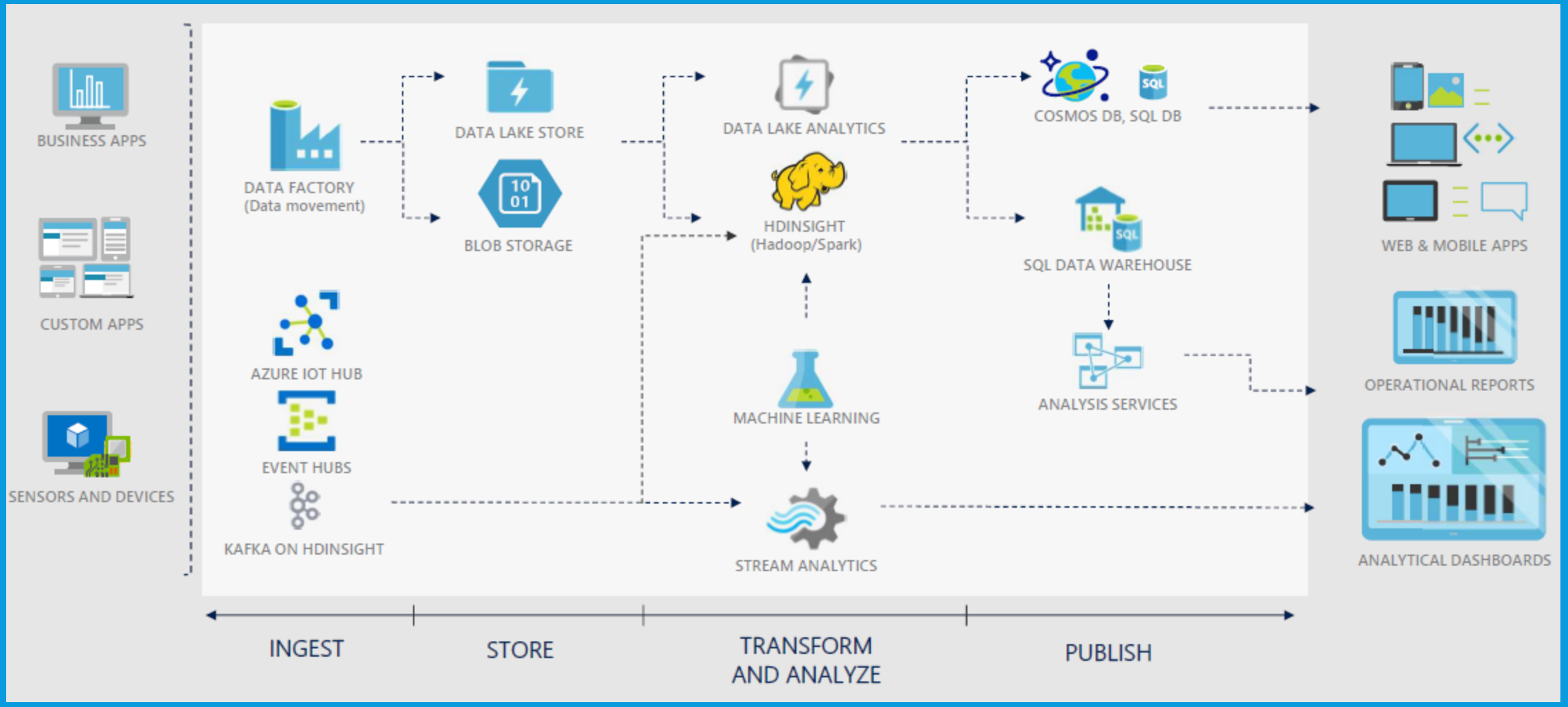
1 - CLOUD DATA WAREHOUSE



2 - ADVANCED ANALYTICS

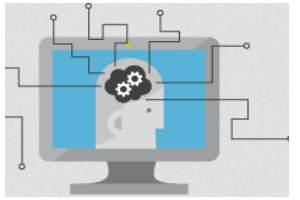


3 - DO IT ALL!


BUSINESS APPS
CUSTOM APPS
SENSORS AND DEVICES
DATA FACTORY
(Data movement)
DATA LAKE STORE
BLOB STORAGE
AZURE IOT HUB
EVENT HUBS
KAFKA ON HDINSIGHT
DATA LAKE ANALYTICS
HDINSIGHT
(Hadoop/Spark)
MACHINE LEARNING
STREAM ANALYTICS
COSMOS DB, SQL DB
SQL DATA WAREHOUSE
ANALYSIS SERVICES
WEB & MOBILE APPS
OPERATIONAL REPORTS
ANALYTICAL DASHBOARDS
INGEST
STORE
TRANSFORM
AND ANALYZE
PUBLISH

MACHINE LEARNING

AZURE MACHINE LEARNING

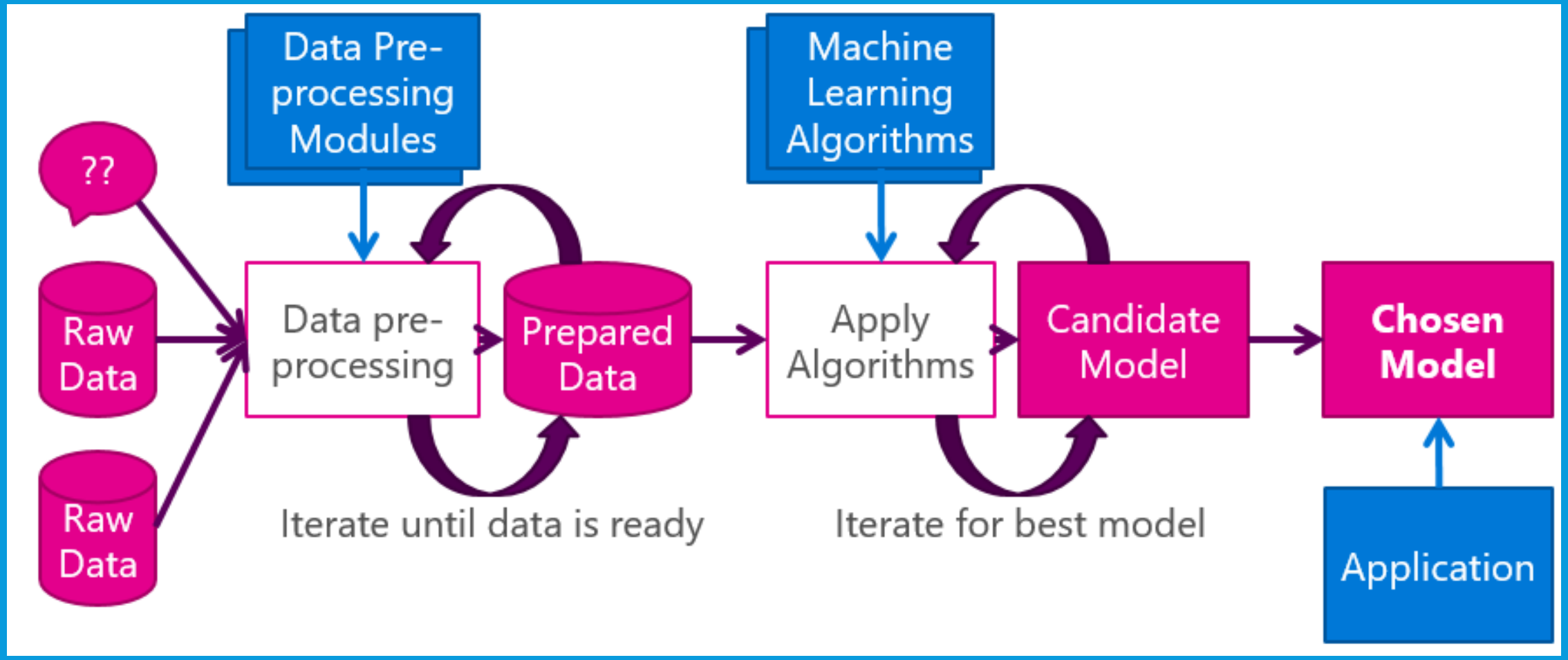


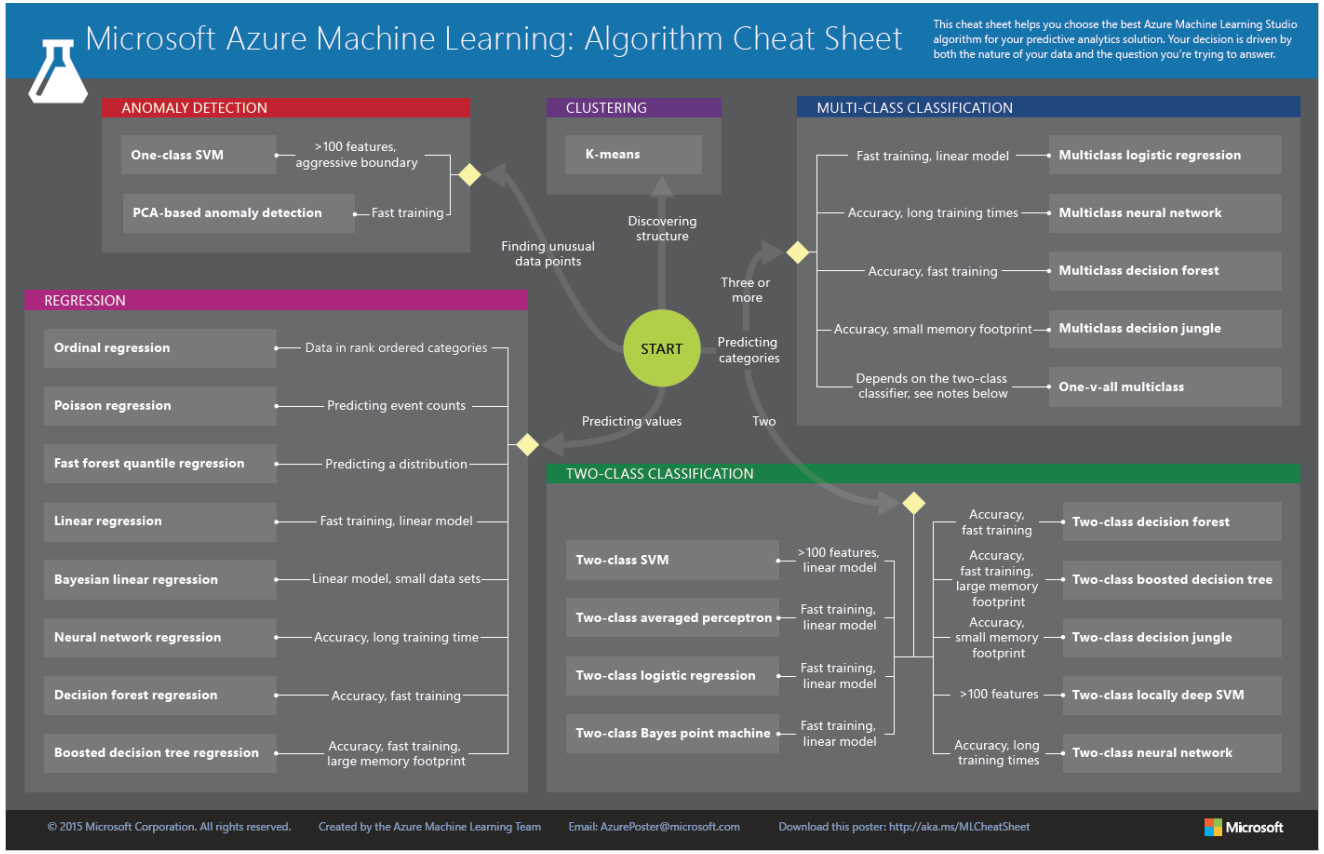
- What it is:
  - A multi-platform environment and engine to create and deploy Machine Learning models and API's
- When to use it
  - When you need to create predictive analytics
  - When you need to share Data Science experiments across teams
  - When you need to create call-able API's for ML functions
  - When you also have R and Python experience on your Data Science team

AZURE MACHINE LEARNING STUDIO

- Web-based drag-and-drop interface for creating ML experiments
- Easiest way to get started with Azure ML
- Get a free account to try it out: **https://studio.azureml.net** or
- add an Azure ML Workspace to your existing Azure account

AZURE ML PROCESS





AZURE ML – CHOOSING EXPERIMENT

COFFEE BREAK

FOLLOWED BY AZURE MACHINE LEARNING DEMO

MACHINE LEARNING DEMO

FLIGHT DELAY PREDICTION MODEL

SUMMARY

- Azure includes many services for working with data allowing you to choose the right tool for the job
- Data loaded into Azure where it is Transformed, Analysed and Prepared for presentation
- ADF coordinates these tasks
- Azure Machine Learning allows you to easily execute predictive analysis on your data
- Azure services are managed, scalable and exist in the one ecosystem

Azure Data Analytics & Machine Learning Bootcamp

- **Overview of Data and Analytics Platform in which you will see:**
  - High-level architecture of Azure data platform
  - Overview of Technologies and Services
- **Data Storage Options in which will learn about:**
  - Structured Data Storage in SQL Azure, SQL Data Warehouse
  - File Storage in Blob Storage, Azure Data Lake
  - NoSQL databases such as Cosmo DB and HDInsight
- **Ingesting Data in which you will learn about**
  - Data Factory
  - PolyBase
  - Related technologies
- **Data Analytics in which you will learn about**
  - Data Lake Analytics
  - Machine Learning
  - HDInsight, Redbricks etc.
- **Streaming Data in which you will lean about**
  - Event Hubs and Stream Analytics
  - IoT including IoT Hub and IoT Edge
- **Visualising and consuming your data in which you will learn about**
  - Power BI dashboards and live streaming tiles
  - Intelligence services such as Cognitive Services, Bot Framework etc.

QUESTIONS