Building Your First Data Science Project in Microsoft Azure

DETERMINING WHICH TOOLS TO USE

Jared Rhodes
INDEPENDENT CONSULTANT

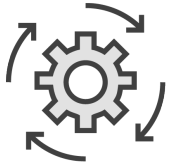@qimata www.jaredrhodes.com

Overview

Introduction to Data Science

Introduction to Data Science

Commodity

Process

Career

Evolving

It is data that will surely fuel Industry 4.0—and will be the banks' most valuable asset for the remainder of this century

Eli Rosner

Uses of Data

Medical

Targeting

Cognition

Personalization

Prediction

Marketplace

Process

Capture

Process

Analyze

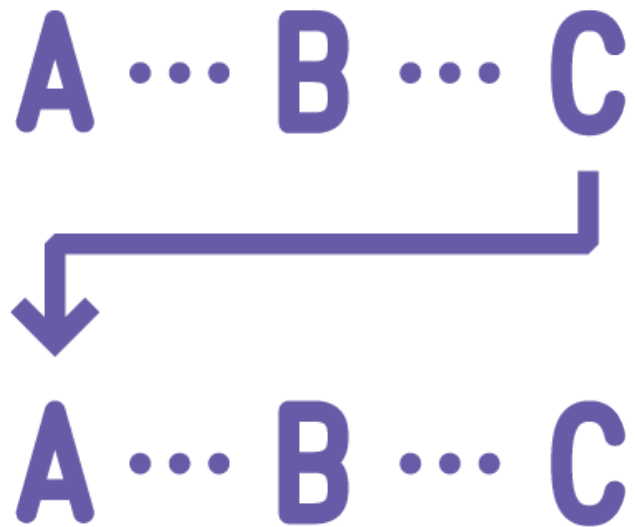Maintain

Communicate

Career

Entering the field

Finding your role

Advancing

Evolving

A ··· B ··· C

A ··· B ··· C

Hype cycle

Industry influx

Requirements convolution

Cloud

Identifying Constraints
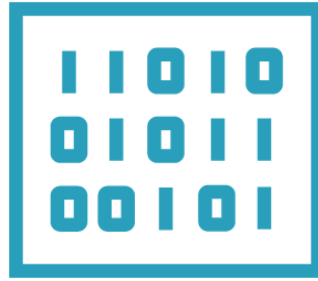
Stages of a Data Science Project
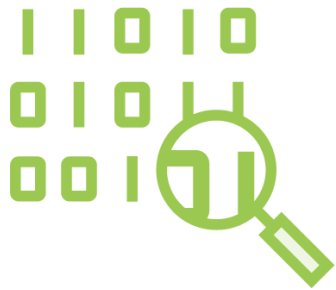
Scoping

Development

Deployment

Scoping

Why?

Determining data need

Locating data

Exploring data

Assessing data

Finalizing

Development

- **Create a data policy**
- **Create a testable hypothesis**
- **Create a test**
- **Re-evaluate hypothesis**

Deploying

Determining deployment environment

Creating the model
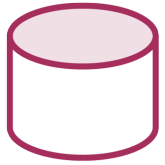
Testing the model

Change control
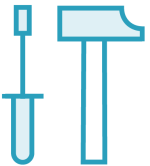
Exposing the model

Repeat

Tools Discussion

Tools

Storage

Reporting
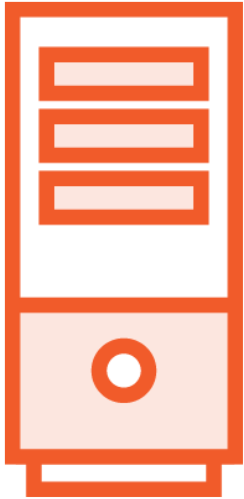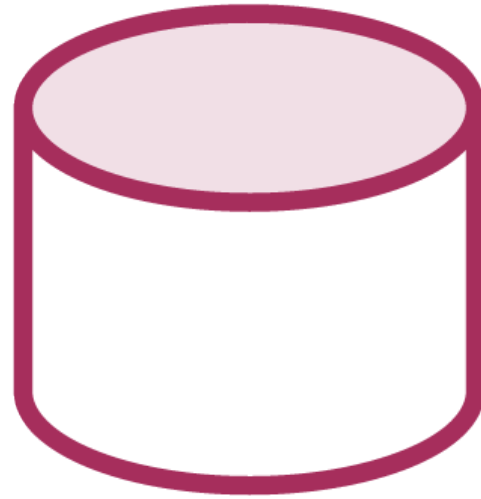
Languages

Data Processors

Storage

Self-Hosted   **Hosted**   **Cloud**

Storage

Hierarchy

Characteristics

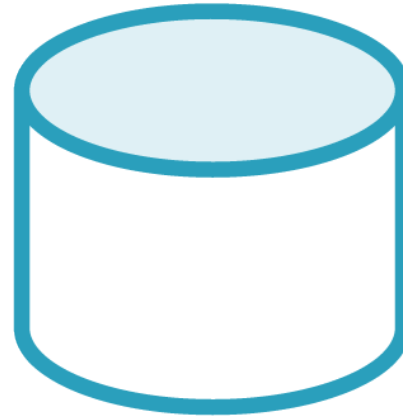**Media**

Storage in Azure

AI & ML   **Analytics**   **Databases**   **Storage**

SAP Business
Intelligence

MicroStrategy
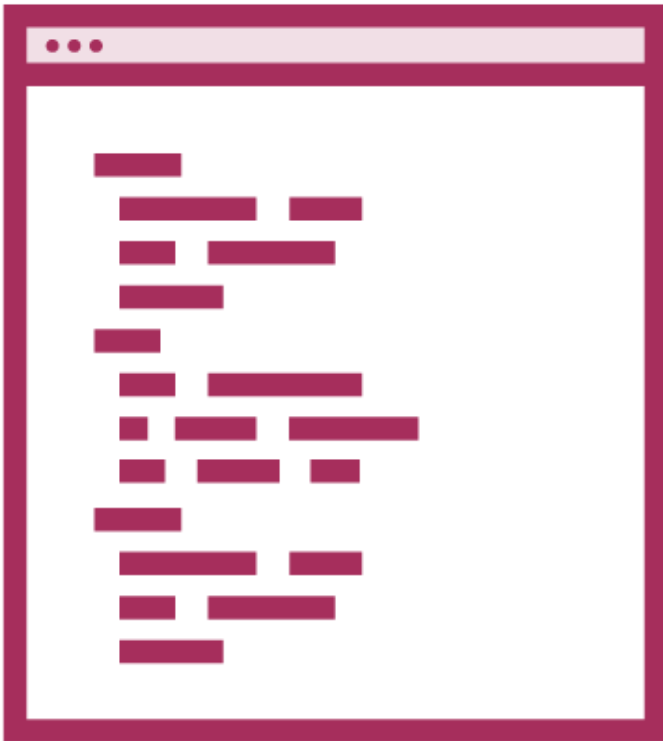
Sisense

SAS Business
Intelligence

Yellowfin BI

Dundas BI

TIBCO Spotfire

Microsoft Power BI

Languages

SQL

Python

R

Scala

SAS

MatLab

Hive

C/C++

Languages

PY | Python

R | R

SQL | SQL

Apache Hadoop

Apache Spark

Apache Storm

Apache Cassandra

SQL Data Warehouse

Azure Data Factory

Azure Stream
Analytics

Review

Introduction to data science

Stages of a data science project

Tools