

## **WHAT YOU WILL BE ABLE TO DO AFTER THIS 2 DAY TRAINING**

Be able to piece apart the components of a deployed solution for predictive analytics in Azure for streaming data, including gaining familiarity with the data science aspects of the solution.

Be able to build back up and deploy a simple predictive analytics solution in Azure for streaming data including a shareable BI dashboard for rapid synthesis of information.



<http://aka.ms/iotedu>



The BI Dashboard from a Power BI-integrated solution template will be shown now. We will be using the Energy Demand Forecasting solution template (not public yet) for our example solution throughout the course.

## DAY 2 SESSIONS

- 1) Logistics – 30 min
    - Azure Passes walk-through
  - 2) Predictive Analytics with Azure Machine Learning with Labs – 2 hours (Including a ~20 min break)
  - 3) Lunch at 12:30 – 1 hr
  - 4) Data at Rest and Orchestrating Data Movement with Demo – 60 min
  - 5) Break – ~20 min
  - 6) Pattern Authoring with Darwin Schweitzer – 90 min
- (note: these are approximate)



<http://aka.ms/iotedu>



This is the Day 1 Agenda

### Course Description

This 4 day workshop entails 2 days of training and a 2-day build-a-thon that solidifies the training topics. The training will take an attendee through an end-to-end solution from the demand forecasting domain. Demand forecasting revolves around predictive analytics, so the emphasis will be on creating and deploying machine learning models, the ingestion of fast-moving-data, dashboarding and orchestrating data flow and storage. The makings of the predictive cloud solution will be drawn from the Cortana Intelligence Suite on Azure, Microsoft's cloud.

The Build-a-thon will give attendees a chance to bring their own ideas and data to expand and broaden the example of the cloud deployed predictive solution from within a proctored environment. In the end, attendees will be able to demonstrate and create their own predictive analytics workflows.

## **CORTANA INTELLIGENCE IN A SENTENCE:**

Cortana Intelligence is a **Platform** and a **Process** to perform advanced analytics from start to finish



<http://aka.ms/iotedu>



### 1. What you can do with CIS:

<https://www.microsoft.com/en-us/server-cloud/cortana-intelligence-suite/why-cortana-intelligence.aspx>

### 2. More about the process:

<https://channel9.msdn.com/Blogs/Seth-Juarez/Understanding-Data-Science-for-building-Predictive-Analytics-Solutions-by-Francesca-Lazzeri>

# PREDICTIVE ANALYTICS FOR IoT

Event Hubs + Streaming Analytics + Storage + SQL + Machine Learning + Data Factory = AWESOME

## PART 1 OF 3: PREDICTIVE ANALYTICS WITH MACHINE LEARNING



<http://aka.ms/iotedu>



# MACHINE LEARNING 101



1. Data Science for Beginners: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-for-beginners-the-5-questions-data-science-answers/>

## MACHINE LEARNING 101



**WIKIPEDIA**  
The Free Encyclopedia

...machine learning explores the study and construction of algorithms that can learn from and make predictions on data – such algorithms overcome following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs.



<http://aka.ms/iotedu>



More from the Wikipedia article on Machine learning: Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms is unfeasible; example applications include spam filtering, optical character recognition (OCR),[5] search engines and computer vision.

# MACHINE LEARNING 101

The first way of thinking about ML, is by the type of information or input given to a system. So, given that criteria there are three classical categories:

1. Supervised learning - we get the data and the labels e.g. linear regression
2. Unsupervised learning - only get the data (no labels) e.g. clustering
3. Reinforcement learning - reward/penalty based information (feedback)

Another way of categorizing ML approaches, is to think of the desired output:

1. Classification (e.g. decision tree)
2. Regression (e.g. linear regression)
3. Clustering (e.g. k-means)
4. Density estimation (e.g. histograms)
5. Dimensionality reduction (e.g. principal component analysis)



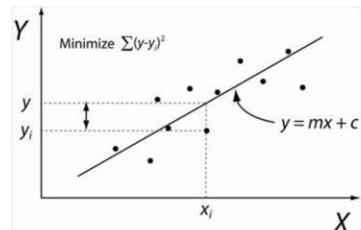
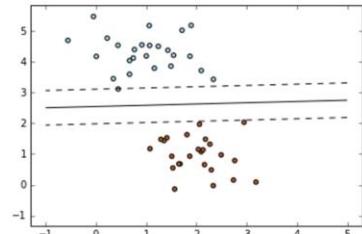
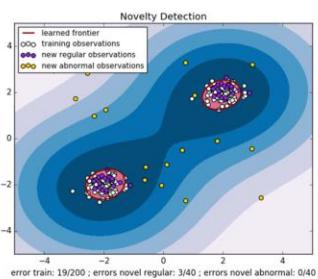
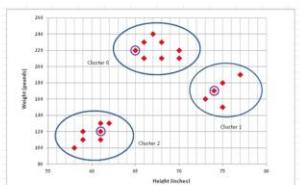
<http://aka.ms/iotedu>



More examples of supervised learning (in python's ML module docs): [http://scikit-learn.org/stable/supervised\\_learning.html](http://scikit-learn.org/stable/supervised_learning.html)

More examples of unsupervised learning (in python's ML module docs): [http://scikit-learn.org/stable/unsupervised\\_learning.html](http://scikit-learn.org/stable/unsupervised_learning.html)

# MACHINE LEARNING 101



<http://aka.ms/iotedu>

Microsoft

Starting upper left, counter-clockwise:

1. Clustering – k means
2. Novelty/Anomaly detection – one-class SVM
3. Classification - SGD
4. Regression - linear

# MACHINE LEARNING 101

Some terms you will encounter as a Machine Learner

Term	Definition
Training set	set of data used to learn a model
Test set	set of data used to test a model
Feature	a variable (continuous, discrete, categorical, etc.) aka column
Target	Label (associated with dependent variable, what we predict)
Learner	Model or algorithm
Fit, Train	learn a model with an ML algorithm using a training set
Predict	w/ supervised learning, give a label to an unknown datum(data), w/ unsupervised decide if new data is weird, in which group, or what to do next with the new data
Accuracy	percentage of correct predictions $((TP + TN) / \text{total})$
Precision	percentage of correct positive predictions $(TP / (FP + TP))$
Recall	percentage of positive cases caught $(TP / (FN + TP))$



<http://aka.ms/iotedu>



## MACHINE LEARNING CAPABILITIES

Which category



How much/many



Which group



Is it odd



Which action

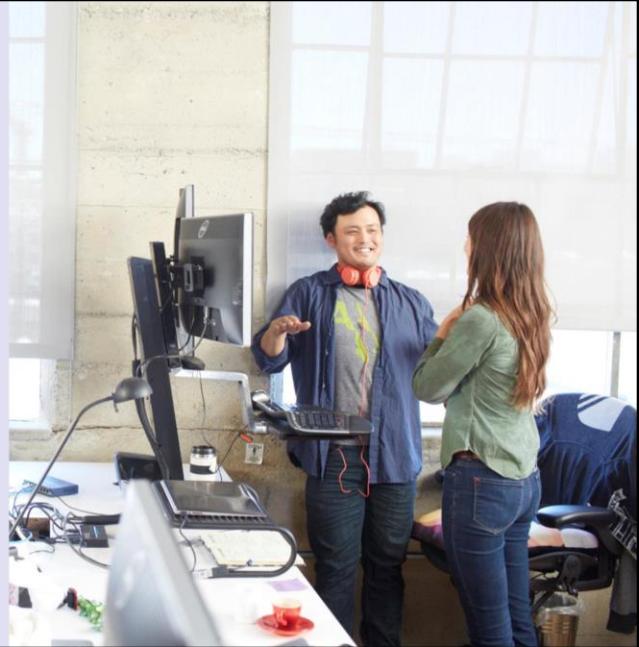


<http://aka.ms/iotedu>

Microsoft

1. Regression: Predict a real value for each item (stock/currency value, temperature). – How much/how many?
2. Classification: Assign a category to each item (Chinese | French | Indian | Italian | Japanese restaurant). – Which Category?
3. Clustering/Recommendation: Partition items into homogeneous groups (clustering twitter posts by topic). – Which Groups?
4. Anomaly: Identify when something unexpected happens. – Is this weird?
5. Reinforcement Learning: Make an appropriate action for some new data. – Which action?

## **DATA SCIENCE: THE BIGGER PICTURE**



1. Data Science for Beginners: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-for-beginners-the-5-questions-data-science-answers/>

## BIGGER THAN ML: DATA SCIENCE



<http://aka.ms/iotedu>



Taking a step back

Data science infographic by Research Hubs:

<http://www.researchhubs.com/post/ai/introduction-to-data-science/index.html>

DS is multidisciplinary with ML as only one aspect of the whole

Pattern recognition

Statistics

Visualizations

Data mining

Database and data process

Neurocomputing (deep learning)

Machine Learning

...this is only one interpretation, but to me it is very accurate

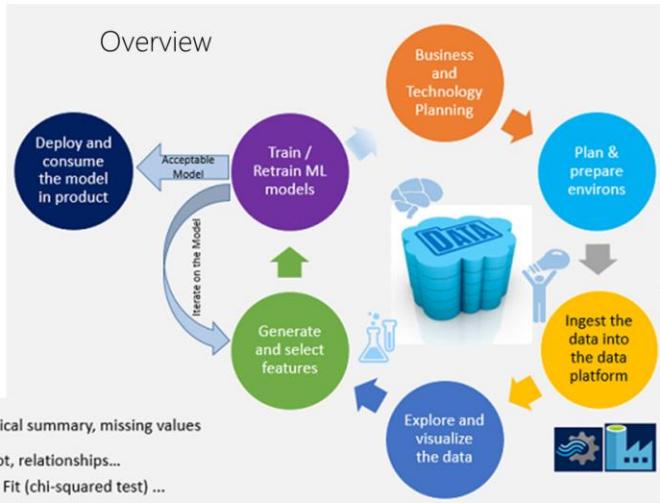
In case you want to learn about data science here's Data Science for Beginners:

<https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-for-beginners-the-5-questions-data-science-answers/>

# TEAM DATA SCIENCE PROCESS

## Considerations

- 🔍 Data dimensions, types, statistical summary, missing values
- 📊 Distribution, histogram, boxplot, relationships...
- 📈 Statistical significance (t-test), Fit (chi-squared test) ...



<http://aka.ms/iotedu>



Adopted by the Azure Data Science team and more info can be found at  
<https://azure.microsoft.com/en-us/documentation/articles/data-science-process-overview/>

# INTRODUCTION TO MACHINE LEARNING TOOLS ON AZURE



## THE AZURE ML ENVIRONMENT

### Development Environment

- Creating Experiments
- Sharing a Workspace



### Deployment Environment

- Publishing the Model
- Using the API
- Consuming in various tools



<http://aka.ms/iotedu>

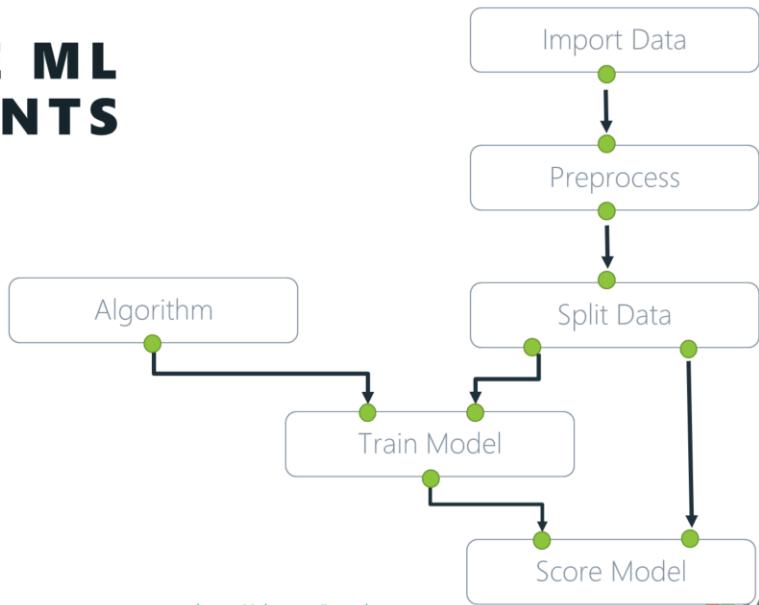
 Microsoft

1. The Azure Machine Learning Studio: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-what-is-ml-studio/>
2. Guided tutorials: <https://azure.microsoft.com/en-us/documentation/services/machine-learning/>
3. Microsoft Azure Virtual Academy course: [https://mva.microsoft.com/en-US/training-courses/microsoft-azure-machine-learning-jump-start-8425?l=ehQZFoKz\\_7904984382](https://mva.microsoft.com/en-US/training-courses/microsoft-azure-machine-learning-jump-start-8425?l=ehQZFoKz_7904984382)

# AZURE ML ELEMENTS

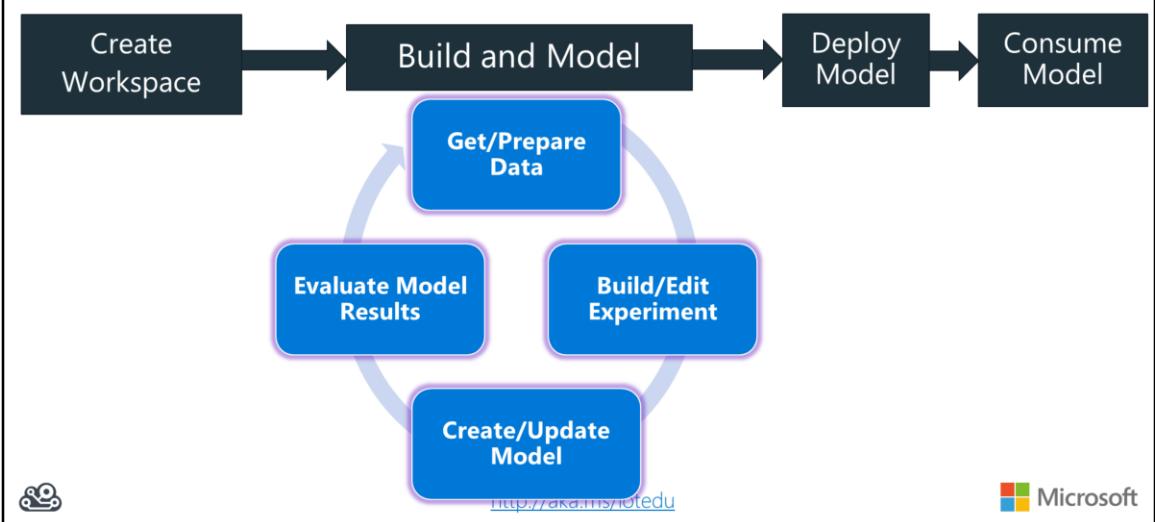


<http://aka.ms/iotedu>



1. Designing an experiment in the Studio: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-what-is-ml-studio/>

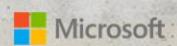
# CREATING AN EXPERIMENT



1. Beginning Series: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-for-beginners-the-5-questions-data-science-answers/>

## EXPLORING DATA AND DATA INGESTION

Microsoft Azure



1. Inspect data – very important
2. Options for data ingestion
3. On-prem data – how to handle
4. Tools for exploration

## FIRST STEP AND MOST IMPORTANT: INSPECTING DATA



Items to check for:

- Number of data points
- Number of attributes or features
- Data types (nominal, ordinal, continuous) are represented correctly
- Amount of missing values
- Well-formedness
- Consistency

<http://aka.ms/iotedu>



1. In reference to machine learning, but applicable to all data usage: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-prepare-data/>

## OPTIONS FOR DATA INGESTION

- PowerShell
- Azure Data Factory
- Azure Automation
- Azure storage SDKs (.NET, Node.js, python, C++, etc.)
- Microsoft Azure Storage Explorer application (blob only right now)
- AzCopy (blob, file, and table only)
- Import/Export service



<http://aka.ms/iotedu>



1. PowerShell in Azure Storage - <https://azure.microsoft.com/en-us/documentation/articles/storage-powershell-guide-full/>
2. Azure Data Factory data movement -  
<https://azure.microsoft.com/en-us/documentation/articles/data-factory-data-movement-activities/>
3. Azure Automation - <https://azure.microsoft.com/en-us/documentation/articles/automation-intro/>
4. Azure storage SDKs – for examples see  
<https://azure.microsoft.com/en-us/documentation/articles/storage-dotnet-how-to-use-blobs/>
5. Azure tools and SDKs in general can be downloaded here -  
<https://azure.microsoft.com/en-us/downloads/>
6. MS Azure Storage Explorer - <http://storageexplorer.com/>
7. AzCopy - <https://azure.microsoft.com/en-us/documentation/articles/storage-use-azcopy/>
8. Import/Export service - <https://azure.microsoft.com/en->

[us/documentation/articles/storage-import-export-service/](#)

## **CONNECT ON-PREM TO <ANYTHING>**

### VPN Gateway

- Send network traffic from virtual networks to on-prem locations
- Send network traffic between virtual networks within Azure
- Site-to-site vs. Point-to-site
- You can connect multiple on-prem locations to a virtual network (Multi-site)
- ExpressRoute can directly connect your WAN to Azure
- Tool-Specific



<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/vpn-gateway-about-vpngateways/>
2. <https://azure.microsoft.com/en-us/documentation/articles/vpn-gateway-vpn-faq/#connecting-to-virtual-networks>
3. <https://azure.microsoft.com/en-us/documentation/articles/expressroute-faqs/>

## **AML: OPTIONS FOR DATA SOURCING**

1. Import from Local
2. Import from Online
3. Import from Experiment (a Import Module or shared dataset in Studio)



<http://aka.ms/iotedu>



1. Getting data: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-import-data/>
2. Importing Data to Azure ML:  
<https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-import-data/>

## **AML: OPTIONS FOR EXPLORING DATA**

- R language or MS R Server
- Azure ML Studio
- Other Tools



<http://aka.ms/iotedu>



1. Data Exploration and Predictive Modeling with R -  
<https://msdn.microsoft.com/en-us/library/mt590947.aspx>
2. Data Exploration with Azure ML -  
<https://blogs.technet.microsoft.com/machinelearning/2015/09/24/data-exploration-with-azure-ml/>
3. Statistics Using Excel – <http://www.excelfunctions.net/Excel-Statistical-Functions.html>
4. Sed, awk, grep (in Windows as well) - <https://www.simple-talk.com/cloud/data-science/data-science-laboratory-system---testing-the-text-tools-and-sample-data/>
5. Python scripts (similar to R data exploration) – would be used from within AML Studio



LAB

## Working in Azure ML Studio – Exploratory Data Analysis



<http://aka.ms/iotedu>



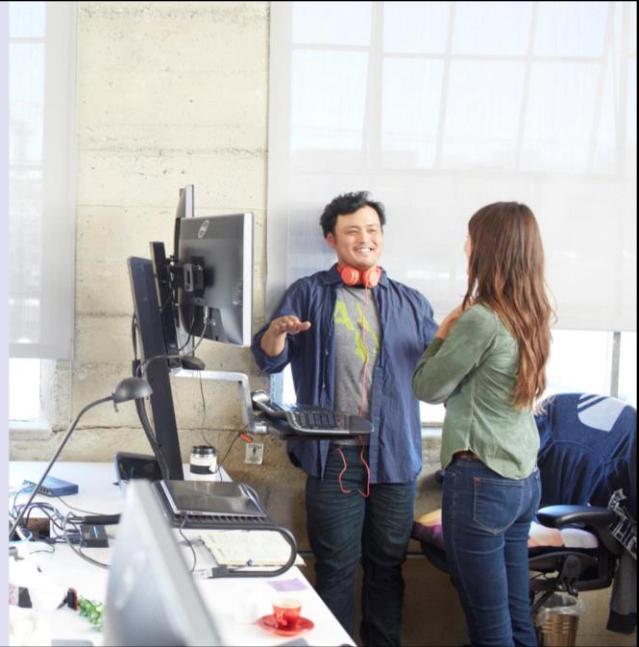
These instructions are in more detail and update on the class website under Labs and LabAML.md or here:  
<https://github.com/michhar/data-pipeline-education/blob/master/Labs/LabAML.md>

Essentially, these are our steps

1. Open Azure ML Studio from <https://studio.azureml.net>
2. Sign-up – this is separate from Azure account, but is completely free and has many free benefits for ML on Azure
3. Let's navigate back to the portal at <https://portal.azure.com>
4. Explore distributions and inspect data in AzureML with Visualize
5. Update the R script in the first module to an EDA

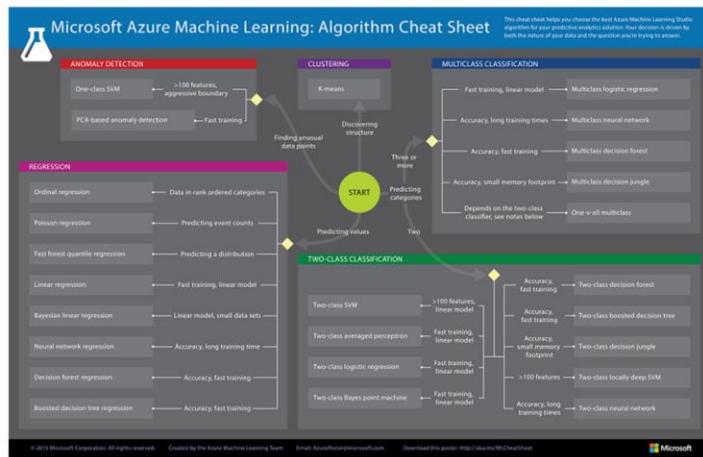
(exploratory data analysis) plot and check this. Use the following code to add at the second to last line in the script.

## **ALGORITHMS YOU'LL ENCOUNTER IN AML**



Algorithm cheatsheet (see next slide as well): <https://azure.microsoft.com/en-in/documentation/articles/machine-learning-algorithm-cheat-sheet/>

# ALGORITHM CHEATSHEET FOR AML



<http://aka.ms/iotedu>



Can download in a more readable format here: <https://azure.microsoft.com/en-in/documentation/articles/machine-learning-algorithm-cheat-sheet/>

# CLUSTERING

Grouping items based on defined Features



<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/>
2. <https://msdn.microsoft.com/en-US/library/azure/dn906012.aspx>
3. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
4. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
5. <https://msdn.microsoft.com/en-us/library/azure/dn905970.aspx>
6. <https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx>

# CLASSIFICATION

Predicting the class or category  
for a single instance of data



<http://aka.ms/ioteuu>

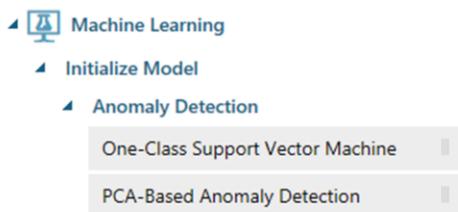


Initialize Model
Classification
Multiclass Decision Forest
Multiclass Decision Jungle
Multiclass Logistic Regression
Multiclass Neural Network
One-vs-All Multiclass
Two-Class Averaged Perceptron
Two-Class Bayes Point Machine
Two-Class Boosted Decision Tree
Two-Class Decision Forest
Two-Class Decision Jungle
Two-Class Locally-Deep Support Vector Machine
Two-Class Logistic Regression
Two-Class Neural Network
Two-Class Support Vector Machine

1. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/>
2. <https://msdn.microsoft.com/en-US/library/azure/dn906012.aspx>
3. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
4. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
5. <https://msdn.microsoft.com/en-us/library/azure/dn905970.aspx>
6. <https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx>

## ANOMALY DETECTION

Selecting items based on unusual or suspicious patterns



<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/>
2. <https://msdn.microsoft.com/en-US/library/azure/dn906012.aspx>
3. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
4. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
5. <https://msdn.microsoft.com/en-us/library/azure/dn905970.aspx>
6. <https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx>

# REGRESSION

Predicting the value of a datum given its history

- ▲ Initialize Model
- ▲ Classification
  - Multiclass Logistic Regression
  - Two-Class Logistic Regression
- ▲ Regression
  - Bayesian Linear Regression
  - Boosted Decision Tree Regression
  - Decision Forest Regression
  - Fast Forest Quantile Regression
  - Linear Regression
  - Neural Network Regression
  - Ordinal Regression
  - Poisson Regression

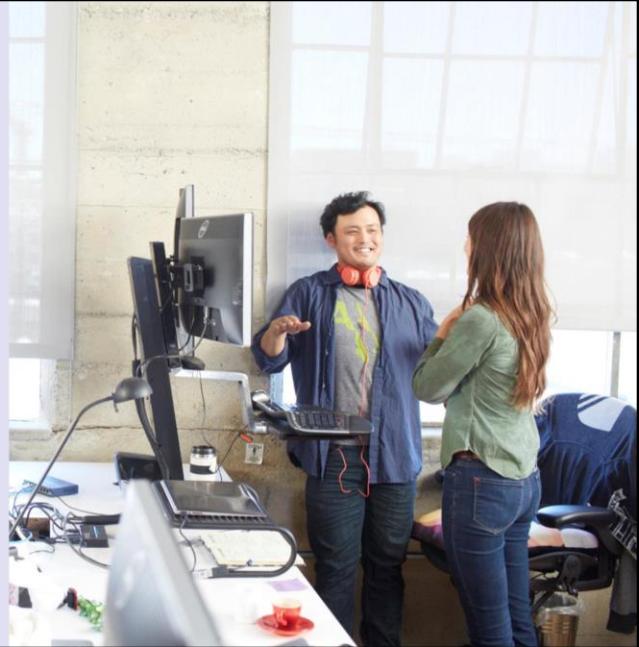


<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/>
2. <https://msdn.microsoft.com/en-US/library/azure/dn906012.aspx>
3. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
4. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
5. <https://msdn.microsoft.com/en-us/library/azure/dn905970.aspx>
6. <https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx>

## **MODEL SCORING AND EVALUATION**



Train and Evaluate your Model: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-walkthrough-4-train-and-evaluate-models/>

## SCORING A MODEL

Apply a trained model to:

- A list of recommended items
- Forecasts for time series models
- Estimates of projected demand, volume, or other numeric quantity, for regression models
- Cluster assignments
- A predicted class or outcome, for classification models
- Probability scores associated with these outputs



<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-choice/>
2. <https://msdn.microsoft.com/en-US/library/azure/dn906012.aspx>
3. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
4. <https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx>
5. <https://msdn.microsoft.com/en-us/library/azure/dn905970.aspx>
6. <https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx>

## EVALUATING A MODEL

### Metrics for Classification Models

- Accuracy, Recall, Precision, F1-Score
- AUC
- Average Log Loss
- Training Log Loss

### Metrics for Regression Models

- Mean absolute error (MAE)
- Root mean squared error (RMSE)
- Relative absolute error (RAE)
- Relative squared error (RSE)
- Coefficient of determination



<http://aka.ms/iotedu>



1. Simple explanation of the ROC Curve: <http://blog.revolutionanalytics.com/2016/08/roc-curves-in-two-lines-of-code.html>
2. <https://msdn.microsoft.com/en-us/library/azure/dn906026.aspx>
3. <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-evaluate-model-performance/>
4. <https://msdn.microsoft.com/library/azure/75fb875d-6b86-4d46-8bcc-74261ade5826>
5. <https://msdn.microsoft.com/library/azure/927d65ac-3b50-4694-9903-20f6c1672089>
6. <https://msdn.microsoft.com/library/azure/e9ad68a7-e91b-4ae6-800e-b5ee7e22cd17>

## AML OPTIONS FOR DATA ACCESS

- Azure ML API
- Code
- Push to Storage
- Power BI / Excel
- ADF
- ASA



<http://aka.ms/iotedu>

The screenshot shows the 'Export Data' dialog box of the 'SET UP WEB SERVICE' application. It includes fields for specifying the data destination (Azure SQL Database), database server name, database name, server user account name, server user account password, and various export options like accepting insecure certificates and specifying column lists. The Microsoft logo is present at the bottom right.

1. Access and read through this page:

<https://support.office.com/en-us/article/Connect-to-Microsoft-Azure-Blob-Storage-Power-Query-f8165faa-4589-47b1-86b6-7015b330d13e?ui=en-US&rs=en-US&ad=US&fromAR=1>

2. Access and read through this page:

<http://social.technet.microsoft.com/wiki/contents/articles/2128.azure-and-sql-database-tutorials-tutorial-1-using-azure-web-role-and-azure-table-service.aspx>

3. Accessing storage using Code: <https://azure.microsoft.com/en-us/documentation/articles/storage-dotnet-how-to-use-blobs/>

4. Working with Azure Storage: <https://azure.microsoft.com/en-us/documentation/services/storage/>



# LAB

Customizing the Azure Machine Learning experiment



<http://aka.ms/iotedu>



In this simple lab:

- Follow along with instructor
- Change the classification algorithm from Boosted Tree to another classifier as you see fit

# PREDICTIVE ANALYTICS FOR IoT

Event Hubs + Streaming Analytics + Storage + SQL + Machine Learning + Data Factory = AWESOME

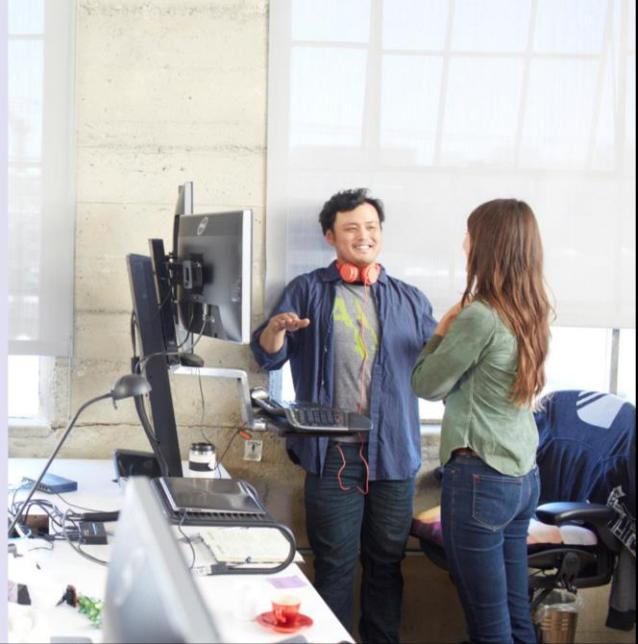
## PART 2 OF 3: DATA AT REST AND ORCHESTRATING DATA FLOW



<http://aka.ms/iotedu>



## DATA AT REST AND AZURE STORAGE OPTIONS



1. Data Storage Options (Building Real-World Cloud Apps with Azure):  
<https://www.asp.net/aspnet/overview/developing-apps-with-windows-azure/building-real-world-cloud-apps-with-windows-azure/data-storage-options>

# STORAGE SCENARIOS

Unstructured data such as media files, logs, binary data, backups

0100110100  
1010101001  
0101010101  
0101010101



Blob

Metadata (e.g. user info), in key-value format, fast and easy to query

{'name': 'Sue',  
'role': 'admin',  
'status': 'active',  
'location': 'WA'}

Table

Messaging between components of your application



Queue

Shared file systems option – when your application is already built to use a SMB protocol



File



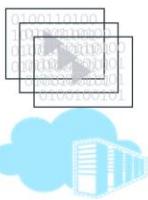
<http://aka.ms/iotedu>



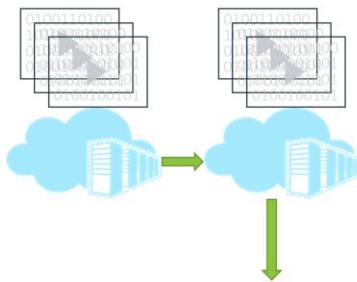
1. <https://channel9.msdn.com/Blogs/Windows-Azure/Azure-Storage-5-Minute-Overview>
2. <https://azure.microsoft.com/en-us/documentation/articles/storage-introduction/>

## Redundancy and Location

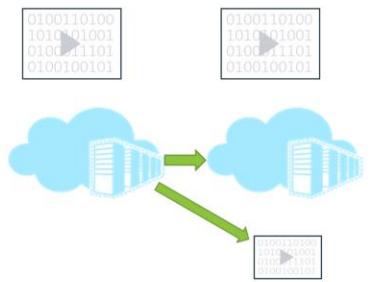
LRS: 3 Copies,  
1 Datacenter



GRS: 6 Copies,  
2 Datacenters



ZRS: 3 Copies,  
2-3 Datacenters



<http://aka.ms/iotedu>

### 1. Locations and Redundancy Overview:

<https://azure.microsoft.com/en-us/documentation/articles/storage-introduction/>

### 2. Affects on Scalability and Performance Targets:

<https://azure.microsoft.com/en-us/documentation/articles/storage-scalability-targets/>

### 3. Pricing Details: <https://azure.microsoft.com/en-us/pricing/details/storage/>

## **OPTIONS FOR CREATING AND MANAGING AZURE STORAGE**

- Azure Portal
- Azure PowerShell
- Azure Command Line Interface (CLI)
- Service Management REST API
- Azure Storage Resource Provider REST API



<http://aka.ms/iotedu>



1. Azure Portal - <https://portal.azure.com/>
2. Azure PowerShell - <https://azure.microsoft.com/en-us/documentation/articles/storage-powershell-guide-full/>
3. AZCOPY - <https://azure.microsoft.com/en-us/documentation/articles/storage-use-azcopy/>
4. Azure CLI - <https://azure.microsoft.com/en-us/documentation/articles/storage-azure-cli/>
5. Service management REST API -  
<http://msdn.microsoft.com/library/azure/ee460799.aspx>
6. Azure Storage Resource Provider REST API -  
<https://msdn.microsoft.com/library/azure/mt163683.aspx>
7. SDKs (e.g. python Azure official SDKs -  
[https://github.com/Azure/azure-storage-python \)](https://github.com/Azure/azure-storage-python)

## **CREATING AND MANAGING AZURE STORAGE**

- Azure Portal
- Azure PowerShell
- Azure Command Line Interface (CLI)
- Service Management REST API
- Azure Storage Resource Provider REST API



<http://aka.ms/iotedu>



1. Azure Portal - <https://portal.azure.com/>
2. Azure PowerShell - <https://azure.microsoft.com/en-us/documentation/articles/storage-powershell-guide-full/>
3. AZCOPY - <https://azure.microsoft.com/en-us/documentation/articles/storage-use-azcopy/>
4. Azure CLI - <https://azure.microsoft.com/en-us/documentation/articles/storage-azure-cli/>
5. Service management REST API -  
<http://msdn.microsoft.com/library/azure/ee460799.aspx>
6. Azure Storage Resource Provider REST API -  
<https://msdn.microsoft.com/library/azure/mt163683.aspx>

## **DATA AT REST AND SQL DB**



1. Data Science for Beginners: <https://azure.microsoft.com/en-us/documentation/articles/machine-learning-data-science-for-beginners-the-5-questions-data-science-answers/>

## AZURE SQL DATABASE

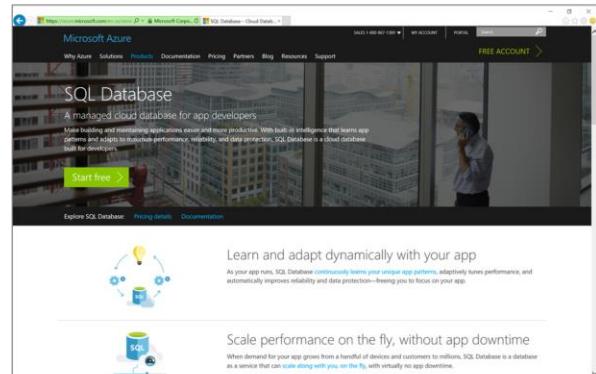
As your app runs, SQL Database continuously learns your unique app patterns, adaptively tunes performance, and automatically improves reliability and data protection...

...SQL Database is a database as a service that can scale along with you, on the fly, with virtually no app downtime....

SQL Database helps you manage and monitor multitenant apps and gain isolation benefits of one-customer-per-database.

Develop your app and connect to SQL Database with the tools and platforms you prefer...

With physical and operational security, Azure SQL Database helps you meet the most stringent regulatory compliances.



<http://aka.ms/iotedu>



# DEMO

## Connecting to SQL Database in Visual Studio

Follow along as in <https://github.com/michhar/data-pipeline-education/blob/master/Labs/LabSQL.md>

### Troubleshooting:

- Make sure you use server name w/o “,1433”
- In portal make sure:
  - Turn on Firewall (Under Server)
  - Transparent data encryption (under Database, not Server)
  - In classic portal “Configure” under SQL Database -> Servers (pick your server) and Allow Client ip
- If you need to reset password use classic portal as described here:  
<http://stackoverflow.com/questions/13790752/password-reset-for-azure-database>

## **NOTE - SQL SERVER NEW ENCRYPTION FEATURE**

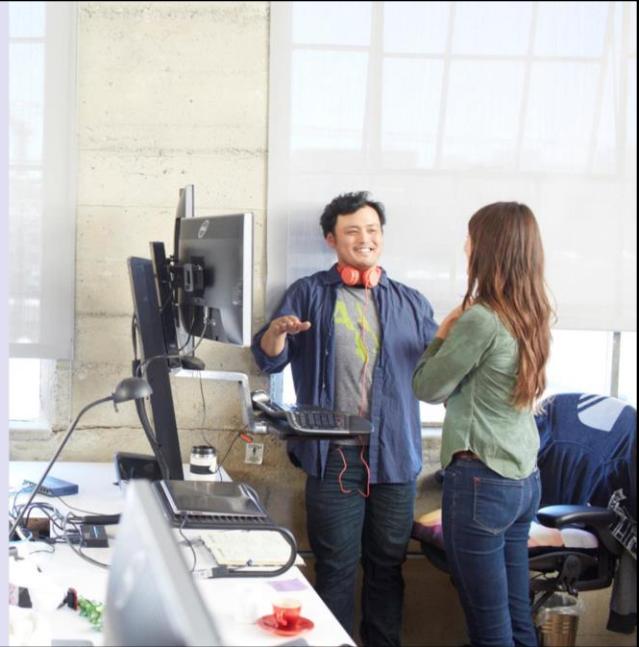
Always Encrypted is a new data encryption technology in Azure SQL Database and SQL Server that helps protect sensitive data at rest on the server, during movement between client and server, and while the data is in use. Always Encrypted ensures that sensitive data never appears as plaintext inside the database system.



<http://aka.ms/iotedu>



# DATA ORCHESTRATION WITH AZURE DATA FACTORY



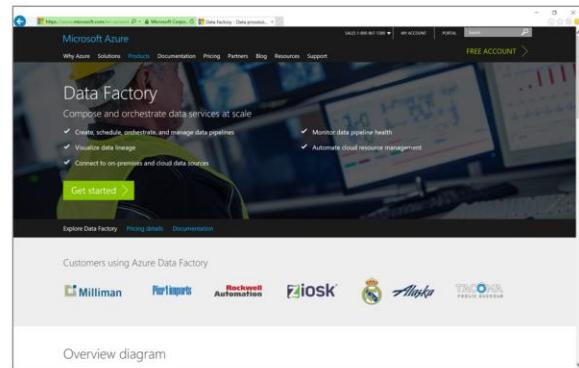
## AZURE DATA FACTORY

Azure Data Factory is a globally deployed **data movement service** in the cloud. Use it to ingest data from multiple on-premises and cloud sources easily.

Schedule, orchestrate, and manage the data transformation and analysis process.

Using data pipelines, transform raw data into finished or shaped data that's ready for consumption by BI tools or applications.

Visualize, monitor, and manage your entire network of **data pipelines** at a glance to identify issues and take action.



<http://aka.ms/iotedu>



## AZURE DATA FACTORY



Create, orchestrate, and manage data movement and enrichment through the cloud

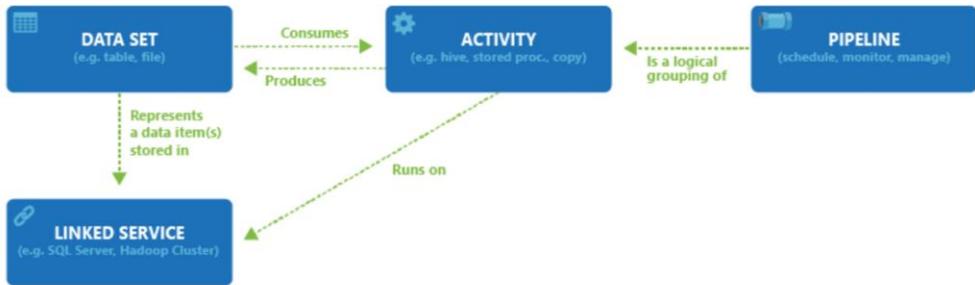


<http://aka.ms/iotedu>



1. Learning Path: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-introduction/>
2. Developer Reference: <https://msdn.microsoft.com/en-us/library/azure/dn834987.aspx>

## ADF COMPONENTS



<http://aka.ms/iotedu>



1. Pricing: <https://azure.microsoft.com/en-us/pricing/details/data-factory/>

## ADF LOGICAL FLOW

Overview diagram



<http://aka.ms/iotedu>



1. Learning Path: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-introduction/>

2. Quick Example:

<http://azure.microsoft.com/blog/2015/04/24/azure-data-factory-update-simplified-sample-deployment/>

## ADF Process

1. Define Architecture: Set up objectives and flow
2. Create the Data Factory: Portal, PowerShell, VS
3. Create Linked Services: Connections to Data and Services
4. Create Datasets: Input and Output
5. Create Pipeline: Define Activities
6. Monitor and Manage: Portal or PowerShell, Alerts and Metrics



<http://aka.ms/iotedu>



1. Full Tutorial: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline/>

## DESIGN PROCESS

1. Define data sources
2. Define processing requirements
3. Define output
4. Plan for management and monitoring



<http://aka.ms/iotedu>



1. <https://azure.microsoft.com/en-us/documentation/articles/data-factory-customer-profiling-usecase/>

## **CREATE THE DATA FACTORY**

Methods: Portal, PowerShell or Visual Studio



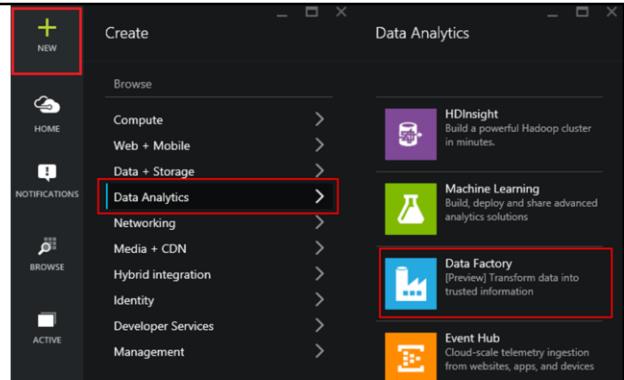
<http://aka.ms/iotedu>



1. Setting Up: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline/>

## USING THE PORTAL

- Use for Exploration
- Use when teaching or in a Demo
- Use when UI preferred or needed



<http://aka.ms/iotedu>



1. Overview: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline/>
2. Using the Portal: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline-using-editor/>

# USING POWERSHELL

- Use in MS Clients
- Use for Automation
- Use for quick set up and tear down

```
Windows PowerShell V2 (C:\)
PS C:\> Get-WmiObject -Namespace root\virtualization -Query "Select * From Msvm_ComputerSystem Where ElementName='TESTVM'
M1">

```

GENUS	VERSION	CLASS	CLASSID
2	1	Msvm_ComputerSystem	CIM_ComputerSystem
			CIM_ManagedElement
			Method Msvm_ComputerSystem.CreationClassName="Msvm_ComputerSystem",Name="3F837600-F140-4B89-982F-8597737C16"
			2F-8E91FF37C16"
			2
			DERIVATION
			1
			GENERALIZATION
			1
			INTERFACE
			1
			PRIM
			1
			QualifiedNodeList
			1
			CreationClassName
			Msvm_ComputerSystem
			Description
			Microsoft Virtual Computer System
			EnabledDefault
			2
			HealthState
			5
			HealthStateDescriptions
			1
			InstallDate
			280905000005614_0000000-0000
			NameFormat
			3F837600-F140-4B89-982F-8597737C16
			OnLineInMilliseconds
			234778
			OtherDedicatedDescriptions
			1
			OtherIdentifyingInfo
			1
			PrimaryPowerContact
			SERUEBS5\\$Administrator
			ProcessorName
			2608
			ProcessId
			12
			ResetCapability
			1
			Roles
			1
			StatusDescription
			2
			TimeOfLastConfigurationChange
			28090500034126_0000000-0000
			TimeOfLastStateChange
			28090500034126_0000000-0000

```
PS C:\>
```

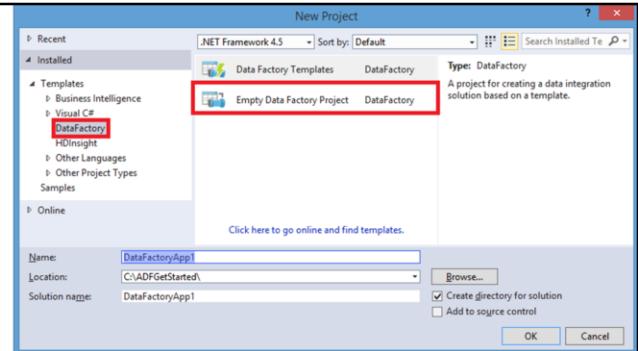


<http://aka.ms/iotedu>



1. Learning Path: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-introduction/>
2. Full Tutorial: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline/>

## USING VISUAL STUDIO



- Use in mature dev environments
- Use when integrated into larger development process



<http://aka.ms/iotedu>



1. Overview: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline/>
2. Using the Portal: <https://azure.microsoft.com/en-us/documentation/articles/data-factory-build-your-first-pipeline-using-editor/>

# PREDICTIVE ANALYTICS FOR IoT

Event Hubs + Streaming Analytics + Storage + SQL + Machine Learning + Data Factory = AWESOME

## PART 3 OF 4: SECURITY OF THINGS



<http://aka.ms/iotedu>



## IOT SECURITY TENANTS

Security of IoT pertains to

- ensuring the integrity of code running on devices
- providing device and user authentication
- defining clear ownership of devices (as well as data generated by those devices)
- being resilient to cyber and physical attacks

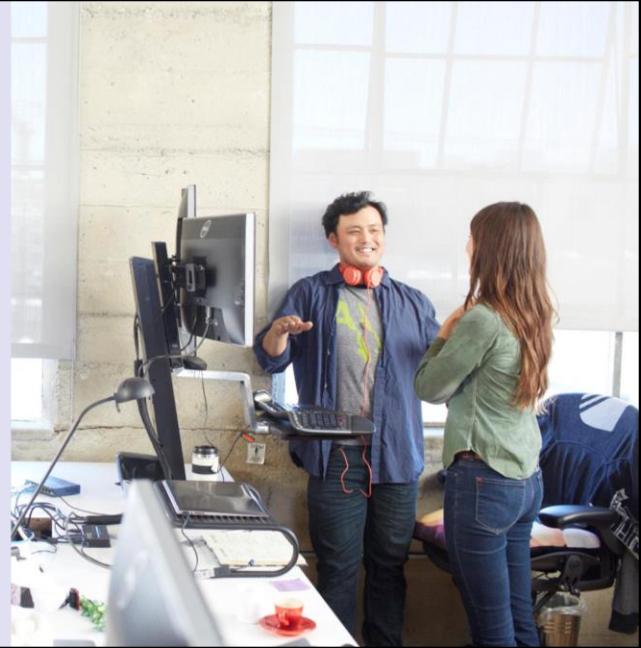


<http://aka.ms/iotedu>



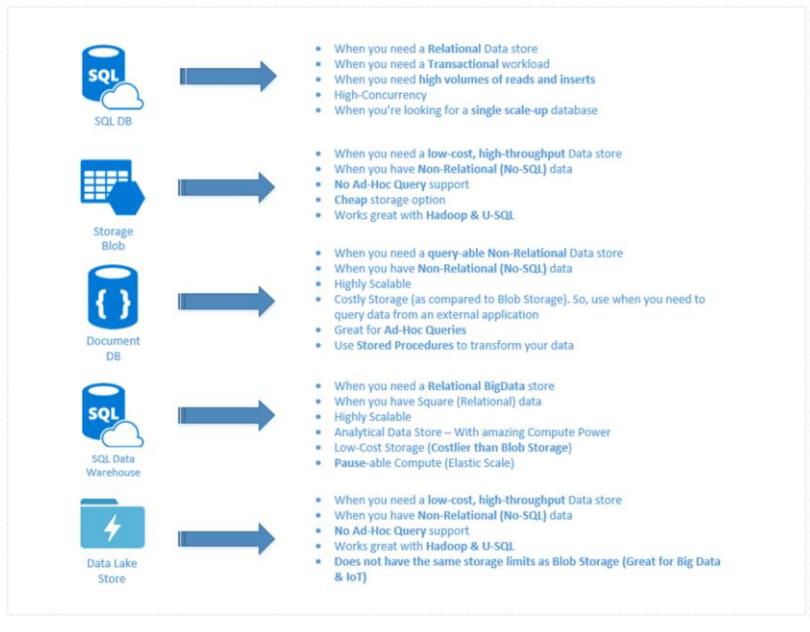
Please see: <https://azure.microsoft.com/en-us/documentation/articles/securing-iot-ground-up/>

## **SOME GENERAL CIS NOTES**



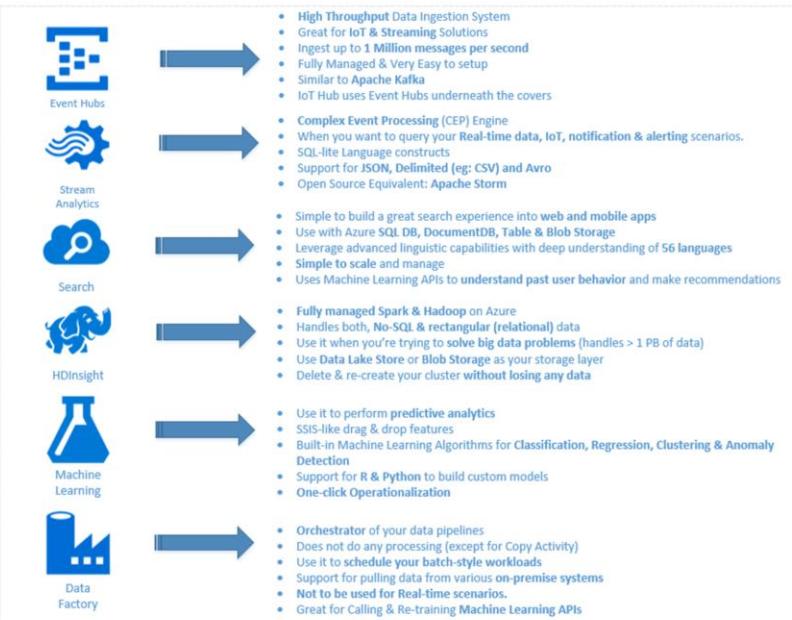
Again, a great foundations course is here: (need your MS account id, but otherwise it's free and open):

## WHAT TO USE FOR DATA



1. A discussion of this graphic: <https://buckwoody.wordpress.com/2016/05/16/the-cortana-intelligence-suite-what-to-use-when/>

## WHAT TO USE FOR COMPUTE



1. A discussion of this graphic: <https://buckwoody.wordpress.com/2016/05/16/the-cortana-intelligence-suite-what-to-use-when/>

## GENERAL RESOURCES

Video Resource for all Azure offerings: <https://azure.microsoft.com/en-us/documentation/videos/home/> (this is what I use mainly to learn!)

General/Azure UserVoice: <https://feedback.azure.com/forums/34192--general-feedback> (can find individual product offerings here as well)

Azure SLAs: <https://azure.microsoft.com/en-us/support/legal/sla/>



<http://aka.ms/iotedu>



# PREDICTIVE ANALYTICS FOR IoT

Event Hubs + Streaming Analytics + Storage + SQL + Machine Learning + Data Factory = AWESOME

## PART 4 OF 4: **CORTANA INTELLIGENCE QUICK START PATTERN AUTHORING - PRESENTED BY DARWIN SCHWEITZER**

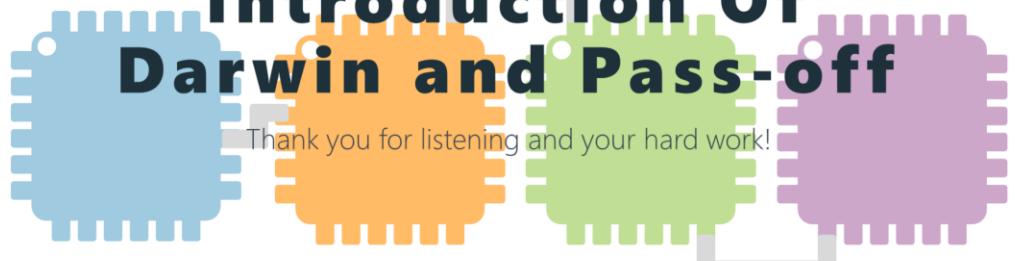


<http://aka.ms/iotedu>



<https://azure.microsoft.com/en-us/documentation/articles/securing-iot-ground-up/>

# Introduction Of Darwin and Pass-off



Thank you for listening and your hard work!



<http://aka.ms/iotedu>

