1. Main page: http://cortanaanalytics.com
2. To use this Module, you need to be able to:
    1. Understand the process of marrying a series of requirements to a series of capabilities within the Cortana Intelligence Suite
    2. Be able to show the solution designed to a business and technical audience
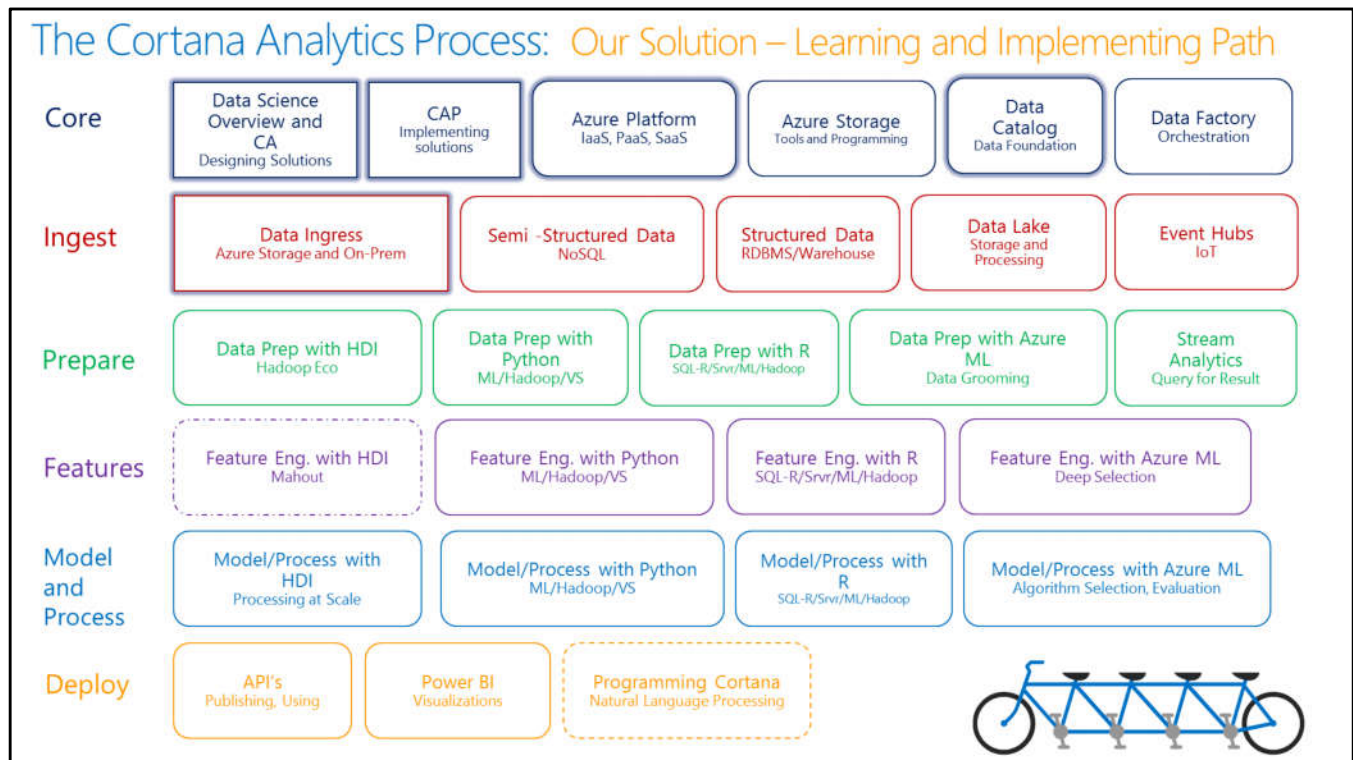   3. Use the Graphical Tools for multiple Cortana Intelligence Suite products
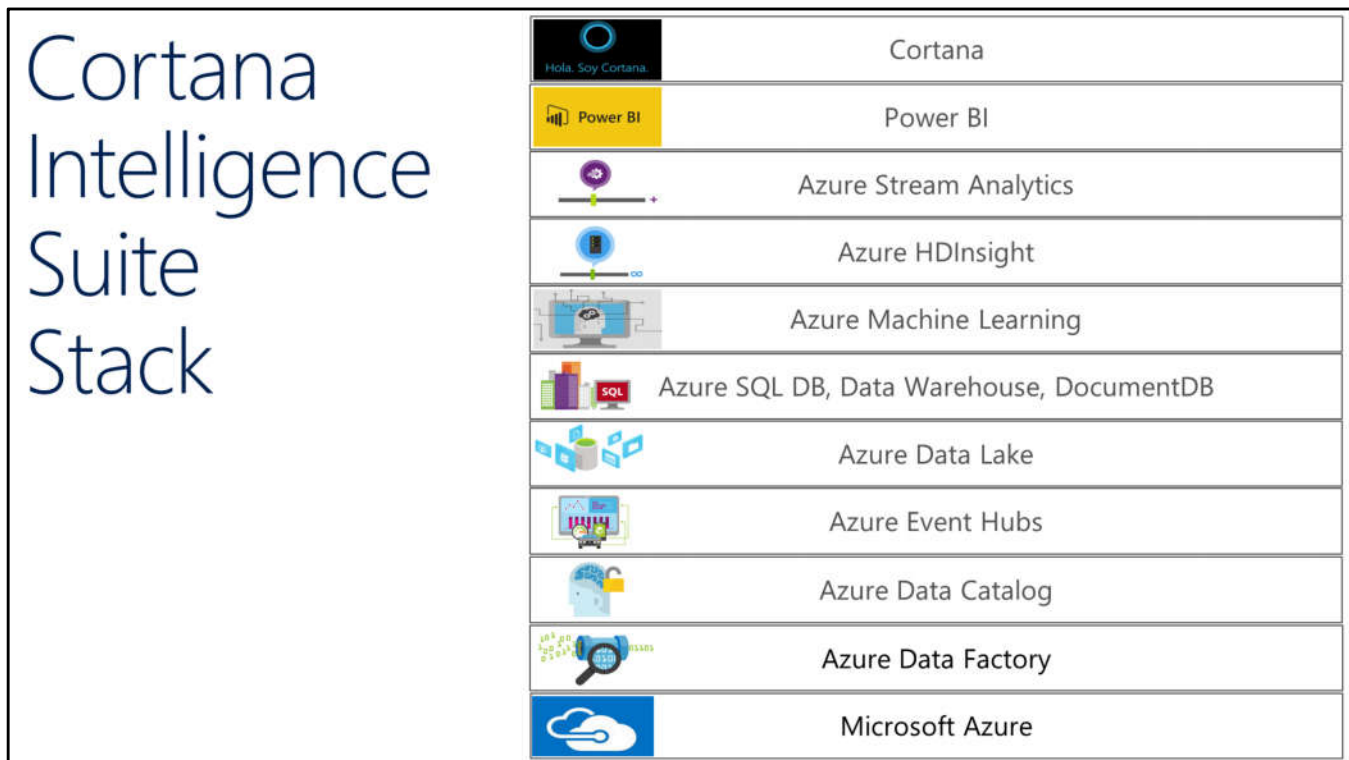
Microsoft Data Science

When you finish this model, you will be able to:
1. Understand how to vet data sources
2. Use Azure Data Catalog to identify, discover and use data in any source
3. Use multiple methods for data ingestion into Azure Storage for use with the Cortana Intelligence Suite
4. Use bridging technologies such as VPN's to leave data on-prem and use it in the Cortana Intelligence Suite
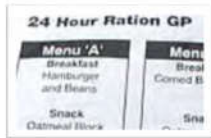
The Cortana Analytics Process: Our Solution – Learning and Implementing Path

1. The Cortana Analytics Process: https://azure.microsoft.com/en-us/documentation/learning-paths/cortana-analytics-process/

1. Platform and Storage: Microsoft Azure – http://microsoftazure.com Storage: https://azure.microsoft.com/en-us/documentation/services/storage/ (Host It)
2. Azure Data Factory: http://azure.microsoft.com/en-us/services/data-factory/ (Move It)
3. Azure Data Catalog: http://azure.microsoft.com/en-us/services/data-catalog (Doc It)
4. Azure Event Hubs: http://azure.microsoft.com/en-us/services/event-hubs/ (Bring It)
5. Azure Data Lake: http://azure.microsoft.com/en-us/campaigns/data-lake/ (Store It)
6. Azure DocumentDB: https://azure.microsoft.com/en-us/services/documentdb/?WT.srch=1&WT.mc_ID=SEM_JQ3fO8dU, Azure SQL Data Warehouse: http://azure.microsoft.com/en-us/services/sql-data-warehouse/ (Relate It)
7. Azure Machine Learning: http://azure.microsoft.com/en-us/services/machine-learning/ (Learn It)
8. Azure HDInsight: http://azure.microsoft.com/en-us/services/hdinsight/ (Big It)
9. Azure Stream Analytics: http://azure.microsoft.com/en-us/services/stream-analytics/ (Stream It)
10. Power BI: https://powerbi.microsoft.com/ (See It)
11. Cortana: http://blogs.windows.com/buildingapps/2014/09/23/cortana-integration-and-speech-recognition-new-code-samples/ and https://blogs.windows.com/buildingapps/2015/08/25/using-cortana-to-interact-with-your-customers-10-by-10/ (Say It)

Data Sourcing and Quality

Defense Catering Group
Stock Number: **99 000 1111**
24 Hour Cold Climate Ration Pack

Army Stores System 3
Stock Number: **99 000 1111**
Valve, Electronic

Inspecting data

Keys to quality source data

- Authority
- Spread
- Consistency
- Types and Units
- Representation

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/

Azure Data Catalog

1. Register data sources
2. Tag the data descriptions
3. Make it easy to find data in context
4. Use the data – keep it secure

1. Full example: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-get-started/

Register Data Sources

Connect
Location, credentials

Select
Objects such as tables and view

Register
Metadata extraction, registration

- SQL Server Table
- SQL Server View
- Oracle Table
- Oracle View
- Teradata Table
- Teradata View
- SSAS Dimension
- SSAS Measure

- SSAS KPI
- SSAS Tabular Table
- SSRS Report
- Azure Storage Blob
- Azure Storage Directory
- HDFS File
- HDFS Directory
- Hive Table

1. Registering a Data Source: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-how-to-register/
2. Types of Data Supported: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-dsr/

Tag the Data

General Properties
Friendly name, Description, Experts, Tags, Connection Info, Datestamps

Columns
Extracted, tags, description

Docs
Formatted longer text

1. Annotating Data Sources: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-how-to-annotate/

Explore Data Sources

Search

Enter terms   *name:finance AND (tags:tag1 OR tags:tag2)*

Filter

By expert, data source, object type, or tags

Annotate

Where allowed, enter tags

1. Discovering Data Sources: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-how-to-discover/
2. Search Term Examples: https://msdn.microsoft.com/library/azure/mt267594.aspx

1. Connecting to data sources: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-how-to-connect/
2. API Concepts and Reference with Examples: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-developer-concepts/

1. Open the Azure Data Catalog site (in a Private Browser View) and log in, using the "Sign In" Worksheet – http://azuredatacatalog.com
2. Follow along with the instructor
3. Add one "Other" datatype
4. Annotate another person's data
5. Browse and search for data
6. Optionally, follow this example: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-get-started/
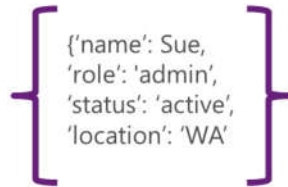
Data Ingestion

Storage Scenarios

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

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

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

Bridging Options for on-prem Data

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

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/

1. Understand how to vet data sources
2. Use Azure Data Catalog to identify, discover and use data in any source
3. Use multiple methods for data ingestion into Azure Storage for use with Cortana Intelligence Suite Components
4. Use bridging technologies such as VPN's to leave data on-prem and use it in the Cortana Intelligence Suite