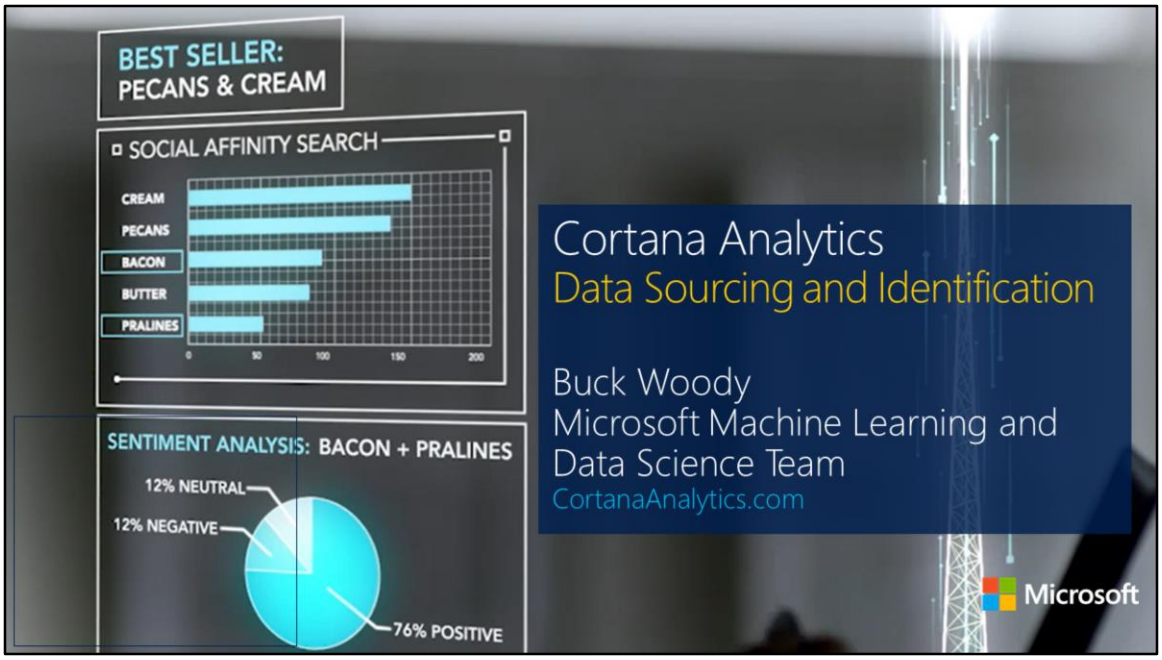
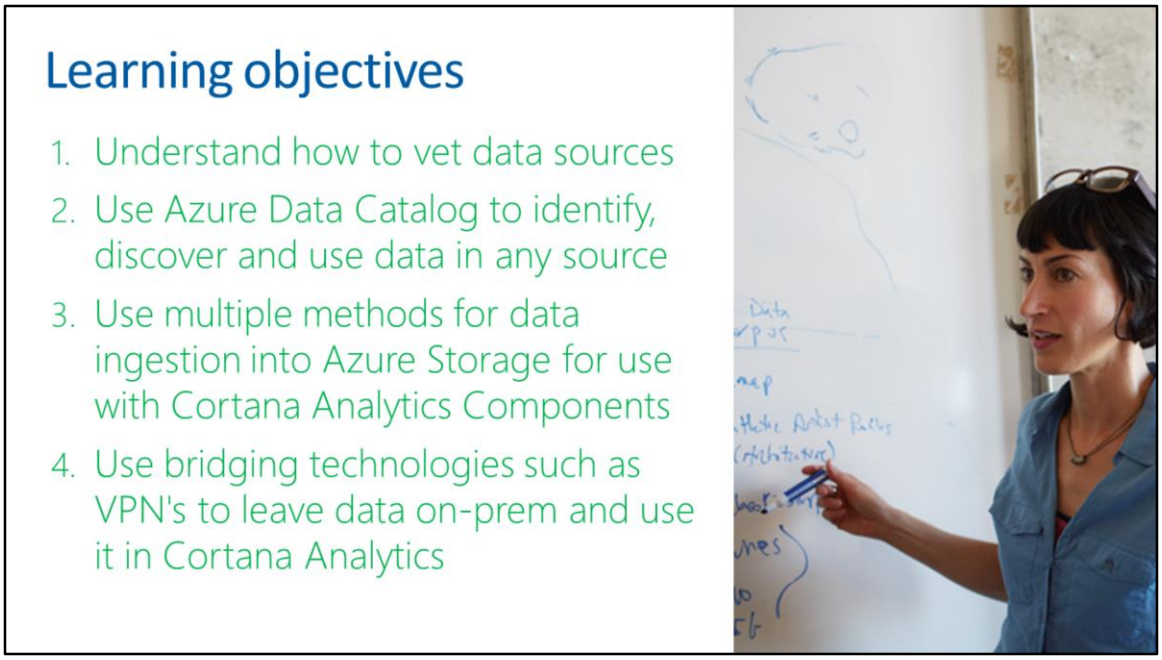


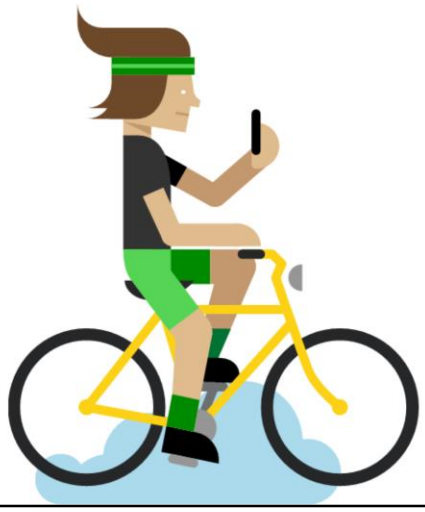
1. Main page: http://cortanaanalytics.com
2. To use this Module, you need to be able to:
    1. Understand the process of marrying a series of requirements to a series of capabilities within the Cortana Analytics suite
    2. Be able to show your solution designed to a business and technical audience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



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 Cortana Analytics Components
4. Use bridging technologies such as VPN's to leave data on-prem and use it in Cortana Analytics

Business Case Review



Business Case

AdventureWorks is a company that makes and sells bicycles. The sales are conducted around the world. We also support our products. But as we've made more sales in the last 10 year, we've farmed out the support function to various companies that take in maintenance and support issues in call centers around the world.

We're growing. And now we want to take our bicycles to several large retailers, but a few of them want to know a lot about our churn rate.

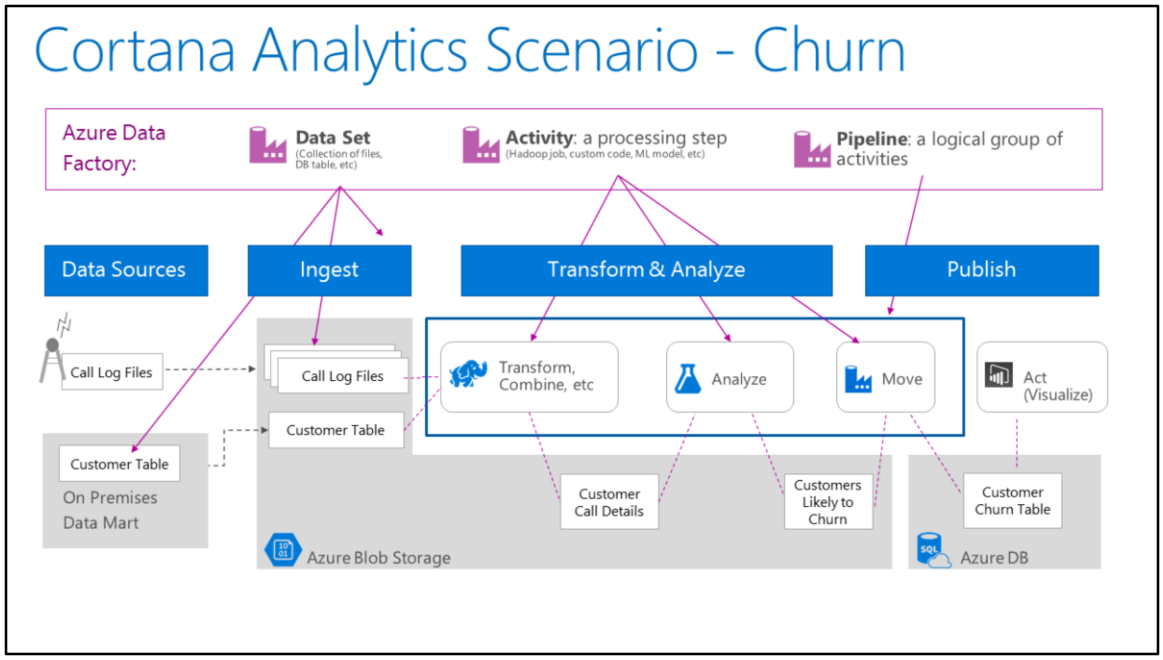
For over 10 years, we've collected a lot of information about our customers and of course we know a lot about our products. But since we've outsourced our call centers, we don't own the databases that hold their data – they will give us an export, though. (They support multiple customers)

We're not sure about our churn rate – we have the data of who has and has not bought again, and we think we can get the data from the call centers for the complaints and repairs, but we need a way to analyze a lot of data that has different formats to find a prediction of who will churn and who will not.

Ideally we want a list of customers we think will churn, in a structured database we could share out to our potential resellers sales staff, so they know how to target at-risk and new clients.

More on our in-house data: https://technet.microsoft.com/en-us/library/ms124501%28v=sql.100%29.aspx

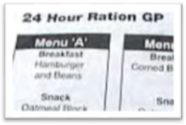
1. AdventureWorks Data Dictionary: https://technet.microsoft.com/en-us/library/ms124438(v=sql.100).aspx



1. Full explanation of this example: https://azure.microsoft.com/en-us/blog/getting-started-with-azure-data-factory-and-azure-machine-learning-4/

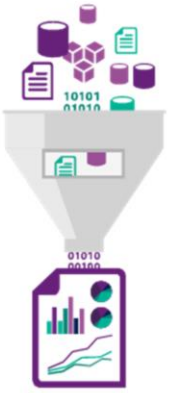
Data Sourcing and Quality

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

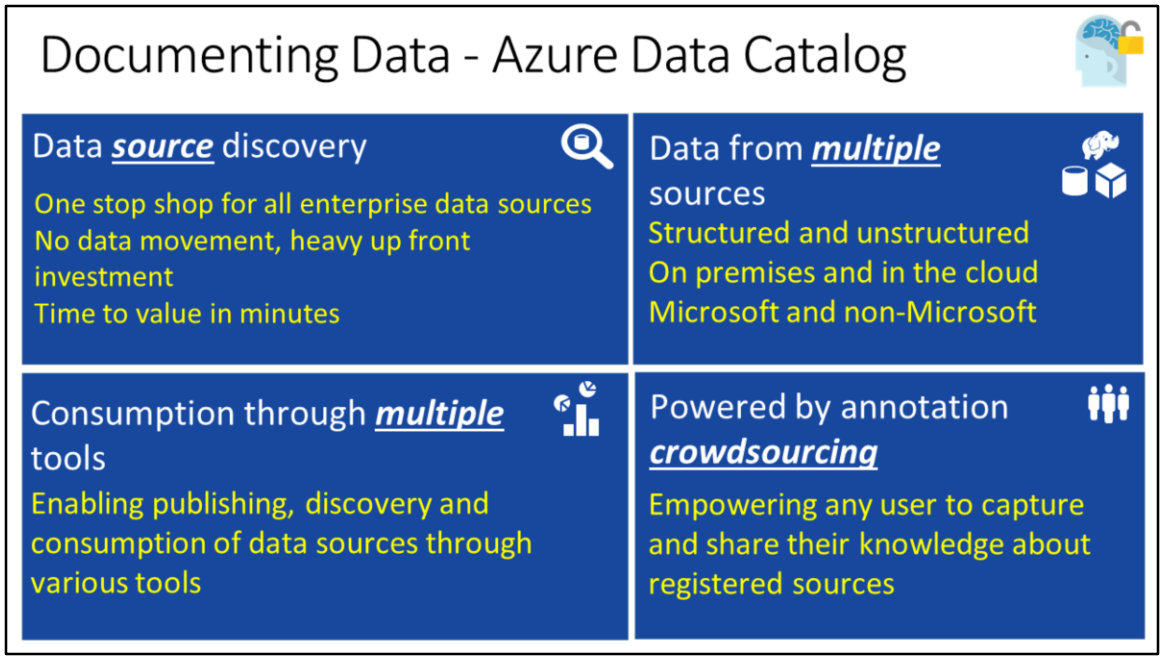
Inspecting data



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/



1. Primary site: http://azure.microsoft.com/en-us/services/data-catalog
2. Full description: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-what-is-data-catalog/
3. Setting up your Azure Account: https://msdn.microsoft.com/en-us/library/azure/mt403304.aspx?f=255&MSPPError=-2147217396

Responsibility Matrix

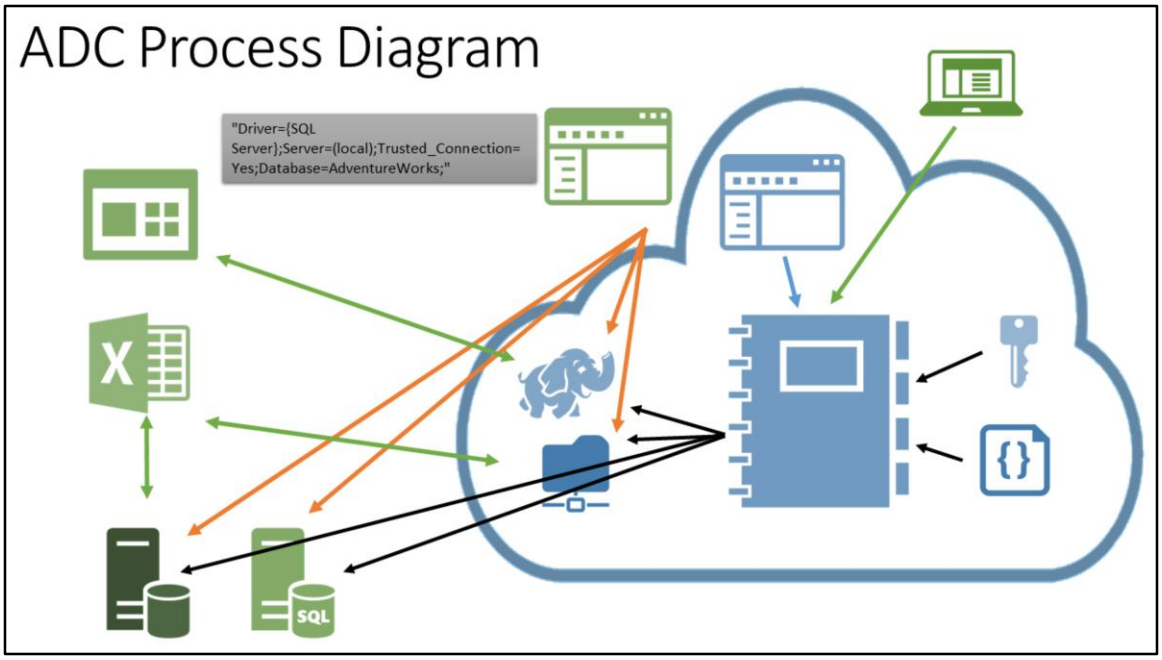
IT Admin	Publisher	Consumer
Govern *Apply policies and control access* *Analyze *Track and monitor usage*	Publish *Register Data Sources* Enrich *Categorize and Annotate*	Discover *Browse and search* Understand *Get context* Enrich *Categorize and Annotate*

1. Process description, including roles: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-terminology/#request-access

Process:

1. Register data using the portal or the tool
2. Enrich data sources with the portal
3. Explore data sources
4. Consume data source

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



1. Components description: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-terminology/

1. 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/

2. Enrich Data Sources

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/

3. 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

3. Consume Data

"Open In"

Preview, Excel, Power BI

Connection Strings and Requesting Access

Display as much as you want public

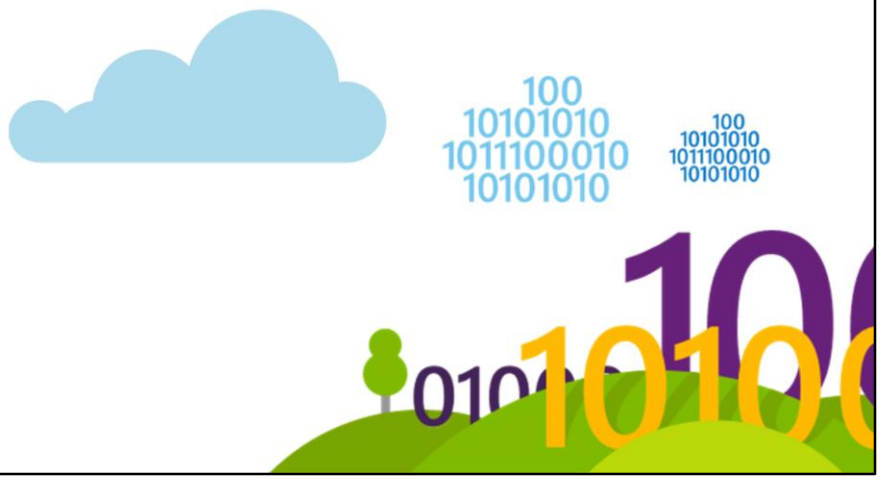
API Access

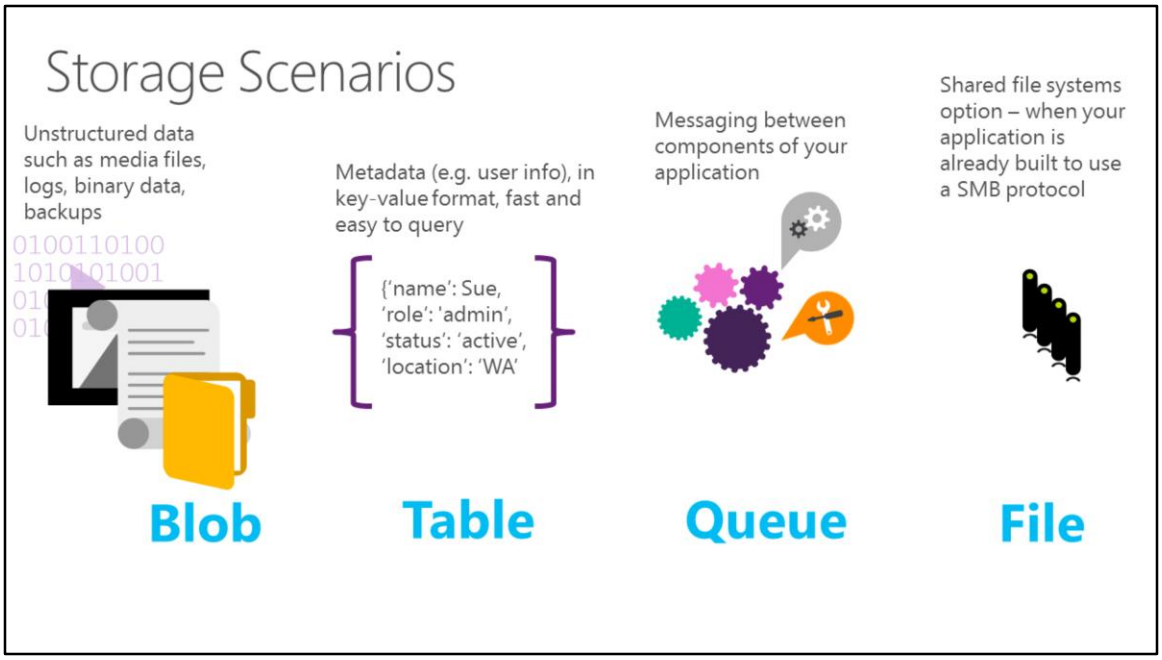
REST API Model

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 View) and log in – http://azuredatacatalog.com
2. Follow along with the instructor
3. Add one “Other” datatype
4. Add one Web Endpoint:
5. Optionally, follow this example: https://azure.microsoft.com/en-us/documentation/articles/data-catalog-get-started/

Data Ingestion





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

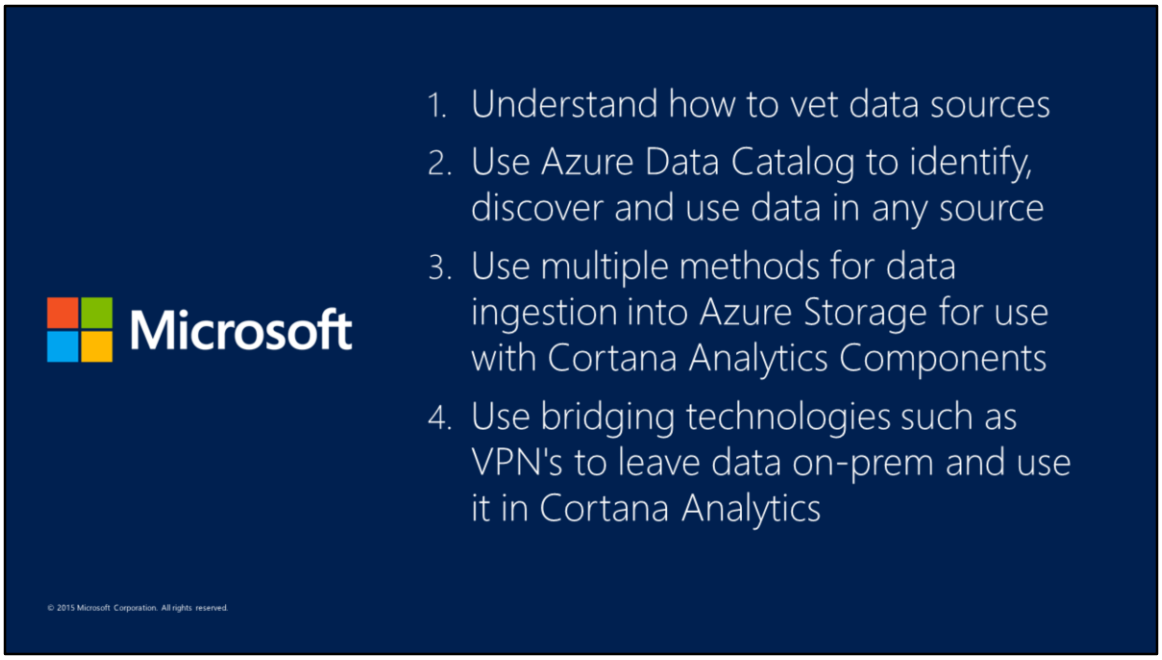
Connect on-prem to <anything>

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. Use this for Q/A time