The Cortana Intelligence Suite
Foundations – Workshop Recap
Microsoft Machine Learning and Data Science Team
CortanaIntelligence.com

1. Main page: http://cortanaanalytics.com
2. To begin this module, you should have:
    1. Basic Math and Stats skills
    2. Business and Domain Awareness
    3. General Computing Background

NOTE: These workbooks contain many resources to lead you through the course, and provide a rich set of references that you can use to learn much more about these topics. If the links do not resolve properly, type the link address in manually in your web browser. If the links have changed or been removed, simply enter the title of the link in a web search engine to find the new location or a corollary reference.
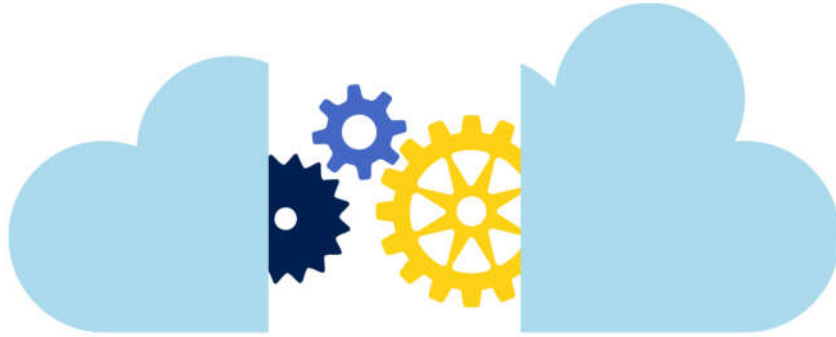
Section 7 Learning Objectives

1. Understand how to use CIS in Advanced Analytic Projects
2. Know where to find more information
3. Know how to extrapolate this process to other business cases
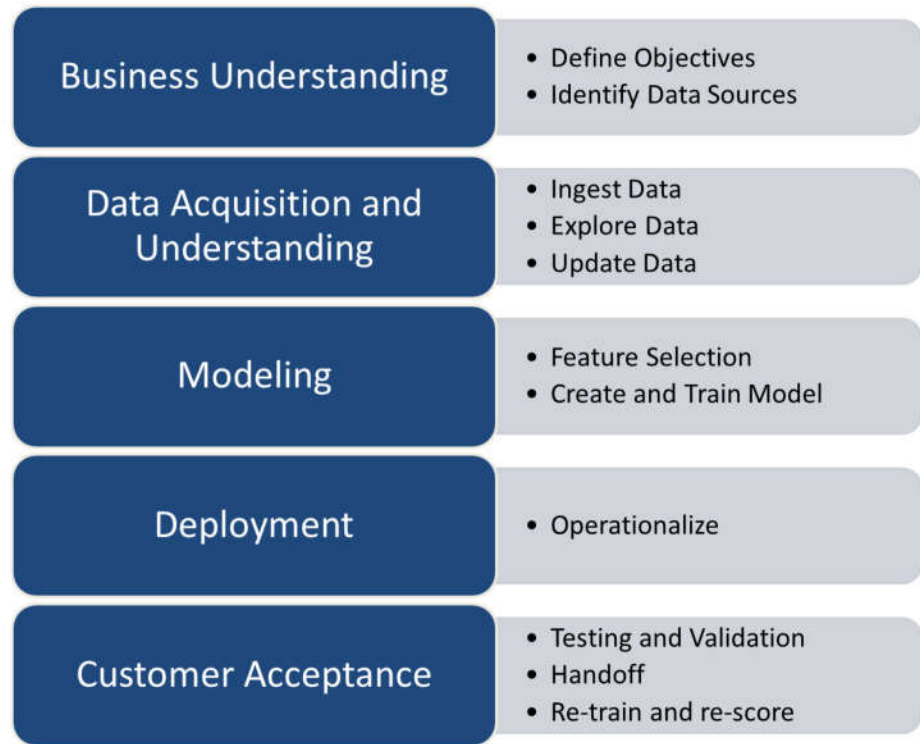
1. At the end of this Module, you will:
   1. Understand how to use CIS in Advanced Analytic Projects
   2. Know where to find more information
   3. Know how to extrapolate this process to other business cases

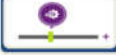# The Data Science Process and Platform

1. This process largely follows the CRISP-DM model: http://www.sv-europe.com/crisp-dm-methodology/
2. It also references the Cortana Intelligence process: https://azure.microsoft.com/en-us/documentation/articles/data-science-process-overview/
3. A complete process diagram is here: https://azure.microsoft.com/en-us/documentation/learning-paths/cortana-analytics-process/
4. Some walkthrough's of the various services: https://azure.microsoft.com/en-us/documentation/articles/data-science-process-walkthroughs/
5. An integrated process and toolset allows for a more close-to-intent deployment

6. Iterations are required to close in on the solution – but are harder tio management and monitor

The Cortana Intelligence Platform

	Cortana, Cognitive Services, Bot Framework
Power BI	Power BI
	Stream Analytics
	HDInsight
	Azure Machine Learning (MRS)
	SQL Data Warehouse (SQL DB, Document DB)
	Data Lake
	Event Hubs
	Data Factory
	Data Catalog
	Microsoft Azure

1. Platform and Storage: Microsoft Azure – http://microsoftazure.com Storage: https://azure.microsoft.com/en-us/documentation/services/storage/ (Host It)
2. Azure Data Catalog: http://azure.microsoft.com/en-us/services/data-catalog (Doc It)
3. Azure Data Factory: http://azure.microsoft.com/en-us/services/data-factory/ (Move 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/ , 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/ (Scale 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/ and https://developer.microsoft.com/en-us/Cortana (Say It)
12. Cognitive Services: https://www.microsoft.com/cognitive-services
13. Bot Framework: https://dev.botframework.com/
14. All of the components within the suite: https://www.microsoft.com/en-us/server-cloud/cortana-intelligence-suite/what-is-cortana-intelligence.aspx
15. What can I do with it? https://gallery.cortanaintelligence.com/

16. Getting Started Quickly: https://caqs.azure.net/#gallery

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 years, 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

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 years, 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.
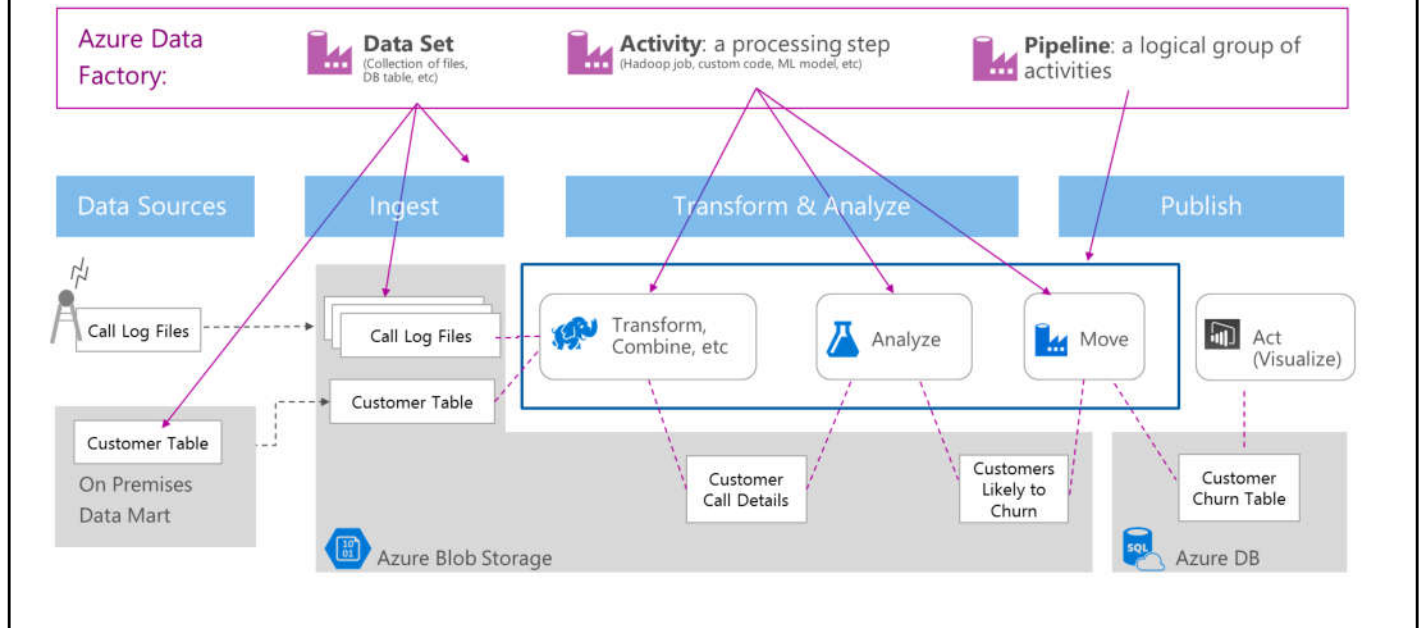
Ideally we want a list of customers we think will churn, in a structured database we could share out to our potential resellers sales staff, so they know how to target at-risk and new clients.
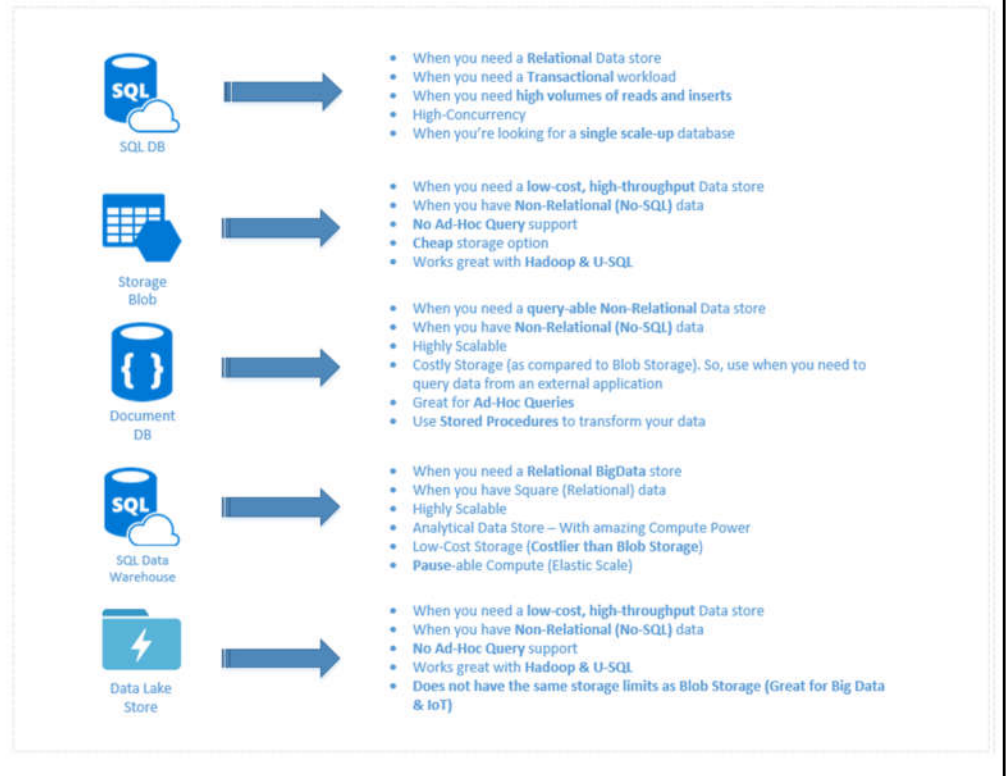
More on our in-house data: https://technet.microsoft.com/en-us/library/ms124501%28v=sql.100%29.aspx

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

1. Video of this process: https://azure.microsoft.com/en-us/documentation/videos/azure-data-factory-102-analyzing-complex-churn-models-with-azure-data-factory/

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

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

Group Project

1. Select one of the scenarios.
2. Create a Business Case, and highlight the "design statements" it contains.
3. Create a Decision Matrix of the options to solve the design statements.
4. Create a Solution Diagram.
5. Mail these to the instructor.
6. Be ready to discuss the reasons you chose each technology in your solution and how the data flow path will work. Be as detailed as you can. You can submit multiple documents or including them all in one – Word, PDF, PowerPoint or Visio are all acceptable tools.

- Select one of the following scenarios.
- Create a Business Case, and highlight the "design statements" it contains.
- Create a Decision Matrix of the options to solve the design statements.
- Create a Solution Diagram.
- Mail these documents to the instructor.
- Be ready to discuss the reasons you chose each technology in your solution and how the data flow path will work. Be as detailed as you can. You can submit multiple documents or including them all in one – Word, PDF, PowerPoint or Visio are all acceptable tools.

1. Using the Contoso Medical Scenario, find the Marketing Partners bullet. How could you use the CIS process and

platform to assist in determining optimal promotions? https://msdn.microsoft.com/en-us/library/ff650330.aspx

2. Contoso auto repairs wants to be able to predict the flow of repairs through their shops nation-wide, to properly staff, equip, set pricing and inventory for their operations. Which datasets do you think they need to collect to have this kind of predictive model? How would you go about documenting and ingesting this data? https://msdn.microsoft.com/en-us/library/ee861194.aspx

3. Contoso University wants to predict which students they should accept from high-schools around the country that will complete their degrees, and offer the best scores to the college. What processes can they use to determine these students, and which CIS platform elements could they use to create a yearly report for the Board of Admissions to give the most numerically accurate number? Describe your solution.

4. (You can also check these resources for more examples of what to turn in: https://blogs.msdn.microsoft.com/shishirs/2017/02/28/cortana-intelligence-suite-retail-industry-solution-how-to-guides/)

Questions?