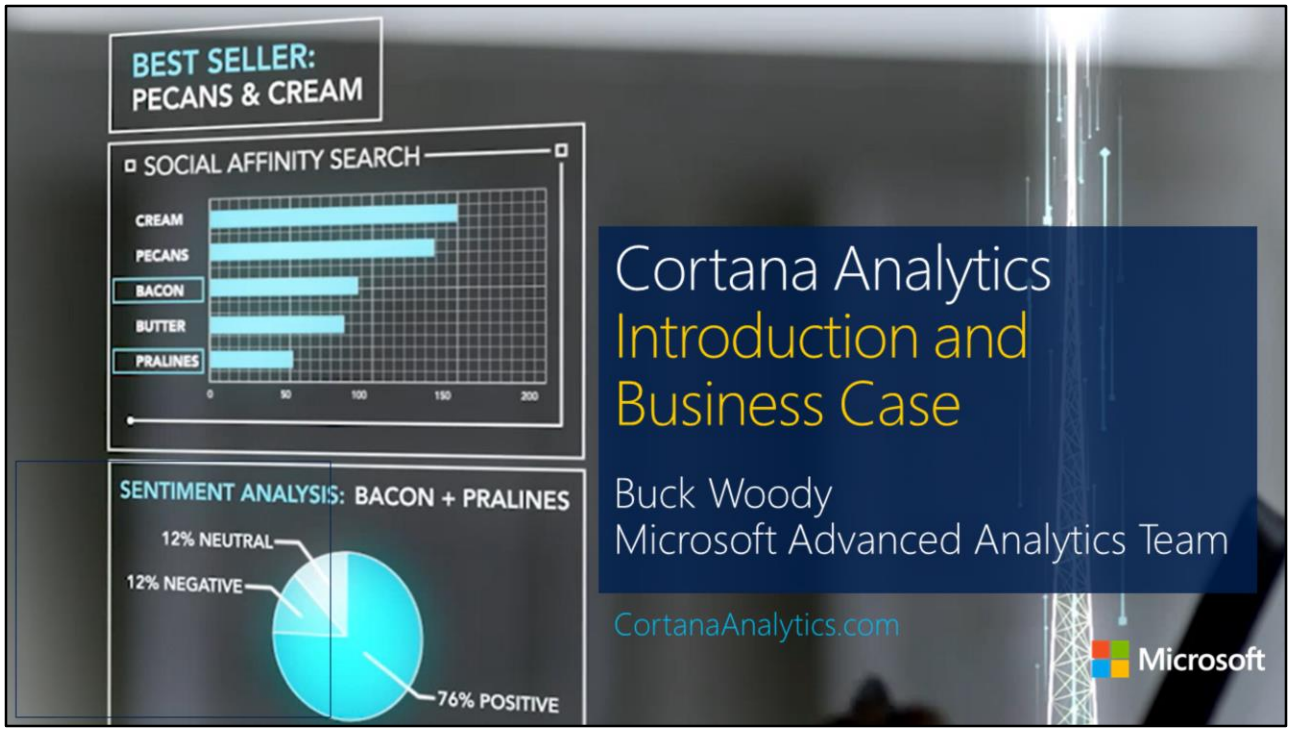
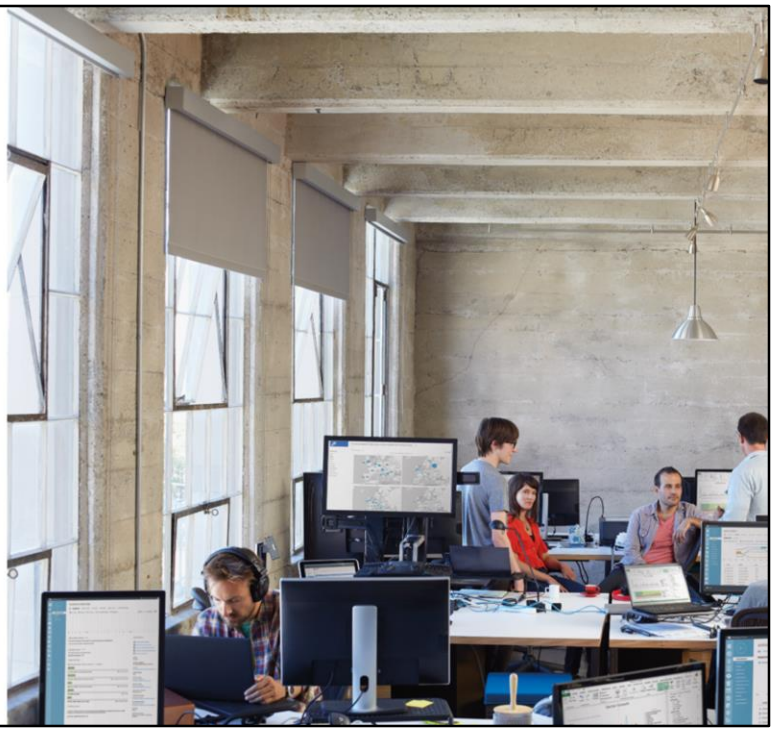


1. Main page: http://cortanaanalytics.com
2. To begin this module, you should have:
    1. Basic Math and Stats skills
    2. Business and Domain Awareness
    3. General Computing Background

Introduction

- Class hours
- Facilities
- Meals and Breaks
- Internet and Azure Access
- Installations
- Feedback



There will be a general feedback form, but after each module you are asked if you can complete the objectives. If you cannot, let the instructor know.

There are a few things you need prior to coming to class:

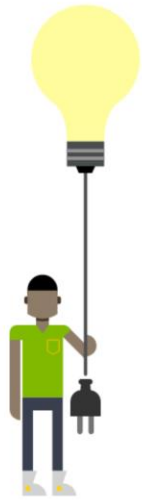
- A background in data technologies, such as working with Relational and Non-Relational data processing systems
- A general level of predictive and classification Statistics
- A general understanding of Machine Learning
- A subscription to Microsoft Azure (this may be provided through your company or as part of your invitation)
- A laptop with Visual Studio installed – the Community Edition (free) is acceptable – Version 2015 preferable (https://www.visualstudio.com/en-us/products/visual-studio-community-vs.aspx)
- Azure SDK and Command-line Tools installed (https://azure.microsoft.com/en-us/downloads/)
- Azure Storage Explorer (http://go.microsoft.com/fwlink/?linkid=698844&clcid=0x409)
- Power BI Desktop (https://powerbi.microsoft.com/en-us/desktop/)

Module 1 Learning Objectives

1. Have a general Workshop Process awareness
2. Understand and interpret problem a Statement for a given Customer Scenario
3. Understand the Data Science Process (at a general level)
4. Understand architecture development basics

1. At the end of this Module, you will:
    1. Have a general Workshop Process awareness
    2. Understand and interpret problem a Statement for a given Customer Scenario
    3. Understand the Data Science Process (at a general level)
    4. Understand architecture development basics

Data Science Overview



- Determining the actual problems to solve
- Identifying and vetting data sources
- Defining the data path
- Create Cleansing and Homogenizing Processes
- Create Feature Selection process (where needed)
- Create Computation Processes (Predictive or Classification)
- Create Output and Presentation Instruments

1. The Data Science Process: http://columbiadatascience.com/2012/09/24/reflections-after-jakes-lecture/

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

The AdventureWorks Scenarios: https://technet.microsoft.com/en-us/library/ms124501%28v=sql.100%29.aspx

Creating Architectures, High Level

- Decompose problem statement
- Develop Requirements
- Develop Constraints
- Create a Logical Solution Flow

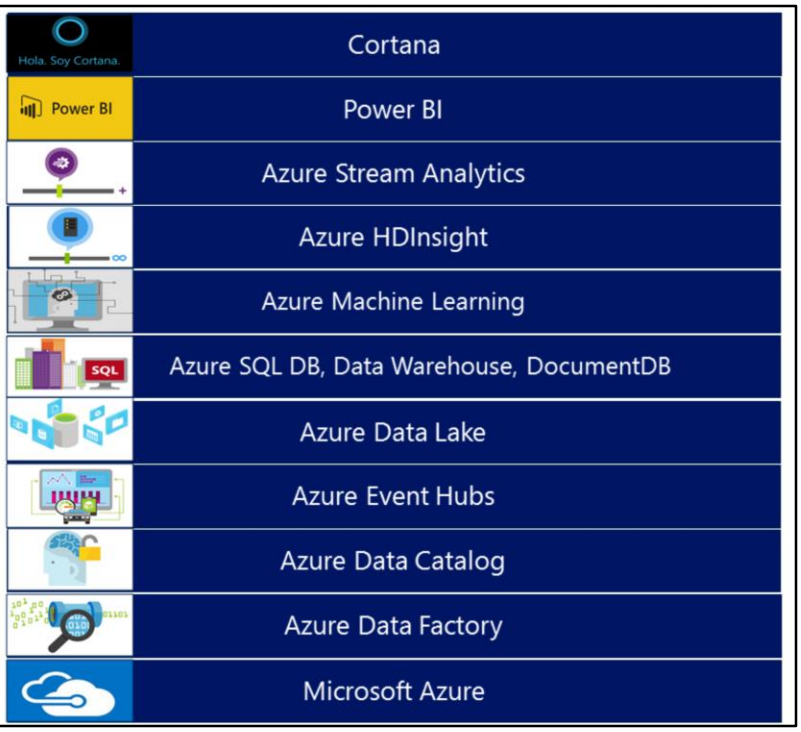
1. General Software Architecture Design Approach - https://msdn.microsoft.com/en-us/library/ee658084.aspx

Cortana Analytics

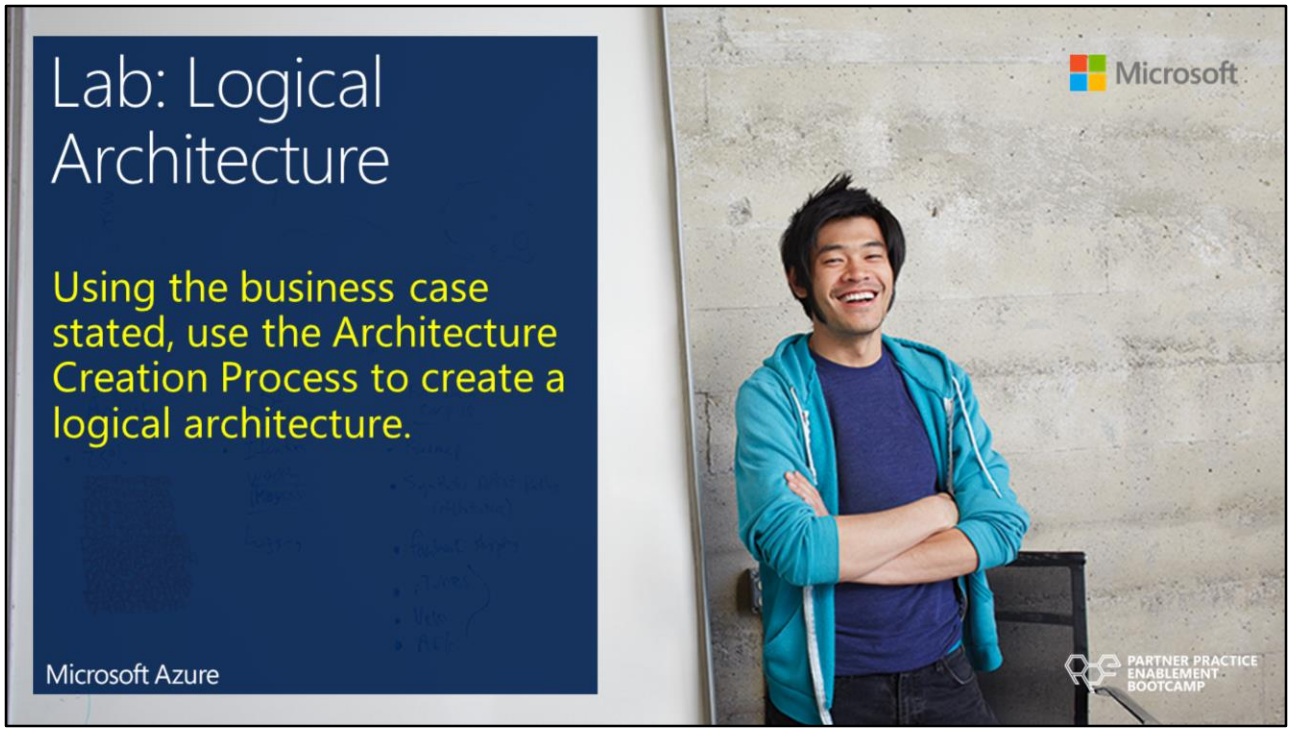
A Suite of Products that allow you to Predict Outcomes, Prescribe Actions and Automate Decisions

1. Suite Overview - https://www.microsoft.com/en-us/server-cloud/cortana-analytics-suite/overview.aspx

Cortana Analytics Stack



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)



1. Focus on the Logical architecture – do NOT include any technologies, just specify the data movement path and the logical operations that will be performed.