**Azure Quick Start Guide** (as of 9/16/16)

*Note: Azure’s new portal reached GA on December 4, 2015*

**Azure Data Catalog**

Overview: Fully managed service in the cloud that allows any user to register, enrich, discover, understand, and consume enterprise-wide data sources

Status: GA on March 30, 2016

Landing page: [www.azuredatacatalog.com](http://www.azuredatacatalog.com)

Documentation: <https://azure.microsoft.com/en-us/documentation/services/data-catalog/>

Getting started: <https://azure.microsoft.com/en-us/documentation/articles/data-catalog-get-started/>

**Azure Data Factory (ADF)**

Overview: Cloud-based data integration service that orchestrates and automates the movement and transformation of structured and unstructured data from nearly any source

Status: GA on Aug 6, 2015

Via new portal: New -> Data + Analytics -> Data Factory

Documentation: <https://azure.microsoft.com/en-us/documentation/services/data-factory/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/data-factory/>

Video overview: <https://channel9.msdn.com/blogs/Cloud-and-Enterprise-Premium/Introduction-to-Azure-Data-Factory>

**Azure Data Lake Analytics (ADLA)**

Overview: Cloud-based big data computation service that dynamically provisions resources so you can run queries on exabytes of data. Includes U-SQL, a new big data query language

Status: Public preview on October 30, 2015 (GA early Fall)

Via new portal: New -> Data + Analytics -> Data Lake Analytics (preview)

Documentation: <https://azure.microsoft.com/en-us/documentation/services/data-lake-analytics/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/data-lake-analytics-self-guided-training/>

Video overview: <https://channel9.msdn.com/blogs/Cloud-and-Enterprise-Premium/Azure-Data-Lake-Analytics-Deep-Dive>

**Azure Data Lake Store (ADLS)**

Overview: Hyperscale repository to store data of all sizes, formats, and speeds in the cloud

Status: Public preview on October 28, 2015 (GA early Fall)

Via new portal: New -> Data + Storage -> Data Lake Store (preview)

Documentation: <https://azure.microsoft.com/en-us/documentation/services/data-lake-store/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/data-lake-store-self-guided-training/>

Video series on Azure Data Lake: <https://channel9.msdn.com/Series/AzureDataLake>

**Azure Event Hub**

Overview: A managed service that can intake and process massive data streams from websites, apps, and devices

Status: GA on October 31, 2014

Via new portal: New -> Data + Analytics -> Event Hub

Documentation: <https://azure.microsoft.com/en-us/documentation/services/event-hubs/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/event-hubs/>

Video overview: [https://channel9.msdn.com/Shows/Cloud+Cover/Episode-160-Event-Hubs-with-Elio-Damaggio](https://channel9.msdn.com/Shows/Cloud%2BCover/Episode-160-Event-Hubs-with-Elio-Damaggio)

**Azure ML (Machine Learning)**

Overview: Cloud-based predictive analytics service to create and deploy predictive models as analytics solutions

Status: GA on Feb 18, 2015

Via new portal: New -> Data + Analytics -> Machine Learning (uses old portal)

Documentation: <https://azure.microsoft.com/en-us/documentation/services/machine-learning/>

Video series: <https://channel9.msdn.com/Blogs/DevRadio/Part-1-An-Introduction-to-Azure-Machine-Learning-and-Azure-ML-Studio>

**Azure Stream Analytics (ASA)**

Overview: An event data processing service providing real-time analytics and insights from apps, devices, sensors, and more

Status: GA on April 16, 2015

Via new portal: New -> Data + Analytics -> Stream Analytics job

Documentation: <https://azure.microsoft.com/en-us/documentation/services/stream-analytics/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/stream-analytics/>

Video overview: <https://channel9.msdn.com/blogs/Cloud-and-Enterprise-Premium/An-Introduction-to-Azure-Stream-Analytics>

**HDInsight**

Overview: Managed Apache Hadoop clusters in the cloud that uses the Hortonworks Data Platform (HDP) Hadoop distribution

Status: GA on Oct 28, 2013, HDInsight Premium (Microsoft R Server) public preview on March 29, 2016

Via new portal: New -> Data + Analytics -> HDInsight

Documentation: <https://azure.microsoft.com/en-us/documentation/services/hdinsight/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/hdinsight-self-guided-hadoop-training/>

Video series on HDInsight: <https://channel9.msdn.com/Series/Big-Data-Analytics-with-HDInsight>

**Power BI**

Overview: Business analytics service that enables anyone to visualize and analyze data

Status: Version 2.0 GA on July 24, 2015

Landing page: [www.powerbi.com](http://www.powerbi.com)

Documentation: <https://powerbi.microsoft.com/en-us/documentation/powerbi-service-get-started/>

Video Series: <https://www.youtube.com/user/mspowerbi/videos>

**SQL Database (SQL DB)**

Overview: Cloud-based relational database-as-a-service based on the SQL Server engine

Status: GA of v12 on February 9, 2015

Via new portal: New -> Data + Storage -> SQL Database

Documentation: <https://azure.microsoft.com/en-us/documentation/services/sql-database/>

Learning guide: <https://azure.microsoft.com/en-us/documentation/learning-paths/sql-database-jr2/>

Video tips and tricks: <https://channel9.msdn.com/events/SQL-Saturday/SQL-Saturday-Moscow-2015/Hall-A-Azure-SQL-Database-Tips-and-Tricks-for-Beginners>

**SQL Data Warehouse (SQL DW)**

Overview: Cloud-based data warehouse service that runs using a massive parallel processing (MPP) architecture

Status: GA on July 14, 2016

Via new portal: New -> Data + Storage -> SQL Data Warehouse (preview)

Documentation: <https://azure.microsoft.com/en-us/documentation/services/sql-data-warehouse/>

Getting started: <https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-get-started-provision/>

Video overview: <https://channel9.msdn.com/events/Ignite/Microsoft-Ignite-New-Zealand-2015/M242>

**Sample architecture:**



The details on the numbered steps in the above diagram (which is cloud solution, but it can access on-prem sources):

1) Copy source data into the Azure Data Lake Store (twitter data example)

2) Massage/filter the data using Hadoop (or skip using Hadoop and use stored procedures in SQL DW/DB to massage data after step #5)

3) Pass data into Azure ML to build models using Hive query (or pass in directly from Azure Data Lake Store)

4) Azure ML feeds prediction results into the data warehouse

5) Non-relational data in Azure Data Lake Store copied to data warehouse in relational format (optionally use PolyBase with external tables to avoid copying data)

6) Power BI pulls data from data warehouse to build dashboards and reports

7) Azure Data Catalog captures metadata from Azure Data Lake Store, SQL DW/DB, and SSAS cubes

8) Power BI and Excel can pull data from the Azure Data Lake Store via HDInsight

9) To support high concurrency if using SQL DW, or for easier end-user data layer, create an SSAS cube