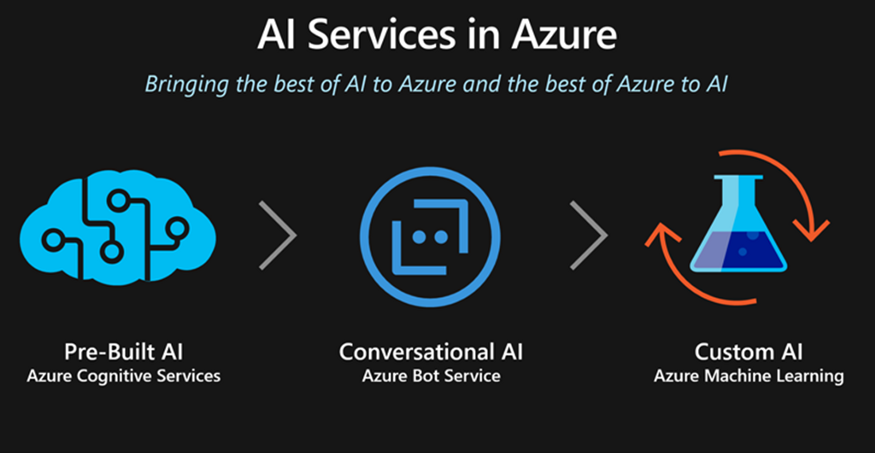
# Lesson 0

Intro to AI

https://medium.com/@caiomsouza/how-to-prepare-for-microsoft-azure-certification-ai-100-designing-and-implementing-an-azure-ai-af60b92b8068



# Lesson 0

Big Data I.S.P.M.

<https://azure.microsoft.com/en-us/solutions/big-data/>

Real-time analytics on big data:

Graphical user interface

Description automatically generated

Advanced analytics on big data:

Diagram

Description automatically generated

# Lesson 1

Storage Options

**Azure Table Storage**

Table Storage is a NoSQL key-value store for rapid development using massive semi-structured datasets. You can develop applications on Cosmos DB using popular NoSQL APIs

Azure Cosmos DB Tables aims for:

**high throughput** (single-digit millisecond latency),

**global distribution** (multiple failover),

SLA-backed predictive performance with automatic indexing of each attribute/property and a pricing model focused on throughput.

Azure Storage Tables is aimed at high capacity on a single region (optional secondary read only region but no failover), indexing by PK/RK and storage optimized pricing;

**Azure Data Lake Storage Gen1**

**Azure File Storage**

**Azure Blob**

Managed storage service that is highly available, secure, durable, scalable, and redundant. Microsoft takes care of maintenance and handles critical problems for you. Flexible pricing options (cold vs hot storage)

[GE] **Example:** You are designing an AI solution that will analyze millions of pictures by using Azure HDInsight Hadoop cluster. You need to recommend a solution for storing the pictures. The solution must minimize costs.

**Azure Blob Storage**

[GE] **Example:** Your company recently deployed several hardware devices that contain sensors. The sensors generate new data on an hourly basis. The data generated is stored on-premises and retained for several years. During the past two months, the sensors generated 300 GB of data. You plan to move the data to Azure and then perform advanced analytics on the data.

**Azure Blob Storage**

[GE] **Example:** You are configuring data persistence for a Microsoft Bot Framework application. The application requires a structured NoSQL cloud data store. You need to identify a storage solution for the application. The solution must minimize costs.

**Azure Table Storage**

[GE] **Example:** You are designing an AI workflow that will aggregate data stored in Azure as JSON documents. You expect to store more than 2 TB of new data daily. You need to choose the data storage service for the data. The solution must minimize costs. Which data storage service should you choose?

**Azure Blob Storage**

**Lesson 2**

Data Workflows

With Azure Data Factory you can use workflows to orchestrate data integration and data transformation processes at scale.

Build data integration, and easily transform and integrate big data processing and machine learning with the visual interface.

Row-level security is supported by SQL Server, Azure SQL Database, and Azure SQL Data Warehouse.

[GE] **Example:** You are designing an AI application that will use an azure Machine Learning Studio experiment. The source data contains more than 200 TB of relational tables. The experiment will run once a month. You need to identify a data storage solution for the application. The solution must minimize compute costs. Which data storage solution should you identify?

**Azure SQL Database**

[GE] **Example:** You plan to deploy two AI applications named AI1 and AI2. The data for the applications will be stored in a relational database.

You need to ensure that the users of AI1 and AI2 can see only data in each user’s respective geographic region. The solution must be enforced at the database level by using row-level security.

Which database solution should you use to store the application data?

**Microsoft SQL Server on a Microsoft Azure virtual machine**

[GE] **Example:** You design an AI workflow that combines data from multiple data sources for analysis. The data sources are composed of:

JSON files uploaded to an Azure Storage account

On-premises Oracle databases

Azure SQL databases

Which service should you use to ingest the data?

**Azure Data Factory**

[GE] Example: You are developing a Computer Vision application.

You plan to use a workflow that will load data from an on-premises database to Azure Blob storage, and then connect to an Azure Machine Learning service.

What should you use to orchestrate the workflow?

**Azure Data Factory**

**Lesson 3**

Data Tools

A. Microsoft Visual Studio Code

B. Azure Notebooks

C. Azure Machine Learning Studio

D. Microsoft Visual Studio

[GE] **Example:** You plan to design an application that will use data from Azure Data Lake and perform sentiment analysis by using Azure Machine Learning algorithms. The developers of the application use a mix of Windows- and Linux-based environments. The developers contribute to shared GitHub repositories. You need all the developers to use the same tool to develop the application. What is the best tool to use? More than one answer choice may achieve the goal.

**Azure Machine Learning Studio**

[GE] **Example:** Your company is building custom models that integrate into microservices architecture on Azure Kubernetes Services (AKS).

The model is built by using Python and published to AKS.

You need to update the model and enable Azure Application Insights for the model. What should you use?

Azure Machine Learning SDK

**Lesson 2**

Azure Kubernetes Service

Azure Kubernetes Service (AKS) is a managed Kubernetes offering that further simplifies container-based application deployment and management.

The kubectl command line tool lets you control Kubernetes clusters. For configuration, kubectl looks for a file named config in the $HOME/.kube directory.

**horizontal pod autoscaler**

automatically adjust the number of pods that run your application.

B. cluster autoscaler

C. AKS cluster virtual 32 node autoscaling

D. Azure Container Instances

[GE] **Example:** You have several AI applications that use an Azure Kubernetes Service (AKS) cluster. The cluster supports a maximum of 32 nodes. You discover that occasionally and unpredictably, the application requires more than 32 nodes. You need to recommend a solution to handle the unpredictable application load. Which scaling methods should you recommend? (Choose two.)

**Horizontal pod autoscaler**

**Cluster autoscaler**

[GE] **Example:** You have a solution that runs on a five-node Azure Kubernetes Service (AKS) cluster. The cluster uses an N-series virtual machine. An Azure Batch AI process runs once a day and rarely on demand. You need to recommend a solution to maintain the cluster configuration when the cluster is not in use. The solution must not incur any compute costs. What should you include in the recommendation?

**Downscale the cluster to one node**

[GE] **Example:** You are developing an application that uses an Azure Kubernetes Service (AKS) cluster. You are troubleshooting a node issue. You need to connect to an AKS node by using SSH.

Solution: You run the kubectl command, and then you create an SSH connection.

Does this meet the goal?

**Yes**

You are developing an application that uses an Azure Kubernetes Service (AKS) cluster. You are troubleshooting a node issue. You need to connect to an AKS node by using SSH.

Solution: You change the permissions of the AKS resource group, and then you create an SSH connection.

Does this meet the goal?

**No** (Instead add an SSH key to the node, and then you create an SSH connection.)

You are developing an application that uses an Azure Kubernetes Service (AKS) cluster. You are troubleshooting a node issue. You need to connect to an AKS node by using SSH.

Solution: You add an SSH key to the node, and then you create an SSH connection.

Does this meet the goal?

**Yes**

**Lesson 3**

Compute Infrastructure

A. graphics processing units (GPUs)

Field-Programmable Gate Arrays (FPGAs)

Central Processing Units (CPUs)

Application-specific integrated circuits (ASICs)

ASICs are custom circuits, such as Google's TensorFlow Processor Units (TPU), provide the highest efficiency. They can't be reconfigured as your needs

change.

[GE] **Example:** You are designing an AI solution in Azure that will perform image classification. You need to identify which processing platform will provide you with the ability to update the logic over time. The solution must have the lowest latency for inferencing without having to batch.

**field-programmable gate arrays (FPGAs)**

**Lesson 4**

Big Data Compute

<https://docs.microsoft.com/en-us/machine-learning-server/r/concept-what-is-compute-context>

The primary reason for shifting compute context is to eliminate data transfer over your network, bringing computations to where the data resides. This is particularly relevant for big data platforms like Hadoop, where data is distributed over multiple nodes, or for data sets that are simply too large for a client workstation.

A. SQL

B. Spark

C. local parallel

D. HBase

E. local sequential

[GE] **Example:**You deploy an infrastructure for a big data workload. You need to run Azure HDInsight and Microsoft Machine Learning Server. You plan to set the RevoScaleR compute contexts to run rx function calls in parallel. What are three compute contexts that you can use for Machine Learning Server? Each correct answer presents a complete solution.

**SQL**

**Spark**

**local parallel**

**Lesson 5**

Real-time Analysis

**Azure HDInsight Storm cluster**

**Azure Notification Hubs**

**Azure HDInsight Hadoop cluster**

Azure HDInsight makes it easy, fast, and cost-effective to process massive amounts of data. You can use HDInsight to process streaming data that's received in real time from a variety of devices.

**Azure HDInsight R cluster**

[GE] **Example:** Your company has recently deployed 5,000 Internet-connected sensors for a planned AI solution. You need to recommend a computing solution to perform a real-time analysis of the data generated by the sensors. Which computing solution should you recommend?

**Azure HDInsight Hadoop cluster**

[GE] **Example:** You deploy an application that performs sentiment analysis on the data stored in Azure Cosmos DB. Recently, you loaded a large amount of data to the database. The data was for a customer named Contoso, Ltd. You discover that queries for the Contoso data are slow to complete, and the queries slow the entire application. You need to reduce the amount of time it takes for the queries to complete. The solution must minimize costs. What is the best way to achieve the goal? More than one answer choice may achieve the goal. Select the BEST answer.

**Change the partitioning strategy**

Throughput provisioned for a container is divided evenly among physical partitions. Increasing request units would also improve throughput, but at a cost.

[GE] **Example:** You plan to implement a new data warehouse for a planned AI solution. You have the following information regarding the data warehouse:

The data files will be available in one week.

Most queries that will be executed against the data warehouse will be ad-hoc queries.

The schemas of data files that will be loaded to the data warehouse will change often.

One month after the planned implementation, the data warehouse will contain 15 TB of data.

You need to recommend a database solution to support the planned implementation.

What two solutions should you include in the recommendation? Each correct answer is a complete solution.

**Apache Hadoop**

**Apache Spark**

[GE] **Example:** Your company has a data team of Scala and R experts.

You plan to ingest data from multiple Apache Kafka streams.

You need to recommend a processing technology to broker messages at scale from Kafka streams to Azure Storage.

What should you recommend?

**Azure HDInsight with Apache Storm**

**Lesson 6**

Azure

When an application needs access to deploy or configure resources through Azure Resource Manager in Azure Stack, you create a service principal, which is a

credential for your application. You can then delegate only the necessary permissions to that service principal.

**[**GE**] Example:** Your company has 1,000 AI developers who are responsible for provisioning environments in Azure. You need to control the type, size, and location of the resources that the developers can provision. What should you use?

**Azure service principals**

[GE] **Example:** You need to build a solution to monitor Twitter. The solution must meet the following requirements:

Send an email message to the marketing department when negative Twitter messages are detected.

Run sentiment analysis on Twitter messages that mention specific tags.

Use the least amount of custom code possible.

Which two services should you include in the solution? Each correct answer presents part of the solution.

**Azure Stream Analytics**

**Azure Logic Apps**

[GE] **Example:** You plan to build an application that will perform predictive analytics. Users will be able to consume the application data by using Microsoft Power BI or a custom website. You need to ensure that you can audit application usage. Which auditing solution should you use?

**Azure Active Directory (Azure AD) reporting**

[GE] **Example:** You are developing a mobile application that will perform optical character recognition (OCR) from photos.

The application will annotate the photos by using metadata, store the photos in Azure Blob storage, and then score the photos by using an Azure Machine Learning model. What should you use to process the data?

**Azure Functions**

[GE] **Example:** You create an Azure Cognitive Services resource.

A data scientist needs to call the resource from Azure Logic Apps by using the generic HTTP connector.

Which two values should you provide to the data scientist? Each correct answer presents part of the solution.

**Resource group name**

**Subscription ID**

[GE] **Example:** You plan to deploy an AI solution that tracks the behavior of 10 custom mobile apps. Each mobile app has several thousand users.

You need to recommend a solution for real-time data ingestion for the data originating from the mobile app users.

Which Microsoft Azure service should you include in the recommendation?

**Azure Event Hubs**

**Lesson 7**

Data Encryption

A. Generate a shared access signature (SAS)

B. Obtain an SSL certificate

C. Add a deployment slot

D. Update the web service

E. Update DNS

F. Create an Azure Key Vault

**[**GE**] Example:** You have an Azure Machine Learning model that is deployed to a web service. You plan to publish the web service by using the name ml.contoso.com. You need to recommend a solution to ensure that access to the web service is encrypted. Which three actions should you recommend? Each correct answer presents part of the solution.

**Obtain an SSL certificate**

**Update the web service**

**Update DNS**

**Lesson 7**

Data privacy, protection, and regulatory requirements

Azure Bot service is compliant with

ISO 27001:2013,

ISO 27019:2014,

SOC 1 and 2,

(PCI DSS) Payment Card Industry Data Security Standard

(HIPAA BAA) Health Insurance Portability and Accountability Act Business Associate Agreement

Compliance Manager for Azure helps you assess and manage GDPR compliance. Compliance Manager is a free, Microsoft cloud services help organizations meet complex compliance obligations, including the GDPR, ISO 27001, ISO 27018, and NIST 800-53. Generally available today for Azure customers, the Compliance Manager GDPR dashboard enables you to assign, track, and record your GDPR compliance activities so you can collaborate across teams and manage your documents for creating audit reports more easily.

**[**GE**] Example:** Your plan to design a bot that will be hosted by using Azure Bot Service. Your company identifies the following compliance requirements for the bot:

Payment Card Industry Data Security Standards (PCI DSS)

General Data Protection Regulation (GDPR)

ISO 27001

You need to identify which compliance requirements are met by hosting the bot in the bot service. What should you identify?

**PCI DSS, ISO 27001, and GDPR**

**[**GE**] Example:** You have an Azure Machine Learning experiment. You need to validate that the experiment meets GDPR regulation requirements and stores documentation about the experiment. What should you use?

A. Compliance Manager (for Azure)

**Lesson 8**

IOT

There is support for running Azure Cognitive Services containers for Text Analytics and Language Understanding containers on edge devices with Azure IoT Edge.

This means that all your workloads can be run locally where your data is being generated while keeping the simplicity of the cloud to manage them remotely,

securely and at scale.

You would have to set up an IoT Edge device and its IoT Hub.

[GE] **Example:** You need to create an IoT solution that performs the following tasks:

Identifies hazards

Provides a real-time online dashboard

Takes images of an area every minute

Counts the number of people in an area every minute

Solution 1: You implement Azure Cognitive Services containers on the IoT devices, and then you configure an Azure Automation call to Azure Cognitive Services that sends the results to an Azure event hub. You configure Microsoft Power BI to connect to the event hub by using Azure Stream Analytics.

Does this meet the goal?

**Yes**

Solution2 : You configure the IoT devices to send the images to an Azure IoT hub, and then you configure an Azure Functions call to Azure Cognitive Services that sends the results to an Azure event hub. You configure Microsoft Power BI to connect to the event hub by using Azure Stream Analytics.

**No**

Solution 3: You configure the IoT devices to send the images to an Azure IoT hub, and then you configure an Azure Automation call to Azure Cognitive Services that sends the results to an Azure event hub. You configure Microsoft Power BI to connect to the event hub by using Azure Stream Analytics.