# **DP200 - Implementing a Data Platform Solution**

## Lab 1 - Azure for the Data Engineer

### Exercise 2: Determine the Azure Data Platform services to use for AdventureWorks.

Use the table below to document the data requirements and data platform technology as identified from the AdventureWorks case study.

|  |  |  |
| --- | --- | --- |
| # | Data requirement | Technology |
| 1 | Adventure works website : data store is made available that will hold the images of the products that are sold on the website. | Blob storage for images comaptible with shared access signature to share the data with applications and allow activity, eg read only. Expiration date policy. |
| 2 | historical reporting and descriptive analytics. In recent times, that server has been struggling to process the reporting data in a timely manner, as a result the organization has evaluated the data warehouse capabilities of Azure Synapse Analytics and want to migrate their on-premises data to this platform. Your team should ensure that access to the data is restricted.  In addition, AdventureWorks would like to take their data analytics further and start to utilize predictive analytics capabilities. This is currently not an activity that is undertaken. The organization understands that a recommendation or a text analytics engine could be built and would like you to direct them on what would be the best technology and approach to take in implementing such a solution that is also resilient and performant.  You are also assessing the tooling that can help with the extraction, load and transforming of data into the data warehouse, and have asked a Data Engineer within your team to show a proof of concept of Azure Data Factory to explore the transformation capabilities of the product | Azure synapse analytics |
| 3 | AI Engineers, but they have requested a platform is provided by the Data Engineer that enables them to store conversation history. | cosmosDB/blob storage (ADLS) |
| 4 | Social media analysis |  |
| 5 | Furthermore, daily summary data can be saved to flat files that include Bicycle model, serial number, registered owner and a summary of the total miles cycled per day and the average speed. (connected bycicle) | Blob storage but will also be connecting with other services eg for app authentication so databricks can use a service principle to connect with data lake storage (ADLS) |

we are identifying the requirements and see how they can be implemented.