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

## Lab 1 - Azure for the Data Engineer

### Exercise 3: Identify the tasks to be performed by the Data Engineer

Pick one of AdventureWorks requirements that you have identified in the last exercise and document it in the box below

|  |
| --- |
| Connected Bicycle |

List out the high-level tasks that you *think* you will conduct as a data engineer to complete the requirement you have selected in the table below.

|  |  |
| --- | --- |
| Task | High level task description |
| 1 | Connect the microcomputer equipped in the Bike in Azure IoT Hub |
| 2 | Connect the microcomputer to a Cosmos DB Database in Azure |
| 3 | Connect Azure Databricks to the Cosmos DB Database |
| 4 | Create a Storage Account to store the flat files containing a summary of the daily data generated by each bicycle |
| 5 | Develop an automated pipeline in Azure Data Factory where you would get data from the Storage Account and generate a daily report to the Storage Account |
| 6 | Develop an automated pipeline in Azure Data Factory where you would train a predictive model using Azure Databricks for the predictive maintenance service |
| 7 | Secure the connection to the Cosmos DB Database with Key Vault/IAM (between the microcomputer and the Database and the Database and ADF) |
| 8 | Enable access and the necessary authorizations using RBAC to each developer/user and for the applications as well |