

# DP 200 - Implementing a Data Platform Solution

## Lab 1 - Azure for the Data Engineer

**Estimated Time:** 60 minutes

**Pre-requisites:** It is assumed that the case study for this lab has already been read.

**Lab files:** The files for this lab are located in the *Allfiles\Labfiles\Starter\DP-200.1* folder.

### Lab overview

The students will take the information gained in the lessons and from the case study to scope out the deliverables for a digital transformation project within AdventureWorks. They will first identify how the evolving use of data has presented new opportunities for the organization. The students will also explore which Azure Data Platform services can be used to address the business needs and define the tasks that will be performed by the data engineer. Finally, students will finalize the data engineering deliverables for AdventureWorks.

### Lab objectives

After completing this lab, you will be able to:

1. Identify the evolving world of data within AdventureWorks.
2. Determine the Azure Data Platform services to use for AdventureWorks.
3. Identify the tasks to be performed by the Data Engineer.
4. Finalize the data engineering deliverables for AdventureWorks.

### Scenario

You have been hired as a Senior Data Engineer to advise the management and employees of AdventureWorks on a digital transformation project. AdventureWorks has been selling bicycles and bicycle parts directly to end-consumer and distributors for over a decade. In the last few years, they have observed how different industries have been taking advantage of recent developments in cloud technology to offer customers a more personalized and engaging service. They want to lead the way in their industry.

You have been asked to work with the IT department to perform a discovery workshop with the organization to identify which data platform technologies can be used to the benefit of both AdventureWorks and their customers. They are specifically wanting to make sure that the technology brings the business closer to their customers on a variety of levels. This can include making personalized offers when making purchases or using customer services, to offering telemetry information on the bikes that they use. There is also a requirement to manage an existing reporting system that is held in a data warehouse.

The first step is to assess the current state of the cloud technologies and make sure that they can perform the project by performing the following analysis:

1. Identified the evolving world of data within AdventureWorks
2. Determined the Azure Data Platform services to use for AdventureWorks
3. Identified the tasks to be performed by the Data Engineer
4. Finalized the data engineering deliverables for AdventureWorks

### Exercise 1: Identify the evolving world of data within AdventureWorks.

**Estimated Time:** 15 minutes

**Group exercise**

The main task for this exercise is as follows:

1. From the case study, identify the data requirements of AdventureWorks and identify if the data structure for the requirement is structured, semi-structured or unstructured.
2. The instructor will discuss the findings with the group.

### Task 1: Identify the data requirements and structures of AdventureWorks.

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab01-Ex01.docx** from the **Allfiles\Labfiles\Starter\DP-200.1** folder.
2. As a group, spend **10 minutes** discussing and listing the data requirements and data structure that your group has identified within the case study document.

### Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

**Result:** After you completed this exercise, you have created a Microsoft Word document that shows a table of data requirements and the associated data structures that meets the requirement.

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

---

Estimated Time: 15 minutes

Group exercise

The main task for this exercise are as follows:

1. Determine the data platform technology that delivers the identified data requirements.
2. The instructor will discuss the findings with the group.

### Task 1: Determine the data platform technology that delivers the identified data requirements

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab01-Ex02.docx** from the **Allfiles\Labfiles\Starter\DP-200.1** folder.
2. As a group, spend **10 minutes** discussing and listing the data platform technologies against a given data requirement that your group has identified within the case study document.

### Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

**Result:** After you completed this exercise, you have created a Microsoft Word document that shows a table of data requirements and the associated data platform technology that meets the requirement.

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

---

Estimated Time: 15 minutes

Individual exercise

The main task for this exercise are as follows:

1. Select one of the requirements and determine the high-level tasks that will perform to meet the data requirement selected.
2. The instructor will discuss the findings with the group.

### Task 1: Determine the high-level tasks that you think you will perform to meet the data requirement.

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab01-Ex03.docx** from the **Allfiles\Labfiles\Starter\DP-200.1** folder.

2. Individually, spend **10 minutes** listing the high-level tasks that you think you will perform to meet the data requirement you have selected.

## Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

**Result:** After you completed this exercise, you have created a Microsoft Word document that shows a data requirements and the high-level tasks that you will perform to meet the data requirement you have selected.

## Exercise 4: Finalize the data engineering deliverables for AdventureWorks.

---

Estimated Time: 15 minutes

Individual exercise

The main task for this exercise are as follows:

1. Finalize the data engineering deliverables for AdventureWorks
2. The instructor will discuss the findings with the group.

### Task 1: Finalize the data engineering deliverables for AdventureWorks

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab01-Ex04.docx** from the **Allfiles\Labfiles\Starter\DP-200.1** folder.
2. Individually, spend **10 minutes** listing the data engineering deliverables for AdventureWorks as reflected by the next five modules of this course.

### Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

**Result:** After you completed this exercise, you have created a Microsoft Word document that shows the data engineering deliverables for AdventureWorks as reflected by the next five modules of this course.