



# Data Technician

Name:

Course Date:

## Table of contents

Day 1: Task 1 .....	3
Day 1: Task 2 .....	6
Day 1: Task 3 .....	8
Day 1: Task 4 .....	9
Day 2: Task 1 .....	11
Day 3: Task 1 .....	14
Day 3: Task 2 .....	16
Day 3: Task 3 .....	19
Day 4: Task 1 .....	21
Day 4: Task 2 .....	24
Day 4: Task 2 (Optional) .....	25
1. Scenario Background.....	25
2. Data Laws and Regulations.....	26
3. Azure Service Recommendations .....	26
4. Data Types and Data Modelling .....	26
5. Data Storage Formats and Structures in Azure.....	26
6. Additional Considerations.....	27
Submission Guidelines: .....	27
Course Notes.....	29
Additional Information.....	30



## Day 1: Task 1

Please research and complete the below questions relating to key concepts of cloud.

Be prepared to discuss the below in the group following this task.

### What can cloud computing do for us in the real-world?

1. Cloud computing is the on-demand availability of computing resources (such as storage and infrastructure), as services over the internet. It eliminates the need for individuals and businesses to self-manage physical resources themselves and only pay for what they use.
2. Enables Remote Work & Collaboration- Teams can work together from anywhere in the world.  
Examples Google Workspace, Microsoft 365.
3. Cloud services allow dynamic scaling up/down depending on demand. Automatically handle more traffic or data without crashing.
4. Cloud providers offer strong security tools, encryption, and backups. Data recovery is easier and faster.
5. Cloud platforms offer powerful tools for processing and analysing massive data sets.
6. Developers can quickly launch new products and test ideas without complex setups. Access to cutting-edge services like AI, machine learning.
7. Cloud systems offer high availability across multiple regions and failover capabilities.
8. Shared cloud infrastructure is more energy-efficient than each company running its own servers.

Cloud computing enables **flexibility, efficiency, and innovation** across nearly every industry — from healthcare and finance to education and entertainment.

### How can it benefit a business?



- 1. Cost Savings- No need for physical infrastructure** (servers, storage, etc.) Reduces costs on maintenance, electricity, and IT staff. **Pay-as-you-go** pricing means you only pay for what you use
- 2. Scalability-** Easily **scale up or down** depending on demand. Useful for seasonal spikes (e.g. holidays, product launches)
- 3. Remote Access & Collaboration-** Employees can **work from anywhere** with internet access. Teams collaborate in real time using cloud tools like Google Workspace or Microsoft
- 4. Security & Backup-** Cloud providers offer **data encryption, firewalls, and regular backups.** Improved disaster recovery options, Faster Time to Market, Quick deployment of apps or services. Access to tools like **AI, machine learning, and data analytics** without heavy setup
- 5. Business Continuity-** Cloud ensures **uptime and reliability**, Reduces downtime due to localized outages
- 6. Environmentally Friendly-** Shared infrastructure reduces energy use, Supports a business's sustainability goals, The **cloud** is the central hub, Devices and users **connect to the cloud** to access data, services, and tools, It symbolizes **mobility, flexibility, and centralization of resources**

### What's the alternative to cloud computing?

On-premises computing refers to running all software, data, and infrastructure **locally within a business's own physical facilities.** Everything — servers, storage, and networking — is managed and maintained by the organization itself.

### What cloud providers can we use, what are their features and functions?

#### 1. Amazon Web Services (AWS)- Most popular & mature cloud platform

##### Key Features:

- **EC2** – Virtual server hosting
- **S3** – Scalable object storage
- **Lambda** – Serverless computing
- **RDS & DynamoDB** – Databases



- **CloudFront** – Content delivery (CDN)
- **SageMaker** – Machine learning tools

**Best for:** Scalable applications, global reach, enterprise workloads

---

## 2. Microsoft Azure- Strong enterprise integration, especially with Microsoft tools

**Key Features:**

- **Azure Virtual Machines** – Compute resources
- **Azure Blob Storage** – Object storage
- **Azure Active Directory** – Identity & access
- **Power BI & Synapse** – Data analytics
- **Azure DevOps** – CI/CD pipelines

**Best for:** Businesses using Microsoft 365, .NET apps, and hybrid cloud setups

---

## 3. Google Cloud Platform (GCP)- Known for AI, data, and open-source friendliness

**Key Features:**

- **Compute Engine** – Virtual machines
- **Cloud Storage** – Object storage
- **BigQuery** – Serverless data warehouse
- **Vertex AI** – AI & machine learning
- **Firebase** – App backend for mobile/web

**Best for:** Data-heavy applications, startups, and ML-driven projects

---

## 4. IBM Cloud- Focus on AI (Watson), legacy system integration

**Key Features:**

- **Watson AI** – NLP and ML services
- **IBM Cloud Foundry** – PaaS services
- **Bare Metal Servers** – Physical servers
- **Blockchain as a Service**

**Best for:** Regulated industries, AI-driven use cases, hybrid cloud

---

## 5. Oracle Cloud- Best for databases and enterprise ERP systems

**Key Features:**

- **Oracle Autonomous Database**



- **OCI (Oracle Cloud Infrastructure)** – IaaS
- **Fusion Cloud ERP**

**Best for:** Businesses using Oracle software and databases

## 6. Alibaba Cloud- Leading cloud provider in Asia

**Key Features:**

- **Elastic Compute Service**
- **Object Storage Service**
- **AI and IoT tools**
- **Global CDN and security**

**Best for:** Asian markets, international e-commerce, scalability

## Day 1: Task 2

Please research the below cloud offerings, explain what they are and examples of use cases.

Cloud Offerings	Explain what it is	When / how might you use this service in the real-world?
<b>IaaS (Infrastructure as a service)</b>	<b>IaaS</b> is a <b>cloud computing</b> service model that provides <b>virtualized computing resources</b> over the internet. Instead of buying and maintaining physical servers, storage, and networking equipment, businesses can <b>rent IT infrastructure</b> on-demand from a cloud provider.	A growing startup wants to launch a web app. Instead of buying servers, they use <b>AWS EC2</b> (Elastic Compute Cloud) to host it. They can scale up the infrastructure automatically when user traffic increases.
<b>PaaS (Platform as a service)</b>	<b>PaaS</b> is a cloud computing model that provides a <b>ready-to-use platform for developing, testing, and deploying applications</b> — without having to manage the underlying infrastructure (servers, storage, networking, etc.). - You focus on building the app — the provider handles everything else.	<ul style="list-style-type: none"> <li>• Build and deploy apps <b>faster</b></li> <li>• Avoid managing infrastructure</li> <li>• Easily scale with user demand</li> <li>• Enable <b>continuous integration</b> and deployment</li> <li>• Focus only on <b>code and features</b></li> </ul>



<b>SaaS (Software as a service)</b>	<p><b>SaaS</b> is a cloud computing model where <b>software applications are delivered over the internet</b> — no installation, no maintenance, no infrastructure needed.</p> <ul style="list-style-type: none"> <li>- You simply <b>access the software via a browser or app</b> — everything else is managed by the provider.</li> </ul>	<ul style="list-style-type: none"> <li>• For <b>everyday business apps</b> like email, docs, finance, sales, and customer support</li> <li>• When you want to avoid software installation or maintenance</li> <li>• When you need to get started <b>quickly</b></li> <li>• When you work with a <b>remote or hybrid team</b></li> <li>• When you prefer <b>monthly or yearly subscriptions</b></li> </ul>
-------------------------------------	--	---



## Day 1: Task 3

Pricing Calculator: In this exercise, you use the Pricing calculator to estimate the cost of running a basic web application on Azure.

Cost of a  
basic  
web  
application

Microsoft Azure Estimate					
Your Estimate					
Service category	Service type	Custom name	Region	Description	Estimated monthly cost
Compute	Virtual Machines		West US	2 D2 v3 (2 vCPUs, 8 GB RAM) x 730 Hours (Pay as you go), Windows (Licence included), OS Only, 0 managed disks – S4; Inter Region transfer type, 5 GB outbound data transfer from West US to East Asia	\$305.14
Databases	Azure SQL Database		West US	Single Database, vCore, General Purpose, Provisioned, Standard-series (Gen 5), Primary or Geo replica Disaster Recovery, Locally Redundant, 1 - 8 vCore Database(s) x 730 Hours, 32 GB Storage, SQL Licence (Pay as you go), RAGRS Back up Storage Redundancy, 0 GB Point-In-Time Restore, 0 x 5 GB Long Term Retention	\$1,567.39
Networking	Application Gateway		West US	Web Application Firewall V1 tier, Medium Instance size: 2 Gateway hours instance(s) x 730 Hours, 1 TB Data processed unit(s), 5 GB Zone unit(s)	\$206.04
Support			Support		\$0.00
			Licensing Program	Microsoft Customer Agreement (MCA)	
			Billing Account		
			Billing Profile		
			Total		\$2,078.56

TCO Calculator: In this exercise, you use the Total Cost of Ownership (TCO) Calculator to compare the cost of running a sample workload in your datacentre versus on Azure.

Datacentre vs Azure



## Day 1: Task 4

Please research the below terms and explain what they are, when they would be appropriate and a real-world example of where it could be implemented (i.e. what type of organisation).

### Public Cloud

A **public cloud** is a cloud environment **owned and operated by a third-party provider** (like AWS, Azure, or Google Cloud) that offers computing services **over the internet**. Resources are **shared** among multiple customers (multi-tenant).

When to use it:

- When you want **cost-effective, scalable, on-demand infrastructure**
- When **data sensitivity or compliance is not a major concern**
- Ideal for **startups, small businesses**, or public websites

Real-world example:

A **start-up e-commerce company** uses **AWS EC2** to host their online store. They scale quickly during sales without investing in servers.

### Private Cloud

A **private cloud** is a cloud environment **used exclusively by a single organization**. It can be hosted **on-premises** or by a third party, but resources are not shared with others.

When to use it:

- When your business handles **sensitive data** (e.g., financial, healthcare, government)
- When you need **complete control** over infrastructure and security
- Useful for industries with **strict compliance and regulatory requirements**

Real-world example:

A **government agency** or **bank** sets up a private cloud on-premises to store confidential citizen or financial data securely.

### Hybrid Cloud

A **hybrid cloud** combines **public and private clouds**, allowing data and applications to move between them as needed. It offers **flexibility**: keep sensitive data in the private cloud, and use the public cloud for less critical workloads or for scaling.



## Community Cloud

### When to use it:

- When you need both **security** (private cloud) and **scalability** (public cloud)
- When you have **existing infrastructure** but want to use cloud capabilities
- Ideal for **large enterprises** undergoing digital transformation

### Real-world example:

A **hospital** stores patient records in a private cloud for compliance but uses the public cloud to run an AI model that predicts appointment no-shows.

A **community cloud** is shared among several organizations that have **similar goals, concerns, or regulations** (e.g., compliance, performance).

- It can be managed by the organizations or a third party.
- Offers a balance between **shared infrastructure and shared governance**.

### When to use it:

- When organizations in the **same industry** want to share infrastructure and costs
- When **data sovereignty, privacy, or compliance** is important

### Real-world example:

A group of **universities or healthcare providers** create a community cloud to share research data or patient records under common regulations (e.g., HIPAA in the US), **(UK GDPR (General Data Protection Regulation))** & **Data Protection Act 2018**)



## Day 2: Task 1

Describe, with examples, the **three** major areas that the Computer Misuse Act deals with.

Area	Description	Example
<b>1.Unauthorised access to computer systems</b>	Accessing a computer system or data without permission (even if no damage is done)	Logging into someone else's email or network without their consent
<b>2.Unauthorised access with intent to commit further offences</b>	Accessing a system without permission <b>to commit another crime</b> , such as fraud or theft	Hacking into a bank system to steal money or customer data
<b>3.Unauthorised modification of computer material</b>	Changing, deleting, or introducing data or software without permission (e.g. spreading viruses or malware)	Creating and spreading ransomware that encrypts files on a victim's computer

The computer misuse act 1990 is an act where an individual can be criminalised because of computer related offense.  
Describe three extra powers that the Police and Justice Act 2006 (Computer Misuse) has added.

Description
<b>1. Criminalising denial-of-service (DoS) attacks</b> Made it illegal to intentionally impair the operation of computers, such as by overloading them with traffic
<b>2. Increased penalties</b> Raised the maximum sentence for serious computer misuse offences (e.g. from 5 years to up to 10 years for intent to damage national security or economic well-being)
<b>3. Making it illegal to produce or distribute hacking tools</b> Outlawed the creation, supply, or possession of software/tools designed to commit cybercrime
<b>The Police and Justice Act 2006 strengthened the original Computer Misuse Act by addressing new forms of cybercrime, increasing punishments, and targeting tools used by hackers.</b>

Look at the below website to answer the questions:



## <https://www.gov.uk/personal-data-my-employer-can-keep-about-me>

**Write down three items of data which a company can store about an employee.**

1. Full Name and Contact Details
  - e.g. address, phone number, email
2. National Insurance Number
  - used for tax and payroll purposes
3. Employment Records
  - e.g. job title, salary, performance reviews, training history

**Give three more examples of data that an employer can only store if they first get the employee's permission.**

1. Medical or Health Information
  - e.g. details about disabilities, mental health conditions, or medical reports
2. Biometric Data
  - e.g. fingerprints or facial recognition used for clocking in/out or access control
3. Racial or Ethnic Background
  - often collected for diversity monitoring or equal opportunity reporting, but consent is required

Conduct further research to answer the below questions.

Question	Answer
<b>Provide one example of: Copyright infringement</b>	Downloading and sharing a movie or song online without the creator's permission.
<b>Provide one example of: Plagiarism</b>	Copying a paragraph from someone else's blog or essay and submitting it as your own work without crediting the original author.
<b>What are two consequences of copyright infringement and software piracy?</b>	<ul style="list-style-type: none"><li>• Legal <b>action or fines</b> – The copyright holder can sue the offender, and courts can issue fines or compensation claims.</li><li>• Criminal <b>charges</b> – In serious cases, piracy can lead to prosecution and even imprisonment.</li></ul>



<b>Give three possible consequences for individuals when using pirated software</b>	<ul style="list-style-type: none"> <li>• <b>Malware or viruses</b> – Pirated software often comes from untrusted sources and can include harmful malware.</li> <li>• <b>No updates or support</b> – Users won't receive software updates, bug fixes, or technical help.</li> <li>• <b>Legal risks</b> – Individuals can be fined or prosecuted for using unlicensed software.</li> </ul>
---	--

Listed below are some laws which we have covered today:

1. Computer Misuse Act 1990
2. Police and Justice Act 2006 (Computer Misuse)
3. Copyright, Designs and Patents Act 1988
4. Copyright (Computer Programs) Regulations 1992
5. The Health and Safety (Display Screen Equipment) Regulations 1992
6. Data Protection Act 2018
7. Consumer Rights Act 2015

- Insert a number in the first column of each row to match each of the statements with one of the above Acts.
- One of statements is incorrect and not illegal. For this statement, write ‘Not illegal’.

<b>Act number</b>	<b>Clause</b>
3	With some exceptions, it is illegal to use unlicensed software
7	Any product, digital or otherwise, must be fit for the purpose it is supplied for
1	Unauthorised modification of computer material is illegal



<b>Not Illegal</b>	It is illegal to create or use a hacking tool for penetration testing
<b>6</b>	Personal data may only be used for specified, explicit purposes
<b>5</b>	Employers must provide their computer users with adequate health and safety training for any workstation they work at
<b>2</b>	It is illegal to distribute hacking tools for criminal purposes
<b>6</b>	It is illegal to distribute an illicit recording
<b>6</b>	Personal data may not be kept longer than necessary
<b>1</b>	Gaining unauthorised access to a computer system is illegal
<b>5</b>	Employers must ensure that employees take regular and adequate breaks from looking at their screens
<b>1</b>	It is illegal to prevent or hinder access (e.g. by a denial-of-service attack) to any program or data held in any computer
<b>6</b>	Personal data must be accurate and where necessary kept up to date

## Day 3: Task 1

Please complete the below lab (3) '*Explore relational data in Azure*' and paste evidence of the completed lab in the box provided.

### Explore relational data in Azure

Learning Path 02 (CSS)

Duration:	2 Hours, 15 Minutes
Lab Series:	DP-900T00-A Microsoft Azure Data Fundamentals [Cloud Slice Provided]
Virtualization Platform:	Hyper-V
RAM:	6.5GB
Cloud Platform:	Azure
Content Version:	2
Is Exam:	No
Status:	Not Running

[Launch](#)



## Complet ed lab

The screenshot shows the Microsoft Azure Query editor (preview) interface. On the left, there's a sidebar with various options like Overview, Activity log, Tags, Diagnose and solve problems, and Query editor (preview). The Query editor (preview) option is selected. In the main area, there's a message about showing limited object explorer here. Below it, the 'Tables' section is expanded, showing a list of tables such as dbo.BuildVersion, dbo.ErrorLog, SalesLT.Address, SalesLT.Customer, SalesLT.CustomerAddress, SalesLT.Product, SalesLT.ProductCategory, SalesLT.ProductDescription, SalesLT.ProductModel, SalesLT.ProductModelProductDesc, SalesLT.SalesOrderDetail, SalesLT.SalesOrderHeader, Views, and Stored Procedures. A query is being run in the 'Query 1' pane:

```
1 SELECT p.ProductID, p.Name AS ProductName,
2          c.Name AS Category, p.ListPrice
3 FROM SalesLT.Product AS p
4 JOIN [SalesLT].[ProductCategory] AS c
5      ON p.ProductCategoryID = c.ProductCategoryID;
```

The results pane shows a table with columns ProductID, ProductName, and Category. The data is as follows:

ProductID	ProductName	Category
771	Mountain-100 Silver, 38	Mountain Bikes
772	Mountain-100 Silver, 42	Mountain Bikes
773	Mountain-100 Silver, 44	Mountain Bikes
774	Mountain-100 Silver, 48	Mountain Bikes
775	Mountain-100 Black, 38	Mountain Bikes
776	Mountain-100 Black, 42	Mountain Bikes

A green status bar at the bottom right indicates "Query succeeded | 0s".



## Day 3: Task 2: Skillable

Complete below exercises in lab environment using AdventureWorks DB.

### -- 1. Customers with last names starting with 'A'

-- Scenario: The marketing department is preparing a campaign targeting customers whose last names begin with the letter 'A'. They need a list of these customers to personalize outreach messages.

```
SELECT CustomerID, FirstName, LastName  
FROM [SalesLT].[Customer]  
WHERE LastName LIKE 'A%';
```

### -- 2. Customers with first name starting with 'A' and last name ending with 'a'

-- Scenario: The CRM team is analyzing customer name patterns for personalization strategies. They're particularly interested in users with a first name starting with 'A' and a last name ending in 'a'.

```
SELECT CustomerID, FirstName, LastName  
FROM [SalesLT].[Customer]  
WHERE FirstName LIKE 'A%' AND LastName LIKE '%a'
```

### -- 3. Count of customers with first name starting with 'A' and last name including 'o'

-- Scenario: A data analyst is generating statistics for a name segmentation report and needs to count how many customers meet both criteria: first name starts with 'A' and last name contains the letter 'o'.

```
SELECT COUNT(*) AS MatchingCustomerCount  
FROM [SalesLT].[Customer]  
WHERE FirstName LIKE 'A%' AND LastName LIKE '%o%';
```

### -- 4. Top 3 black products of size 'S' ordered by list price

-- Scenario: The merchandising team is curating a list of 'Black' colored products in size 'S' for a limited-time promotion. They want to showcase the top 3 most affordable options first.

```
SELECT TOP 3 Name, Color, Size, ListPrice  
FROM [SalesLT].[Product]  
WHERE Color = 'Black' AND Size = 'S' AND ListPrice > 0  
ORDER BY ListPrice ASC;
```

### -- 5. Average standard cost of products

-- Scenario: The finance team is calculating the average standard cost across all products to assess baseline production expenses.

```
SELECT AVG(StandardCost) AS AverageStandardCost  
FROM [SalesLT].[Product]  
WHERE StandardCost > 0;
```



## -- 6. Difference between average standard cost and average list price

-- Scenario: The pricing strategy team is examining the average markup between the standard cost and the list price to evaluate profitability trends.

```
SELECT AVG(ListPrice - StandardCost) AS AverageMarkup  
FROM [SalesLT].[Product]  
WHERE StandardCost > 0 AND ListPrice > 0;
```

In Percentage-

```
SELECT AVG((ListPrice - StandardCost) / StandardCost * 100.0) AS AverageMarkupPercentage  
FROM [SalesLT].[Product]  
WHERE StandardCost > 0 AND ListPrice > 0;
```

## -- 7. Product with highest profit margin

-- Scenario: Management wants to identify the most profitable product by finding the item with the highest difference between list price and standard cost.

```
SELECT TOP 1 Name, ProductID, ListPrice, StandardCost,  
       (ListPrice - StandardCost) AS Profit  
  FROM [SalesLT].[Product]  
 WHERE ListPrice > 0 AND StandardCost > 0  
 ORDER BY (ListPrice - StandardCost) DESC;
```

## -- 8. Number of products per category

-- Scenario: Inventory control needs a summary of how many products exist in each product category to manage stock levels effectively.

```
SELECT pc.Name AS CategoryName, COUNT(p.ProductID) AS ProductCount  
  FROM SalesLT.Product p  
 JOIN SalesLT.ProductCategory pc ON p.ProductCategoryID = pc.ProductCategoryID  
 GROUP BY pc.Name  
 ORDER BY ProductCount DESC;
```

## -- 9. Average list price by category

-- Scenario: The product pricing team is reviewing average list prices by category to adjust pricing strategies and identify outliers.

```
SELECT pc.Name AS CategoryName, AVG(p.ListPrice) AS AverageListPrice  
  FROM SalesLT.Product p  
 JOIN SalesLT.ProductCategory pc ON p.ProductCategoryID = pc.ProductCategoryID  
 WHERE p.ListPrice > 0  
 GROUP BY pc.Name  
 ORDER BY AverageListPrice DESC;
```



## -- 10. Orders with customer information

-- Scenario: The operations team needs a comprehensive report of customer orders, including order details and basic customer info, to review sales performance.

```
SELECT soh.SalesOrderID, soh.OrderDate, soh.TotalDue, sod.ProductID, sod.OrderQty,
sod.UnitPrice, sod.LineTotal, c.CustomerID, c.FirstName, c.LastName, c.CompanyName
FROM SalesLT.SalesOrderHeader soh
JOIN SalesLT.SalesOrderDetail sod ON soh.SalesOrderID = sod.SalesOrderID
JOIN SalesLT.Customer c ON soh.CustomerID = c.CustomerID
ORDER BY soh.OrderDate DESC, soh.SalesOrderID;
```

## -- 11. Products priced above average list price

-- Scenario: The sales team wants to identify premium products that are priced above the average list price for highlighting in high-end marketing campaigns.

```
SELECT p.ProductID, p.Name, p.ListPrice
FROM SalesLT.Product p
WHERE p.ListPrice > (SELECT AVG(ListPrice)
    FROM SalesLT.Product WHERE ListPrice > 0)
ORDER BY p.ListPrice DESC;
```

## -- 12. Total quantity sold per product

-- Scenario: The business intelligence team is evaluating product performance by analyzing the total quantity sold for each product across all orders.

```
SELECT p.ProductID, p.Name AS ProductName,
    SUM(sod.OrderQty) AS TotalQuantitySold
FROM SalesLT.SalesOrderDetail sod
JOIN SalesLT.Product p ON sod.ProductID = p.ProductID
GROUP BY p.ProductID, p.Name
ORDER BY TotalQuantitySold DESC;
```

## -- 13. Total sales per customer

-- Scenario: The customer success team wants to calculate the total value of orders placed by each customer to identify high-value clients for loyalty programs.

```
SELECT c.CustomerID, c.FirstName, c.LastName, c.CompanyName,
    SUM(soh.TotalDue) AS TotalOrderValue
FROM SalesLT.SalesOrderHeader soh
JOIN SalesLT.Customer c ON soh.CustomerID = c.CustomerID
GROUP BY c.CustomerID, c.FirstName, c.LastName, c.CompanyName
ORDER BY TotalOrderValue DESC;
```



## Day 3: Task 3: Skillable

Please complete the below lab (4) 'Explore non-relational data in Azure' and paste evidence of the completed lab in the box provided.



### Explore non-relational data in Azure

Learning Path 03 (CSS)

Duration:	2 Hours, 15 Minutes
Lab Series:	DP-900T00-A Microsoft Azure Data Fundamentals [Cloud Slice Provided]
Virtualization Platform:	Hyper-V
RAM:	6.5GB
Cloud Platform:	Azure
Content Version:	2
Is Exam:	No
Status:	Not Running

Launch

The screenshot shows the Microsoft Azure Storage browser interface. On the left, a sidebar lists various storage-related options: Overview, Activity log, Tags, Diagnose and solve problems, Access Control (IAM), Data migration, Events, Storage browser (which is selected and highlighted in grey), Partner solutions, Resource visualizer, Data storage (Containers and File shares), and Outposts. The main pane displays the 'shreya2025 | Storage browser' for the 'shreya2025' storage account. It shows a list of blob containers: shreya2025, Favorites, Recently viewed, Blob containers, File shares, and Queues. Under the 'Tables' section, there is a table named 'products'. The table has two items: one with PartitionKey 1 and RowKey 1, and another with PartitionKey 1 and RowKey 2. A message at the bottom of the table area says 'Showing all 2 items'. At the bottom of the main pane, there is a footer bar with icons for Search, Home, My resources, and Support.

Complet  
ed lab

The screenshot shows the Microsoft Azure Data Explorer interface for an Azure Cosmos DB account named "shreya". The left sidebar contains navigation links such as Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Cost Management, Quick start, and Data Explorer (which is selected). The main area displays a table of items under the "SampleDB" database and "SampleContainer" container. The table has columns for id, ... (ellipsis), ... (ellipsis), and egory (category). A specific item is selected, showing its details in the right pane. The details pane includes a code editor with a JSON snippet and line numbers, and a preview pane showing the item's properties.

	id	...	... egory
1	B267655B-A7C1-41E3		
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

```
1   {
2     "id": "B267655B-A7C1-41E3"
3     "categoryId": "AA2BAE74-D"
4     "categoryName": "Clothing"
5     "sku": "TG-W091-L"
6     "name": "Women's Tights,"
7     "description": "The produc"
8     "price": 74.99,
9     "tags": [
10       {
11         "id": "9C89E562-1"
12         "name": "Tag-128"
13       }
14     ],
15     "_rid": "Ux4NAPeTmxoAAAAA"
16     "_self": "dbs/Ux4NAA=/co"
17     "_etag": "'45039c1a-0000"
18     "_attachments": "attachme"
19     "_ts": 1755093450
20 }
```



The screenshot shows the Azure Cosmos DB Data Explorer interface. The left sidebar menu is open, showing various options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Cost Management, Quick start, Data Explorer (selected), Mirroring in Fabric (Preview), Resource visualizer, Settings, Integrations, Containers, Monitoring, Automation, and Help. The main area displays a query results page. A query is written in the Query Editor:

```
1 SELECT *
2 FROM c
3 WHERE CONTAINS(c.name, "Helmet")
```

The Results pane shows the following JSON data:

```
[{"id": "47ED1C3B-C205-4507-94EE-3B69A744B261", "categoryId": "14A1AD5D-59EA-4B63-A189-67B077783B0E", "categoryName": "Accessories, Helmets", "sku": "HL-U509", "name": "Sport-100 Helmet, Black", "description": "The product called \"Sport-100 Helmet, Black\"", "price": 34.99, "tags": [{"id": "12A06E6F-458F-42DF-9641-F1376CDD87B1", "name": "Tag-22"}]}
```

## Day 4: Task 1: MS Fabric using Just IT

Please follow the link below to complete the lab using your Just IT account in MS Fabric. There are 3 modules to complete.

[Data Factory end-to-end tutorial introduction and architecture - Microsoft Fabric | Microsoft Learn](#)



The screenshot shows the Microsoft Data Studio interface. At the top, there are navigation links for Home, Reporting, and Help, along with a SQL analytics endpoint dropdown. Below the header is a toolbar with various icons. A message bar at the top states: "This warehouse has a default Power BI semantic model. To automatically add objects, go to warehouse settings. To manually add objects, use Manage default semantic model. Learn more".

The main area is divided into several sections:

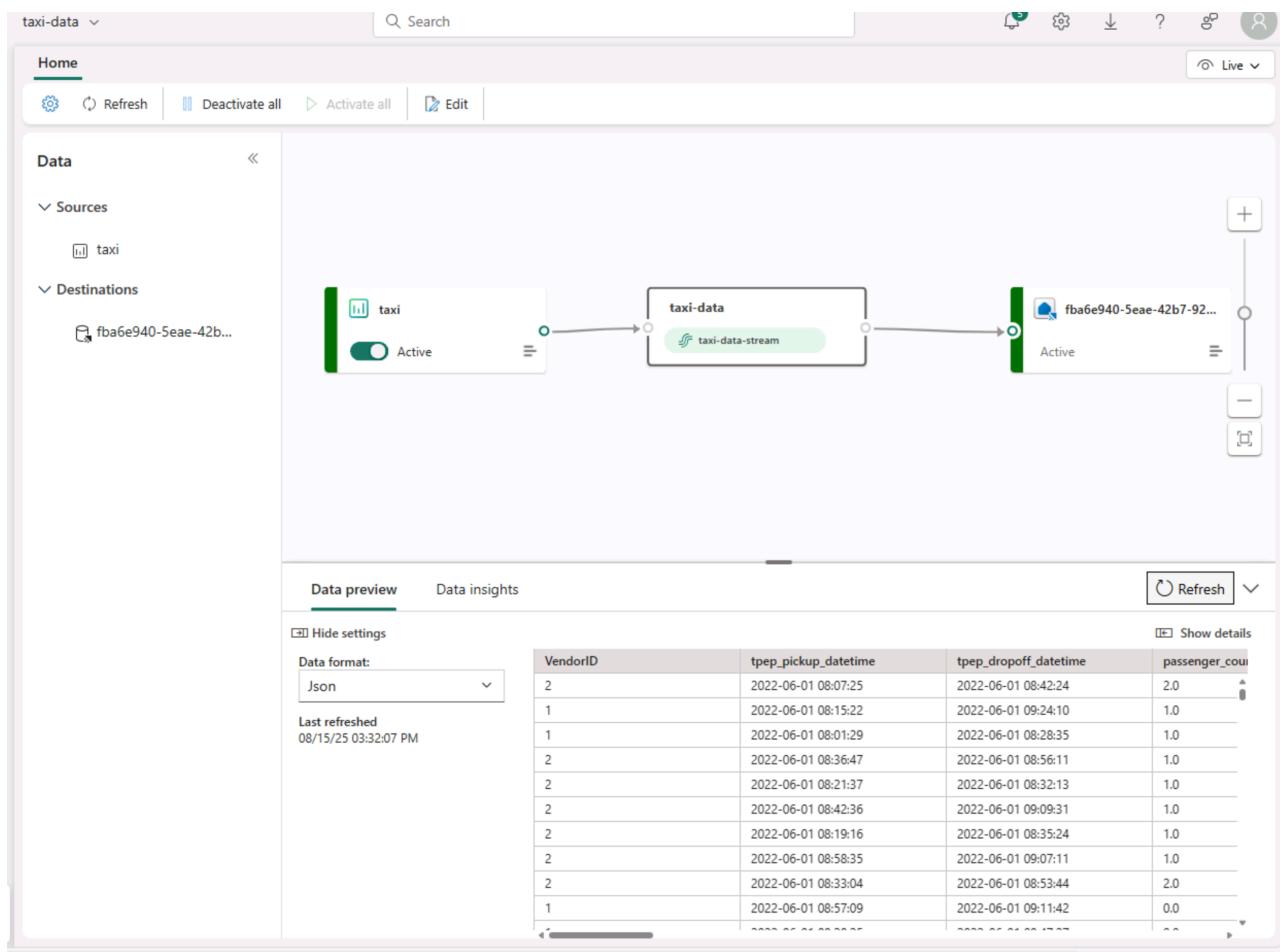
- Explorer:** On the left, it shows the warehouse structure under "Data2025". It includes "Schemas" (dbo, INFORMATION\_SCHEMA, queryinsights, sys), "Security", and "Queries" (My queries).
- SQL query 1:** This section contains the query code:

```
1 SELECT DATENAME(dw,lpepPickupDatetime) AS Day,
2       AVG(tripDistance) AS AvgDistance
3   FROM taxi_rides
4  GROUP BY DATENAME(dw,lpepPickupDatetime)
```
- Results:** This section displays the query results in a table format. The columns are "Day" and "AvgDistance". The data is as follows:

Day	AvgDistance
Monday	2.87683938944497
Wednesday	2.82952699686272
Saturday	2.98862075182174
Sunday	3.07944628039239
Thursday	2.85611154546786
Tuesday	2.82074476088204
Friday	2.85009385216234

At the bottom, there are status indicators: "Succeeded (12 sec 580 ms)", "Copilot completions: On", and "Copilot completions: Ready". It also shows "Columns: 2 Rows: 7".





The screenshot shows the Eventhouse interface. At the top, there's a navigation bar with 'Eventhouse', 'Database', 'Queryset', and various icons. Below the navigation is a toolbar with 'Live view', 'New', 'Get data', 'Query with code', 'KQL Queryset', 'Notebook', 'Real-Time Dashboard', 'Data policies', 'OneLake', and 'Share'.

**Left Sidebar:**

- Event2025:**
  - System overview
  - Databases
  - Monitoring
- KQL databases:**
  - Event2025
  - Event2025\_queryset

**Central Area:**

A KQL query is being typed into a code editor:

```

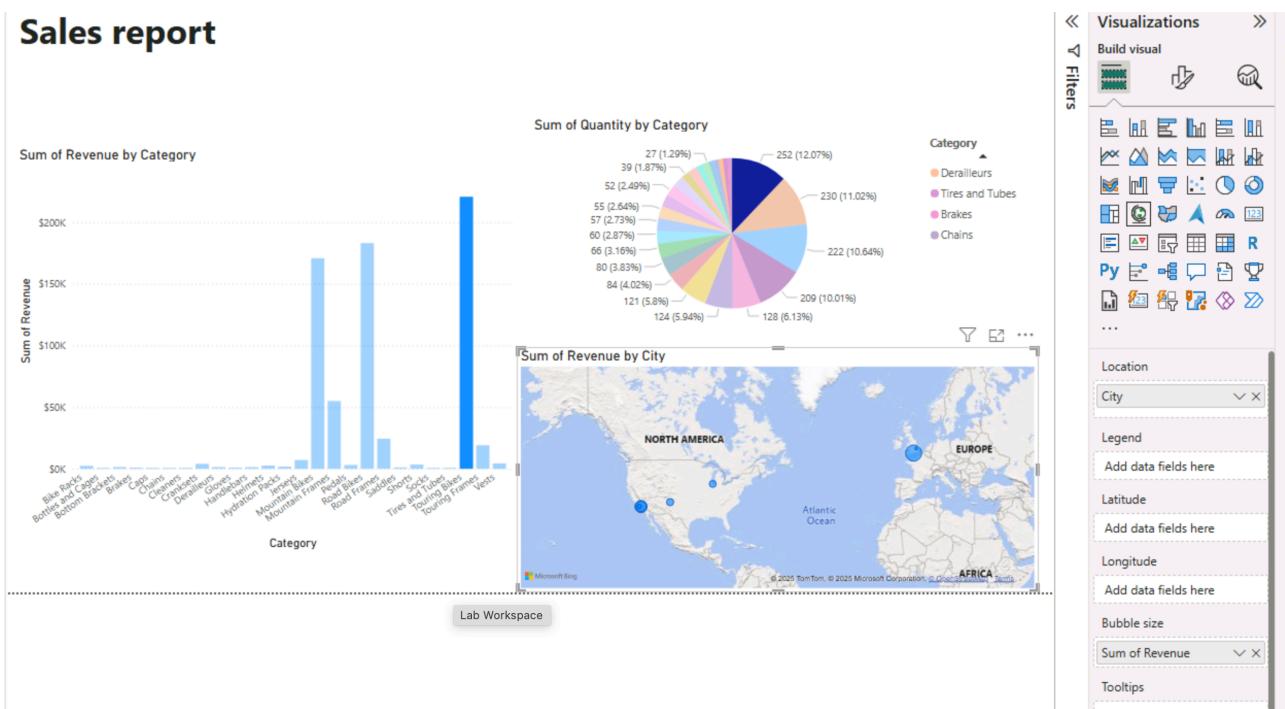
1 //*****
2 // Here are two articles to help you get started with KQL:
3 // KQL reference guide - https://aka.ms/KQLguide
4 // SQL - KQL conversions - https://aka.ms/sqlcheatsheet
5 *****
6
7 // Use "take" to view a sample number of records in the table and check the data.
8 taxi
9 | summarize PickupCount = count() by bin(todatetime(tpep_pickup_datetime), 1h)
10
11 // See how many records are in the table.
12 YOUR_TABLE_HERE
13 | count
14
15 // This query returns the number of ingestions per hour in the given table.
16 YOUR_TABLE_HERE
17 | summarize IngestionCount = count() by bin(ingestion_time(), 1h)
18

```

The results of the query are displayed in a table titled 'Table 1'.

tpep_pickup_datetime	PickupCount
2022-06-01 05:00:00.000	5
2022-06-01 06:00:00.000	19
2022-06-01 07:00:00.000	38
2022-06-01 08:00:00.000	1,852
2022-06-01 09:00:00.000	5,758
2022-06-01 10:00:00.000	6,010
2022-06-01 11:00:00.000	6,258
2022-06-01 12:00:00.000	6,655
2022-06-01 13:00:00.000	5,773
2022-06-01 14:00:00.000	32

## Sales report



## Day 4: Task 2: Skillable

In your teams, complete the Azure DP-900 practice exam and paste your result below – this is open book and please research and discuss your answers as a team.

<b>6</b> 🔗	<p>Practice Assessment: DP-900T00-A Microsoft Azure Data Fundamentals Practice Assessment for Microsoft Certifications for DP-900T00-A</p>	<p>Additional Details</p> <p>Required: No Available Instructor-Led: Yes Available Self-Paced: Yes</p>
---------------	--	---



**Result**

Your practice assessment results | Microsoft Learn

learn.microsoft.com/en-us/credentials/certifications/azure-data-fundamentals/practice/results?assessmentId=24&practice-assessme

## Overall Results

To be better prepared for the exam, aim to achieve a score of 80% or higher in multiple attempts.

Score: 64%

Show My Answers

## Performance by assessment section

To further strengthen your skills in the following areas, refer to the Customized Learning Material section below.

Describe core data concepts

Identify considerations for relational data on Azure

Describe considerations for working with non-relational data on Azure

Describe an analytics workload on Azure

## Day 4: Task 2 (Optional)

### 1. Scenario Background

"Paws & Whiskers" is a growing pet shop that aims to improve its business by analysing sales, customer information, and inventory data. Currently, the data is collected manually or stored in spreadsheets. Management is interested in transitioning to Microsoft Azure to streamline data storage, analysis, and reporting, enabling them to make data-driven decisions.



## **2. Data Laws and Regulations**

Identify and explain the data laws and regulations relevant to handling customer data within the proposal. Ensure you cover the following points:

- **GDPR Compliance:** Highlight the importance of adhering to the General Data Protection Regulation (GDPR), particularly as it relates to storing and processing customer information.
- **Data Protection Act (DPA) 2018:** Outline how the DPA 2018 may affect the way "Paws & Whiskers" collects and stores data, ensuring compliance with UK laws on data privacy.
- **Other Industry Standards:** Research any additional data protection standards or regulations that may apply to pet shop data, particularly if they involve sensitive or payment information.

## **3. Azure Service Recommendations**

Recommend Microsoft Azure services that would suit the company's data analysis needs and explain why these services are suitable. Your recommendations should include:

- **Data Storage:** Identify suitable storage options, such as **Azure Blob Storage** or **Azure SQL Database**, and discuss the benefits of each for storing large datasets, including inventory, sales transactions, and customer details.
- **Data Analysis Tools:** Recommend tools such as **Azure Machine Learning** for customer behaviour analysis or **Azure Synapse Analytics** for analysing sales trends.
- **Data Integration and Automation:** Explain how services like **Azure Data Factory** could automate data collection and integration processes, improving efficiency.

## **4. Data Types and Data Modelling**

Define the types of data "Paws & Whiskers" will need to work with and describe your approach to data modelling:

- **Data Categories:** Identify key data types, such as customer demographics, transaction history, pet inventory, and product categories.
- **Data Modelling Approach:** Outline how you would structure this data using a relational model or a data warehouse approach, considering factors like tables, entities, relationships, and primary keys.

## **5. Data Storage Formats and Structures in Azure**

Discuss how you would store data within Azure and the formats you would recommend:

- **Data Formats:** Specify recommended formats (e.g., CSV for raw data imports, JSON for structured data, Parquet for analytics) and explain why these formats are suitable for specific data types.
- **Data Security and Encryption:** Include recommendations for securing data using Azure's built-in encryption features and access controls to ensure compliance with data privacy regulations.



## **6. Additional Considerations**

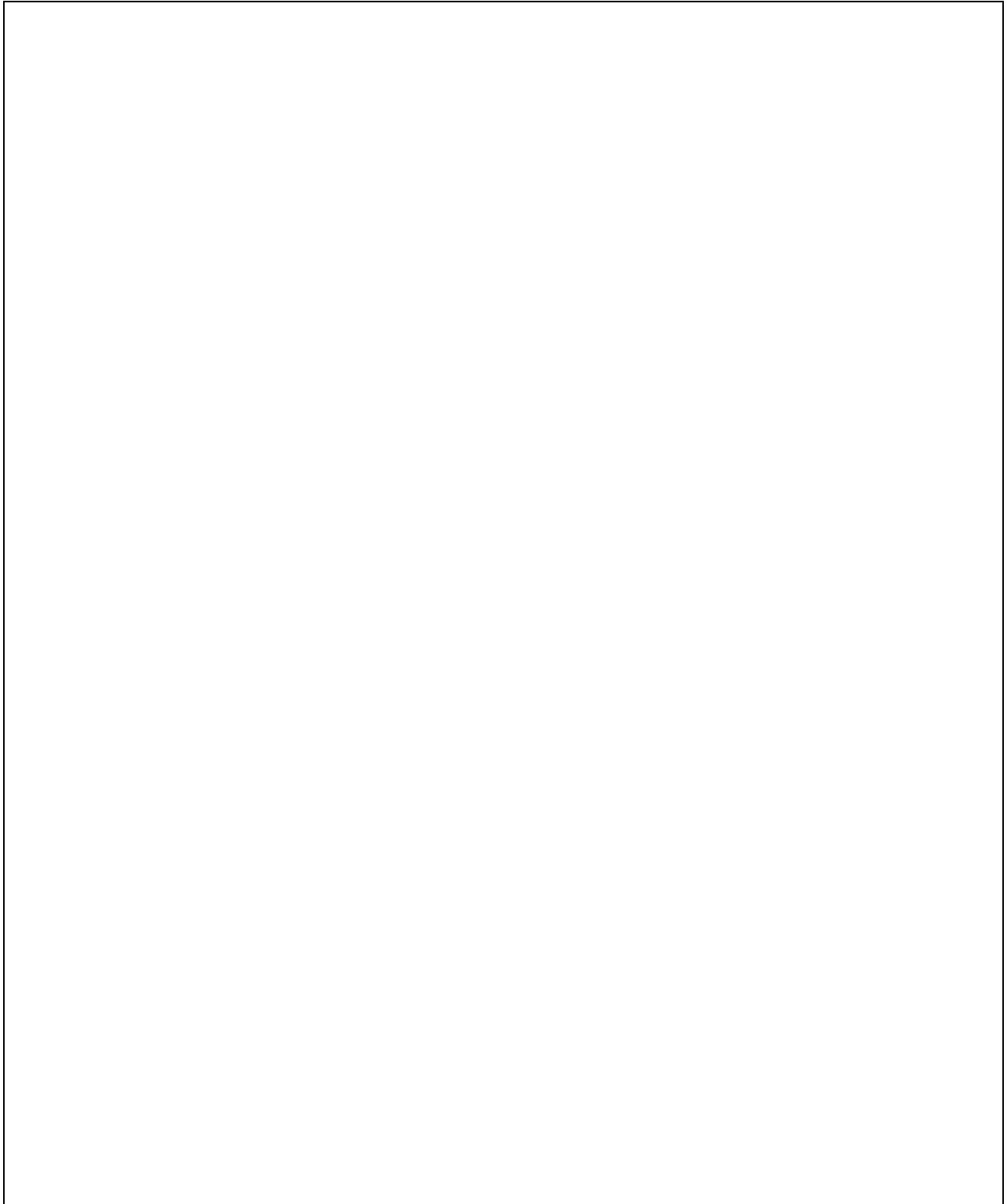
Provide any other considerations that might enhance data handling and efficiency in Azure, such as:

- **Backup and Disaster Recovery:** Outline a backup plan using **Azure Backup** or **Azure Site Recovery** to safeguard against data loss.
- **Data Visualisation:** Discuss potential use of **Power BI** within Azure for creating dashboards that provide management with real-time insights into sales and customer trends.
- **Future Scalability:** Comment on how Azure services can scale as the business grows, accommodating larger datasets and more complex analyses.

## **Submission Guidelines:**

1. **Structure:** Ensure your report is well-organised, with sections for each task (e.g., Data Laws, Azure Services, Data Types, etc.).
2. **Formatting:** Include headings, bullet points where appropriate, and any visuals or diagrams that support your explanations.
3. **References:** Cite any resources or regulations referenced in the report.
4. **Length:** Aim for 1500-2000 words.





## Course Notes

It is recommended to take notes from the course, use the space below to do so, or use the revision guide shared with the class:



We have included a range of additional links to further resources and information that you may find useful, these can be found within your revision guide.

## **END OF WORKBOOK**

**Please check through your work thoroughly before submitting and update the table of contents if required.**

**Please send your completed work booklet to your trainer.**

