



Data Technician

Table of contents

Day 1: Task 1.....	3
Day 1: Task 2.....	5
Day 1: Task 3.....	5
Day 1: Task 4.....	6
Day 2: Task 1.....	9
Day 3: Task 1.....	13
Day 3: Task 2: Skillable	14
Day 3: Task 3: Skillable	17
Day 4: Task 1: MS Fabric using Just IT	19



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?	<p>Hospitals and GPs store patient data in the cloud allowing patient records to be shared between healthcare providers. Netflix use cloud servers to store video library and deliver it globally, using cloud computing allows them to scale instantly during high-demand periods.</p>
How can it benefit a business?	<p>Cloud computing reduces requirement for business to invest in physical storage and so can reduce costs. Additionally, businesses can more easily scale their resources with demand. Cloud computing also makes collaboration easier with employees and allows real-time access to data and applications from anywhere with internet connectivity, beneficial for remote working.</p>
What's the alternative to cloud computing?	<p>Cloud computing allows people and companies to store data online allowing data to be accessed at any time and thereby reduce the need for physical storage systems. Additionally, often secure data storage and recovery is offered alongside cloud storage systems making backup and recovery easier. Cloud computing also allows companies to store and analyse large data sets without need for physical storage. Lastly cloud computing simplifies the setup and testing of development platforms which reduces the need for IT resources and infrastructure.</p>
	<p>Three alternatives are:</p> <ol style="list-style-type: none">1. On-premises infrastructure where all servers, networking and storage are hosted and maintained within company premises.



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

2. Edge computing where data is processed closer to where it's generated e.g. local servers instead of being sent to cloud.
3. Colocation where company owns and maintains their servers, but they are housed in a third-party data centre.

Amazon Web Services (AWS): Storage is S3 (files) and EBS (disks) and databases are RDS (SQL) and DynamoDB (noSQL). Useful for web app hosting, big data processing and analytics, ML and AI workloads and disaster recovery.

Microsoft Azure: blob storage and disk storage, SQL and Cosmos DB databases. Useful for enterprise apps, hybrid cloud, AI and ML model development big data and analytics pipelines.

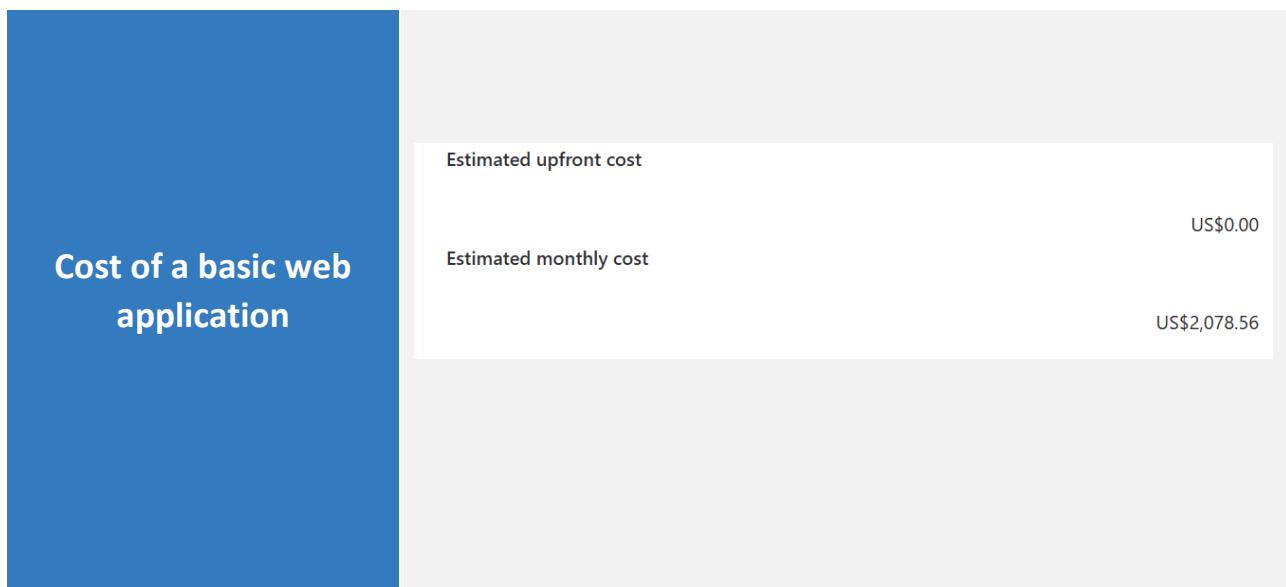
Google Cloud Platform: Storage is cloud storage (object), persistent disks and firestore. Databases are cloud SQL, cloud spanner (global RDBMS), Bigtable (NoSQL) and Firestore.



Day 1: Task 2

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

[Exercise - Estimate workload costs by using the Pricing calculator - Training | Microsoft Learn](#)



Please review this learning path if you finish early.

[Introduction to Cloud Infrastructure: Describe Cloud Concepts - Training | Microsoft Learn](#)

Day 1: Task 3

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?
IaaS (Infrastructure as a service)	IaaS is a cloud computing model where a provider delivers virtualised computing resources over the internet e.g. virtual machines, firewalls and storage. The cloud provider provides the physical servers, and the client controls the OS, applications and data.	A company could use IaaS to host websites without buying servers. Additionally, it could be used in development and testing, where developers could create temporary environments using virtual machines to test apps.
PaaS (Platform as a service)	PaaS is a cloud computing model that provides a ready-to-use platform for building, deploying and managing applications. The provider handles servers, storage, networking and OS. The client focuses on application code, data and configuration.	An example is for analytics and AI, where built-in tools can be used for data analytics and machine learning e.g. Azure Synapse Analytics and Google AI Platform. Also, it can be used by companies to build enterprise applications such as using the Salesforce platform to develop a customer relationship management (CRM) application.
SaaS (Software as a service)	SaaS is a cloud computing model where software applications are hosted by a provider and delivered over the internet. The provider manages the infrastructure, platform and maintenance. SaaS can be used on any device with internet.	Email is an example of SaaS, users can access it by logging in and can be used anywhere with internet. Another example is Microsoft 365 where documents can be accessed online, and files can be worked on collaboratively from anywhere with internet connectivity.

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 computing model where computing resources are owned and managed by a third-party provider and shared among multiple users. Users rent resources from providers such as Microsoft, Amazon Web Services and Google Cloud Platform. It is appropriate to use when cost efficiency is required as you won't need to buy and maintain your own hardware and if fast-growing demand is required as it is easily scalable. A public cloud may be used by startups or small businesses as it is more cost effective and easy to scale, also universities may use Microsoft for online learning platforms.
Private Cloud	A private cloud is a cloud computing model that is used only by one organisation and not shared as with public cloud. It can be hosted on-site or by a third party, but the infrastructure is dedicated only to one company. This is appropriate for companies where high security and privacy is required and where constant scaling isn't required. Examples would be government agencies in order to maintain security of national data and also banks which would need to protect financial data and comply with regulations such as GDPR.
Hybrid Cloud	A hybrid cloud is a combination of public and private clouds giving organisations the scalability and cost-effectiveness of the public cloud and also the security and privacy of the private cloud. This would be appropriate in companies where sensitive data needs to be kept private but the public cloud can still be used for scalability. An example could be healthcare organisations where patient data would be kept in the private cloud but public cloud could be used for running analytics.
Community Cloud	A community cloud is a shared cloud environment used by multiple organisations with common interests, only approved members can use it and costs are shared among users. It can be managed on-site or by a third party. It is appropriate when several companies need to



collaborate securely, joint infrastructure is required for joint projects or research and also there are shared compliance requirements. An example would be civil service departments, different government agencies could share a secure cloud platform and also universities working on a joint research project may choose to share a community cloud to analyse and store data.



Day 2: Task 1

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

Area	Description	Example
Unauthorised access to computer material.	It is an offense to gain unauthorised access to any computer material. This includes accessing data or systems without permission, as well as using an unauthorised method to gain access to a computer.	Logging to someone's email without permission
Unauthorised access to computer materials with intent to commit a further crime.	Accessing a system with the intent to commit a crime such as fraud or theft.	Hacking into a bank account to steal money or hacking in a shop's database to steal customer information (credit/debit card information, names and addresses etc.)
Unauthorised acts with intent to impair the operation of a computer.	Carrying out actions which are designed to damage, disrupt or modify computers or data.	Creating and/or spreading malware and deleting or corrupting business files to disrupt operations.

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



Increased the punishment for offence 1 (Unauthorised access to computer material) to 2 years imprisonment from six months initially.

Extended scope of offence 3 (Unauthorised acts with intent to impair the operation of a computer) to include Denial of Service attacks.

Added offence 3a which is making, supplying or obtaining anything that could be used in computer misuse offences such as creating a malware program that can be used to gain access to a system.

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.

Name

Address

National Insurance number

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

Race and ethnicity

Trade union membership

Health and medical conditions

Conduct further research to answer the below questions.



Question	Answer
Provide one example of: Copyright infringement	Downloading or streaming pirated movies and music e.g. using illegal websites to watch films.
Provide one example of: Plagiarism	Submitting someone else's work as your own e.g. using an assay writing service to write an essay for you at university
What are two consequences of copyright infringement and software piracy?	There can be reputational damage, as individuals and companies involved may sustain damage to their reputation within their industry and among customers. Additionally, there are financial consequences as software piracy and copyright infringement deprive creators of revenue.
Give three possible consequences for individuals when using pirated software	There can be legal consequences such as fines or criminal charges. Additionally, pirated software can contain malware and lastly as software is pirated users can't access updates or technical support.

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'.

Act number	Clause
4	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
Not Illegal	It is illegal to create or use a hacking tool for penetration testing
6	Personal data may only be used for specified, explicit purposes
5	Employers must provide their computer users with adequate health and safety training for any workstation they work at
2	It is illegal to distribute hacking tools for criminal purposes
3	It is illegal to distribute an illicit recording
6	Personal data may not be kept longer than necessary



1	Gaining unauthorised access to a computer system is illegal
5	Employers must ensure that employees take regular and adequate breaks from looking at their screens
1	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
6	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](#)



Completed lab

The screenshot shows the Microsoft Azure Query editor (preview) interface. On the left, there's a sidebar with options like Overview, Activity log, Tags, Diagnose and solve problems, and Query editor (preview). The main area shows a query window titled 'Query 1' with the following SQL code:

```
3 p.Name AS ProductName,
4 c.Name AS Category,
5 p.ListPrice
6 FROM SalesLT.Product AS p
7 INNER JOIN SalesLT.ProductCategory AS c
8 ON p.ProductCategoryID = c.ProductCategoryID;
9
```

The results pane shows two rows of data:

ProductID	ProductName	Category	ListPrice
771	Mountain-100 Silver...	Mountain Bikes	339
772	Mountain-100 Silver...	Mountain Bikes	339

A message in the sidebar says: "Showing limited object explorer here. For full capability please click here to open Azure Data Studio."

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 * 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 * 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(CustomerID) 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* FROM [SalesLT].[Product] WHERE Color='Black' 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) FROM [SalesLT].[Product];
```

-- 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)-AVG(StandardCost) FROM [SalesLT].[Product];
```

-- 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, ListPrice, StandardCost, (ListPrice - StandardCost) AS Profit
FROM [SalesLT].[Product] ORDER BY Profit 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 [SalesLT].[ProductCategory].Name, COUNT(*) AS Product_Count
FROM [SalesLT].[ProductCategory] INNER JOIN [SalesLT].[Product] ON
[SalesLT].[ProductCategory].ProductCategoryID=[SalesLT].[Product].ProductCategoryID
GROUP BY [SalesLT].[ProductCategory].Name;
```

-- 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 [SalesLT].[ProductCategory].Name, AVG(ListPrice) FROM
[SalesLT].[ProductCategory] INNER JOIN [SalesLT].[Product] ON
[SalesLT].[ProductCategory].ProductCategoryID=[SalesLT].[Product].ProductCategoryID
GROUP BY [SalesLT].[ProductCategory].Name;
```

-- 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 [SalesLT].[SalesOrderHeader].SalesOrderId, OrderDate,  
[SalesLT].[Customer].CustomerID, FirstName, LastName, TotalDue, Status, ProductID,  
OrderQty FROM [SalesLT].[SalesOrderHeader] INNER JOIN [SalesLT].[Customer] ON  
[SalesLT].[SalesOrderHeader].CustomerID=[SalesLT].[Customer].CustomerID INNER JOIN  
[SalesLT].[SalesOrderDetail] ON  
[SalesLT].[SalesOrderHeader].SalesOrderID=[SalesLT].[SalesOrderDetail].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 ProductID, Name, ListPrice FROM [SalesLT].[Product] WHERE ListPrice>(SELECT  
AVG(ListPrice) FROM [SalesLT].[Product]) ORDER BY 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 [SalesLT].[SalesOrderDetail].ProductID, Name, sum(OrderQty) AS  
Total_Quantity_Sold FROM [SalesLT].[SalesOrderDetail] INNER JOIN  
[SalesLT].[Product] ON  
[SalesLT].[SalesOrderDetail].ProductID=[SalesLT].[Product].ProductID GROUP BY  
[SalesLT].[SalesOrderDetail].ProductID,Name ORDER BY Total_Quantity_Sold 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 [SalesLT].[Customer].CustomerID, FirstName, LastName, Sum(TotalDue) AS  
Total_Value FROM [SalesLT].[Customer] INNER JOIN [SalesLT].[SalesOrderHeader] ON  
[SalesLT].[SalesOrderHeader].CustomerID=[SalesLT].[Customer].CustomerID GROUP BY  
[SalesLT].[Customer].CustomerID,FirstName, LastName ORDER BY Total_Value 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

Completed lab

The screenshot shows the Microsoft Azure Data Explorer interface for a Cosmos DB account named 'iba3'. The left sidebar navigation includes 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Diagnose and solve problems', 'Quick start', 'Data Explorer' (which is selected), 'Mirroring in Fabric', 'Resource visualizer', 'Containers', 'Monitoring', 'Automation', and 'Help'. The main area displays a query editor window with the following SQL code:

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

The results pane shows a table with one row of data:

id	categoryID	categoryName	sku	name	description	price	tags	ETag
F0EF81CB-5867-414F-BB03-07FF8A4B8B00	14A1A050-59EA-4B63-A1B9-67B077783B0E	Accessories, Helmets	Sport-100	Sport-100 Helmet, Red	The product called \Sport-100 Helmet, Red\	34.99	[empty]	\\"6501ef8a-0000-1000-0000-690b75fe0000\\"

Below the results, there is a 'Query Stats' section and a note: 'Add or remove favorites by pressing Ctrl + Shift + F'.

Part Two: Explore Azure Cosmos DB

The screenshot shows the Microsoft Azure Storage browser interface for a storage account named 'iba1'. The left sidebar contains navigation links such as 'Diagnose and solve problems', 'Access Control (IAM)', 'Data migration', 'Events', 'Storage browser' (which is selected), 'Storage Mover', 'Partner solutions', 'Resource visualizer', 'Data storage' (with 'Containers', 'File shares', 'Queues', and 'Tables' sub-options), 'Security + networking', 'Data management', and 'Settings' (with 'Configuration', 'Data Lake Gen2 upgrade', 'Resource sharing (CORS)', and 'Advisor' sub-options). The main pane displays a table named 'products' under the 'Tables' section. The table has columns: PartitionKey, RowKey, Timestamp, Name, and Price. There are two rows of data:

PartitionKey	RowKey	Timestamp	Name	Price
1	1	2025-11-06T11:52:08.319Z...	Widget	2.99
1	2	2025-11-06T11:53:24.072Z...	Knlnak	1.99

At the bottom right of the interface, there are language and date/time settings: ENG US, 3:53 AM, 11/6/2025.

Part one: Explore non-relational data in Azure



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](#)

Module 1: Create a pipeline with Data Factory

The screenshot shows the Microsoft Fabric Data Factory interface. A pipeline named 'Copy job' is displayed, containing a single activity named 'Copy job_b6'. The pipeline status is 'Succeeded'. Below the pipeline, a table shows the details of the most recent run, which completed successfully at 11/15/2025, 3:54:09 PM, with a duration of 43m 55s. The table includes columns for Activity name, Activity status, Run start, Duration, Input, and Output.

Module 2: Transform data with a dataflow in Data Factory

The screenshot shows the Microsoft Fabric Data Factory interface. A dataflow named 'nyc_taxi_with_discounts' is being previewed. The preview table displays 1000 rows of transformed data from the 'nyc_taxi_with_discounts' source table. The columns include vendorID, pickupLocation, passengerCount, tripDistance, dropoffLat, paymentType, fareAmount, extra, mtaTax, improvement, and tipAmount. The data shows various taxi trips with their respective details and discounts applied.



Module 3: Transform data with a dataflow in Data Factory

