

Data Technician

Name: Mohammad Ruhan Ahmed

Course Date:

Table of contents

Day 1: Task 1	3
Day 1: Task 2	6
Day 1: Task 3	8
Day 2: Task 1	9
Day 3: Task 1	14
Day 3: Task 2	15
Day 3: Task 3	16
Day 4: Task 1	19
Day 4: Task 2	21
1. Scenario Background	21
2. Data Laws and Regulations	21
3. Azure Service Recommendations	22
4. Data Types and Data Modelling	22
5. Data Storage Formats and Structures in Azure	22
6. Additional Considerations	22
Submission Guidelines:	23
Course Notes	30
Additional Information	32



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?

The on-demand delivery of IT resources over the internet, with pay-as-you-go pricing.

Instead of buying, owning and maintaining physical data centres and servers, you can access technology services – such as computing power, storage and databases – on an as-needed basis from a cloud provider.

Some other examples include:

- Data storage (e.g., Google Drive, Dropbox)
- Software as a Service (SaaS) (e.g., Microsoft 365, Salesforce)
- Disaster recovery & backup solutions
- Streaming services (e.g., Netflix, Spotify)
- Remote collaboration and productivity (e.g., Zoom, Slack)
- Smart city infrastructure and IoT systems
- Healthcare data management and telemedicine
- AI and machine learning model training

How can it benefit a business?

- Scalable: You can scale resources up or down based on demand.
- Cost-efficient: Pay-as-you-go pricing avoids large, unnecessary costs on hardware.
- Remote accessible: Enables work-from-anywhere setups.
- Fast deployment: Applications and services can be launched in minutes.



- Reliable uptime: Cloud services often offer 99.9%+ uptime guarantees for businesses.
- Security & compliance: Providers often have advanced security measures and compliance certifications.
- AI: Tools for AI, analytics, and big data are readily available.

What's the alternative to cloud computing?

Alternatives include on-premises, where all the servers, data and networking are managed within the business premises – or Colocation (CoLo), where businesses rent a space to a third-party data centre privately.

This allows complete control over the data, however it is expensive to maintain and usually less scalable than cloud computing

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

Features	AWS	Azure	GCP
Compute	EC2 (virtual machines), Lambda (serverless), ECS/EKS (containers /Kubernetes)	Azure Virtual Machines, Azure Functions (serverless), Azure Kubernetes Service (AKS)	Compute Engine (VMs), Cloud Functions (serverless), GKE (Google Kubernetes Engine)
Storage	S3 (object storage), EBS (block storage), Glacier (archival)	Blob Storage, Disk Storage, Archive Storage	Nearline /Coldline (archival)
Database	RDS (SQL), DynamoDB (NoSQL),	Azure SQL Database, Cosmos DB	BigQuery (analytics), Firestore



	Redshift (data warehousing)	(NoSQL), Data Lake	(NoSQL), Cloud SQL (MySQL /PostgreSQL)
AI/ML	SageMaker, Rekognition, Lex	Azure Machine Learning, Cognitive Services, Bot Framework	Vertex AI, AutoML, Vision AI, Natural Language
Security	IAM (user/role management), Shield (DDoS protection), KMS (encryption keys)	Azure Active Directory (identity), Sentinel (SIEM), Key Vault	IAM (user/role management), Identity-Aware Proxy, Security Command Center



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?
IaaS (Infrastructure as a service)	<p>Provides virtualized computing resources over the internet.</p> <p>E.g. Servers, storage, and networking.</p>	<p>Used when you want full control over the infrastructure without buying physical hardware.</p> <p>For example, hosting a custom website or application on virtual machines (e.g., using Azure Virtual Machines or AWS EC2).</p> <p>Other use cases are:</p> <ul style="list-style-type: none">• Disaster recovery• Ecommerce• IoT / event processing• Startups• Software development
PaaS (Platform as a service)	<p>Provides a platform with tools to develop, test, and deploy applications – without managing the underlying infrastructure.</p>	<p>Ideal for developers building web or mobile apps quickly.</p> <p>For example, deploying a web app using Azure App Service or Google App Engine.</p> <p>Other use cases are:</p> <ul style="list-style-type: none">• API development• IoT• Agile development and DevOps
SaaS (Software as a service)	<p>Delivers ready-to-use software over the internet, allowing users to</p>	<p>Used when you need off-the-shelf tools without managing updates or infrastructure.</p>



	access and use them without needing to install or manage the software on their own devices	<p>For example, using Microsoft 365 for email, documents, and collaboration.</p> <p>Most personal/employee productivity applications are available as SaaS</p>
--	--	--



Day 1: Task 3

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	<p>A public cloud offers services – like storage, computing power, or applications – are delivered over the internet by third-party providers (e.g., AWS, Microsoft Azure, Google Cloud). Infrastructure is shared between multiple users (tenants).</p> <p>It's used when scalability and cost-efficiency are priorities, workloads are variable or unpredictable, or when there's no need for strict (stringent) data security or compliance</p> <p>E.g. A startup building a mobile app might use Amazon Web Services (AWS) to quickly deploy and scale their backend services without investing in physical infrastructure.</p>
Private Cloud	<p>A Private cloud is a cloud environment dedicated to a single organization. It can be hosted on-premises or by a third party but offers greater control over security and customization.</p> <p>It's used when data privacy, control, and compliance are critical – and for organizations with steady workloads, or legacy applications requiring specific configurations</p> <p>E.g. A financial institution like a bank might use a private cloud to store sensitive customer data and run transaction processing systems due to strict regulatory requirements.</p>

Hybrid Cloud

A hybrid cloud is a combination of public and private clouds, allowing data and applications to move between them. It offers flexibility and more deployment options, and used when organizations need to keep some data private, but want to take advantage of public cloud scalability

Also, it's used for disaster recovery, burst workloads, or transitioning to the cloud gradually

E.g. A hospital might keep patient records on a private cloud for privacy but use the public cloud to run analytics on anonymized data.

Community Cloud

A Community cloud infrastructure is shared by several organizations with similar concerns (e.g., security, compliance, jurisdiction). It's managed internally or by a third party – when organizations in the same sector need a shared environment, or for collaboration, cost-sharing, and compliance alignment

E.g. Multiple universities in a region could use a community cloud to collaborate on research and share computing resources, data, and applications securely.

Day 2: Task 1



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

Area	Description	Example
Unauthorised access to computers	Accessing a computer system or data without permission.	A person logs into someone else's email account without their knowledge, even if they don't change or steal anything.
Unauthorised access with intent to commit or facilitate crime	Accessing a system illegally to commit another crime, like fraud or theft.	A hacker breaks into a company's payroll system to steal employee bank details for identity theft.
Unauthorised modification of computer material	Deliberately changing or deleting data – or installing malware into computer systems.	A hacker installs ransomware on the organisation's network, locking access to files until money is paid.

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
1. Making, Supplying or Obtaining articles: It's an offence to make, supply, adapt or offer to supply an article with the intention to commit/assist a Computer Misuse Act offence.
2. Increased Penalties for Unauthorised access: Increased the penalties for the offence of unauthorised access to computer material (already covered under Computer Misuse Act).
3. Expanded scope including DoS & DDoS: Penalties added for the offence of Denial and Distributed Denial of Service.

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. National Insurance number
2. Tax code
3. Employment history with the organisation

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

2. Genetics

3. Health and medical conditions

Conduct further research to answer the below questions.

Question	Answer
Provide one example of: Copyright infringement	Downloading and sharing a movie online without a licence or creator's permission
Provide one example of: Plagiarism	Copying text from a website or another student/colleague and pasting it in your school/work without crediting them
What are two consequences of copyright infringement and software piracy?	<ol style="list-style-type: none">1. Legal action2. Loss of reputation/business
Give three possible consequences for individuals when using pirated software	<ol style="list-style-type: none">1. Malware infections (viruses/keyloggers/ransomware)2. Lack of updates/support3. Legal penalties (facing fines)

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
2	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

Complete
d lab



[Get Data in Power BI Desktop](#) (Expected Duration 2 hours)

PL-300T00-A Microsoft Power BI Data Analyst [Cloud Slice Provided], Lab 01

Required: Yes

Started: 22 May 2025 10:25 (UTC)

Ended: 22 May 2025 10:47 (UTC)

Launch

Day 3: Task 2

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



[Explore non-relational data in Azure](#) (Expected Duration 2 hours, 15 minutes)

DP-900T00-A Microsoft Azure Data Fundamentals [Cloud Slice Provided], Learning Path 03 (CSS)

Required: Yes

Started: 11 June 2025 10:22 (UTC)

Ended: 11 June 2025 11:02 (UTC)

Launch

8 of 10 launch attempts remaining

Day 3: Task 3

Please complete the below lab (5) 'Explore data analytics in Azure' and paste evidence of the completed lab in the box provided.



Explore data analytics in Azure

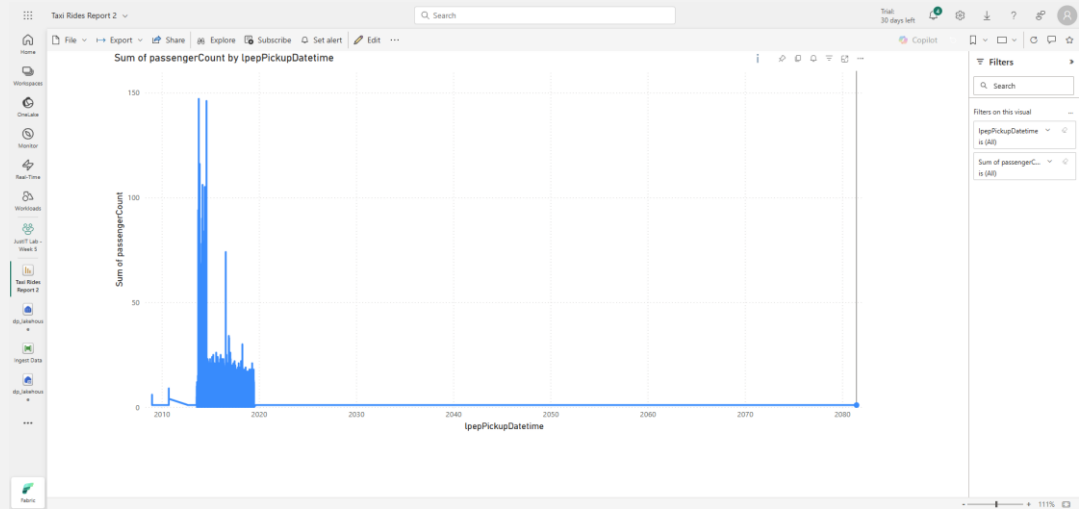
Learning Path 04 (CSS)

Duration: 3 Hours
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



Exercise – Explore data analytics in Microsoft Fabric



SQL query 1

```

1 SELECT DATENAME(dw, lpepPickupDatetime) AS Day,
2        AVG(tripDistance) AS AvgDistance
3 FROM taxi_rides
4 GROUP BY DATENAME(dw, lpepPickupDatetime)
    
```

Day	AvgDistance
Monday	2.87663938944486
Tuesday	2.98862075182174
Wednesday	2.85009385216235
Thursday	2.82052069660273
Friday	3.07544628050239
Saturday	2.85811154546786
Sunday	2.82074476080204

Succeeded (8 sec 192 ms) | Copilot completions: Off | Columns: 2 Rows: 7

Exercise – Explore Microsoft Fabric Real-Time Intelligence

my_event_house

my_event_house.queryset

```

1 // Use "Table" to view a sample number of records in the table and check the data.
2 taxi
3 | limit 100
4
5 // See how many records are in the table.
6 taxi
7 | count
8
9 // This query returns the number of ingestions per hour in the given table.
10 // | summarize IngestionCount = count() by bin(ingestion_time(), 1h)
11
12 // Show the number of taxi pickups for each hour:
13 taxi
14 | summarize PickupCount = count() by bin(todatetime(tpep_pickup_datetime), 1h)
    
```

VendorID	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	RatecodeID	store_and_fwd_flag	PULocationID	DOLocationID	payment_type	fa
2	2022-06-01 08:41:34.0000	2022-06-01 08:10:24.0000	2	4.41	1	N	262	68	1	
2	2022-06-01 08:59:43.0000	2022-06-01 08:05:53.0000	3	0.98	1	N	125	68	1	
2	2022-06-01 09:00:03.0000	2022-06-01 08:25:06.0000	1	5.56	1	N	151	68	1	
2	2022-06-01 09:00:44.0000	2022-06-01 08:08:11.0000	2	1.01	1	N	239	238	1	
2	2022-06-01 09:05:08.0000	2022-06-01 09:15:42.0000	1	1.51	1	N	262	238	1	
2	2022-06-01 09:05:13.0000	2022-06-01 09:35:03.0000	1	3.49	1	N	236	234	1	
2	2022-06-01 09:05:45.0000	2022-06-01 09:43:22.0000	1	8.23	1	N	231	146	1	
2	2022-06-01 09:05:59.0000	2022-06-01 09:25:45.0000	1	3.49	1	N	262	161	1	
1	2022-06-01 09:06:49.0000	2022-06-01 09:09:36.0000	0	0.5	1	N	24	238	1	
2	2022-06-01 09:07:45.0000	2022-06-01 09:18:51.0000	1	1.22	1	N	163	229	2	



Tab

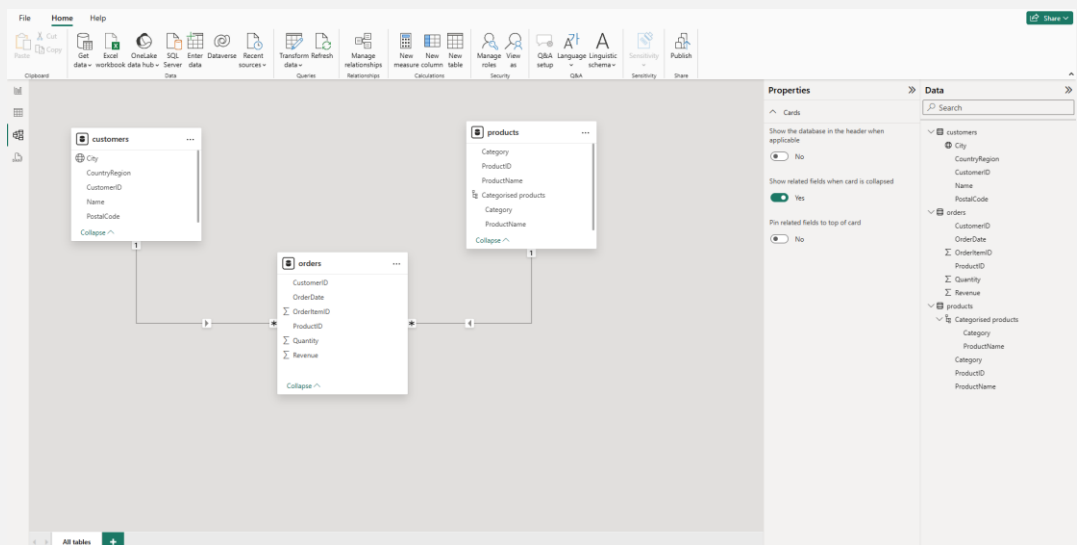
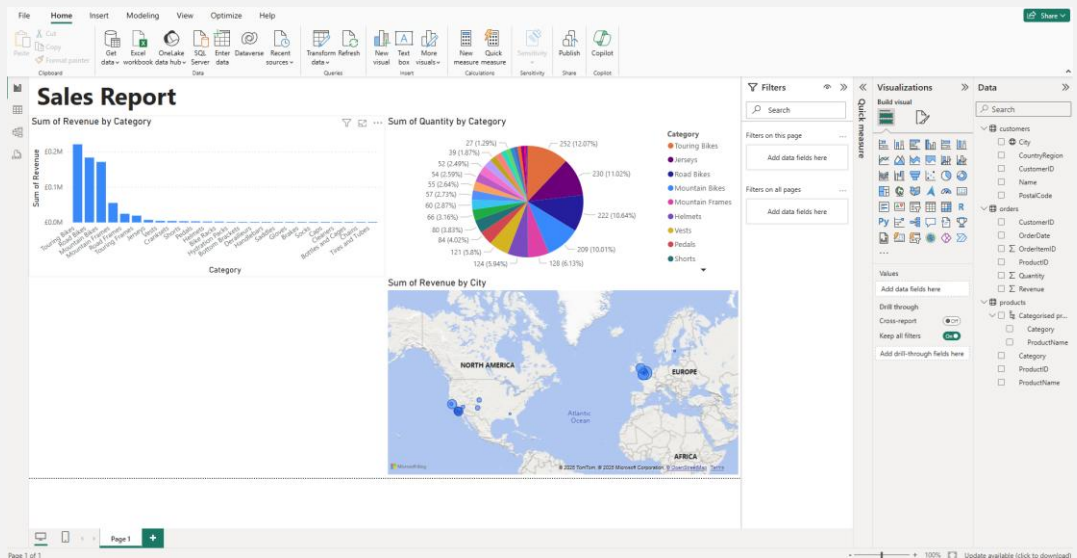
```

5 // Use "take" to view a sample number of records in the table and check the data.
6 taxi
7 | take 100
8
9 // See how many records are in the table.
10 taxi
11 | count
12
13 // This query returns the number of ingestions per hour in the given table.
14 // | summarize IngestionCount = count() by bin(ingestion_time(), 1h)
15
16 // Show the number of taxi pickups for each hour:
17 taxi
18 | summarize PickupCount = count() by bin(todatetime(spep_pickup_datetime), 1h)

```

time	PickupCount
2022-06-01 05:00:00.0000	5
2022-06-01 06:00:00.0000	22
2022-06-01 07:00:00.0000	1,125
2022-06-01 08:00:00.0000	5,339
2022-06-01 09:00:00.0000	5,893
2022-06-01 10:00:00.0000	1,410
2022-06-01 11:00:00.0000	6

Exercise – Explore fundamentals of data visualization with Power BI



File Home Help Table tools

Name orders

Table tools

Mark as date table

Manage relationships

New Quick New

measure measure columns table

Structure

Calendars

Calculations

OrderItemID OrderDate CustomerID ProductID Quantity Revenue

71774110562	1/1/2020 0:00	29647	836	1	£356.90
71774110563	1/1/2020 0:00	29647	832	1	£356.90
71776110567	1/2/2020 0:00	30072	907	1	£63.90
71780110630	1/7/2020 0:00	30113	990	1	£333.99
71780110621	1/8/2020 0:00	30113	926	1	£149.87
71780110622	1/8/2020 0:00	30113	743	1	£88.76
71780110626	1/13/2020 0:00	30113	937	1	£48.59
71780110638	1/15/2020 0:00	30113	985	1	£67.80
71780110631	1/16/2020 0:00	30113	992	1	£32.99
71780110640	1/27/2020 0:00	30113	810	1	£72.16
71780110642	1/29/2020 0:00	30113	925	1	£149.87
71780110644	1/31/2020 0:00	30113	880	1	£32.99
71782110669	3/2/2020 0:00	29485	954	1	£1436.44
71782110678	3/12/2020 0:00	29485	978	1	£445.41
71782110679	3/13/2020 0:00	29485	955	1	£1436.44
71782110680	3/14/2020 0:00	29485	948	1	£63.90
71782110681	3/15/2020 0:00	29485	994	1	£32.99
71782110682	3/16/2020 0:00	29485	970	1	£728.91
71782110688	3/22/2020 0:00	29485	957	1	£1436.44
71782110684	3/26/2020 0:00	29485	963	1	£445.41
71782110695	3/29/2020 0:00	29485	953	1	£728.91
71782110699	3/4/2020 0:00	29485	916	1	£31.58
71782110700	3/5/2020 0:00	29485	886	1	£206.05
71782110701	3/6/2020 0:00	29485	892	1	£602.35
71782110702	3/7/2020 0:00	29485	893	1	£602.35
71782110705	3/10/2020 0:00	29485	859	1	£14.69
71783110719	3/24/2020 0:00	29937	813	1	£72.16
71783110723	3/26/2020 0:00	29937	832	1	£356.90
71783110724	3/29/2020 0:00	29937	799	1	£672.29
71783110735	4/9/2020 0:00	29937	788	1	£672.29
71783110736	4/10/2020 0:00	29937	801	1	£672.29
71783110746	4/20/2020 0:00	29937	875	1	£3.39
71784110759	5/3/2020 0:00	29736	959	1	£445.41
71784110784	5/26/2020 0:00	29736	900	1	£206.05
71784110785	5/29/2020 0:00	29736	893	1	£602.35
717961110218	6/9/2020 0:00	29680	891	1	£602.35
71796111021	6/12/2020 0:00	29680	916	1	£31.58
71796111076	4/20/2020 0:00	29680	936	1	£716.91

Data

Search

- customers
 - City
 - CountryRegion
 - CustomerID
 - Name
 - PostalCode
- orders
 - CustomerID
 - OrderDate
 - OrderItemID
 - ProductID
 - Quantity
 - Revenue
- products
 - Categorised products
 - Category
 - ProductID
 - ProductName

Day 4: Task 1

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.

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



Practice Assessment Results: June 12, 2025



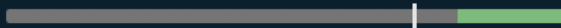
Practice Assessment for Exam DP-900: Microsoft Azure Data Fundamentals

It took you 8 minutes to complete this assessment.

Overall Results

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

Score: 72%

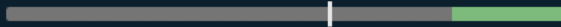


Show My Answers

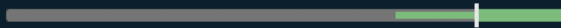
Performance by assessment section

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

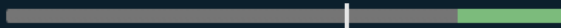
Describe core data concepts



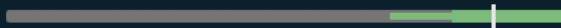
Identify considerations for relational data on Azure



Describe considerations for working with non-relational data on Azure



Describe an analytics workload on Azure



Day 4: Task 2

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.



Data Laws and Regulations

When handling customer data, Paws & Whiskers must comply with both European and UK-specific data protection laws. GDPR and DPA are the most important laws that govern how customer data must be collected, stored, processed, and protected.

General Data Protection Regulation (GDPR)

GDPR applies to all businesses processing personal data of EU and UK citizens, including those operating in the UK.

Key Principles:

- **Lawfulness, Fairness, and Transparency:**

The company must clearly inform customers about what data is being collected and why.

- **Purpose Limitation:**

Data must only be collected for specific, explicit, and legitimate purposes. E.g. customer email addresses collected for order confirmations should not be used for marketing without consent.

- **Data Minimisation:**

Only the necessary data should be collected. E.g. the pet shop does not need to collect a customer's date of birth unless it serves a valid business purpose (e.g., birthday discounts).

- **Accuracy:**

The company must ensure customer data is accurate and up to date.

- **Storage Limitation:**

Personal data should only be kept for as long as necessary.

- **Integrity and Confidentiality:**

Adequate security measures must be in place to protect customer data from breaches. E.g. using Azure's built-in encryption, secure access controls, and auditing tools.

Customer Rights Under GDPR:

- Right to access personal data
- Right to rectification
- Right to erasure ("right to be forgotten")
- Right to data portability
- Right to object to data processing



Failure to comply with GDPR can lead up to €20 million or 4% of global turnover in fines, whichever is higher. (GDPR.eu, 2025)

Data Protection Act (DPA) 2018

The DPA 2018 is the UK's implementation of GDPR and shows how personal data must be handled domestically (within the UK). It includes additional provisions and exceptions tailored to the UK's legal environment. (Information Commissioner's Office, 2025)

- **Lawful Basis for Processing:**

The company must identify a legal reason for processing personal data, such as consent, performance of a contract (e.g., fulfilling an order), or legitimate interest (e.g., fraud prevention).

- **Consent Management:**

Consent must be freely given, specific, informed, and unambiguous. Azure services like Azure Active Directory (for authentication) and Azure API Management (for consent-based API integration) can be used to help track and manage user consent.

- **Accountability and Governance:**

The pet shop must maintain records of data processing activities and demonstrate compliance. Tools like Microsoft Purview and Azure Policy can support this by enabling data classification and policy enforcement.

- **Data Breach Notification:**

If a data breach occurs, the company must notify the Information Commissioner's Office (ICO) within 72 hours. Azure Security Center can help detect and respond to breaches in real-time.

- **Mandates:**

- Clear consent for processing personal data.
- Rights to access, rectification, and erasure for individuals.
- Notification of data breaches within 72 hours.

- **Impacts:**

- Customer data collection via websites or loyalty programs must include privacy notices.
- Employee handling of customer data should receive GDPR training.

Other Industry Standards

- PCI DSS (Payment Card Industry Data Security Standard):
- Relevant if “Paws & Whiskers” processes card payments online or in-store.
- Azure offers PCI DSS-compliant services, including secure payment gateways.
- ISO/IEC 27001:
- Azure is certified, ensuring best practices in information security management.

Azure Service Recommendations

To meet business needs, the following Azure services are recommended:

Data Storage

Service	Purpose	Benefits
Azure Blob Storage	Store unstructured or semi-structured data (CSV, images, logs)	Low-cost, scalable, suitable for raw sales data or customer uploads
Azure SQL Database	Structured data (e.g., customer info, transactions)	Relational storage with built-in high availability, backups, and indexing
Azure Table Storage	NoSQL storage for simple structured data (e.g., pet food SKUs)	High-speed lookups and low cost

Data Analysis Tools

Tool	Use Case	Benefits
Azure Synapse Analytics	Complex sales, inventory, and trend analysis	Integrates data warehousing and big data; real-time analytics
Azure Machine Learning	Predict customer preferences or pet product recommendations	Enables creation of predictive models from customer behaviour data
Azure Databricks	Advanced analytics for large datasets	Supports Spark-based processing for large-scale, fast computations



Data Integration and Automation

Tool	Purpose	Benefits
Azure Data Factory	Automate data import from spreadsheets to Azure SQL or Blob	Scalable, supports data transformation and scheduling
Logic Apps	Workflow automation (e.g., trigger email alerts on low inventory)	No-code setup, integrates with external services like Gmail or Excel

Data Types and Data Modelling

Data Categories

- **Customer Demographics:** Name, address, email, phone number, pet preferences.
- **Transaction History:** Purchase date, items bought, price, payment method.
- **Pet Inventory:** Animal type, breed, age, health status, availability.
- **Product Categories:** Food, accessories, grooming, healthcare products.

Data Modelling Approach

A relational data model is suitable for structured datasets:

Entity	Attributes	Primary Key	Relationships
Customers	CustomerID, Name, Email, Phone	CustomerID	One-to-many with Transactions
Transactions	TransactionID, Date, CustomerID, TotalAmount	TransactionID	Many-to-one with Customers
Products	ProductID, Name, CategoryID, Price	ProductID	Many-to-one with Categories
Categories	CategoryID, Name	CategoryID	One-to-many with Products
Pets	PetID, Species, Breed, Availability	PetID	Optional link to Transactions (if pets sold)



Data Storage Formats and Structures in Azure

Data Formats

Format	Use Case	Reason
CSV	Raw imports from spreadsheets	Widely supported, easy for manual data input/export
JSON	Structured customer or transaction data	Ideal for API-based data exchange
Parquet	Sales and inventory analytics	Optimised for querying in Synapse or Databricks, columnar format reduces storage and enhances speed

Data Security and Encryption

- **Azure Defender for SQL/Storage:** Alerts on anomalous access.
- **Encryption at Rest:** Azure Storage encrypts all data using AES-256 by default.
- **Encryption in Transit:** Uses HTTPS and TLS.
- **Role-Based Access Control (RBAC):** Grants users access only to necessary data.
- **Azure Key Vault:** Stores and manages encryption keys securely.

Additional Considerations

Backup and Disaster Recovery

- **Azure Backup:** Schedule daily backups of Azure SQL, Blob storage, and VMs.
- **Azure Site Recovery:** Enables recovery of entire infrastructure in another region.
- Backup retention policies and geo-redundancy ensure data is recoverable after incidents.

Data Visualisation

- **Power BI Integration:**
 - Connects to Azure SQL and Synapse.
 - Dashboards can display:
 - Top-selling products
 - Customer demographics trends
 - Stock level alerts
 - Enables mobile-friendly, real-time visual insights for store managers.

Future Scalability

- Azure services offer **pay-as-you-grow** pricing.
- **Elastic pools** for Azure SQL accommodate traffic spikes.
- **Autoscaling** for Azure Synapse and Blob ensures performance remains high even with growing datasets.
- Easy integration with AI, IoT (for in-store sensors), and mobile applications in future phases.



References

GDPR.eu. (2025, June 12). *General Data Protection REgulations (GDPR) Compliance Guidelines*. Retrieved from <https://gdpr.eu/>

Information Commissioner's Office. (2025, June 12). *UK GDPR guidance and resources*. Retrieved from <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/>



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.

