

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: |
| Course Date: |
|  |

**Table of contents**

[Day 1: Task 1 3](#_Toc1514505634)

[Day 1: Task 2 3](#_Toc2125927931)

[Day 1: Task 3 4](#_Toc1965065538)

[Day 2: Task 1 5](#_Toc2141651249)

[Day 3: Task 1 9](#_Toc1229522580)

[Day 3: Task 2 10](#_Toc508692763)

[Day 3: Task 3 11](#_Toc1233463339)

[Day 4: Task 1 12](#_Toc1556426903)

[Day 4: Task 2 13](#_Toc214156810)

[1. Scenario Background 13](#_Toc248266112)

[2. Data Laws and Regulations 14](#_Toc1387014804)

[3. Azure Service Recommendations 14](#_Toc1952198484)

[4. Data Types and Data Modelling 14](#_Toc1792190821)

[5. Data Storage Formats and Structures in Azure 14](#_Toc385598743)

[6. Additional Considerations 15](#_Toc561077662)

[Submission Guidelines: 15](#_Toc1001523541)

[Course Notes 17](#_Toc977988415)

[Additional Information 17](#_Toc1081373283)

# Day 1: Task 1

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

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

|  |  |
| --- | --- |
| What can cloud computing do for us in the real-world? | * **We use it for Streaming Services** like Netflix, YouTube, and Spotify use the cloud to deliver content. * **Email & Storage** – Gmail, Google Drive, and Dropbox allow easy access to files from any device. * **Remote Work & Collaboration** – Tools like Microsoft 365, Zoom, and Slack enable teams to work together from anywhere. |
| How can it benefit a business? | **1. Cost Savings**  * **No Need for Expensive Hardware –** Businesses don’t have to buy and maintain physical servers. * **Pay-As-You-Go Model** – Companies only pay for the computing resources they use, reducing IT expenses. * **Lower Maintenance Costs –** Cloud providers handle updates, security, and maintenance.  **2. Scalability & Flexibility**  * **Easily Scale Up or Down** – Businesses can increase or decrease resources based on demand. * **Supports Remote Work** – Employees can access work systems from anywhere, improving productivity.  **3. Data Security & Backup**  * **Automatic Data Backup** – Prevents data loss due to cyberattacks, system failures, or human error. * **Advanced Security Features** – Encryption, multi-factor authentication (MFA), and threat detection protect business data. * **Disaster Recovery** – Ensures business continuity even during cyberattacks or system failures.  **4. Collaboration & Productivity**  * **Cloud-Based Collaboration Tools –** Apps like Microsoft 365, Google Workspace, and Slack allow employees to work together in real time. * **Access from Any Device** – Employees can work on projects from laptops, tablets, or phones.  **5. Faster Innovation & Deployment**  * **Quick Setup for New Services** – Businesses can launch new applications without long setup times. * **AI & Big Data Analytics** – Cloud platforms offer tools for analysing customer behaviour and improving decision-making.  **6. Competitive Advantage**  * **Access to Advanced Technologies** – Small businesses can use AI, machine learning, and data analytics without huge investments. * **Improved Customer Service** – Cloud-based CRM systems help manage customer interactions more effectively.  **7. Sustainability & Environmental Benefits**  * **Reduced Carbon Footprint** – Cloud providers use energy-efficient data centres. * **Less Physical Waste** – Businesses use fewer on-site servers and IT equipment. |
| What’s the alternative to cloud computing? | **1. On-Premises (Traditional IT Infrastructure)**  * Businesses own and maintain their own servers, storage, and networking hardware. * **Pros:**   + Full control over security and data.   + No reliance on third-party providers.   + Faster access to data since it's stored locally. * **Cons:**   + High upfront costs for hardware and maintenance.   + Requires in-house IT expertise.   + Limited scalability compared to cloud solutions.  **2. Hybrid Computing** (Mix of Cloud & On-Premises)  * A combination of cloud computing and on-premises infrastructure. * **Pros:**   + Offers flexibility—businesses can keep sensitive data on-premises while using cloud resources for scalability.   + Cost-efficient, as businesses only use cloud services when needed. * **Cons:**   + Complex to manage and integrate different environments.   + Security risks if not properly configured.  **3. Private Cloud**  * A cloud environment dedicated to a single organization, either hosted on-site or by a third-party provider. * **Pros:**   + More control over data security and compliance.   + Can be optimized for specific business needs. * **Cons:**   + Expensive to set up and maintain.   + Less scalability compared to public cloud solutions. |
| What cloud providers can we use, what are their features and functions? | 1. **Amazon Web Services (AWS)** **Features & Functions:**   * **Compute:** EC2 (Elastic Compute Cloud) allows scalable computing capacity. * **Storage:** S3 (Simple Storage Service) for object storage and EBS (Elastic Block Store) for block storage. * **Databases:** RDS (Relational Database Service), DynamoDB (NoSQL). * **Networking:** VPC (Virtual Private Cloud), Direct Connect, Route 53 (DNS). * **Analytics:** AWS Redshift for data warehousing, AWS Athena for querying data in S3. * **Machine Learning:** SageMaker for building, training, and deploying ML models. * **Security:** AWS Identity and Access Management (IAM), encryption, and compliance. * **Developer Tools:** CodeBuild, CodeDeploy, CodePipeline for CI/CD.   **Best For:** Large-scale, customizable infrastructure for enterprises. 2. **Microsoft Azure** **Features & Functions:**   * **Compute:** Virtual Machines, Azure Kubernetes Service (AKS). * **Storage:** Azure Blob Storage, Azure Disk Storage, Azure Files. * **Databases:** Azure SQL Database, Cosmos DB (NoSQL), MySQL, PostgreSQL. * **Networking:** Virtual Network, Load Balancer, Azure DNS. * **Analytics:** Azure Synapse Analytics, Azure Databricks. * **Machine Learning:** Azure Machine Learning service, Cognitive Services (AI APIs). * **Security:** Azure Active Directory (AD), Key Vault, Security Center. * **Developer Tools:** Azure DevOps, Visual Studio Code integration.   **Best For:** Enterprises using Windows-based technologies, hybrid cloud, and integration with Microsoft products. 3. **Google Cloud Platform (GCP)** **Features & Functions:**   * **Compute:** Google Compute Engine, Kubernetes Engine. * **Storage:** Google Cloud Storage, Persistent Disk. * **Databases:** Cloud SQL, Cloud Bigtable, Firestore. * **Networking:** VPC, Cloud Load Balancing, Cloud CDN. * **Analytics:** BigQuery for data analytics, Dataflow for stream processing. * **Machine Learning:** AI Platform for training and deployment, TensorFlow on GCP. * **Security:** Identity and Access Management (IAM), Data Loss Prevention API. * **Developer Tools:** Cloud Build, Cloud Functions, Firebase for mobile/web apps.   **Best For:** Data-intensive operations, machine learning, and integration with open-source tools. |

# 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) | **IaaS (Infrastructure as a Service)** – Provides virtualized computing resources over the internet, such as servers, storage, and networking. Users manage their own applications and operating systems. | Example: **Microsoft Azure Virtual Machines, Amazon EC2, Google Compute Engine** |
| PaaS (Platform as a service) | **PaaS (Platform as a Service)** – Provides a platform for developers to build, test, and deploy applications without managing the underlying infrastructure. It includes tools, databases, and operating systems. | Example: **Google App Engine, Microsoft Azure App Services, Heroku** |
| SaaS (Software as a service) | **SaaS (Software as a Service)** – Delivers software applications over the internet, which users can access without installation or maintenance. Everything is managed by the provider. | Example: **Google Drive, Microsoft 365, Dropbox, Zoom, Salesforce** |

# 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 | **What it is:** A public cloud is a cloud computing model where services (like storage, computing power, or applications) are delivered over the internet by third-party providers and shared across multiple organizations (tenants). Examples include **AWS (Amazon Web Services), Microsoft Azure, and Google Cloud Platform (GCP)**.  **When it’s appropriate:**   * For businesses that need **scalability, cost-efficiency, and minimal infrastructure management**. * For startups or small businesses that cannot afford large upfront IT investments. * For applications with **variable workloads** (e.g., seasonal traffic spikes).   **Real-world example:**   * A **streaming service like Netflix** uses AWS (public cloud) to handle massive, fluctuating user demand efficiently without maintaining its own data centers. |
| Private Cloud | **What it is:** A private cloud is a dedicated cloud environment used exclusively by a single organization. It can be **hosted on-premises (in-house data centers) or by a third-party provider**, but access is restricted.  **When it’s appropriate:**   * For industries with **strict security and compliance requirements** (e.g., finance, healthcare, government). * For businesses that need **full control over their infrastructure and data**. * For large enterprises with **predictable workloads** that justify the investment.   **Real-world example:**   * **Banks like JPMorgan Chase** use private clouds to ensure high security for sensitive financial data and comply with regulations like GDPR or HIPAA. |
| Hybrid Cloud | **What it is:** A hybrid cloud combines **public and private clouds**, allowing data and applications to move between them. This provides flexibility, security, and scalability.  **When it’s appropriate:**   * For businesses that want to **keep sensitive data private** but use public cloud for less critical workloads. * For **disaster recovery** (e.g., backup in public cloud while primary systems run on private cloud). * For organizations undergoing **digital transformation** (gradually migrating to the cloud).   **Real-world example:**   * **Toyota** uses a hybrid cloud model, keeping core manufacturing systems on a private cloud while using AWS for customer-facing apps and analytics. |
| Community Cloud | **What it is:** A community cloud is a shared cloud infrastructure used by **multiple organizations with common concerns** (e.g., security, compliance, or industry requirements). It can be managed internally or by a third party.  **When it’s appropriate:**   * For **collaborative industries** like healthcare, education, or government. * When organizations need to **share resources but maintain higher security than a public cloud**.   **Real-world example:**   * **Hospitals in a healthcare network** may use a community cloud to share patient data securely while complying with HIPAA regulations. |

# 

# 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**** | * Knowingly accessing a computer system **without permission** (even if no harm is done). * **Section 1** of the Act. | * An employee uses a colleague’s login to view confidential files. |
| ****Unauthorised Access with Intent to Commit Further Offences**** | * Illegally accessing a system **with the intent to commit fraud, theft, or other crimes**. * **Section 2** of the Act. | A hacker accesses a company’s payroll to **divert salaries** to their own account |
| ****unauthorised Modification of Computer Material**** | * Deliberately **deleting, altering, or corrupting data** or systems (e.g., deploying malware). * **Section 3** of the Act. | * A disgruntled employee **deletes critical company databases** before resigning. |

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 |
| Broadens **Section 3** of the Computer Misuse Act to criminalise **any deliberate act** (beyond just modification) that:   * Impairs a computer’s operation. * Prevents/hinders access to data. * Impacts data reliability. |
| * **Section 37** criminalises creating, adapting, or distributing tools **specifically designed for hacking**, even if not used yet. * Targets the **cybercrime supply chain** (e.g., malware developers). |
| * UK courts can prosecute **offences committed abroad** if:   + The attacker is a UK national.   + The attack targets UK systems. * Closes loopholes for overseas hackers targeting the UK. |

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. |
| ****Personal Identification Data****  * **Examples:**   + Full name   + Date of birth   + National Insurance number   + Home address * **Purpose:** Payroll, tax compliance, and employment records. |
| ****Employment Details****  * **Examples:**   + Job title and role   + Salary/payroll information   + Work email and phone number   + Contract terms * **Purpose:** HR management, performance reviews, and legal compliance. |
| ****3. Health & Safety Records****  * **Examples:**   + Workplace injury reports   + Occupational health assessments   + COVID-19 test results (if required for workplace safety) * **Purpose:** Duty of care under health/safety laws (e.g., **Health and Safety at Work Act 1974**). |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| Biometric Data – Fingerprints, facial recognition data, or retina scans used for security or attendance tracking.  Medical Records – Health conditions, disability status, or vaccination records, unless legally required.  Financial Information – Bank account details for purposes beyond payroll, such as credit checks or loans. |
|  |
|  |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | Copying text from a website into an essay without crediting the author. |
| Provide one example of: Plagiarism |  |
| What are two consequences of copyright infringement and software piracy? | **Legal Trouble** – You could be fined or sued.  **Bad Reputation** – People may lose trust in you or your business. |
| Give three possible consequences for individuals when using pirated software | **Viruses & Hacking** – Pirated software can contain harmful viruses.  **No Updates or Support** – You won’t get security fixes or help if something goes wrong.  **Legal Risk** – You could face fines or punishment for using illegal software. |

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** |
| 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 |
| 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.



|  |  |
| --- | --- |
| Completed lab |  |

# 

# 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.



|  |  |
| --- | --- |
| Completed lab |  |

# 

# 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.



|  |  |
| --- | --- |
| Completed lab | Not working – will try again to see if I can get it to work |

# 

# 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.



|  |  |
| --- | --- |
| Result | Ongoing - |

# 

# 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.

|  |
| --- |
| **Paws & Whiskers: Data Management & Azure Transition**  **1. Scenario Background**  Paws & Whiskers is a growing pet shop that wants to improve business operations by using data analytics. Currently, sales, customer information, and inventory are stored manually or in spreadsheets, making decision-making inefficient. The company plans to move to Microsoft Azure for better data storage, analysis, and reporting.  **2. Data Laws and Regulations**  Handling customer data requires compliance with legal and industry standards to protect privacy and security. The GDPR (General Data Protection Regulation) ensures customer data is collected, stored, and processed securely, requiring businesses to minimize data collection, maintain accuracy, and allow customer access or deletion. The UK Data Protection Act (DPA) 2018 aligns with GDPR and mandates securing data, limiting unauthorized access, and respecting customer rights. Additionally, PCI DSS (Payment Card Industry Data Security Standard) applies to online payments by enforcing encryption, while ISO 27001 provides guidelines for managing data security.  **3. Azure Service Recommendations**  To improve efficiency, Azure SQL Database provides a secure and scalable solution for storing customer, sales, and inventory data, while Azure Blob Storage is ideal for unstructured data like images and receipts. Azure Synapse Analytics helps analyse sales and inventory trends, and Azure Machine Learning predicts customer buying behaviour. To streamline operations, Azure Data Factory automates data collection and integration from multiple sources, improving efficiency and reducing manual workload.   1. **Data Types and Data Modelling**   Organizing data properly improves efficiency and insights. The key data types include Customer Data (names, contact details, purchase history, preferences), Sales Data (transactions, payment methods, sales trends), Inventory Data (product details, stock levels, supplier information), and Pet Information (breed, age, health records, if applicable). A Relational Model will be used to structure this data, linking customers, sales, and inventory through unique IDs (Primary Keys). Additionally, a Data Warehouse approach in Azure Synapse will integrate multiple data sources for advanced analysis and reporting.   1. **Data Storage Formats and Security in Azure**   Efficient data storage and security are essential for protecting business and customer information. CSV is useful for importing and exporting raw data, JSON is ideal for structured customer data and API integrations, and Parquet is optimized for large-scale analytics. To enhance security, Azure Storage Encryption protects data at rest and in transit, Role-Based Access Control (RBAC) restricts access based on user roles, and Multi-Factor Authentication (MFA) adds an extra layer of security to database access.   1. **Additional Considerations**   **Backup and Disaster Recovery**   * **Azure Backup:** Protects data with daily backups. * **Azure Site Recovery:** Ensures recovery in case of failure.   **Data Visualization**   * **Power BI:** Creates dashboards for real-time sales, customer insights, and inventory tracking.   **Future Scalability**   * Azure services can grow with the business, handling larger datasets and complex analytics. |

|  |
| --- |
| **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:

|  |
| --- |
| **Additional Information** |

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.**