

**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 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? | Cloud computing allows users to store, access, and process data over the internet instead of on local servers or personal computers. It enables remote work, real-time collaboration, scalable computing power, data backup, and easy access to applications from anywhere. |
| How can it benefit a business? | Businesses can benefit from cloud computing by reducing IT infrastructure costs, improving flexibility and scalability, enhancing data security, enabling remote work, and increasing operational efficiency. It also allows companies to focus on core activities while cloud providers handle maintenance and updates. |
| What’s the alternative to cloud computing? | The main alternative is on-premises computing, where businesses use local servers and data centers to store and process data. Other options include hybrid computing (a mix of cloud and on-premises) and edge computing, which processes data closer to the source instead of relying entirely on cloud servers. |
| What cloud providers can we use, what are their features and functions? |  **Amazon Web Services (AWS):** Offers computing power, storage, databases, AI, and machine learning services.   **Microsoft Azure:** Provides cloud solutions for businesses, including virtual machines, AI, and IoT services.   **Google Cloud Platform (GCP):** Specializes in big data, AI, and machine learning, with strong integration with Google services.   **IBM Cloud:** Focuses on AI-driven solutions, hybrid cloud options, and security.   **Oracle Cloud:** Best for enterprise applications and database management |

# 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 provides virtualized IT infrastructure over the internet, including servers, storage, and networking. Users can scale resources as needed without investing in physical hardware. | Companies hosting web applications without managing physical servers. A common example is using **Amazon Web Services (AWS) EC2** to host websites or applications. |
| PaaS (Platform as a service) | PaaS offers a development platform that includes tools, middleware, and databases, allowing developers to build, test, and deploy applications without worrying about the underlying infrastructure. | Businesses developing web and mobile applications without managing servers, such as using **Google App Engine** or **Microsoft Azure App Services**. |
| SaaS (Software as a service) | SaaS provides cloud-hosted software that users can access via a web browser without installing anything on their devices. | Services like **Google Workspace (Gmail, Drive, Docs)** or **Salesforce**, which allow users to access tools from anywhere. |

# 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 | A cloud computing model where resources (such as servers and storage) are owned and managed by a third-party provider and shared among multiple customers over the internet. It offers scalability, cost-effectiveness, and easy access.  **When it's appropriate:** Suitable for startups, small businesses, and enterprises that need scalable and flexible computing resources without the cost of maintaining physical infrastructure.  **Real-world example:** **E-commerce businesses** using **Amazon Web Services (AWS), Microsoft Azure, or Google Cloud** for hosting websites, storing customer data, and running applications. |
| Private Cloud | A cloud environment dedicated to a single organization, either hosted on-premises or by a third-party provider. It provides enhanced security, control, and customization.  **When it's appropriate:** Best for organizations that handle sensitive data and require strict security and compliance, such as financial institutions and government agencies.  **Real-world example:** **Banks or healthcare organizations** using a **private cloud** to store confidential customer data while ensuring compliance with regulations like GDPR or HIPAA. |
| Hybrid Cloud | A combination of public and private cloud services, allowing organizations to store sensitive data in a private cloud while leveraging public cloud resources for scalability.  **When it's appropriate:** Ideal for businesses that need both security and flexibility, such as companies with variable workloads or those transitioning to the cloud gradually.  **Real-world example:** **A retail company** using a **private cloud** for handling payment processing while using a **public cloud** for hosting its website during high-traffic shopping seasons. |
| Community Cloud | A cloud infrastructure shared by multiple organizations with similar security, compliance, or operational requirements. It is typically managed by a group of organizations or a third-party provider.  **When it's appropriate:** Suitable for industries with shared regulatory needs, such as government departments, research institutions, or healthcare providers.  **Real-world example:** **Universities and research institutions** using a **community cloud** to share computing resources for collaborative research projects. |

# Day 2: Task 1

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

|  |  |  |
| --- | --- | --- |
| Area | Description | Example |
| |  | | --- | | **Unauthorized Access** | | Accessing a computer system or network without permission | Hacking into someone’s email account without their consent. |
| unauthorized Access with Intent | Accessing a system with intent to commit further crimes (e.g., theft or fraud). | Using unauthorized access to steal sensitive customer data. |
| Impairing Computer Operation | Disrupting or damaging the operation of a computer system or data. | Launching a DDoS attack to take down a website. |

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 |
| **Possession of hacking tools** – Making it illegal to possess tools for hacking. |
| **Extended offenses** – Criminalizing attempts to access systems even if successful access doesn't occur. |
| **Dealing with denial of service attacks** – Strengthening the law around attacks that impair services. |

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. |
| **Three items of data a company can store about an employee**:   * Name * Address * Date of birth |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| 1. **Three items of data that an employer can only store with the employee's permission**:    * Race and ethnicity    * Religion    * Health and medical conditions |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | Using a movie or music file without permission, such as downloading or distributing pirated content. |
| Provide one example of: Plagiarism | Copying and submitting someone else’s essay or article as your own work without attribution. |
| What are two consequences of copyright infringement and software piracy? | Legal action, such as fines or imprisonment  Damage to reputation |
| Give three possible consequences for individuals when using pirated software | Malware infections  Legal consequences  Loss of software support and updates |

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 |
| 2 | 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 |
| 1 | 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 | A screenshot of a computer  AI-generated content may be incorrect. |

# 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 | A screenshot of a computer  AI-generated content may be incorrect. |

# 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 | A screenshot of a computer  AI-generated content may be incorrect. |

# 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 |  |

# 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.

|  |
| --- |
| **Title: Data Management Proposal for "Paws & Whiskers" Pet Shop** **1. Introduction** This report outlines the recommended approach for "Paws & Whiskers" to transition to Microsoft Azure, focusing on data storage, analysis, and compliance with relevant data protection regulations. The goal is to improve the pet shop’s ability to make data-driven decisions through secure, efficient, and scalable cloud solutions. **2. Data Laws and Regulations****GDPR Compliance**  * **Overview:** The General Data Protection Regulation (GDPR) is a regulation in the European Union (EU) that governs how companies collect, store, and process personal data. GDPR applies to businesses handling customer data, even if the company is not based in the EU, as long as the business processes data of EU residents. * **Key Aspects:**   + **Data Minimization:** Only collect the data necessary for specific purposes.   + **Data Subject Rights:** Customers have the right to access, correct, or delete their personal data.   + **Consent:** Explicit consent must be obtained for data processing activities.   + **Storage Limitations:** Data should not be stored longer than necessary. * **Relevance to "Paws & Whiskers":** Since the company will store customer information, including personal details (e.g., name, address, payment information), they must ensure that data is securely stored and processed in compliance with GDPR.  **Data Protection Act (DPA) 2018**  * **Overview:** The DPA 2018 is the UK's implementation of GDPR and provides additional provisions for data protection within the UK. It governs how personal data is collected, stored, and processed, with a focus on the rights of data subjects. * **Key Aspects:**   + **Data Processing Principles:** Personal data must be processed lawfully, transparently, and for specific purposes.   + **Security Measures:** Adequate security measures must be in place to protect personal data.   + **Data Breach Notification:** Any data breaches must be reported to the relevant authorities within 72 hours. * **Relevance to "Paws & Whiskers":** As the business operates in the UK, adherence to the DPA 2018 is crucial, especially when storing and processing sensitive customer data.  **Other Industry Standards**  * **PCI DSS (Payment Card Industry Data Security Standard):** If "Paws & Whiskers" processes payment data (e.g., credit card information), they must comply with PCI DSS to ensure secure transactions. * **Consumer Protection Regulations:** Specific regulations in the pet retail industry may apply to handling customer data, especially when it involves sensitive or health-related information regarding pets.  **3. Azure Service Recommendations****Data Storage**  * **Azure Blob Storage:**   + **Benefits:** Ideal for storing large volumes of unstructured data (e.g., images, documents, logs) in a highly scalable manner. This is suitable for storing inventory images or documents. * **Azure SQL Database:**   + **Benefits:** A relational database service for structured data like customer records, transactions, and product inventories. It supports complex queries and offers high availability, making it a suitable choice for storing customer information, sales transactions, and pet inventory.  **Data Analysis Tools**  * **Azure Synapse Analytics:**   + **Benefits:** This integrated analytics service would help in analyzing sales trends, customer behavior, and inventory management. It combines big data and data warehousing, enabling scalable analysis of business data. * **Azure Machine Learning:**   + **Benefits:** This tool could be used to analyze customer behavior and predict future sales trends, providing "Paws & Whiskers" with insights into customer preferences and purchasing patterns.  **Data Integration and Automation**  * **Azure Data Factory:**   + **Benefits:** Azure Data Factory would automate data collection, integration, and transformation processes, making it easier to handle data from various sources like spreadsheets, databases, and cloud services. It ensures data flows efficiently into the company's data warehouses for analysis.  **4. Data Types and Data Modelling****Data Categories**  * **Customer Data:** Includes personal information such as name, address, phone number, email, and purchase history. * **Sales Transactions:** Contains details of sales such as transaction ID, date, product sold, price, and customer details. * **Pet Inventory:** Data on available pets, their breeds, ages, health status, and prices. * **Product Data:** Details about products sold in the shop, such as pet food, accessories, and toys, including product categories, prices, and stock levels.  **Data Modelling Approach**  * **Relational Model:** A relational database schema will be used to structure customer, sales, pet inventory, and product data into tables.   + **Entities:** Customer, Sales, Pet, Product.   + **Relationships:** A customer can have multiple sales transactions; a sale can involve multiple products; each product belongs to a specific category.   + **Primary Keys:** Unique identifiers for each entity (e.g., CustomerID, ProductID, TransactionID) will be used to establish relationships and ensure data integrity.  **5. Data Storage Formats and Structures in Azure****Data Formats**  * **CSV (Comma-Separated Values):** Recommended for importing raw data (e.g., bulk uploads of customer or transaction data). * **JSON (JavaScript Object Notation):** Suitable for storing structured data like customer preferences or transaction details. * **Parquet:** A columnar storage format ideal for analytics due to its compression and performance advantages when processing large datasets.  **Data Security and Encryption**  * **Encryption in Transit and at Rest:** Azure provides built-in encryption to protect data both while it's being transferred and while stored. Data should be encrypted using Azure Storage Service Encryption (SSE) and Azure SQL Transparent Data Encryption (TDE). * **Access Controls:** Implement Azure role-based access control (RBAC) to ensure that only authorized users can access specific datasets.  **6. Additional Considerations****Backup and Disaster Recovery**  * **Azure Backup:** Ensure that data is regularly backed up using Azure Backup, which provides secure and scalable solutions for disaster recovery. * **Azure Site Recovery:** Use Azure Site Recovery to replicate critical workloads and protect against data center outages.  **Data Visualisation**  * **Power BI:** Azure integrates seamlessly with Power BI, enabling "Paws & Whiskers" to create real-time dashboards for monitoring sales trends, inventory levels, and customer behavior. This would empower the management team with actionable insights.  **Future Scalability**  * **Scalable Services:** Azure’s flexible architecture allows the business to scale its data storage and analysis capabilities as the business grows. With services like Azure Synapse Analytics and Azure SQL Database, the company can easily manage larger datasets and more complex analyses as its customer base expands.  **7. Conclusion** The implementation of Microsoft Azure provides "Paws & Whiskers" with the tools needed to securely store, analyze, and integrate business data. By adhering to relevant data protection laws and leveraging Azure’s advanced services, the company can enhance operational efficiency and make informed, data-driven decisions to support business growth. **References**  * General Data Protection Regulation (GDPR) (EU) 2016/679 * Data Protection Act 2018 (UK) * Azure Documentation: <https://docs.microsoft.com/en-us/azure/> * Payment Card Industry Data Security Standard (PCI DSS) |

|  |
| --- |
| **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.**