

**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 individuals and businesses to store and access data remotely, collaborate in real-time, and scale computing resources based on demand. It supports hosting websites, running applications, performing big data analytics, and enhancing security with automated backups and disaster recovery. |
| How can it benefit a business? | Cloud computing helps businesses by reducing infrastructure costs, improving scalability, enhancing collaboration, and ensuring data security. It enables faster deployment of applications, supports remote work, and offers business continuity through disaster recovery solutions. Businesses can focus on growth without worrying about maintaining hardware. |
| What’s the alternative to cloud computing? | Alternatives to cloud computing include **on-premises infrastructure**, where businesses manage their own servers; **colocation services**, which involve renting space in data centres; **hybrid cloud**, which combines on-premises and cloud solutions; and **edge computing**, which processes data closer to the source for faster response times. |
| What cloud providers can we use, what are their features and functions? | **AWS (Amazon Web Services):** Offers a vast range of services, including computing, storage, and AI, with a global presence.  **Microsoft Azure:** Integrates well with Microsoft products and provides hybrid cloud solutions.  **Google Cloud:** Excels in AI/ML and big data analytics.  **IBM Cloud:** Specializes in AI-driven business solutions and enterprise applications.  **Oracle Cloud:** Focuses on database management and enterprise software solutions. |

# 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 is a cloud computing model that provides virtualized computing resources such as servers, storage, and networking over the internet. It allows businesses to rent infrastructure on a pay-as-you-go basis, eliminating the need for costly on-premises hardware. | IaaS can be used when businesses need to scale their infrastructure without investing in physical hardware. It is ideal for start-ups and growing businesses that require flexibility and cost-efficiency. Businesses can use IaaS to host websites and applications, ensuring high availability and performance. For example, Handlingtraffic spikes during sales for E-commerce. Supporting online learning platforms for education purposes. |
| PaaS (Platform as a service) | **Platform as a Service (PaaS)** is a cloud computing model that provides developers with a complete platform, including infrastructure, development tools, operating systems, databases, and runtime environments, to build, test, and deploy applications. PaaS allows developers to focus on coding and application development without worrying about managing the underlying infrastructure such as servers, storage, and networking. | Developers can use PaaS to build applications in programming environments like Python, Java, or Node.js without worrying about server maintenance.  Companies can use PaaS to automate software deployment processes, improving speed and reducing manual errors.  Organizations can leverage PaaS to develop, test, and scale applications efficiently with built-in monitoring and analytics tools.  Businesses can integrate PaaS with third-party services such as databases, machine learning tools, and DevOps pipelines to streamline workflows. |
| SaaS (Software as a service) | **Software as a Service (SaaS)** is a cloud computing model that delivers software applications over the internet. Instead of purchasing and installing software on individual computers or servers, users can access SaaS applications via a web browser on a subscription or pay-as-you-go basis. The service provider manages all infrastructure, security, updates, and maintenance. | Businesses can use SaaS for customer relationship management (CRM) to manage client interactions, such as using Salesforce.  Companies can adopt SaaS accounting solutions, like QuickBooks or Xero, for financial management.  Employees can collaborate using productivity tools, such as Google Workspace or Microsoft 365, to edit documents and hold virtual meetings.  Organizations can use SaaS for HR management, such as hiring, payroll processing, and performance tracking through tools like Workday.  Individuals can use SaaS solutions like Netflix or Spotify for entertainment purposes. |

# 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 | The **public cloud** is a cloud computing model where computing resources such as servers, storage, and applications are owned and operated by third-party providers and made available to the general public over the internet. Resources are shared among multiple users (multi-tenancy), and users pay on a subscription or pay-as-you-go basis.   * When businesses need cost-effective, scalable solutions without investing in infrastructure. * Ideal for startups, small businesses, or organizations with fluctuating workloads. * When collaboration across locations is required with minimal IT overhead. * Suitable for hosting applications with high demand, such as web applications and streaming services.   **Real-World Example:**   * **Industry:** E-commerce companies like **Shopify**, which host their platforms on AWS or Google Cloud to handle large-scale traffic and seasonal spikes without managing hardware. * **Cloud Providers:** Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP) |
| Private Cloud | A **private cloud** is a cloud computing environment that is exclusively dedicated to a single organization. It can be hosted on-premises or by a third-party provider but remains isolated from public use. Private clouds offer greater control, security, and compliance compared to public clouds.   * When businesses handle sensitive data and require strict security and regulatory compliance (e.g., healthcare, finance). * Suitable for enterprises that want full control over their infrastructure. * When organizations need custom configurations and specialized IT requirements. * When performance and privacy are top priorities.   **Real-World Example:**   * **Industry:** Banks and financial institutions such as **JPMorgan Chase**, which use private cloud solutions to ensure data security and compliance with regulations such as GDPR and PCI-DSS.   **Cloud Providers:** VMware, OpenStack, Microsoft Azure Stack. |
| Hybrid Cloud | A **hybrid cloud** combines both public and private cloud environments, allowing businesses to benefit from the flexibility of the public cloud while maintaining control over sensitive data in a private cloud. It enables data and applications to be shared between both environments, providing scalability and security.   * When an organization needs to keep sensitive operations private but leverage public cloud services for scalability. * Suitable for disaster recovery and business continuity planning. * When businesses want a phased approach to cloud adoption. * Ideal for organizations that require workload portability between environments.   **Real-World Example:**   * **Industry:** Healthcare providers such as **Mayo Clinic**, which store confidential patient data on private cloud while using public cloud for running analytics and processing non-sensitive workloads. * **Cloud Providers:** AWS Hybrid Cloud, Microsoft Azure Hybrid, Google Anthos. |
| Community Cloud | A **community cloud** is a shared cloud environment designed for a group of organizations that have common requirements such as security, compliance, or business goals. It provides shared infrastructure but with greater privacy and customization than the public cloud.   * When multiple organizations within the same industry need to collaborate and share resources securely. * Suitable for government agencies, educational institutions, or healthcare organizations that follow the same regulations. * When cost-sharing and resource optimization are important factors. * When organizations with similar security and privacy concerns want to share infrastructure.   **Real-World Example:**   * **Industry:** Government organizations, such as the **European Union institutions**, which share a community cloud infrastructure to ensure compliance with regional regulations. * **Cloud Providers:** IBM Government Cloud, Microsoft Government Community Cloud, AWS GovCloud. |

# Day 2: Task 1

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

|  |  |  |
| --- | --- | --- |
| Area | Description | Example |
| Unauthorized Access to Computer Material | Gaining access to a computer or its data without permission. This includes hacking into systems or accessing files you are not authorized to view. | Logging into someone else's email account by guessing their password without their consent. |
| Unauthorized Access with Intent to Commit or Facilitate a Crime | Accessing a computer system unlawfully with the intention to commit further crimes, such as fraud, theft, or data theft. | Hacking into a bank's database to steal customer information for identity theft or fraudulent activities. |
| Unauthorized Acts with Intent to Impair or with Recklessness | Deliberately or recklessly causing damage to computer systems or data, such as through malware, viruses, or denial-of-service (DoS) attacks. | Spreading ransomware to lock users out of their systems and demanding payment to restore access. |

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 |
| The act makes it illegal to create, supply, or distribute tools or software designed to commit computer-related crimes, such as hacking, distributing malware, or conducting unauthorized access. |
| The act raises the maximum penalties for certain offenses, particularly for actions causing significant harm to systems or data, such as serious hacking or cyberattacks. |
| The act explicitly criminalizes denial-of-service (DoS) attacks, where individuals or organizations intentionally disrupt access to a system by overwhelming it with traffic or requests. |

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. |
| National Insurance Number |
| Tax code |
| Name |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| Race and Ethnicity |
| Political membership or opinions |
| Religion |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | Downloading and sharing a movie or song online without the permission of the copyright owner. |
| Provide one example of: Plagiarism | Copying paragraphs from an academic article and presenting them as your own work without proper citation. |
| What are two consequences of copyright infringement and software piracy? | 1. **Legal Penalties:** Individuals or organizations may face lawsuits, heavy fines, or even imprisonment for violating copyright laws. (Example: Paying damages to a software company for using unlicensed software.) 2. **Financial Losses for Creators and Businesses:** Software companies, musicians, authors, and filmmakers lose revenue when their work is copied or shared without authorization. |
| Give three possible consequences for individuals when using pirated software | 1. **Security Risks:** Pirated software may contain malware or viruses that can compromise personal data, damage devices, or expose sensitive information to hackers. 2. **Lack of Updates and Support:** Users of pirated software often cannot access official updates, bug fixes, or customer support, which may lead to performance issues or vulnerabilities. 3. **Legal Action:** Individuals caught using pirated software may face lawsuits, fines, or even criminal charges for violating intellectual property laws. |

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 |
| 2 | 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.



|  |  |
| --- | --- |
| 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 |  |

# 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.

|  |
| --- |
| **Paws & Whiskers: Transitioning to Microsoft Azure for Data Management** 1. Scenario Background "Paws & Whiskers" is a growing pet shop that wants to enhance its business operations by leveraging data for informed decision-making. Currently, sales, customer information, and inventory data are collected manually or stored in spreadsheets, making data management inefficient and prone to errors. This approach limits the company's ability to gain insights, forecast demand, and improve customer experience. To address these challenges, the company plans to transition to **Microsoft Azure**, which offers cloud-based solutions tailored to business needs, ensuring better data security, accessibility, and scalability. 2. Data Laws and Regulations To ensure compliance with legal and industry standards, "Paws & Whiskers" must adhere to the following data protection laws and regulations: **General Data Protection Regulation (GDPR)**  * **Data Collection and Processing:** GDPR mandates that businesses must collect customer data lawfully, with clear consent, and process it transparently. * **Data Storage:** Personal information must be securely stored, with access restricted to authorized personnel. * **Right to Access and Erasure:** Customers have the right to access their data or request its deletion upon request. * **Data Breach Reporting:** In the event of a data breach, "Paws & Whiskers" must notify the relevant authorities within 72 hours. * **Data Minimization Principle:** Businesses should only collect and retain data necessary for the intended purpose.  **Data Protection Act (DPA) 2018**  * As the UK's implementation of GDPR, **DPA 2018** reinforces customer data protection requirements. * Ensures businesses take measures to safeguard personal information and handle it responsibly. * Stipulates penalties for non-compliance, which could result in fines or legal action. * Covers special categories of personal data that require extra protection, such as payment information.  **Other Industry Standards**  * **Payment Card Industry Data Security Standard (PCI DSS):** If "Paws & Whiskers" processes credit/debit card payments, compliance with PCI DSS is necessary to protect financial transactions. * **Consumer Rights Act 2015:** Ensures that customer data used for transactions is accurate and reliable. * **ISO/IEC 27001:** A globally recognized standard for information security management, ensuring data security and risk management best practices.  3. Azure Service Recommendations To support "Paws & Whiskers" in managing its data efficiently, the following Azure services are recommended: **Data Storage**  * **Azure SQL Database:** A **fully managed relational database** suitable for storing customer transactions, inventory details, and sales records. It ensures **data consistency, security, and scalability**, making it ideal for structured business data. * **Azure Blob Storage:** Ideal for storing unstructured data, such as images of products and scanned invoices, ensuring cost-effective and scalable storage. * **Azure Cosmos DB:** A globally distributed, multi-model database service that allows for real-time data access, especially useful for applications that require fast and scalable data retrieval, such as an online pet shop.  **Data Analysis Tools**  * **Azure Synapse Analytics:** Provides **advanced data analytics and reporting** to track customer purchasing patterns and sales trends, supporting better business insights. * **Azure Machine Learning:** Can be used to analyse customer behaviour, enabling personalized marketing campaigns and sales predictions, improving customer retention. * **Azure Stream Analytics:** Helps process real-time data, which can be beneficial for monitoring inventory stock levels dynamically, ensuring stock availability.  **Data Integration and Automation**  * **Azure Data Factory:** Allows for **automated data integration** by consolidating information from spreadsheets into a centralized database, improving efficiency. * **Azure Logic Apps:** Automates routine tasks like sending order confirmations or inventory updates, reducing manual workload. * **Azure Event Grid:** Manages event-driven architecture to streamline communication between different services, enabling seamless automation.  4. Data Types and Data Modelling**Data Categories**  * **Customer Data:** Name, contact details, purchase history, and preferences, which help in targeted marketing and personalized services. * **Sales Data:** Transaction records, payment methods, order status, and customer loyalty data, ensuring accurate financial reporting. * **Inventory Data:** Product details, stock levels, supplier information, and product categories, helping with supply chain management. * **Employee Data:** Work schedules, sales performance, and staff training records, supporting workforce planning.  **Data Modelling Approach**  * A **relational database model** is recommended, using structured tables to organize data efficiently, supporting better reporting and analysis. * **Tables and Relationships:**   + Customers (**Primary Key: CustomerID**)   + Transactions (**Primary Key: TransactionID**, **Foreign Key: CustomerID**)   + Inventory (**Primary Key: ProductID**)   + Employees (**Primary Key: EmployeeID**)   + Suppliers (**Primary Key: SupplierID**, **Foreign Key: ProductID**)  5. Data Storage Formats and Structures in Azure**Data Formats**  * **CSV:** Suitable for importing/exporting structured data, such as sales records and customer lists. * **JSON:** Used for storing structured data in web applications and APIs, ensuring flexibility. * **Parquet:** Optimized for data analytics in Azure Synapse, providing efficient storage and query performance, useful for large datasets. * **XML:** Used for configurations and system integrations, facilitating interoperability between systems.  **Data Security and Encryption**  * **Azure Key Vault:** Protects sensitive data by securely storing encryption keys, preventing unauthorized access. * **Role-Based Access Control (RBAC):** Ensures only authorized personnel access specific data, reducing security risks. * **Transparent Data Encryption (TDE):** Secures stored data from unauthorized access, ensuring compliance with regulations. * **Azure Defender for SQL:** Provides advanced threat protection to detect vulnerabilities and potential cyber threats, enhancing security.  6. Additional Considerations**Backup and Disaster Recovery**  * **Azure Backup:** Automates daily backups to protect against data loss, ensuring business continuity. * **Azure Site Recovery:** Ensures business continuity in case of system failures, minimizing downtime. * **Geo-Redundant Storage (GRS):** Provides extra redundancy by replicating data across multiple Azure regions, ensuring high availability.  **Data Visualization**  * **Power BI:** Generates real-time dashboards for monitoring sales, customer trends, and inventory levels, enabling data-driven decision-making. * **Azure Data Explorer:** Allows for real-time exploration of large datasets to identify trends quickly, enhancing strategic planning.  **Future Scalability**  * Azure's **scalable infrastructure** ensures that as the business grows, it can handle increased data volume and analytics complexity. * **Elastic Pooling in Azure SQL Database:** Helps in cost-effective scaling by allowing multiple databases to share resources, optimizing performance. * **Azure Auto-Scaling Services:** Dynamically adjusts computing resources based on workload demand, ensuring efficient resource utilization.  Conclusion By transitioning to Microsoft Azure, "Paws & Whiskers" can modernize its data storage, streamline operations, and leverage analytics for better decision-making. Compliance with **GDPR, DPA 2018, PCI DSS, and ISO/IEC 27001** ensures customer data security, while **Azure SQL Database, Synapse Analytics, and Data Factory** optimize storage, analysis, and automation. Azure’s scalability, security features, and backup solutions future-proof the business, making data management more efficient and reliable. With robust data protection, automation, and analytics capabilities, "Paws & Whiskers" is well-positioned for long-term growth and success in an increasingly data-driven world. |

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