

**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 can allow billions of people to access the same service/resources at the same time over the internet. |
| How can it benefit a business? | It can reduce costs and by providing the same services, the business can avoid confusion because people can see other employees/actors’ changes.  centralising data and digital services on remote server where all employees and stakeholder can access data and information. |
| What’s the alternative to cloud computing? | On-premises server computing. Some services like Tableau, Odoo offer this capability at a different fee structure. |
| What cloud providers can we use, what are their features and functions? | Amazon web services [AWS], Microsoft Azure, Google Cloud Platform. They all share the same features by offering different types of servers and configurations Virtual Machines, virtual servers, dedicated or shared servers. |

# 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) | This provides a full virtual computing service for servers, storage, networking, security. | Typically, when a large organisation needs to create a complex solution and needs control over networking, processing, AI features, and analytics resources, they would use most features on services like AWS, GCP, Azure, etc. |
| PaaS (Platform as a service) | This is a service where an organisation would use that allows them to develop and manage a software service on top of a managed hardware and software service. | Usually, start-up companies or mid-size organisations that don’t have the resources (budget) for security and system admins, can use this service to develop their solution while using the platforms tools to easily deploy. |
| SaaS (Software as a service) | These are services delivered over a cloud that offer simple software solutions with one aim. | Cloud storage services like OneDrive, iCloud, Google Drive are SaaS, where you can store files that you want to share with a group of friends or colleagues. Developer use services like Github to share code with their team members, organisations, and services. |

# 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 | This is the usual cloud service where an organisation offers cloud services to the general public. It has multiple tenants for example AWS has clients/tenants like Netflix, Apple, etc. using the same service and individuals like me can also use this service in the capacity that I can afford. |
| Private Cloud | In this setup, a client will have dedicated hardware and software resources. An organisation may rent space from a 3rd party provider for exclusive access to these resources and connect to it over the internet through a virtual private network (VPN). Defence and border organisations typically use this to add an immediate high-level of security. |
| Hybrid Cloud | In a hybrid setup, an organisation will use both private and public clouds. This is useful for organisations that deal with public records and need to keep certain activities private (like staff/employee, ERPs etc.) Most large organisations operate like this, Microsoft, Apple, NHS, Government departments. |
| Community Cloud | A community cloud is a shared cloud that can be used by organisations with common goals or similar functions. Charities can use the same cloud resources to reduce costs and have data security. |

# 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 | Even if no harm is done, accessing a computer system without proper authorization is still illegal. | Hacking or using someone else’s credentials. |
| Unauthorised access with intent to commit further offences. | This refers to entering a computer system to steal data or destroy a device or network. | Recently, a disgruntled IT staff vandalised the network at the British museum, which prevented visitors from having access to some of the displays and hurt revenue for the museum. |
| Unauthorised modification of computer material | It is a criminal offense to delete, alter, or corrupt data without proper authorization, or introducing malware, with the intent to cause damage. | Installing a virus or malware or taking of data to change the system. |

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 |
| Making, Supplying, or Obtaining Articles for Use in Computer Misuse Offences  This expanded the scope beyond just the act of unauthorized access itself, targeting the preparatory stages of such offenses. |
| Impairment of Operation of a Computer:  This addresses attacks like denial-of-service (DoS) attacks and the spread of malware that disrupt computer systems. |
| Increased Penalties  This allows the act to let users and potential hackers know how that activities are serious crimes and they will be prosecuted for committing offences. |

Look at the below website to answer the questions:

<https://www.gov.uk/personal-data-my-employer-can-keep-about-me>

|  |
| --- |
| Write down three items of data which a company can store about an employee. |
| Name |
| Date of birth |
| National insurance number |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| Genetics |
| Health and medical conditions |
| Religion |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | If someone uses a part of a film or song without the owner’s permission in their video. |
| Provide one example of: Plagiarism | When a journalist copies a news report and doesn’t change it or reference it and presents the report as their own. |
| What are two consequences of copyright infringement and software piracy? | Copyright: legal and financial consequences. Piracy: legal consequences and security consequences. |
| Give three possible consequences for individuals when using pirated software | Cyber Security consequences: hacking, viruses, worms, etc. Criminal charges may be laid on the individual. An individual may be fired from their job and face reputational damage. |

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 |
| 1 | Any product, digital or otherwise, must be fit for the purpose it is supplied for |
| 2 | Unauthorised modification of computer material is illegal |
| 2 | It is illegal to create or use a hacking tool for penetration testing |
| 7 | 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 |
| 1 | It is illegal to distribute an illicit recording |
| 6 | Personal data may not be kept longer than necessary |
| 2 | 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  Description automatically generated |

# 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 computer screen with a white box  Description automatically generated |

# 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 | Lab didn’t work even after many tries.  A screenshot of a computer  Description automatically generated |

# 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 | A screenshot of a computer  Description automatically generated  I failed badly, but I will continue with more study and more attempts to get 50% of the certification exam. |

# 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.

|  |
| --- |
| Executive Summary Paws & Whiskers, a growing UK pet shop, currently manages sales, customer, and inventory data manually or via spreadsheets. This report outlines the challenges of the current system, defines business objectives, and presents a cost-effective solution using Microsoft Azure to centralise data storage, streamline analysis, and support data-driven decision-making—all within a small to medium-sized technology budget. Introduction To make data-driven decisions, Paws & Whiskers will have to centralise and digitise their data. Microsoft Azure presents them with this opportunity. However, to make this move, we would have to consider all the dimensions of migrating to Azure.   Data Laws and Regulations According to Microsoft’s trust centre, Azure meets all data governance and compliance standards like UK GDPR, Data Protection Act, Privacy and Electronic Communications Regulations for marketing, Data minimisation, etc.  However, those compliances are between Azure, the service provider, and the client, in this case Paws and Whiskers. To ensure full adherence to these data laws and regulations between Paws and Whiskers and its customers, a data policy document will be created guiding the organisation on handling privacy concerns following British values, safeguarding, and “prevent” policies. Actionable steps Frequent internal audits (monthly) to ensure full compliance will be done. An annual report of the internal audits will be published in the company’s trust centre on the website for transparency.  There will be an implementation of access controls and data encryption to keep data secure.  Additionally, the data privacy policy will have contingency plans for what to do in case of a data breach. Although there will be full adherence to implementing robust security measures, considering the UK's National Cyber Security Centre (NCSC) guidance.   Microsoft Azure Service Recommendations The objective is to find the most suitable configuration of Azure services that fulfils these factors:   * Centralise Data Storage: Transition from disparate spreadsheets to a unified cloud-based system. * Streamline Data Processing: Automate data ingestion and transformation to reduce manual effort. * Improve Reporting & Analysis: Deploy interactive dashboards for real-time business insights. * Enhance decision-making: Provide management with timely and accurate data to drive sales, inventory management, and customer engagement. * Maintain Regulatory Compliance: Ensure data protection in line with UK GDPR and DPA standards. * Affordable cost: pay-as-use model platform can improve    Databases, Data Modelling, and Data TypesData Ingestion: Use Azure Data Factory to import data from spreadsheets and other sources.  Schedule regular ETL (Extract, Transform, Load) processes to clean and integrate data.  This is needed when converting legacy data from various sources, such as Excel and other digitised records. Data Storage:  * **Implement Azure SQL Database (UK South or UK West Region):**   + For structured transactional data (sales, customer information), as a centralised repository for structured data. * **Azure Blob Storage (UK South or UK West Region):**   + For storing unstructured data like product images, documents, and backups.   + **Azure Data Lake Storage Gen2 (UK South or UK West Region):**     - For storing various data types (inventory, logs, potentially images) in a scalable and cost-effective manner.  Data analysis and reporting:  * Leverage Power BI for creating interactive dashboards and reports, enabling management to view sales trends, customer insights, and inventory levels in real time. * Integrate Power BI with Azure SQL Database to allow seamless access to updated data.  Security and Compliance: Azure’s built-in security features (encryption, access controls, and regular audits) ensure data protection and support UK GDPR, DPA, and PECR compliance requirements. Data Types and Modelling  * **Objective:** Define data types, create data models, and ensure data consistency, considering UK-specific data requirements. * **Data Types:** * **Structured Data:** sales transactions, customer details (name, address, postcode), product information. * **Semi-structured Data:** Inventory data in JSON, product reviews. * **Unstructured data:** product images, customer feedback documents, logs. * **Data Modelling:** * **Relational Modelling (Azure SQL Database)**: Design normalised relational schemas for sales and customer data, considering UK address formats and postcode handling.    Data Storage and Formats in Microsoft Azure **Objective:**  Choose appropriate storage formats and optimise storage for performance and cost, keeping data optimised for delivery to the UK.   * **Storage Formats:** * **Azure SQL Database:** Relational tables with appropriate data types; INT, NVARCHAR, etc. * NVARCHAR was chosen over VARCHAR because it is unicode and offers support for multiple language character types. For example, customers can spell their names according to their cultures. * **Azure Data Lake Storage Gen2:** Parquet for efficient analytics; CSV, JSON for data ingestion. * Parquet because it is an open-source cloud data format and is gaining support as a standard for future analysis. * **Azure Blob Storage:** Object storage for various file formats. * **Storage Optimisation:** * **Partitioning:** Partition data based on Azure blob storage types to improve storage and query performance. * **Tiering:** Utilise Azure's storage tiers (hot, cool, and archive) for cost optimisation based on data access frequency.  5. Implementation Roadmap **Phase 1:**  Assessment and planning requirements  Define:   * key performance indicators (KPIs). * Reporting requirements. * Data Audit: Review raw data. * Cost Estimation: Determine Azure services and pricing (leveraging the pay-as-you-go model ideal for small to medium-sized businesses).   **Phase 2:**  Data Migration:   * Use Azure Data Factory to ingest historical data * Setup Set Up Azure Environment: * Provision of Azure SQL Database and Azure Data Factory resources. * Configure network security and identity management. * Initial Dashboard Creation: Develop basic Power BI reports to validate data accuracy and provide early insights.   **Phase 3:**  Testing and training testing:   * Validate the data integration processes. * Put security controls. * Report accuracy. * User Training: Train staff how to use new dashboards and manage data within Azure. * Feedback from management and staff to fine-tune dashboards and data processes.   **Phase 4:**   * Go-live and continuous improvement full roll-out: * Transition all business processes to the new system. * Ongoing monitoring to review performance and adjust the solution as needed. * Scalability planning for future enhancements as business needs grow.  6. Additional ConsiderationsFinancial Considerations Budget Alignment  * The proposed solution leverages Azure’s cost-effective, scalable services tailored for small to medium-sized businesses. * Pay-As-You-Go Model: minimises upfront capital expenditure and scales with business growth. * Potential Savings: Reduced manual labour and improved operational efficiency can lead to long-term cost savings.  For future options: Scalability  * **Data Modelling Options:**   + **Dimensional Modelling (Azure Synapse Analytics):** Create star or snowflake schemas for analytical reporting.   + **Data Lake Schema (Azure Data Lake Storage Gen2):** Implement a data lake architecture (e.g., medallion architecture) to organise and process data.  Risk Assessment and Mitigation  * Key Risks * Data migration errors * Mitigation thorough testing with a phased rollout. * User Adoption Challenges: * Mitigation: Comprehensive training sessions and user-friendly dashboard design.  7. Conclusion Transitioning to Microsoft Azure offers Paws & Whiskers a robust, scalable, and secure platform to transform its data operations. By centralising data storage, automating data processing, and deploying real-time reporting with Power BI, the business can gain actionable insights, improve efficiency, and support strategic decision-making—all while staying within a modest budget and maintaining compliance with UK data protection standards. Recommendation: Start with a phased implementation plan, starting with a pilot phase to validate the new system, followed by a full-scale rollout to ensure a smooth transition from manual processes to a fully automated Azure-based solution. References <https://learn.microsoft.com/en-us/compliance/regulatory/gdpr-dpia-azure>  <https://www.microsoft.com/en-us/trust-center/privacy/gdpr-overview>  <https://ico.org.uk/media2/migrated/4026838/dpia-microsoft-office-365-dec-2022-005.pdf>  Google Gemini Flash 2.0 for structure writing  Prompt  You are a senior data analyst, look at this scenario and come up with a framework to organise the following; data laws and regulations, Microsoft azure service recommendations, data types and modelling, data storage and formats in Microsoft Azure, and additional recommendations in the areas of data visualisation, backup, disaster recovery, scalability, and elasticity. ""Paws & Whiskers" is a growing pet shop that aims to improve its business by analysing sales, customer information, and inventory data. Currently,…”  ChatGP for content  Prompts:   * How does Microsoft Azure adhere to UK GDPR, DPA, PECR, and follow the UK's Information Commission Office * You are a senior data analyst, design a report to for this scenario taking into account that it's a UK business with a technology budget of a small to medium-sized company. "Paws & Whiskers" is a growing pet shop that aims to improve its business by analysing sales, customer information, and inventory data…” |

|  |
| --- |
| **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:

I learned a lot from the lectures of how the practical cloud environment might look and work. I intend to do more studying, setup projects (data gathering), and port them into Azure cloud and Fabrics to play with them on a free trial in 2 to 3 months’ time when I build my skills up in Power BI, Python, SQL and NoSQL databases.

I finally understand how Core values apply to data, applications and the cloud.

**Samples of notes**

**Core Value**

* Prevent
* Safeguarding
* British Values

MS Azure Resource Manager Terms

* Resource Provider
* Services
* Resource group = container
* **DATA WAREHOUSE**
* Data Lake
* Data Lakehouse

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