

**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 is renting computer resources via the internet. This can give a lot of benefits.  A lot of apps back in the day required to save data in physical storage on mobile or pc. Nowadays, the most of apps' data gets saved in cloud storage which app can be used and have access from any device any place. Gives no need of data transfer process. |
| How can it benefit a business? | It can benefit from price reductions because there is no need to purchase physical storage and network devices. Also, fewer human resources are needed as the cloud service platform provides support or maintenance service depending on the type of product. Also, the platform can add or reduce resources on demand, which will help control budgets or handle emergency situations. |
| What’s the alternative to cloud computing? | An alternative to cloud computing could be on-premises which is to build its infrastructure on site. Own data storage centre. Also, there is a hybrid solution which is a mixture of on-premises and cloud computing. |
| What cloud providers can we use, what are their features and functions? | There are many cloud service providers and AWS, Microsoft and Google are the most well-known companies. Features are very similar to any providers. Can build servers, storage and systems by customising your resources then can use different types of modules and apps provided to perform your tasks. |

# 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) | Renting resources for IT infrastructure from cloud service providers such as server, storage and network resources. Applications, OS and system will be managed by the user. | AWS EC2, renting infra or expand resources for a better web hosting environment to manage system for higher traffic. |
| PaaS (Platform as a service) | Developers can purchase OS, development tools or database types of platforms. Infrastructure will be managed by the service provider, which developers can focus on development tasks. | Many application developers use Google App Engine so they can develop and distribute the apps and updates. |
| SaaS (Software as a service) | This is a service to use complete software as a service via the internet. Does not need separate installation or management. Has great accessibility and great UX. Service providers manage all the updates and system management. | Companies use O365 and Slack type of software to provide easy access for employees as work tools. |

# 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 | It is a cloud environment open to public users. Could service providers manage the infrastructure and anybody can access the service and use resources  Kaggle – Kaggle opened its resources to public users to download and upload useful dataset and studies regarding data analytics and machine learning. |
| Private Cloud | A cloud environment for only a specific group. It allows access for authorised users only and stores credential data. It gives a strong security level.  HSBC - HSBC stores their customer data in private cloud storage created by their own data centre. |
| Hybrid Cloud | A cloud environment with both private and public. A large set of data will be stored in the public sector, and credential data will be stored in the private sector. The most common type of environment used. Have strong flexibility but require a high level of managing skills.  NHS - NHS stores patients' data in their private cloud and have other data in Azure Public Cloud for own data analytics purpose. |
| Community Cloud | A cloud environment is managed by multiple organisations for a common purpose. Has a stronger security level than the public and can share the price which gives less budget pressure than private.  Universities – Oxford, Cambridge and other universities create community clouds to share their studied data and resources for subjects they are studying together. |

# 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 | Accessing systems without proper authorisation. This is an illegal act. | Log in some else’s email or sns account via stealing password. |
| Unauthorised Access with Intent to Commit Further Offences | Intended action committed after unauthorised access to the systems. | Access bank system with intent of steal money from digital accounts. |
| Unauthorised Acts with Intent to Impair Computer Operation | Take illegal harmful act to computer systems on purpose. | DDOS attack websites to harm the business operation. |

The Computer Misuse Act 1990 is an act where an individual can be criminalised because of a computer-related offense. Describe three extra powers that the Police and Justice Act 2006 (Computer Misuse) has added.

|  |
| --- |
| Description |
| Consumer Right Act 2005   * Satisfactory quality – no faulty or damaged * Fit for purpose * As described – matching any description given to you or samples shown to you at the time of purchase |
| Data Protection Acts 1998   * The ethical use of personal data * Keeping every individual’s personal data secure |
| GDPR   * To protect not just the UK but to protect the data of European citizens and regulate organisations’ data managing protocol. * Specific regulations under the law such as consent is essential to contain or use personal data of EU citizens, in any case of data breach have 72 hours to report. |

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 |
| Address |
| Date of Birth |

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

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | A youtuber uses a famous song in his video without permission and makes profits from the video. This is an example of infringing copyright law. |
| Provide one example of: Plagiarism | A web designer used an illustration which is created by someone else, use it for own work without permission or referencing. |
| What are two consequences of copyright infringement and software piracy? | Legal penalty – both cases can be legally charged and can be punished under related laws.  Financial and Reputational Damage – an individual or organisation can have financial loss or damage to its reputation by both consequences. Companies can lose trust from their customers. |
| Give three possible consequences for individuals when using pirated software | Legal Consequences – use of software with crack to avoid license issue can be legally charged by the company.  Security Risks – illegal software often contains viruses or malware, which can cause data breaches.  Lack of Updates and Support – illegal software cannot be supported by company for the latest updates and features. |

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** |
| 7 | 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 | For the account issue, the demonstration for fabric was shown and watched. |

# 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 | There has been a constant error since Thursday. |

# Day 4: Task 2

#### **1. Scenario Background**

"Paws & Whiskers" is a growing pet shop that aims to improve its business by analysing sales, customer information, and inventory data. Currently, the data is collected manually or stored in spreadsheets. Management is interested in transitioning to Microsoft Azure to streamline data storage, analysis, and reporting, enabling them to make data-driven decisions.

#### **2. Data Laws and Regulations**

Identify and explain the data laws and regulations relevant to handling customer data within the proposal. Ensure you cover the following points:

* **GDPR Compliance**: Highlight the importance of adhering to the General Data Protection Regulation (GDPR), particularly as it relates to storing and processing customer information.
* **Data Protection Act (DPA) 2018**: Outline how the DPA 2018 may affect the way "Paws & Whiskers" collects and stores data, ensuring compliance with UK laws on data privacy.
* **Other Industry Standards**: Research any additional data protection standards or regulations that may apply to pet shop data, particularly if they involve sensitive or payment information.

#### **3. Azure Service Recommendations**

Recommend Microsoft Azure services that would suit the company’s data analysis needs and explain why these services are suitable. Your recommendations should include:

* **Data Storage**: Identify suitable storage options, such as **Azure Blob Storage** or **Azure SQL Database**, and discuss the benefits of each for storing large datasets, including inventory, sales transactions, and customer details.
* **Data Analysis Tools**: Recommend tools such as **Azure Machine Learning** for customer behaviour analysis or **Azure Synapse Analytics** for analysing sales trends.
* **Data Integration and Automation**: Explain how services like **Azure Data Factory** could automate data collection and integration processes, improving efficiency.

#### **4. Data Types and Data Modelling**

Define the types of data "Paws & Whiskers" will need to work with and describe your approach to data modelling:

* **Data Categories**: Identify key data types, such as customer demographics, transaction history, pet inventory, and product categories.
* **Data Modelling Approach**: Outline how you would structure this data using a relational model or a data warehouse approach, considering factors like tables, entities, relationships, and primary keys.

#### **5. Data Storage Formats and Structures in Azure**

Discuss how you would store data within Azure and the formats you would recommend:

* **Data Formats**: Specify recommended formats (e.g., CSV for raw data imports, JSON for structured data, Parquet for analytics) and explain why these formats are suitable for specific data types.
* **Data Security and Encryption**: Include recommendations for securing data using Azure’s built-in encryption features and access controls to ensure compliance with data privacy regulations.

#### **6. Additional Considerations**

Provide any other considerations that might enhance data handling and efficiency in Azure, such as:

* **Backup and Disaster Recovery**: Outline a backup plan using **Azure Backup** or **Azure Site Recovery** to safeguard against data loss.
* **Data Visualisation**: Discuss potential use of **Power BI** within Azure for creating dashboards that provide management with real-time insights into sales and customer trends.
* **Future Scalability**: Comment on how Azure services can scale as the business grows, accommodating larger datasets and more complex analyses.

### **Submission Guidelines:**

1. **Structure**: Ensure your report is well-organised, with sections for each task (e.g., Data Laws, Azure Services, Data Types, etc.).
2. **Formatting**: Include headings, bullet points where appropriate, and any visuals or diagrams that support your explanations.
3. **References**: Cite any resources or regulations referenced in the report.
4. **Length**: Aim for 1500-2000 words.

|  |
| --- |
| **Data Laws and Regulations**  **GDPR**  This regulation protects personal data in the EU region. The company must have consent from the client to store, analyse, or use it for marketing purposes.  **Data Protection Act 2018**  DPA 2018 is a data law that the company should be aware to collecting and using data in the UK. It is important to know that collecting customer information for sales operations is fine, but using the data for other purposes always needs clients' consent.  **ISO 27001**  This is the standard level of international data security. ISO 27001 can be used as an option to enhance customer data protection. Asure supports ISO 27001, so Azure services can improve data security.  **Data Storage**   * Azure SQL Database - Azure SQL Database is ideal for structured data such as customer details or order details. Since Paws & Whiskers has original data in the spreadsheet, can turn them into query able database system. Use of SQL based database give benefit such as getting “The best sold product last season” or “The most repeat customer”. Azure SQL Database provides a basic security level for GDPR and DPA 2018.   **Data Analysis Tool**   * Azure Synapse Analytics - Azure Synapse Analytics is recommended for the company as it is an integrated analytics service that combines big data and daya warehousing, assist for perfect analysing sales trends and inventory patterns. Sales or inventory data stored in the Azure SQL database can be used with a query to extract specific data for business analysis. Also, it can sync to PowerBI to create a visualised dashboard with live data updates. The service provides great scalability, so there will be no performance down along with business scale growth.   **Data Integration and Automation**   * Azure Data Factory - This service provides cloud-based ETL sync data to Azure storage and analysis services automatically. Can sync sales system with Azure SQL database, send data collected from sales system to analysis system lively. These automated jobs, which can be set up with Azure Data Factory, will provide less manual work for employees.   **Data Types and Data Modelling**  **Data Types**   * Customer Demographics - name, address, contact number, email, age and gender. * Order History - purchase date, product name, purchase quantity, price, payment method, etc. * Stock Inventory - product ID, product name, quantity, supplier, Batch date, etc * Product Categories - Category name, brand, price range, etc   **Data Modelling**  Build a relational model between entities. Building a relation data model is efficient for using Azure SQL Database.   * Customer ↔ Order * Order History ↔ Inventory * Product ↔ Inventory   **Data Storage Formats and Structures in Azure**  **Data Formats**  CSV   * Beneficial for importing customer data or sales history data to Azure from a spreadsheet or external system. * CSV files can be easily read by any type of platform, such as Excel or POS machines. This means it is advantageous for migrating primary stage data created to the system.   **JSON**   * JSON is structured data, unlike CSV. * Can save data by customer, which can contain all information needed such as customer → order history → product * It is more flexible to update or edit the data. Easy to add attributes to the dataset.   **Additional Considerations**  **Back-up and Disaster Recovery**  Azure Backup   * It is important to have constant back up and various back up times to save data for businesses under any situation that might cause chaos such as an external attack or server downtime. * Set multiple back up rules such as everyday → no edit for 7 days → no edit for 30 days to have secured archive structure. * Azure Backup is easy to set up and provides many automated task functions   **Azure Site Recovery**   * This is the system to copy data to other region server and create back up under disaster situation. * Set back up region not same as main region set. * To have Rocovery Time Objective as 1 hour to regenerate the system back after disaster.   **Data Visualisation**  **PowerBI**   * Can use the PowerBI integration function to connect the Azure SQL Database and Synapse Analytics with PowerBI to create a visualised dashboard. * The dashboard can present business core values such as sales trends, customer analysis, product analysis and Inventory status. * Snyc PowerBI and the Azure database will provide live updates of data to the dashboard, giving easy insight for essential business aspects to be checked and reviewed. * It will save time to make decisions and allow easy cooperation between different teams in the company.   **Future Scalability**  **Expand Storage Volume**   * Always can give more resources to storage. It is on demand service so can add anytime under budget situation of the company.   **Azure Synapse Analytics resource upgrade**   * Can upgrade resources of Synapse Analytic service to bigger warehouse unit to complete more complicated query tasks. |

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