

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: Joshua Blofield |
| Course Date: 17/11/2025 |
|  |

**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)

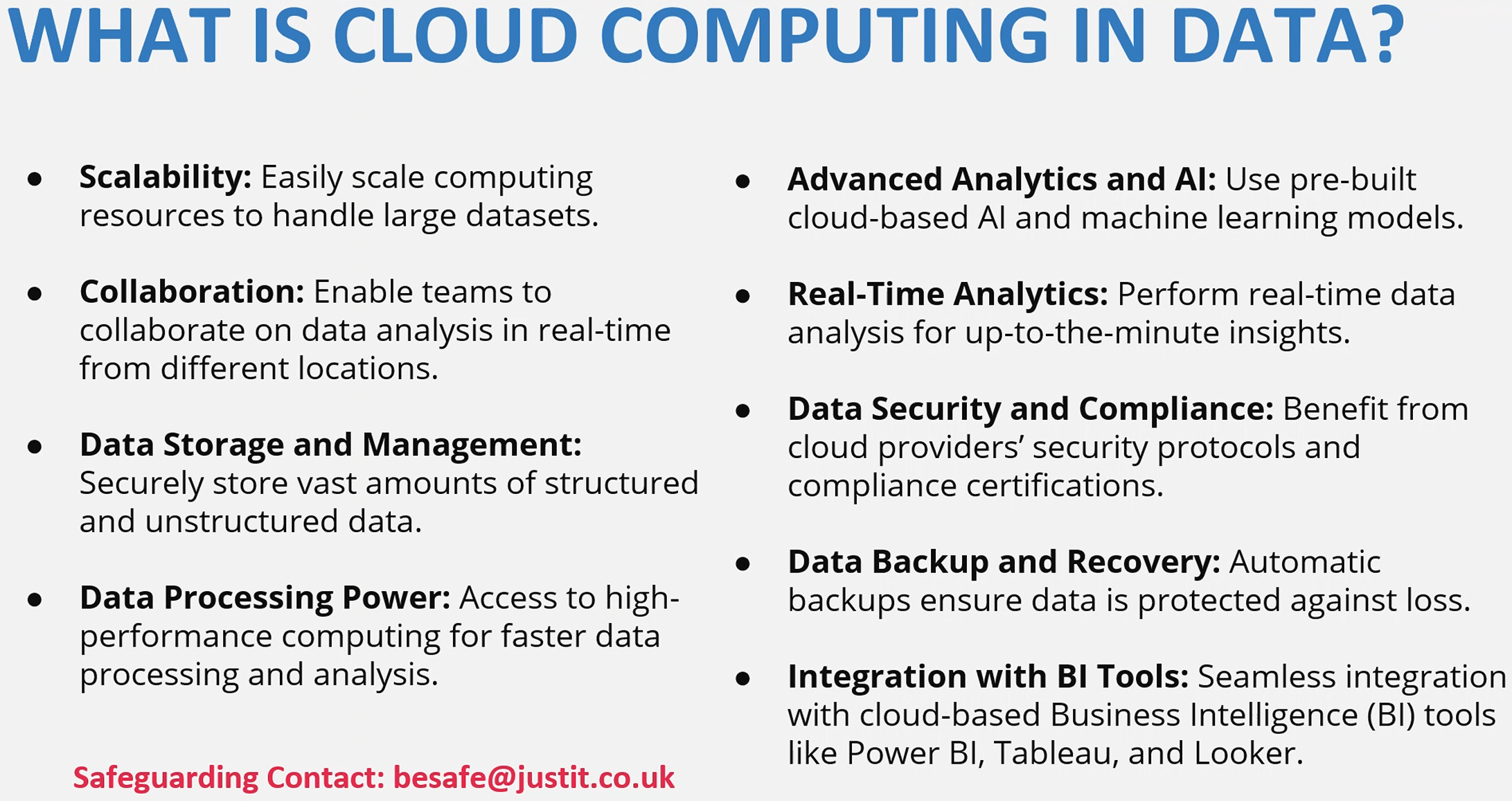
[Additional Information 17](#_Toc1081373283)

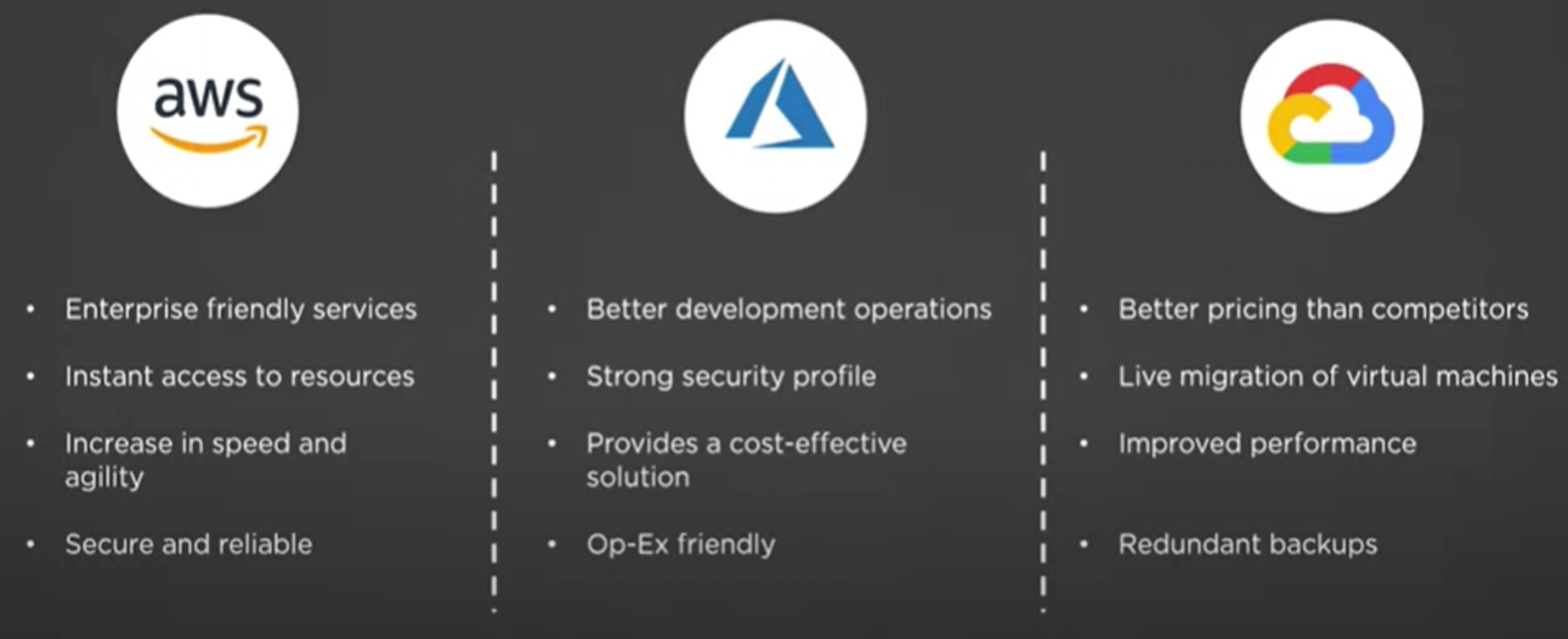
# Day 1: Task 1

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 you to do anything you can do with a regular computer or server. This is done over the internet meaning a stable internet connection is required. However, because it is usually operated by a huge corporation it will be a lot more powerful than what you might be able to reasonably afford. The main benefits are listed below;   * Store data online (Google Drive, iCloud, Dropbox) * Run applications without installing them (Gmail, Office 365) * Access systems from anywhere using any device * Stream services (Netflix, Spotify) * Support remote work and online learning * Scale resources easily (more storage or power when needed) * Backup and disaster recovery to protect data |
| How can it benefit a business? | Because you can use more powerful equipment than you can reasonably afford it would allow you to use more complex or advanced computer models than you might be able to use at every stage of your business. And because it is usually a pay-as-you-go model you can scale your model accordingly to your needs without worry of spending a large amount of capitol on hardware. |
| What’s the alternative to cloud computing? | The alternative is in house or privately owned servers that you own or rent. |
| What cloud providers can we use, what are their features and functions? | |  |  |  |  | | --- | --- | --- | --- | | Features | AWS | Azure | GCP | | Pricing | $200 free | No free opt | $300 free | | Global Network | Availability zones | Virtual network | VPC (Virtual private cloud) | | Compute Services | Create and manage virtual machines, run apps without underlying infrastructure | Create and manage virtual machines, run apps without underlying infrastructure | Create and manage virtual machines, run apps without underlying infrastructure | | AI and machine learning | Provides ready-made intelligence for applications and workflows | Provides applications for creating SMART applications and bots | Vertex AI is integrated | | Networking features | VPC, VPN, multiple gateways for connection | Create personal private network in the cloud, load balancing, VPN gateway | VPC, load balancing, CDN (Content Delivery Network) | |





# 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 is virtualised IT infrastructure for example servers, storage and networking. For example, Microsoft Azure. | It could be used if a company wants to launch a new web application but do not want to purchase all the hardware (servers, networks, etc) they can just rent virtual servers instead. |
| PaaS (Platform as a service) | This is like IaaS however PaaS supplies the operating system, databases and development tools. This means developers only need to focus on writing code instead of server management | The best use case for this would be a company that wants to create and build apps quickly without needing to worry about infrastructure. |
| SaaS (Software as a service) | This is a subscription model where software applications are delivered via the internet meaning no local download is required. | Any business that requires emails, document collaboration or cloud storage would benefit from am SaaS model such as Office 365. |

# 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 public cloud is a shared environment where IT resources are hosted by a provider such as AWS, Microsoft Azure or Google Cloud. These are usually generic environments that are available to anyone who is willing to pay the subscription. |
| Private Cloud | A private cloud is a cloud computing model where a company might rent an environment with specific features required. This might be required to meet strict controls or for security purposes, for example banks or government agencies. |
| Hybrid Cloud | This is a cloud computing model that combines both public and private cloud models. This allows you to keep certain areas under strict protection (private) while also keeping the ability to handle large amounts of traffic. An example would be a retailer using a private cloud for core functions while using a public cloud for a seasonal sale. |
| Community Cloud | A community cloud is a cloud computing model where infrastructure is shared by multiple organizations with common goals, compliance requirements, or security needs. It’s typically used in sectors like healthcare, government, or education, where collaboration and strict regulatory standards make a shared, specialized environment more efficient. |

# 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 a computer, account or file without permission usually done via hacking. | This covers unauthorised logins or viewing data. |
| Unauthorised access with intent to commit or facilitate further offences | Accessing systems with the specific purpose of committing another crime. | Fraud or theft of data. |
| Unauthorised acts with intent to impair, or recklessness as to impairing, the operation of a computer (and related serious‑damage offences) | Actions that damage, disrupt, or risk serious damage to computers, networks, or the information they hold; this also covers creating, supplying, or obtaining tools intended for these offences. | Creating or supplying tools for hacking or with bad intentions |

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 owning and distributing hacking tools if it is intended for bad intentions. |
| DoS attacks were made illegal with up to 10 years in jail. |
| It increased the penalties applied for committing any of the crimes listed in the computer misuse act (1990). |

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 |
| Political membership or opinions |
| Sexual history or orientation |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | Nintendo creating legal cases against any fan made Pokémon games with similar mechanics and style to the original games. |
| Provide one example of: Plagiarism | The CSGO skin the M4A4 howl was originally created using stolen artwork so valve removed the skin from the market and reworked it once the original artist flagged the issue. |
| What are two consequences of copyright infringement and software piracy? | 1. Legal penalties such as fines or prison 2. Financial loss for creators, companies and industries |
| Give three possible consequences for individuals when using pirated software | 1. Legal penalties (fines or prison) 2. Security risks (malware, spyware or ransomware) 3. Loss of support functionality (no updates, bug fixes or customer support) |

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

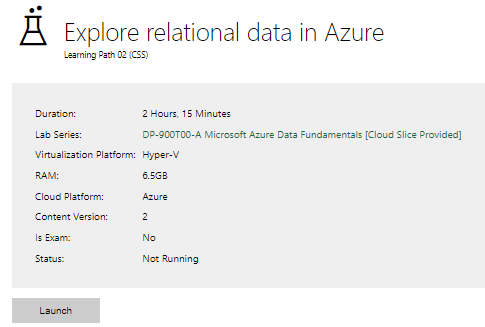
7. Consumer Rights Act 2015

* 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 |
| 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 |
| 1 | Gaining unauthorised access to a computer system is illegal |
| 5 | Employers must ensure that employees take regular and adequate breaks from looking at their screens |
| 2 | It is illegal to prevent or hinder access (e.g. by a denial-of-service attack) to any program or data held in any computer |
| 6 | Personal data must be accurate and where necessary kept up to date |

# Day 3: Task 1

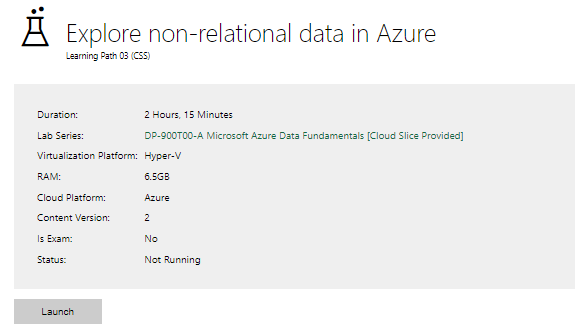
Please complete the below lab (3) *‘Explore relational data in Azure’* and paste evidence of the completed lab in the box provided.



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

# Day 3: Task 2

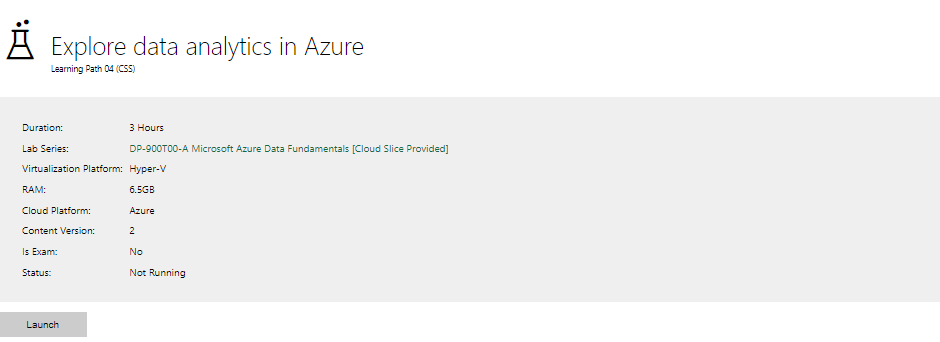
Please complete the below lab (4) *‘Explore non-relational data in Azure’* and paste evidence of the completed lab in the box provided.



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

# Day 3: Task 3

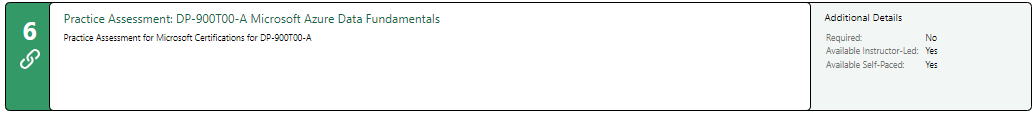
Please complete the below lab (5) ‘Explore data analytics in Azure’ and paste evidence of the completed lab in the box provided.



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

# 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.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Executive Summary A growing pet shop wants to improve its business model by analysing sales, customer information and inventory data and would like a fully online automated system. I have been tasked with providing a potential solution to them explaining the functionality and why some decisions have been made referencing standard practices and regulations. The main service recommendations I have are using the Azure package (Azure SQL Database, Azure Table Storage, Azure Synapse, Azure Data Factory, Azure Backup and Power BI) to create two separate models for the data for operational data I would create a relational model whilst for analytical data I would use a data warehouse star schema model. This would allow for easy data import and storage at a low cost and creates a package that people can be easily trained to use to create analysis and visualisations despite not having much experience meaning a quick transfer. Table of Content                                                Data laws The main laws and regulations that would need to be upheld and why are the following;  UK GDPR: The UK’s retained version of the EU GDPR, governing personal data processing. In this case as you will be tracking customers information you need to ensure you only take information that's necessary, implement fit for purpose security, you need transparency on what data you are collecting and you need to be able to provide access to the information or delete it upon request.  Data Protection Act 2018 (DPA 2018): Implements and supplements the UK GDPR.  Data (Use and Access) Act 2025 (DUAA): Amends the UK GDPR and DPA 2018, introducing targeted reforms to simplify compliance, encourage innovation, and clarify rules on cookies, legitimate interests, and automated decision-making. This is supplementary to the GDPR act which will allow for your business to implement more creative AI and machine learning models with customer data to prevent an accidental breach of the GDPR act.  Privacy and Electronic Communications Regulations (PECR): Regulates electronic marketing, cookies, and tracking technologies. This means that to contact customers about offers or sales you must first have their permission to contact them, and your company must also get permission for any tracking cookies that your website may use. Personal Data Definition Personal data is defined broadly under most major privacy regimes. Under the UK GDPR and EU GDPR, it encompasses any information relating to an identified or identifiable natural person, such as names, contact details, purchase history, and online identifiers. All the data that will be required by the business will come under this type there are more sensitive data types but those will not be required for your business model. The special categories of personal data (or “sensitive personal information”) receive heightened protection. These include data revealing racial or ethnic origin, political opinions, religious beliefs, health information, biometric data, and, under the CCPA/CPRA, financial account details, precise geolocation, and more. Azure Service Recommendations For your business model of a store looking to optimise and expand I believe that the best services would be Azure SQL Database, Azure Table Storage, Azure Synapse, Azure Data Factory, Azure Backup and Power BI. This allows you to store each of the main areas you want to optimise to have the best possible environment. Sales and Customer Information The most optimal service for tracking your sales information would be the Azure SQL Database which you can also pair with Synapse to aid with the analysis. The Azure SQL Database would be the best fit because it has high grade security measures (Including Back-ups), it allows for scalability so if you rapidly expand or shrink it can match your business needs and it is very efficient as it is a Pay-As-You-Go model you only pay for what you use. Inventory Information For inventory I believe that Azure Table Storage would be the best choice for your business. This is because it allows for you easily grow your catalogue due to its schema-less (The data doesn’t need to be in a fixed structure) design, furthermore it allows for massive datasets at a much cheaper price than an SQL database. Allowing for a large catalogue that only needs minimal interaction to expand at a low cost. Analysis and Visualisation Due to all the products listed being a part of the Microsoft Azure package it allows for you to move and transform the data cleanly using the Azure Data Factory. Due to this you can export your data directly into Azure Synapse to analyse a large amount of data to provide business insights. Which you can then use visualisation tools like Power BI to create charts and dashboards to better understand the insights provided. Storage Summary Table  |  |  |  | | --- | --- | --- | | **Data Type** | **Best Storage Options** | **Key Benefits** | | Sales Data | Azure SQL Database / Synapse | Structured, transactional, analytics ready | | Customer Information | Azure SQL Database / Cosmo DB | Compliant, encrypted, flexible schema | | Inventory Data | Azure Table Storage / SQL Database | Scalable, cost-effective |  Analysis Summary Table  |  |  |  | | --- | --- | --- | | **Date Type** | **Best Analysis Tools** | **Use Cases** | | Sales Data | Synapse, Power BI, Machine Learning | Revenue trends, forecasting, KPI, Dashboards | | Customer Information | SQL, Power BI | Compliance, churn prediction | | Inventory Data | Synapse, Power BI | Stock tracking, demand forecasting, anomaly detection |  Data types and modelling To create the framework for the all the data we would need to create a data table framework that we can use to base your entire data model on below I created an example of a potential framework your business could use. This model breaks down the customer information, sales and warehouse and breaks down all the potential subcategories for each category.  **Relational model example**  This model is a nice visualisation to help with understanding the general idea of what we need to do with the data however I don't believe that relational model would be the best for this scenario. I believe the best model would also include a Data warehouse approach using a fact table (the main table all the other tables link to) and several dimensional tables (additional data tables that use the fact table as a reference). I feel this approach would be best for your business model as it is easy to understand and read (easier to train your staff that are not used to online data), it is scalable so as your business grows you can easily expand it, It is very easy to use the structure to meet government compliance and finally it is very good for analysing your data with power BI.  On the page below I have provided an example of what the table structure may look like with a data warehouse model.  *Fact Table;*  *FactSales*   |  |  | | --- | --- | | Sales\_ID | PK | | Customer\_ID | FK | | Product\_ID | FK | | Sale\_Time | DATE/TIME | | Quantity\_Sold | INTEGER | | Price | DECIMAL | | Sale\_Status | ENUM (Pending, Complete, Returned) | | Total\_Sale\_Amount (Price \* Quantity\_Sold) | DECIMAL |   *Dimension Tables;*  *DimCustomer*   |  |  | | --- | --- | | Customer\_ID | PK | | Name | TEXT | | Email | TEXT | | Opt\_In\_Marketing | BOOLEAN |   DimProduct   |  |  | | --- | --- | | Product\_ID | PK | | Product\_Name | TEXT | | Brand | TEXT | | Product\_Category | ENUM (Food, Toy, Habitat, Etc) | | Animal\_Type | ENUM (Canid, Feline, Reptile, Aquatic, Etc) |   DimWarehouse   |  |  | | --- | --- | | Warehouse\_ID | PK | | Product\_ID | FK | | Stock\_quantity | INTEGER | | Use\_By\_Date | DATE/TIME |   DimDate   |  |  | | --- | --- | | Date\_ID | PK | | Date | DATE/TIME | | Time | DATE/TIME | | Weekend | BOOLEAN |   **Data warehouse star schema model** Data storage, formats and structures in Azure Based on the example I provided above on how the data could be structured in Azure we would still need to decide on which format the data should be stored in. And there are 2 main options you could choose from and they are; Format 1 Format 1 would have two separate data types as it would separate your operational data and your analytical data. This is because it would allow for improved performance as your analytics and operational data are segregated even if you have an abnormal spike in data that could slow down the system it would only impact the operational side. Operational data This could be done by having the operational data stored in a relational model, therefore you could use the CSV file format for raw data imports as it is easy to use and export to an SQL server that was mentioned previously. Analytical warehouse The data stored separately for analysis could use Azure Synapse (Microsoft data warehouse heavily integrated with the Microsoft package) to store the data, transform the data and analyse the data. This can be done with a simple star schema which is demonstrated above as all the dimension tables link back to the fact table. The best file format for this would be parquet files as they allow for the quick transfer of large amounts of data and are more compatible with analysis which will save you both time and money. Format 2Hybrid approach The other option would be a hybrid approach where you store all your data in the same place. This can be done using a Data lake. If you were to use this model you would need to import your raw data using CSV files for easy data ingestion and then transform the data into parquet files to store in the data lake. This model has the advantage of needing less programs, but I feel like it adds a lot of complexity as opposed to the other model hence my recommendation of format 1 being the best choice for your business model. Data Security and Encryption Due to the recommendations I have made all of them are classed as IaaS (Infrastructure as a service) this means you don’t need to purchase large amounts of storage or technology. Furthermore, this also means that a lot of the security is handled by Microsoft however you still need to have proper procedures for security and encryption. To do this you should provide a list of staff and their roles so I could provide a basic model showing an example of who should have access to what data and the actions they would be able to take. For example, your IT team would need full read and write access while a standard employee would only need read access to sales and write to inventory. To further protect your data, you can encrypt the entire database if you wish to also use the Azure Key Vault to have the highest level of protection for your data. Additional considerations This was not mentioned previously but Azure Backup would be necessary for an emergency. As a loss of data through either an attack or unforeseen event could cause massive damage to the business and possibly result in legal action. Due to the listed regulations earlier it is best practice to have a backup you can roll back to if the worst-case scenario happens and you lose all your data.  As touched on briefly multiple times in my proposal Power BI is a Microsoft product that allows for the creation of many charts and tools to aid in data visualisation. This would allow for you to create dashboards to track key metrics such as KPIs, sales goals and even compare brands and what is selling best to allow for a more agile business model.  Like I mentioned previously the Azure cloud services models biggest strength is the fact it is scalable so it will do exactly what you want and keep up at any business size. For example, if you triple in size, you can just increase the database size and pay more for that extra space and you can do the same in reverse if needed. This means you do not need to purchase expensive servers or databases only computers. References: [Data protection: The UK's data protection legislation - GOV.UK](https://www.gov.uk/data-protection) - 18/12/2025 - 15:23 - GDPR Regulations  [Data (Use and Access) Act 2025](https://www.legislation.gov.uk/ukpga/2025/18/contents) - 18/12/2025 - 15:26 – Data Use and Access Act  [The Privacy and Electronic Communications (EC Directive) Regulations 2003](https://www.legislation.gov.uk/uksi/2003/2426/contents) - 18/12/2025 - 15:29 – Privacy and electronic communications regulations  [Cloud Products | Microsoft Azure](https://azure.microsoft.com/en-us/products/?msockid=216f5f03b8616f7f205049a5b9ca6edd) - 19/12/2025 - 12:30 – Microsoft cloud product list |

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