

**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 the on-demand delivery of IT resources over the Internet with pay-as-you-go pricing, instead of buying, owning, and maintaining physical data centres and servers. Nowadays, cloud computing is everywhere, from people scrolling on social media, to watching videos and TV shows online, and even checking and sending emails.  In real life, cloud computing is mostly known as the technology that allows us to store and share files online, which can be accessed from any device, at any time, without the need of a physical USB driver. We can sync the files across multiple devices and even back them up safely, in case our devices are lost or damaged.  Moreover, cloud computing also allows us to stream video or music online, since big platforms like Netflix and Youtube rely heavily on cloud computing to store and deliver extremely big libraries of content to users around the world. |
| How can it benefit a business? | Cloud computing benefits businesses by cutting costs, improving collaboration, enhancing security, boosting flexibility and scalability, and ensuring recovery.  Cloud computing reduces upfront hardware costs and energy bills thanks to its pay-as-you-go pricing. It easily scales resources to meet demand, avoiding over-provisioning or any potential capacity issues.  It also enables real-time document editing and access to work from any device, at any point, while also securing offset data backups, ensuring business continuity.  Cloud computing also offers advanced security features, encryption, and compliance certifications. |
| What’s the alternative to cloud computing? | There exist multiple alternatives to cloud computing, and some of them are:   * Edge computing, which process data locally for speed; * Private clouds, which have a dedicated infrastructure for more control; * Hybrid cloud, which is a mix of private and public cloud; * On-premises, which are local servers and storage, allowing for more control; * Local storage, which are external drives. |
| What cloud providers can we use, what are their features and functions? | |  |  |  |  | | --- | --- | --- | --- | | Features | AWS | Azure | GCP | |  |  |  |  | |  |  |  |  | |  |  |  |  |   AWS (Amazon Web Services) provide over 200 cloud services for computing, storage, databases, analytics, and many more. AWS offers a vast network for data servers worldwide for low latency and high availability, with on-demand and scalable resources provided. It allows you to only pay for what you use, therefore eliminating large upfront hardware costs, while also being flexible in letting you choose the preferred operating system, programming language, database, and more. AWS also has a robust physical and digital security. Key functions of AWS include virtual servers (EC2), storage (S3), managed databases (RDS), networking, developer tools, and serverless computing (Lambda) for agility and cost savings.  Microsoft Azure offers a vast array of cloud services for computing, storage, networking, analytics, and even AI, to allow business to build, deploy, and manage applications globally. Azure Functions is a serverless computing service, hosted on the Microsoft Azure public cloud, and the idea behind it is to eliminate infrastructure considerations for the user, letting them focus on the code itself and the events that will execute it, for a lower cost.  GCP (Google Cloud Platform) is Google’s suite of cloud computing services, offering infrastructure, data analytics, machine learning, and developer tools, that allow businesses to build, run, and scale applications on Google’s global infrastructure without managing the physical hardware. GCP offers powerful data & AI/ML services (like Vertex AI, BigQuery), flexible Compute & Modern infrastructure (GKE, Computer Engine, serverless), powerful security & identity and access management, and also hybrid or multi-cloud options for building, deploying, and scaling applications. |

# 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 where providers offer visualised computing resources (servers, storage, networks, and operating systems) over the internet. The provider manages the underlying physical hardware, data centers, and visualisation, and the user manages to operating system, applications, and data, all on a virtual infrastructure. | IaaS is used in real life for scalable web hosting (e.g. e-commerce sales), running development and testing environments, disaster recovery, big data analytics, and hosting complex applications. Amazon Web Service is the most widely used IaaS provider, with Microsoft Azure and GCP coming after.  A real-life example of how and when to use IaaS is media and entertainment. For example, Netflix delivers high-quality streaming content with the help of IaaS, as it handles massive media files and user traffic. |
| PaaS (Platform as a service) | PaaS is a layer above IaaS and a layer below SaaS. It is a cloud computing model that provides developers with a platform to build, deploy, and manage applications, reducing coding time and simplifying app development, due to its abstract infrastructure complexities. In IaaS, the provider manages the infrastructure, and also the operating systems, databases, middleware, and software tools, while the user is only concerned with developing, testing, deploying, and hosting applications. | A real-life example of PaaS is online baking. The specific application of PaaS in the financial service is Payment as a Service, where the financial institutions and banks are allowed to modernise their offerings and adapt quickly to market demands without the burden of building and maintaining an in-house infrastructure. |
| SaaS (Software as a service) | SaaS is a cloud-based model where software applications are delivered over the internet on a subscription basis, rather than being installed locally. The provider manages all the underlying infrastructure, maintenance, and updates, allowing users to access the software via a web browser or app, paying typically on a subscription basis. | SaaS is widely used in real life, from streaming services like Netflix and Spotify, to using Google Drive and Dropbox. Essentially, every feature or app that is subscription-based uses SaaS. |

# 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 cloud computing environment where a third-party cloud service provider owns, manages, and operates the infrastructure, and offers services over the internet to multiple customers simultaneously. All hardware, software, and other supporting infrastructure are shared among different organisations, although their data and workloads remain logically separated and secure.  Public cloud is ideal for applications that have unpredictable traffic, as resources are easily scalable to meet demand. It is also suitable for organisations that prefer operational expenses, paying for what they use, over large upfront expenditures. Public cloud provides quick access to a wide range of services and tools that accelerate the development, testing, and deployment of new applications.  A real-life example of a public cloud usage are email services, like Gmail, Microsoft Outlook, and Yahoo Mail. They operate on the public cloud, which allow users access to their email and store data from any device with internet access. |
| Private Cloud | A private cloud is a cloud computing environment where all hardware and software resources are dedicated exclusively to a single organisation. It provides many of the same benefits of a public cloud computing (scalability, elasticity, and self-service), but it also offers an added control, security, and performance of the dedicated environment. A private cloud can be physically located in the company’s own on-site data centre, or hosted by a third-party service provider on dedicated infrastructure.  A private cloud is suitable when an organisation needs strict security and compliance requirements; when the companies deal with confident information, or mission-critical applications.  A real-life example of a usage of a private cloud is a healthcare provider, as a hospital system needs to manage sensitive patient health records. The company needs to comply with regulations as HIPAA, while allowing doctors and administrators flexible, on-demand access to the data within the hospital’s secured network. |
| Hybrid Cloud | A hybrid cloud is a computing architecture that integrates on-premises private cloud with public cloud services, allowing data and applications to move seamlessly between them to balance flexibility, control, and cost.  A hybrid cloud is appropriate when an organisation in a highly regulated industry needs to store sensitive data in a private, on-premises environment to comply with data residency laws while running fewer sensitive applications in the public cloud.  A real-life example of an organisation that uses a hybrid model is Netflix, that uses AWS for massive data storage and analytics, but operates on its own private content delivery network to ensure low-latency, high-quality streaming directly to users globally. |
| Community Cloud | A community cloud is a cloud deployment model where a shared infrastructure is provisioned for exclusive use by a specific community of consumers from organisations with joint concerns, such as mission, security requirements, policy, and compliance considerations. It is a hybrid form between the privacy and control of a private cloud and the collaborative environment of a public cloud. Community clouds are multi-tenant platforms enabling different organisations from the same common business to work on a shared platform without compromising security or compliance. Community clouds are more expensive than public clouds but also more secure, where each member of the cloud is allocated a fixed amount of data storage and bandwidth, making scalability somewhat more difficult than with private and public clouds.  A real-life example of community cloud I the ‘government cloud’, a community cloud specifically made for government bodies. Government processes and services require constant communication and data transactions between multiple departments, as they all operate on similar infrastructure, with services and resources shared across them. Since government bodies are subject to various privacy, legal, and security concerns, which means that public clouds cannot be used in this instance. |

# 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 (basic hacking) | It means getting into a computer system or data without permission, an action often called hacking. | A real-life example would be hacking into someone else’s social media accounts, or even guessing their password to an account to view private files. |
| Unauthorised access with intent | It means hacking into a system with the goal of committing or facilitating another crime, which can either be theft, fraud, or blackmail. | A real-life instance would still be hacking, however, with an intent to do more harm. For example, hacking into a banking system in order to steal money. |
| Unauthorised modification of computer material | It means intentionally altering, damaging, or deleting data or programs, including introducing viruses, malware, or ransomware. | A real-life example of this area, however, not very common, would be someone purposefully hacking a school system’s database in order to change exam results, which in the end would alter students’ grades. |

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 |
| Unauthorised Acts with Intent to Impair Operation (DDoS attacks) was introduced as an offence for deliberately impairing a computer’s operation or preventing legitimate access, even temporarily, criminalising DDoS attacks. An example where this offence can be punishable would be flooding a company’s website with traffic in order to make it crash. |
| Making, supplying, or obtaining hacking tools was introduced as a new offence targeting creating, adapting, supplying, or even offering to supply tools like malware and password crackers intended to be used for computer misuse, or obtaining them for such purposes. Quite self-explanatory, if a hacking group hacks multiple businesses to steal data, they will now face stricter legal action. |
| Increased penalties and broader scope is not a new type of offence, but rather a stricter one, which significantly increased maximum penalties for existing offences and clarified that “unauthorised access” includes a broad range of actions, strengthening enforcement against serious cybercrime. Again, self-explanatory, the punishments will be harsher in case of breaching the Computer Misuse Act. |

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 |
| Work experience |

|  |
| --- |
| Give three more examples of data that an employer can only store if they first get the employee’s permission. |
| Race and ethnicity |
| Religion |
| Sexual history or orientation |

Conduct further research to answer the below questions.

|  |  |
| --- | --- |
| Question | Answer |
| Provide one example of: Copyright infringement | An example of copyright infringement is downloading music illegally. Moreover, even using copyrighted songs without permission or license is an infringement. |
| Provide one example of: Plagiarism | Plagiarism is a famous breach that has an extensive range of forms. A common example of plagiarism is copying text word-for-word without proper citation or quotation marks, essentially using somebody else’s ideas without giving proper credit, whether it is intentional or not. |
| What are two consequences of copyright infringement and software piracy? | Two consequences of copyright infringement and software piracy are severe legal penalties for individuals and businesses (whether civil penalties or criminal charges), and significant security or performance risks (malware, data theft, system instability) for users, alongside broader economic harm to creators and industries. |
| Give three possible consequences for individuals when using pirated software | Three possible consequences for individuals using pirated software are legal penalties, increased cybersecurity risks, and the loss of essential support and functionality. I will explain these into more detail below.   1. **Legal penalties** – unauthorised use or distribution of copyrighted software is punished with anything from substantial fines (thousands to hundreds of thousand) to imprisonment, which will be recorded permanently. 2. **Increased cybersecurity risks** – pirated software often comes embedded with malware and viruses, which can damage the computer system and even lead to theft of sensitive personal information. 3. **Loss of support and functionality** – using pirated software means the users do not have access to official customer support, updates, or security patches, which will affect the software’s performance and eventually lead to potential data loss and costly repairs. |

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

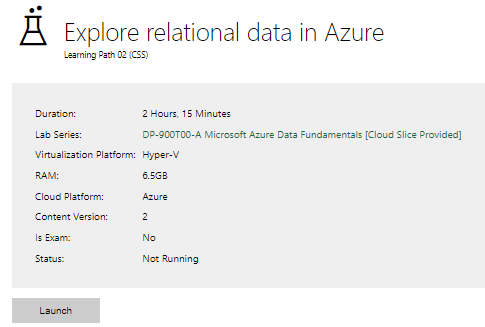
7. Consumer Rights Act 2015

* 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 |
| 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

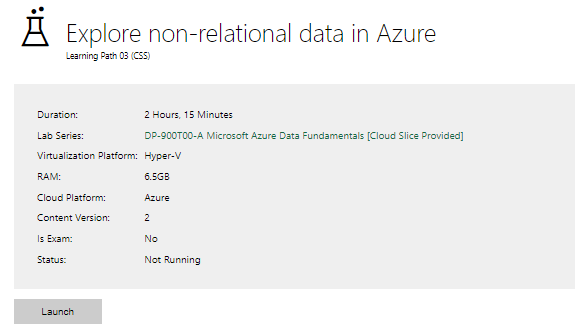
Please complete the below lab (3) *‘Explore relational data in Azure’* and paste evidence of the completed lab in the box provided.



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

# Day 3: Task 2

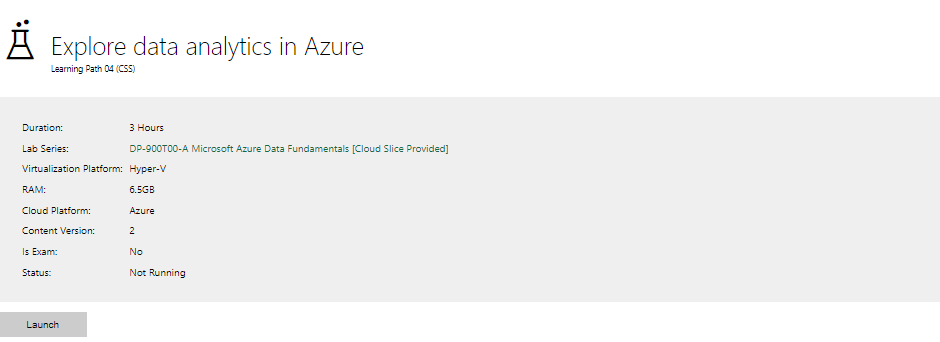
Please complete the below lab (4) *‘Explore non-relational data in Azure’* and paste evidence of the completed lab in the box provided.



|  |  |
| --- | --- |
| Completed lab | STORAGE          COSMOS DB |

# Day 3: Task 3

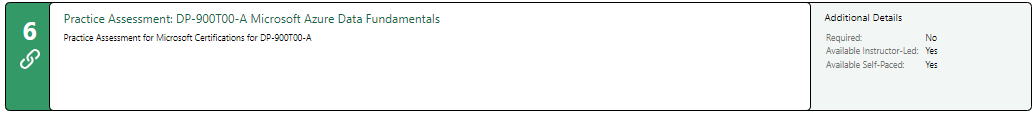
Please complete the below lab (5) ‘Explore data analytics in Azure’ and paste evidence of the completed lab in the box provided.



|  |  |
| --- | --- |
| 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 AZURE SERVICE REPORT  Contents  [BUSINESS OVERVIEW 1](#_Toc216963864)  [DATA LAWS AND REGULATIONS 1](#_Toc216963865)  [AZURE SERVICE RECOMMENDATIONS 2](#_Toc216963866)  [DATA TYPES AND DATA MODELLING 3](#_Toc216963867)  [DATA STORAGE FORMATS AND STRUCTURES IN AZURE 4](#_Toc216963868)  [ADDITIONAL CONSIDERATIONS 5](#_Toc216963869)  [SUMMARY AND CONCLUSION 6](#_Toc216963870) BUSINESS OVERVIEW Paws & Whiskers is a pet shop that is currently growing and wishes to improve its business. The owner wishes to transition to Microsoft Azure to streamline data storage, analysis, and reporting, which will enable them to make data-driven decisions. In order to do that, we need to explain what the Data Laws and Regulations are, what Azure services are recommended, what data types to work with and how to model them, how to store them in Azure, and additional considerations regarding back-up, scalability, and/or visualization. DATA LAWS AND REGULATIONS First and foremost, we want to comply with current Data Laws and Regulations. The UK’s core protection laws are GDPR (General Data Protection Regulation) and DPA (Data Protection Act). The UK GDPR is a version of the EU GDPR adapted post-Brexit and it sets core principles, rights, and obligations for personal data processing. The GDPR requires businesses to obtain explicit consent for collecting personal data, implement data minimization and only collect necessary information, use the data in a way that is relevant and limited to what is necessary, and keep it up to date, also making sure that if the data is no longer needed, it will be removed immediately [[1]](#footnote-1). The DPA 2018 supplements UK GDPR, providing additional UK-specific guidance, as such: ensuring that the data collected is handled appropriately to protect people’s privacy [[2]](#footnote-2). Moreover, since this business offers paid products and/or services, we need to make sure that it complies with the PCI DSS (Payment Card Industry Data Security Standard), which is a set of rules for businesses explaining how to protect themselves and the customers when handling payments [[3]](#footnote-3), by having firewalls and secure configurations, making sure to change the default settings, protecting against malware, using unique IDs, MFA, and strong passwords, encrypting data in transit and protecting stored data, and last, training the staff and have a security policy in place in case of a data breach [[4]](#footnote-4). AZURE SERVICE RECOMMENDATIONS To ensure compliance with UK Data Protection Laws, Paws & Whiskers can leverage Microsoft Azure’s suite of data services, that provide secure, scalable storage, automated data integration, and powerful capabilities. Azure offers a multitude of storage options: Queue, Table, Blob, Azure Files, and Disk Storage. For this instance, we will use Azure Blob Storage, which is a type of data storage for unstructured data (text, images, audio, video, backup, etc.), which is intended to store large files [[5]](#footnote-5). Blob Storage is a go-to choice for organizations because it is highly scalable, cost-effective and accessible, easy to integrate by connecting to other Azure services like Functions, Synapse, Data Factory, Data Lake Gen2, and complies with data regulations [[6]](#footnote-6).  In addition to the Blob Storage, we recommend using Azure SQL Database, which leverages all the benefits of a cloud platform, offering better availability, durability, and scalability [[7]](#footnote-7). Azure SQL Database makes it easy for its users to create databases, organize data into tables, run queries, all possible with an internet connection. It also provides easy maintenance, as Microsoft handles all the updates and backups, and it also offers built-in security features, complying with the GDPR and DPA laws.  Azure SQL Database is integrated with other Azure services like Azure Machine Learning and Azure Data Lake Storage Gen2 for advanced analytics and insights. Azure Data Lake Storage Gen2 is a service built upon Azure Storage, designed to service multiple petabytes of data while sustaining many gigabits of turnout, which means we can handle extremely large datasets. ADLS Gen2 offers the addition of a hierarchical namespace, which organizes files into a hierarchy of directories for efficient data access. ADLS Gen2 also allows to use technologies like Databricks or Hadoop to process and analyze the data as per the business’ needs.  Once the data is centralized, we can move onto Azure Synapse Analytics, which performs large-scale analytical queries and reporting [[8]](#footnote-8), and then onto Azure Databricks, which cleans and aggregates our datasets. The Azure Machine Learning tool will be used to analyze customer behaviours, predict product demands, and support marketing strategies, which will benefit the business by allowing it to make smart predictions. In the end, Power BI can turn our data processing efforts into analytics and reports for real-time insights, providing a complete solution for data-driven decision-making.  Ultimately, we need to make sure our data is secure and complies with the UK GDPR and DPA, and that will be done using the built-in Azure tools, such as Azure Storage Service Encryption, Transparent Data Encryption, and Azure Active Directory authentication. DATA TYPES AND DATA MODELLING It is essential that now we define the types of data Paes & Whiskers will manage. The business will handle multiple categories of data, that include, but are not limited to, customer demographics (name, contact details, location), transaction history (products, dates it was bought, quantities, and sales amounts), product categories (whether they fall under pet food, accessories, toys), and possibly inventory. It is recommended that we use a relational database model, where the tables are linked via primary and foreign keys in order to maintain a relationship between customers, sales, and products. A star schema is recommended for analytical purposes, and Azure Synapse Analytics can implement this with fact tables that encapsulate sales transactions, and dimension tables for customers, products, and time periods. This approach ensures that the data is well-organized, query efficient, and scalable, allowing for accurate trend analysis and reports, and predictive insights with the help of Azure ML service. Below is a suggestion for the tables mentioned, along with their primary and foreign keys.  A screenshot of a computer  AI-generated content may be incorrect. DATA STORAGE FORMATS AND STRUCTURES IN AZURE Based on the image above, we can see that there are four tables in the proposed database. The Customers, Products, and Inventory tables serve as dimension tables in this case, as they describe entities and provide context for analysis, and Sales is a fact table, as it stores measurable, quantitative data that we may want to analyze, referencing Customer and Product via the foreign keys to give additional context. All four tables allow Azure Synapse Analytics to perform large-scale queries and analytics.  Regarding data storage formats, CSV files are ideal for initial data imports from spreadsheets, JSON works well with structured application data, and Parquet is recommended for analytics, because it provides columnar storage and faster query performance.  Moreover, to ensure compliance and security, Azure’s built-in features allow for protecting sensitive customer and transaction data while supporting GDPR and DPA requirements. Possible features that can be used are Storage Service Encryption, Transparent Data Encryption, and even role-based access control.  Overall, the diagram below showcases the proposed Azure Data Platform for Paws & Whiskers, showing how data flows from sources to storage, processing, analytics, and vizualisation, while maintaining security and regulatory compliance.    Last but not least, data vizualisation can be easily generated using Power BI, thanks to its interactive dashboards which showcase real-time insights on sales, inventory levels, and customer behaviours. Dashboards can easily be customized for key performance indicators, trends, and alerts for low stock or high demand products. ADDITIONAL CONSIDERATIONS What else needs to be considered are backup and disaster recovery, future scalability, and automation and efficiency.  In regards of backup and recovery, Azure has a built-in system, Azure Backup, designed to protect SQL databases, Blob Storage, and critical files. Also, Azure Site Recovery ensures that the business can continue in case of outages or disasters.  In case of incredibly large data, Blob Storage, SQL Database, Synapse Analytics, and Databricks, all scale horizontally and vertically, allowing the platform to handle big data volumes and high transaction loads. Azure ML can easily analyze large datasets for predictive insights as the business grows, furthermore looking for customer behaviours, product demands, and even marketing strategies.  Before, the business was using manual and long work, but thanks to the Azure Data Factory pipeline, it reduces the manual work by automating data ingestion and transformation, enabling for consistent and timely data for analytics and reporting. SUMMARY AND CONCLUSION By moving to Microsoft Azure, Paws & Whiskers transforms its data into a strategic asset. Azure enables efficient storage, advanced analytics, and real-time reporting, since it is secure, scalable, and compliant with the UK Data Laws. Azure enables the business to transform raw data into actionable insights, with real-time dashboards, predictive analysis, detailed reporting, all allowing the business to make informed decisions on inventory management, customer engagement, and marketing strategies.  The integration of cloud technologies ensures that the business stays agile, compliant, and empowered to harness the full potential of its data, allowing for sustainable growth and long-term success in the competitive pet retail market. |

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

1. https://www.gov.uk/data-protection [↑](#footnote-ref-1)
2. https://www.virtual-college.co.uk/resources/the-data-protection-act-2018 [↑](#footnote-ref-2)
3. https://www.virtual-college.co.uk/resources/the-data-protection-act-2018 [↑](#footnote-ref-3)
4. https://drata.com/blog/pci-compliance-checklist [↑](#footnote-ref-4)
5. https://www.dev4side.com/en/blog/azure-storage [↑](#footnote-ref-5)
6. https://www.logicmonitor.com/blog/what-is-azure-blob [↑](#footnote-ref-6)
7. https://www.stitchdata.com/resources/azure-sql-database/ [↑](#footnote-ref-7)
8. https://www.datacamp.com/tutorial/azure-synapse [↑](#footnote-ref-8)