**Assessment Week 9 Tomi Jolkkonen**

**Scenario: engineering team for a financial services company is migrating data pipeline from local env to Azure**

**- Synapse Analytics for ETL and warehousing**

**- Azure Data Lake Gen2 for staging**

**Pipeline processes sensitive customer data incl. PII, financial transactions, resulting tables w. transactions w. relations to customers, data is business-critical**

**Our job is to ensure that pipeline is secure, compliant w. regulations and governed correctly**

****Deliverable:** This Word document explaining tasks, understand key points**

**Secure Data Pipeline Migration to Microsoft Azure**

**-----------------------------------------------------------------------------------------------------------------------**

## ****Task 1: Identity and Access Management (IAM) on Azure****

How to securely access to data pipeline in Azure by configuring IAM and what to consider regarding Synapse and Data Lake:

### ****1.1 Azure Synapse Analytics****

* **Role-Based Access Control (RBAC):** Assign least privilege roles, for example Synapse **Administrator, Contributor and User** based on job responsibilities
* Use **Azure Active Directory Managed Identities** for authentication instead of putting in credentials
* **Multi-Factor Authentication (MFA):** Enforce MFA for all users accessing Synapse
* Restrict Synapse access via p**rivate endpoints**, no public internet exposure

### ****1.2 Azure Data Lake Gen2****

* Implement Access Control Lists at the file and folder level
* Use **Azure Active Directory groups** to control access instead of individual users
* Restrict access to trusted networks with storage firewall
* **Use** **data masking** to hide PII (Personally Identifiable Information)

## ****Task 2: Network Security in Azure****

How to secure network communication between the data storage, processing, and warehouse:

### ****2.1 Secure Network Architecture****

* Separate data storage, processing, and analytics into different **subnets, this way we segmentate virtual network, so if attack happens, hacker cannot move horisontally**
* Use private endpoints to prevent data from being exposed to the public internet (**Azure Private Link**)
* Implement **NSGs** (Network Security Group) to allow only necessary traffic between services
* Use **Azure´s Firewall** to filter traffic and set up security rules

### ****2.2 Data Transfer Security****

* Ensure all data transfers use encrypted protocols such as **TLS 1.2+ (Transport Layer Security)**
* **Use VPN Gateway** to Azure connecttions
* Enable **Azure DDoS protection** against network attacks

## ****Task 3: Data Encryption****

What to consider regarding sensitive data to be secure in the Azure Cloud

### ****3.1 Data Encryption at Rest****

* Enable SSE (Storage Service Encryption in Azure) with Key Vault where keys are saved
* Activate **TDE** (Transparent Data Encryption) that can be used in Synapse Analytics SQL pools
* Use column-level encryption for sensitive fields, for example PII fields

### ****3.2 Data Encryption in Transit****

* Use **HTTPS and TLS 1.2+** for all communication for encryption during data transporting from one place to another
* Encrypt virtual machine disks using **BitLocker**

## ****Task 4: GDPR Compliance****

We collect and process data from customers in EU. How to ensure GDPR in our data pipelines:

### ****4.1 Data Subject Rights****

* **Right to Access:** Implement a process for customers to request their personal data
* **Right to Deletion:** Implement a process for customers to request for their data´s deletion
* **Right to Portability:** Implement a process for customers to get their data securely in structured formats, for example JSON or CSV

### ****4.2 Data Minimization****

* Store only necessary data and implement **automatic deletion** policies
* Mask or tokenize PII where possible, to anonymize data

### ****4.3 Legal and Regulatory tasks****

* Ensure that data processing agreements (DPA) align with GDPR requirements
* Store EU customer data inside **EU regions in Azure**
* Use **Azure Purview** to monitor and audit GDPR compliance

### ****Bonus: Further Analysis, Monitoring, Auditing****

* Use **Microsoft Defender for Cloud** for real-time automated security monitoring (or 3rd party apps)
* Implement an incident response process with threat modeling, vulnerability assessment, Defense-in-Depth Strategy and create a structured Security Plan Document with **Security Incident and Event Management (SIEM)** process
* Enable **Azure Monitor, Log Analytics and Azure Sentinel** for tracking and investigation of security events, logs and audits