1.

Imagine the Pokémon are subject to the GDPR. How would you pseudonymize the data to no longer make it re-identifiable?

**Answer:** We can use any pseudonymize function to achieve this example.

import gocept.pseudonymize

*pokemon\_df['Pokemon\_name'] = pokemon\_df['Pokemon\_name'].apply(lambda x: gocept.pseudonymize.text(x, 'secret'))*

There are other encryption methods as well like with key hash functions using the following library.

import hashlib

import boto3

2.

Imagine you had to build a system in the cloud that would continiously deliver updates to the investor with updates about Pokémon. Draw an architecture for exposing new changes to the existing Pokémon to the investor.

**Answer** : There are many ways to this one of the processes explained below.

1. We can use delta query and Microsoft graph API. We can use Azure event hub and azure functions for notifications towards the investor application

Azure Signal IR Services

Below is the complete architecture.

Function Apps

Input hub messages

MS Graph change tracking

Event Hubs

Delta Query

Investor application

Place Notification in event hub

Pokemon Data Source

3.

Make a interactable dashboard using the data where users can get a detailed page about the Pokémon, including showing the image from the url of the sprite.

**Answer:** This could be achieved by a Power Bi report or Azure Synapse and can be connected with datalake house. We can use databricks notebook too but will not have many features.

4.

If a spark compatible framework was not already chosen for the primary requirements, consider how the code would change to be executed on a **spark engine.**

**Answer:** I can change the code according to the platform and can execute in databricks notebook.