**Azure Data factory Project on Covid-19 Prediction and Reporting**

**Overview:**

Solution Architecture Overview,

A diagram of a software development

AI-generated content may be incorrect.

1. Ingestion Layer

* Data Sources:
  + ECDC COVID-19 Data: Pulled via HTTP Connector.
  + Population Data: Retrieved from Azure Blob Storage.
* Storage:
  + Both datasets are stored in Azure Data Lake Storage Gen2 for scalable and secure data storage.

2. Transformation & Analysis Layer

* Processing:
  + Data is transformed and cleaned using ADF pipelines.
  + Analytical operations are performed to prepare the data for modelling.
* Storage post-transformation:
  + Transformed data is stored in:
    - Azure SQL Database for structured querying.
    - Azure Data Lake Storage Gen2 for further processing or archival.

3. Publishing & Reporting Layer

* Machine Learning:
  + ML models are trained using the processed data to predict COVID-19 trends.
* Visualization:
  + Results and insights are published to Power BI, enabling interactive dashboards and reports.

**Web source data from ECDC website:**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a chat

AI-generated content may be incorrect.**