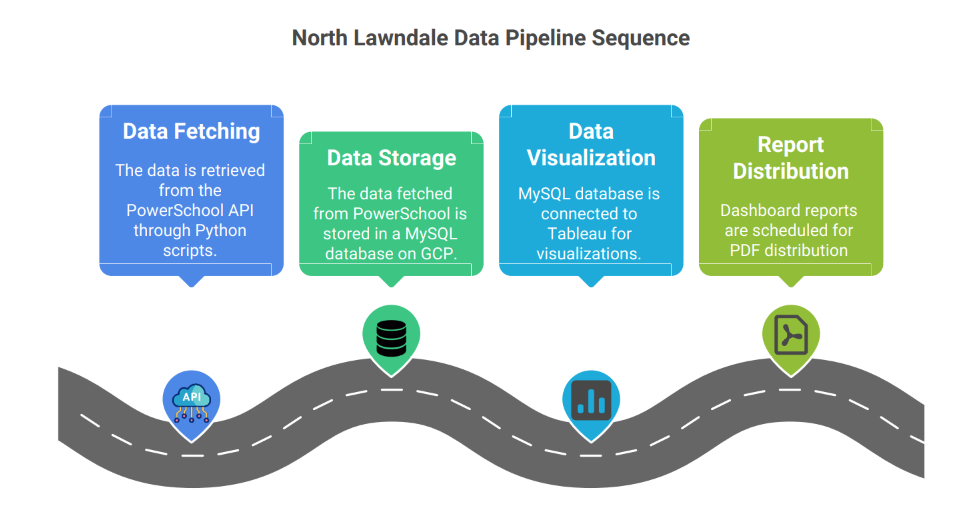
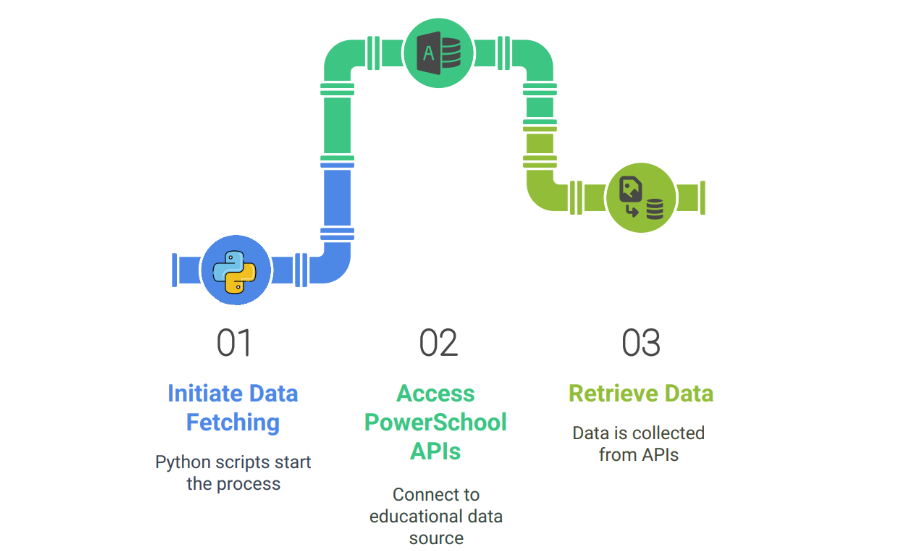
North Lawndale Data Pipeline

This document outlines the dataflow pipeline for North Lawndale, detailing the interactions between various components including Python scripts, Google Cloud Platform (GCP), MySQL database, Tableau, and the scheduled pdf distribution of dashboard reports. The dataflow diagram illustrates how data is fetched, processed, and presented to end-users.

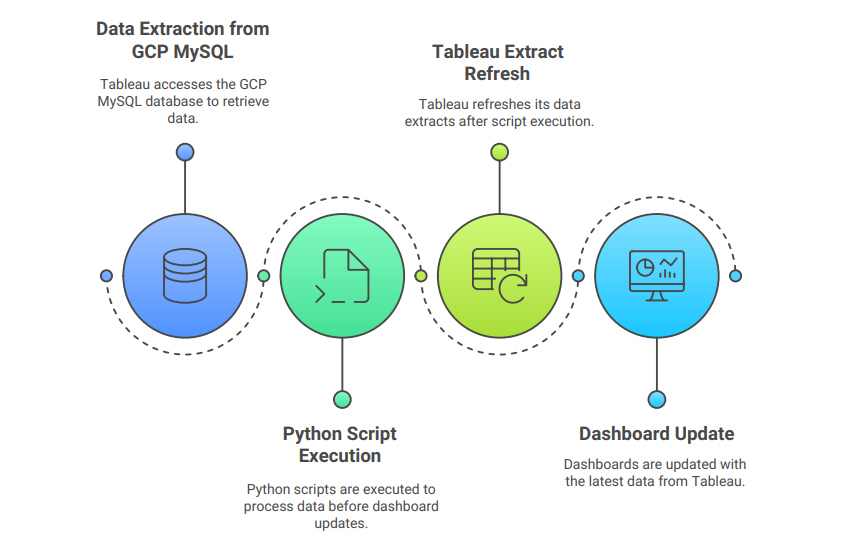


The **data flow pipeline for North Lawndale** operates as follows:

* **Python scripts** stored on a **Google Virtual Machine (Linux)** fetch data from **PowerSchool APIs**.



* The fetched data is then stored in a **GCP MySQL database**.
* This **MySQL database** is directly connected to **Tableau Server**, where dashboards are created and managed.
* The **Virtual Machine** and **MySQL database** are active only for **2 hours** during which:
  + Python scripts run and populate the database.
  + Once data fetching is completed, **Tableau Extract Jobs** are triggered to refresh the Tableau tables.
  + After the extract jobs are completed, the **Tableau Dashboards** automatically update with the latest data.
* Every **Friday**, **PDF copies** of these refreshed dashboards are sent to the **North Lawndale teachers**.



This pipeline ensures automated, consistent, and up-to-date data delivery to the teachers.