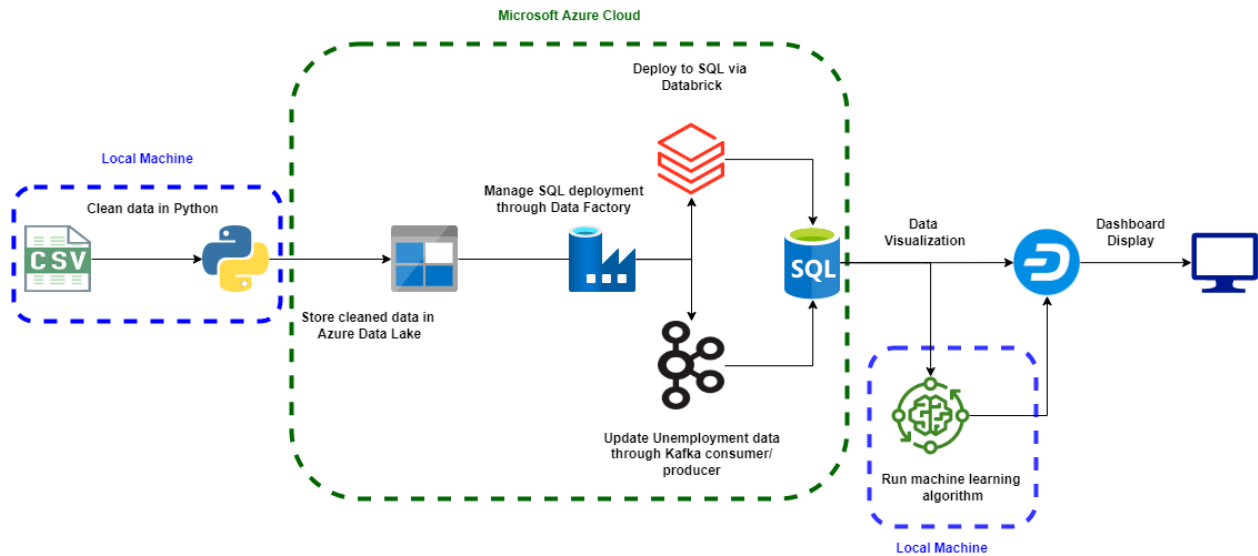


Payment Protection Plan Data Platform Diagram



1. Download data sets as CSV files
2. Clean data on your local machine with Python
3. Store cleaned data in an Azure Data Lake container
4. Deploy the normalization process through a Data Factory pipeline which pushes data to SQL by:
 - a. Running Spark Notebook files from your data brick which import the cleaned CSVs and push them to your SQL database
 - b. Deploying a Kafka producer/consumer that pushes to your SQL database
5. Pull a subset of your data to your local machine to run your machine learning algorithm with Python
6. Feed the machine learning data and all data from your SQL database to create a Dash dashboard