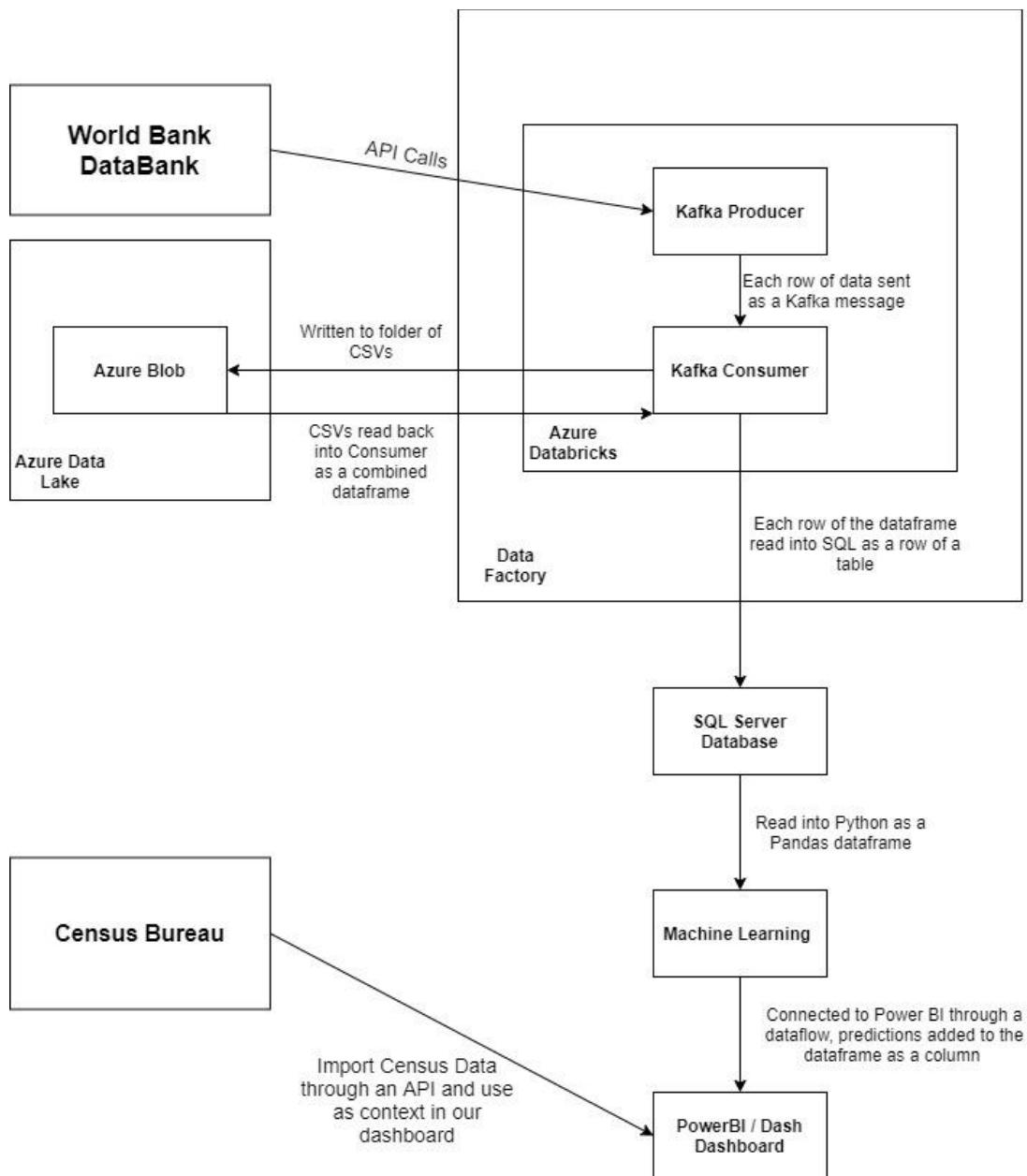# Group 1 Data Platform

Step-by-Step Explanation

1. First, we retrieve data from the World Bank using its API.
2. From here, we route the data through an Azure Data Factory, which automates the following steps:
   a. We route that data through a Kafka Producer.
   b. The producer sends messages containing the data to a Kafka Consumer.
   c. Using the consumer, we save the data as a list of CSV documents to an Azure Data Lake.

      d.   Next, we pull the data back from the data lake back into a SQL database, where our data is stored in tabular form.

3. We access the data from the SQL frame and process it for machine learning, using Pandas and the Sci-Kit Learn Python module.
4. Meanwhile, we access the US Census Bureau API to gather data for extra contextual support.
5. Finally, with machine learning models and predictions in hand, we stream the data from SQL to data dashboard to visualize our results. We also incorporate Census data into this dashboard to buttress the story we're telling.