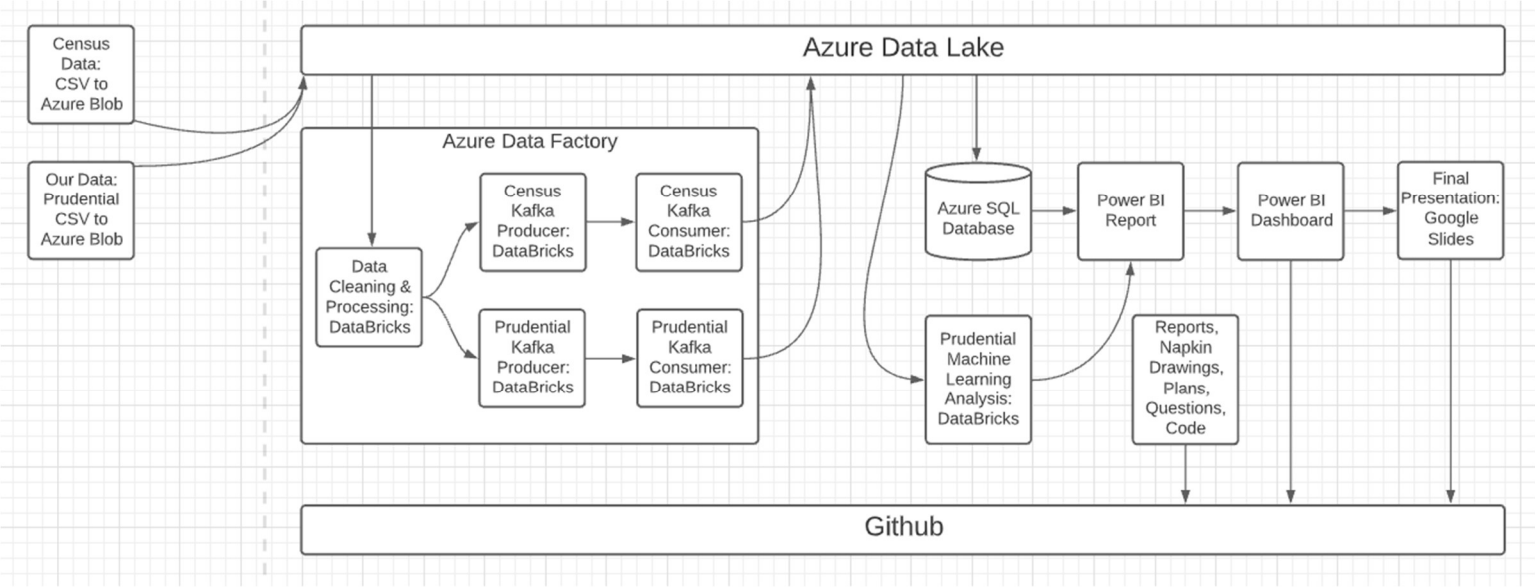
Data Platform Report



1. Download the census and prudential datasets as CSV files to your local pc.
2. Upload both the census and prudential CSV files to the Azure data lake as blobs.
3. Create an Azure data factory to house the ETL, producer, and consumer databricks.
4. Load the census and prudential blobs into the ETL databrick to be cleaned and transformed.
5. Load the cleaned census and prudential blobs into their own separate Kafka producer databrick.
6. Use one Kafka consumer databrick for each dataset to retrieve the data that the producer has created.
7. Take the consumed data for each dataset and send it back into the data lake as a CSV.
8. Load the consumed data for each dataset from the data lake into the Azure SQL database. Also, load the prudential data into the machine learning databrick.
9. Load the census and prudential tables from the SQL database into Power BI. Also, load the machine learning model into Power BI.
10. Load the Power BI reports into a Power BI dashboard.
11. Use the Power BI dashboard in the final presentation.
12. Upload the final presentation slides, Power BI dashboard, reports, napkin drawings, plans, questions, and code to Github.