Project Proposal

Brazil's temperature may influence bus tickets, lead to traffic accidents, and crime events

Student NetID: Amit Porje (aap9114), Victor Guillemot (vbg5833), Sneha Singh(ss15639), Minghe Yang(my2306).

Abstract

We're assuming that increasing temperatures will have a negative influence on the sale of bus tickets, which leads to more vehicles on the road, and thus more traffic accidents. Since high temperature is extreme weather, crime events will raise up.

Most of the data are already in the .csv format so regularization should not be a problem. All data can be stored in HDFS, and we planned to combine data using the date field to correlate with each other.
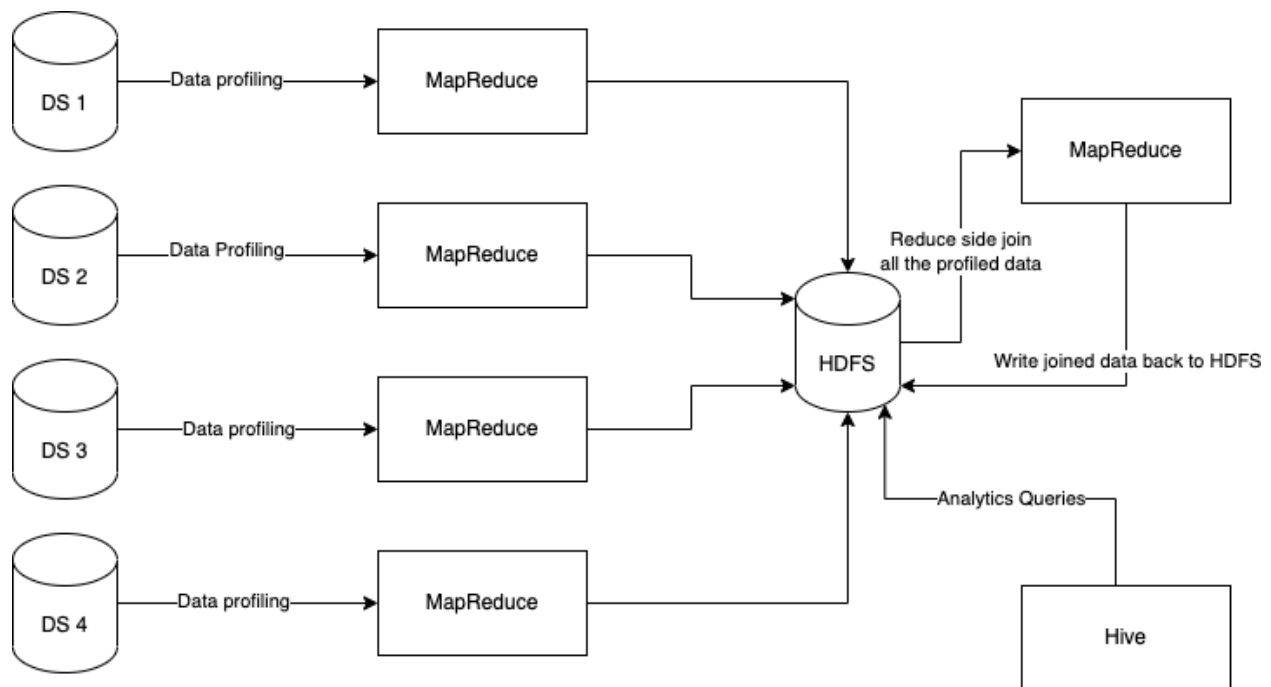
Related Dataset with members responsible for cleaning and profiling

1. Brazil Highway Traffic Accidents:271MB - Sneha Singh
   https://www.kaggle.com/datasets/mcamera/brazil-highway-traffic-accident

2. Climate Weather Surface of Brazil – Hourly:2GB - Amit Porje
   https://www.kaggle.com/datasets/PROPPG-PPG/hourly-weather-surface-brazil-southeast-region

3. Brazil Interstate Bus Tickets:431MB - Minghe Yang
   https://www.kaggle.com/datasets/danlessa/brazil-interstate-bus-travels

4. Crime Data in Brazil:934MB - Victor Guillemot
   https://www.kaggle.com/datasets/inquisitivecrow/crime-data-in-brazil

Potential additional datasets

1. [Data of Brazilian Import and Export data:2GB]
   https://www.kaggle.com/datasets/juniorfazzio/data-of-brazilian-import-and-export-data

2. [Brazil Weather, Automatic Stations (2000-2021):2GB]
   https://www.kaggle.com/datasets/saraivaufc/automatic-weather-stations-brazil

Design Diagram

Legend:

DS - Datasource