US GEOGRAPHIC: TRAFFIC AND GEOGRAPHIC DATA ALMANAC

THE IDEA

THERE ARE RELATIONSHIP BETWEEN ACCIDENTS, WEATHER AND AIR-POLLUTION. OUR SYSTEM IS TRYINGS TO FIND THESE RELATIONSHIP AND TRENDING AMONG US TRAFFIC DATA IN ORDER TO DRAW A COMPREHENSIVE CONCLUSION. ADDITIONALLY, WE EVALUATED CERTAIN ENVIRONMENTAL AND SOCIO-ECONOMIC FACTORS TO ENHANCE OUR FINDING.

FIVE SCHEMAS

ACCIDENTS
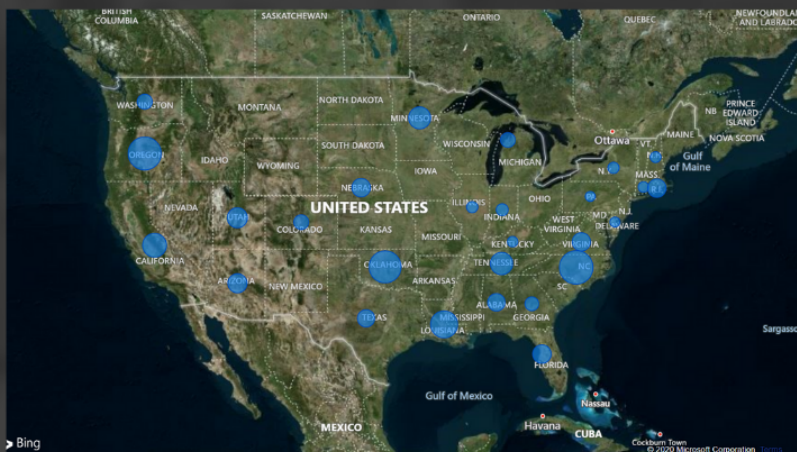~3.9 M

AUTO
~1.3 M

COUNTY
~12 K

POLLUTION
~4.0 M

WEATHER
~4.5 M

TUPLES PROCESSED

13,845,773

CHALLENGES FACED

- THE FIRST ISSUE WE FACE WAS THE INCONSISTENCIES BETWEEN THE DIFFERENT DATASETS AND THE MAPPING.
- THE MAIN ISSUES WE FACED ARE ALL IN DB/SCHEMA DESIGN, WE FACED THE ISSUE OF WHICH ATTRIBUTES TO KEEP AND WHICH TO DROP.
- ADDITIONALLY, WE FACED MANY CHALLENGES DETERMINING THE KEYS FOR THE SCHEMAS/TABLES

USA