DS 2002 F24

ENERGY ANALYSIS

Norah Masrour
Sam O'Brien
Masa Murray

DATA SELECTION

A look into which data sets we chose, why we chose them, and where we went with them.
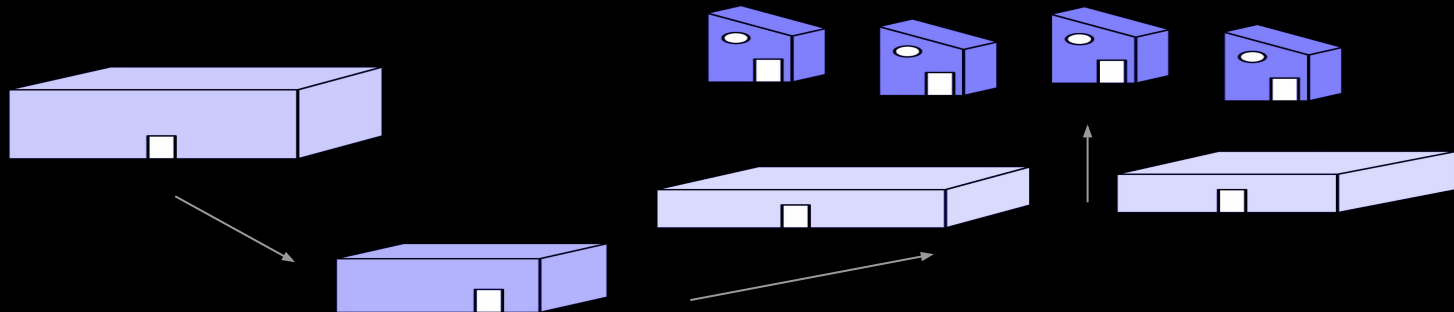
Cummulative_Sum_Year_Full_Data

Year, Technology Type, Project Name, Technology Sub-Type, Country, Power, Storage Capacity
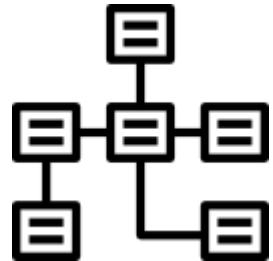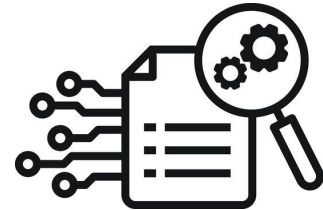
Global_Map_Full_Data_Data

Latitude, Longitude, Project Name, Country, Rated Capacity, Rated Power, Discharge Duration at Rated Power, Storage Capacity
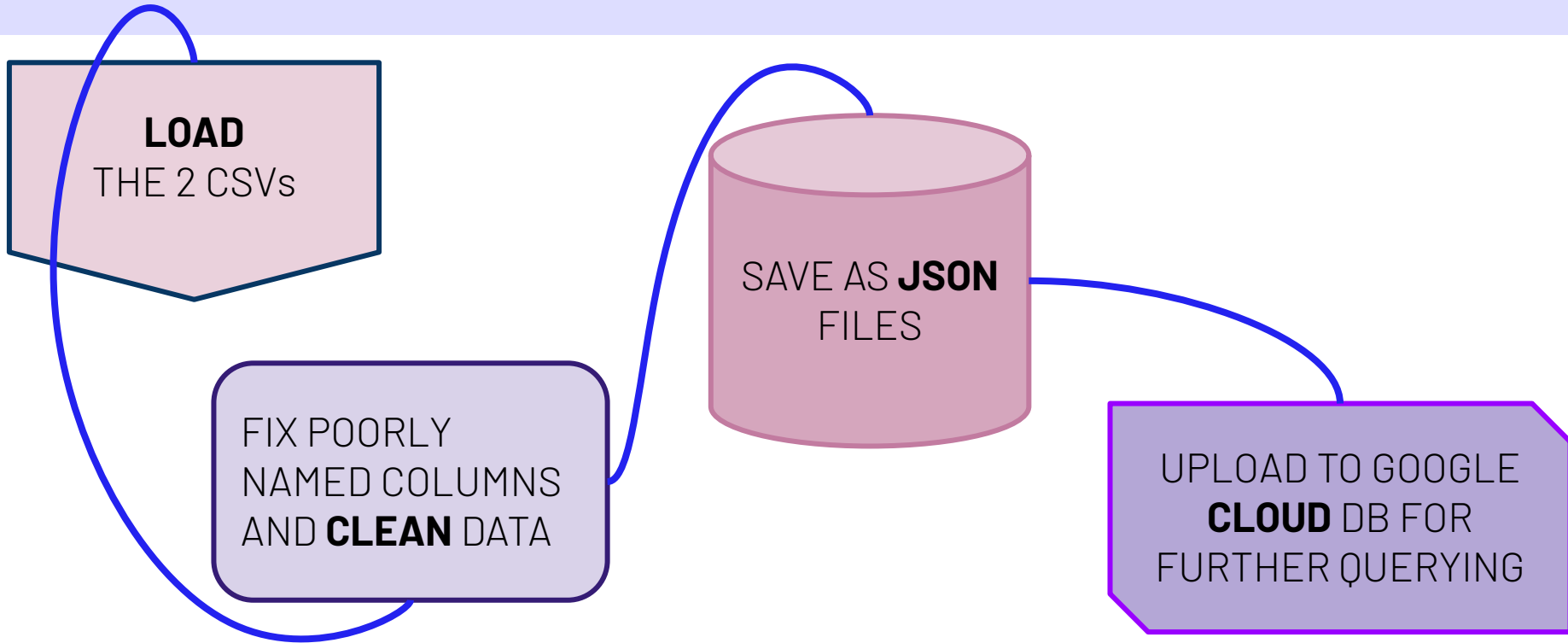
WHY?

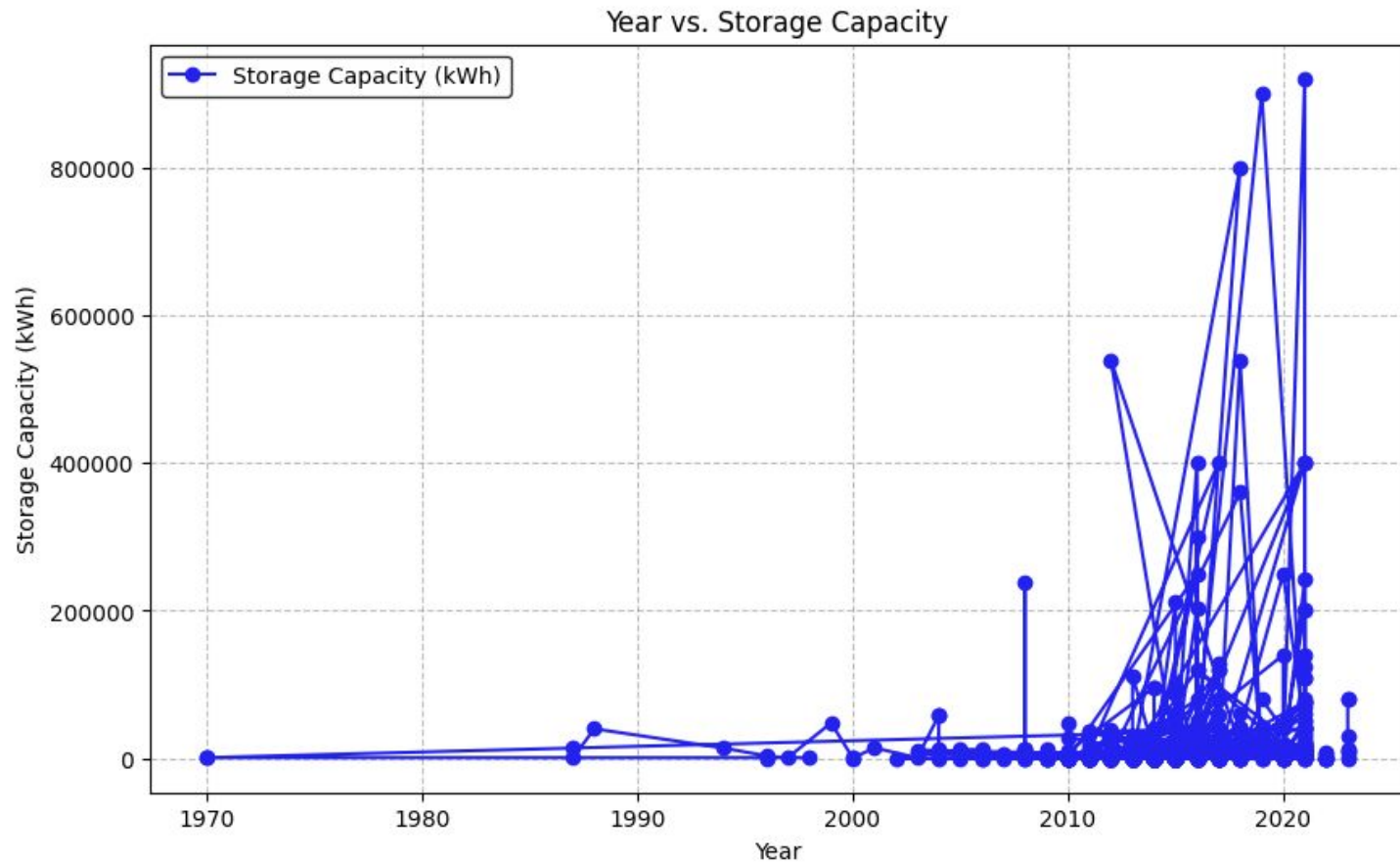# CHALLENGES

- Auto-uploading to **Cloud**
- Data Comprehension
  - Missing **Context**
  - Metrics Complexity
- Determining Ideal **Storage Schema**
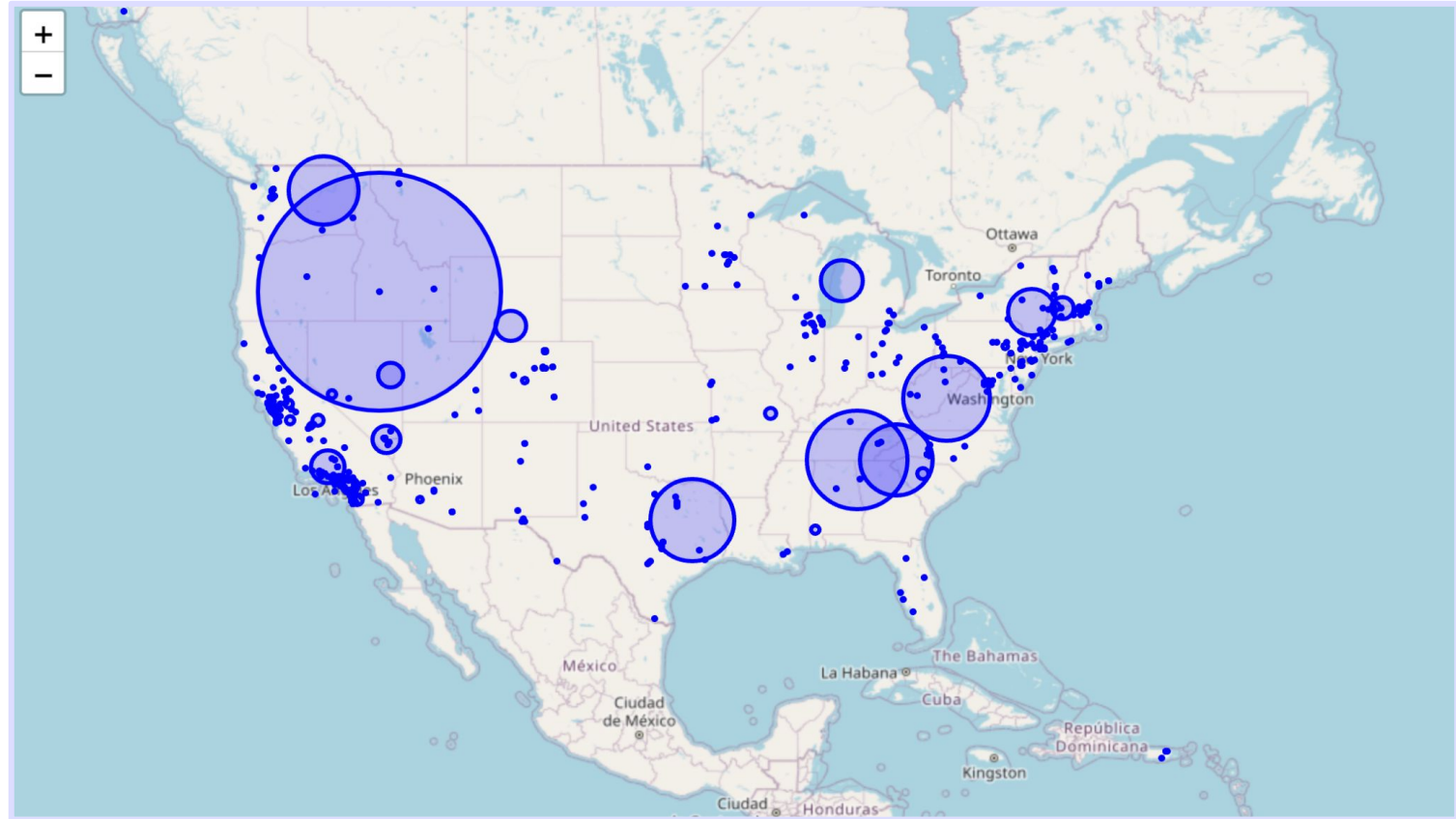 - Simple? Complex?

ETL path and logic

LOAD
THE 2 CSVs

FIX POORLY NAMED COLUMNS AND **CLEAN** DATA

SAVE AS **JSON** FILES

UPLOAD TO GOOGLE **CLOUD** DB FOR FURTHER QUERYING

DATA PLOTTING

Year vs. Storage Capacity

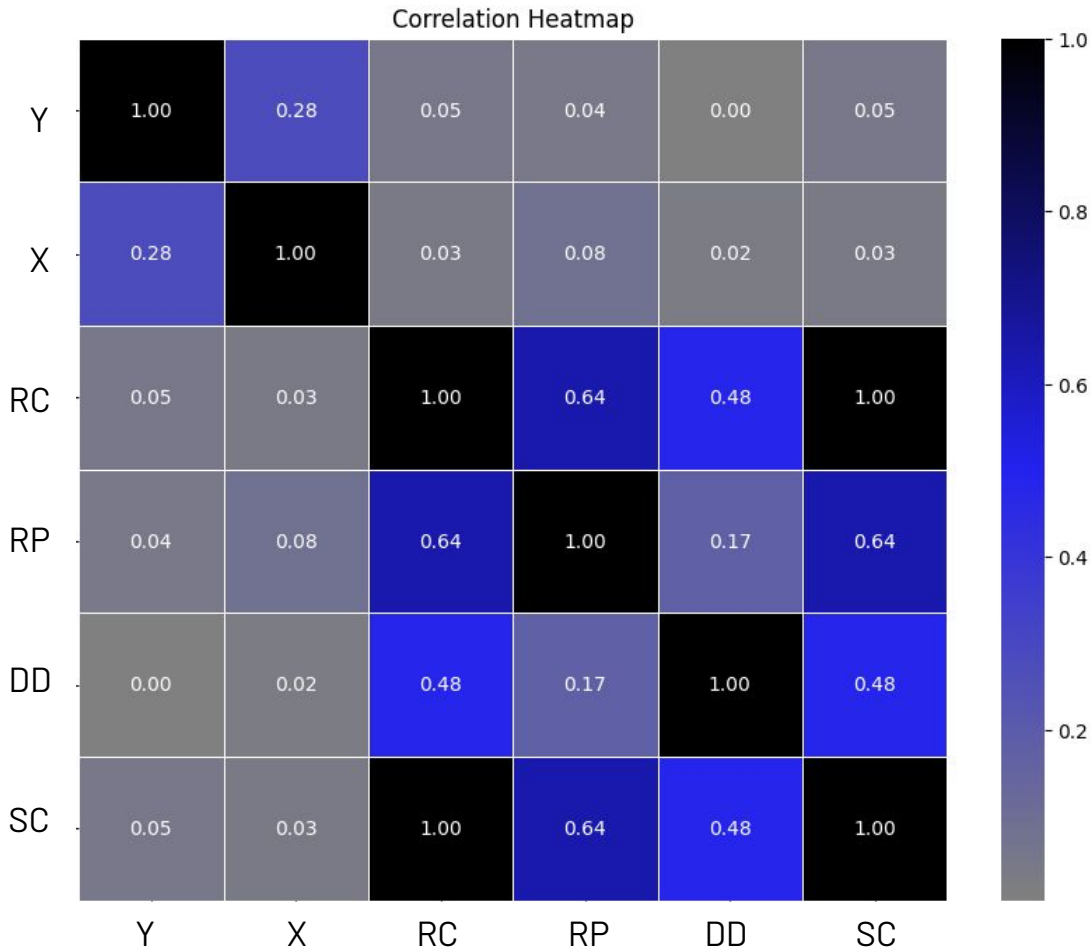# DATA MAPPING

DATA CORRELATIONS

Y - Latitude

X - Longitude

RC - Rated Capacity (kWh)

RP - Rated Power (kW)

DD - Discharge Duration at Rated Power (hrs)

SC - Storage Capacity

Correlation Heatmap

	Y	X	RC	RP	DD	SC
Y	1.00	0.28	0.05	0.04	0.00	0.05
X	0.28	1.00	0.03	0.08	0.02	0.03
RC	0.05	0.03	1.00	0.64	0.48	1.00
RP	0.04	0.08	0.64	1.00	0.17	0.64
DD	0.00	0.02	0.48	0.17	1.00	0.48
SC	0.05	0.03	1.00	0.64	0.48	1.00

LESSONS

- Impact of data **visualization**
- Logical **pipelines**
- Informative connectedness to trends with **context**
- Understanding how to analyze **correlations** between factors