

# **ANALYSIS ON CLIMATE CHANGE**

Group - 2

01

# Introduction

Aim of this project

# AIM

The primary aim of this project is to create a model that will be useful in predicting the patterns in environmental change in different areas of the world in the coming years and to utilize this data to study weather conditions, explicitly as to temperature, precipitation, and irregularities.

We will also analyze the connection between the current shift in weather patterns and the historical production of carbon emissions by various countries. By expanding our understanding of climate change, extreme weather, and regional variations, the project aims to advance climate science.

This will assist with asset portion, debacle planning, and environment variation independent direction, eventually encouraging strength and maintainability.

02

# Datasets

Datas used for this project

# Cities

station_id	city_name	country	state	iso2	iso3	latitude	longitude
41515	Asadabad	Afghanistan	Kunar	AF	AFG	34.86600004	71.15000459
38954	Fayzabad	Afghanistan	Badakhshan	AF	AFG	37.12976076	70.57924719
41560	Jalalabad	Afghanistan	Nangarhar	AF	AFG	34.44152692	70.43610347
38947	Kunduz	Afghanistan	Kunduz	AF	AFG	36.72795066	68.87252966
38987	Qala i Naw	Afghanistan	Badghis	AF	AFG	34.98300013	63.13329964
38915	Sheberghan	Afghanistan	Jawzjan	AF	AFG	36.65798077	65.73830237
13577	Peshkopi	Albania	Dibër	AL	ALB	41.6833021	20.43330349
13461	Shkodër	Albania	Shkodër	AL	ALB	42.06845156	19.51884965
13615	Tirana	Albania	Durrës	AL	ALB	41.32754071	19.81888301
60620	Adrar	Algeria	Adrar	DZ	DZA	27.86999005	-0.2899670831
60369	Algiers	Algeria	Alger	DZ	DZA	36.7630648	3.05055253
60360	Annaba	Algeria	Annaba	DZ	DZA	36.92000612	7.759980834
60468	Batna	Algeria	Batna	DZ	DZA	35.56995933	6.170000365
60525	Biskra	Algeria	Biskra	DZ	DZA	34.85997683	5.73002722
60444	Bordj Bou Arrerîj	Algeria	Bordj Bou Arrerîj	DZ	DZA	36.05900401	4.629996466
60571	Béchar	Algeria	Béchar	DZ	DZA	31.61110537	-2.230003704
60402	Béjaïa	Algeria	Béjaïa	DZ	DZA	36.76037762	5.070015827

# Countries

country	native_name	iso2	iso3	population	area	capital	capital_lat	capital_lng	region	continent
Afghanistan	افغانستان	AF	AFG	26023100	652230	Kabul	34.526011	69.177684	Southern and Cr Asia	
Albania	Shqipëria	AL	ALB	2895947	28748	Tirana	41.326873	19.818791	Southern Europe	Europe
Algeria	الجزائر	DZ	DZA	38700000	2381741	Algiers	36.775361	3.060188	Northern Africa	Africa
American Samoa	American Samoa	AS	ASM	55519	199	Pago Pago	-14.275479	-170.70483	Polynesia	Oceania
Angola	Angola	AO	AGO	24383301	1246700	Luanda	-8.82727	13.243951	Central Africa	Africa
Anguilla	Anguilla	AI	AIA	13452	91	The Valley	41.559572	-98.880548	Caribbean	North America
Antigua and Barb	Antigua and Barb	AG	ATG	86295	442	Saint John's	47.861701	-52.715149	Caribbean	North America
Argentina	Argentina	AR	ARG	42869500	2780400	Buenos Aires	-34.607566	-58.437089	South America	South America
Armenia	Հայաստան	AM	ARM	3009800	29743	Yerevan	40.177612	44.512585	Middle East	Asia
Aruba	Aruba	AW	ABW	101484	180	Oranjestad	12.526874	-70.035684	Caribbean	North America
Australia	Australia	AU	AUS		7692024	Canberra	-35.297591	149.101268	Australia and Ne Oceania	
Austria	Österreich	AT	AUT	8527230	83871	Vienna	48.208354	16.372504	Western Europe	Europe
Azerbaijan	Azərbaycan	AZ	AZE	9552500	86600	Baku	40.376443	49.832675	Middle East	Asia
Bahrain	البحرين	BH	BHR	1316500	765	Manama	26.223504	50.582244	Middle East	Asia
Bangladesh	ভাৰত	BD	BGD	157486000	147570	Dhaka	23.769357	90.378814	Southern and Cr Asia	
Belarus	Беларусь	BY	BLR	9475100	207600	Minsk	63.902334	27.561879	Eastern Europe	Europe
Belgium	België	BE	BEL	11225469	30528	Brussels	50.846557	4.351697	Western Europe	Europe
Belize	Belize	BZ	BLZ	349728	22966	Belmopan	17.250199	-88.770018	Central America	North America

# Daily Weather

station_id	city_name	date	season	avg_temp_c	min_temp_c	max_temp_c	precipitation_mm	snow_depth_mm	avg_wind_dir_deg	avg_wind_speed	peak_wind_gust	avg_sea_level_f	sunshine_total_f	index_level_0
41515	Asadabad	6/30/1957, 5:00	Summer	27	21.1	35.6	0	-	-	-	-	-	-	0
41515	Asadabad	7/1/1957, 5:00	Summer	22.8	18.9	32.2	0	-	-	-	-	-	-	1
41515	Asadabad	7/2/1957, 5:00	Summer	24.3	16.7	36.6	1	-	-	-	-	-	-	2
41515	Asadabad	7/3/1957, 5:00	Summer	26.6	16.1	37.8	4.1	-	-	-	-	-	-	3
41515	Asadabad	7/4/1957, 5:00	Summer	30.8	20	41.7	0	-	-	-	-	-	-	4
41515	Asadabad	7/5/1957, 5:00	Summer	30.2	22.8	41.1	0	-	-	-	-	-	-	5
41515	Asadabad	7/6/1957, 5:00	Summer	31	24.4	39.4	0	-	-	-	-	-	-	6
41515	Asadabad	7/7/1957, 5:00	Summer	30.9	24.4	38.9	0	-	-	-	-	-	-	7
41515	Asadabad	7/8/1957, 5:00	Summer	26.1	21.1	34.4	2	-	-	-	-	-	-	8
41515	Asadabad	7/9/1957, 5:00	Summer	26	-	35.6	0.3	-	-	-	-	-	-	9
41515	Asadabad	7/10/1957, 5:00	Summer	26.3	17.2	36.1	2	-	-	-	-	-	-	10
41515	Asadabad	7/11/1957, 5:00	Summer	28.8	21.7	36.7	0	-	-	-	-	-	-	11
41515	Asadabad	7/12/1957, 5:00	Summer	27.2	21.1	36.1	0	-	-	-	-	-	-	12
41515	Asadabad	7/13/1957, 5:00	Summer	28	20.6	36.1	0.3	-	-	-	-	-	-	13
41515	Asadabad	7/14/1957, 5:00	Summer	28.6	21.1	37.2	0	-	-	-	-	-	-	14
41515	Asadabad	7/15/1957, 5:00	Summer	31.7	22.8	41.7	0	-	-	-	-	-	-	15
41515	Asadabad	7/16/1957, 5:00	Summer	33.1	23.3	46.1	0	-	-	-	-	-	-	16
41515	Asadabad	7/17/1957, 5:00	Summer	33.3	26.1	41.1	0	-	-	-	-	-	-	17
41515	Asadabad	7/18/1957, 5:00	Summer	30.1	25	35.6	1	-	-	-	-	-	-	18
41515	Asadabad	7/19/1957, 5:00	Summer	27.6	21.1	34.4	3	-	-	-	-	-	-	19
41515	Asadabad	7/20/1957, 5:00	Summer	28.8	22.2	35	0	-	-	-	-	-	-	20
41515	Asadabad	7/21/1957, 5:00	Summer	27.4	21.7	33.9	0.3	-	-	-	-	-	-	21

# Cumulative greenhouse gas emission

CNTR_NAME	ISO3	Gas	Component	Year	Data	Unit
Afghanistan	AFG	3-GHG	Fossil		1851	0.000454704256 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1852	0.000913130773 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1853	0.001375296506 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1854	0.00184121966 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1855	0.002310915871 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1856	0.002784399916 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1857	0.003261685266 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1858	0.003742784294 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1859	0.004227706791 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1860	0.004716465226 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1861	0.005209240717 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1862	0.005706055556 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1863	0.00620693253 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1864	0.006711894976 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1865	0.007220965284 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1866	0.007734169016 Pg-CO[2]*-e[100]
Afghanistan	AFG	3-GHG	Fossil		1867	0.008251530954 Pg-CO[2]*-e[100]

# Real time data

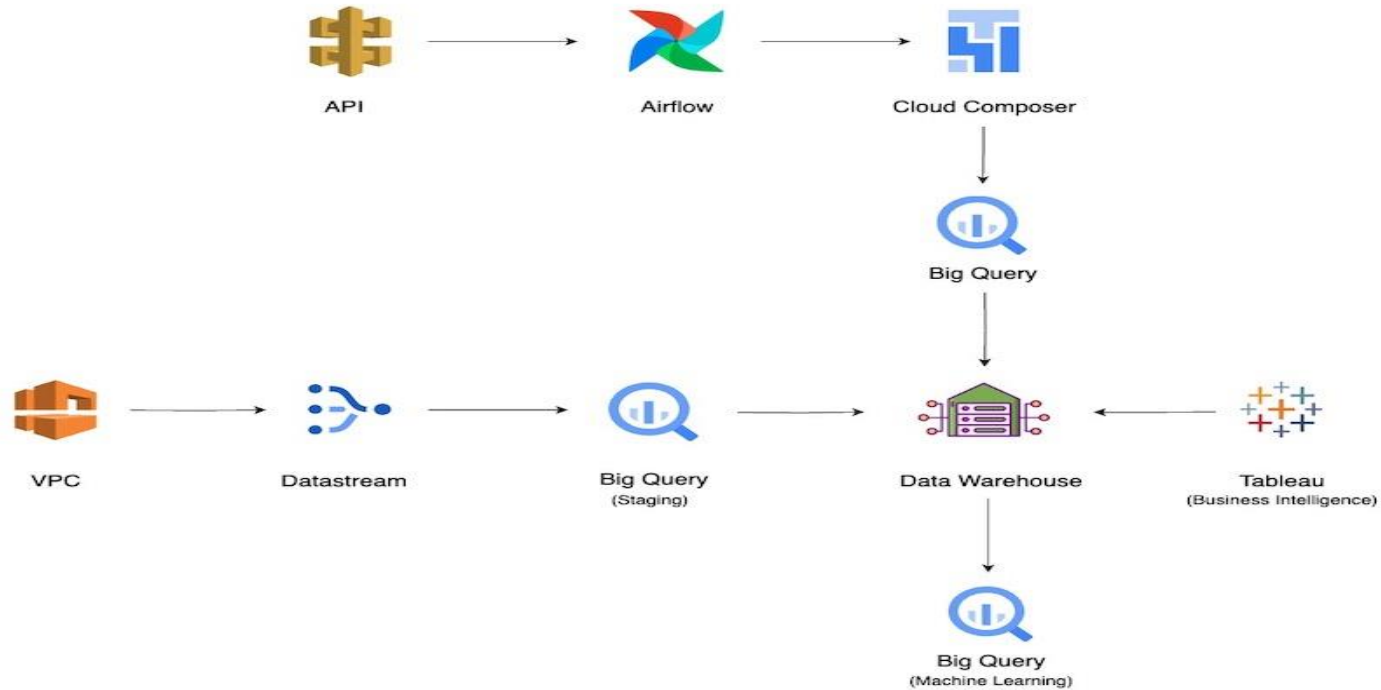
city_name	date	season	avg_temp_c	min_temp_c	max_temp_c	precipitation_mm	avg_wind_dir_deg	avg_wind_speed_km	peak_wind_gust_kmh	avg_sea_level_pres	sunshine_total_min
Helena	2023-12-03	Winter	2.43	-2.49	4.54	0	249	26.712	66.744	1013	528.1333333
Montpelier	2023-12-03	Winter	3.53	1.34	3.53	0.99	169	8.244	29.808	1014	544.5333333
Bismarck	2023-12-03	Winter	3.51	-2.36	4.28	0	295	18.036	25.74	1006	526.6
Saint Paul	2023-12-03	Winter	1.79	-1.3	2.18	0	141	9.108	24.588	1010	539.85
Cheyenne	2023-12-03	Winter	-0.34	-3.81	3.3	0	267	63.036	95.4	1014	563.9833333
Madison	2023-12-03	Winter	15.53	9.59	16.23	0.55	272	17.676	27.828	1014	598.35
Denver	2023-12-03	Winter	3.75	2.67	5.49	0	284	23.616	57.708	1013	572.0666667
Des Moines	2023-12-03	Winter	5.36	-1.36	5.42	0	212	11.952	15.84	1009	561.2666667
Indianapolis	2023-12-03	Winter	7.1	2.48	7.42	0	242	23.004	42.372	1008	571.9666667
Lincoln	2023-12-03	Winter	4.94	-2.95	5.97	0	345	17.748	29.592	1007	566
Boise	2023-12-03	Winter	3.49	1.78	5.44	13.19	112	13.392	21.564	1020	548.6166667
Albany	2023-12-03	Winter	4.79	3.87	5.31	18.82	149	14.364	39.132	1011	555.2166667
Topeka	2023-12-03	Winter	7.87	-1.03	8.46	0	323	22.536	42.156	1008	575.9666667
Columbus	2023-12-03	Winter	10.98	5.88	11.79	0.56	249	26.532	53.244	1007	570.8833333
Springfield	2023-12-03	Winter	9.68	-0.82	10.9	0	262	28.764	51.48	1012	585.8166667
Jefferson City	2023-12-03	Winter	6.04	3.41	9.42	0	268	19.548	37.008	1013	578.55
Frankfort	2023-12-04	Winter	32.31	16.67	34.36	0	316	18.972	33.984	1011	825.3166667

03

# Cloud Architecture

How is this architecture build

## Pipeline using the google cloud services:





04

# ELT/ETL processes

How is the data processed

## 4.1

# ELT process for Batch Data

## Batch Data

- Kaggle
- Cloud SQL
- Data stream
- Big query
  - Dump Database (Raw)
  - Staging Database (Transformations)
- Datawarehouse
- Query Analysis and Tableau

4.2

## ETL process for Realtime

# Realtime

- API()
- Airflow
- Big query
  - Dump Database
  - Staggering Database
  - (After Transformations)Final

## Database

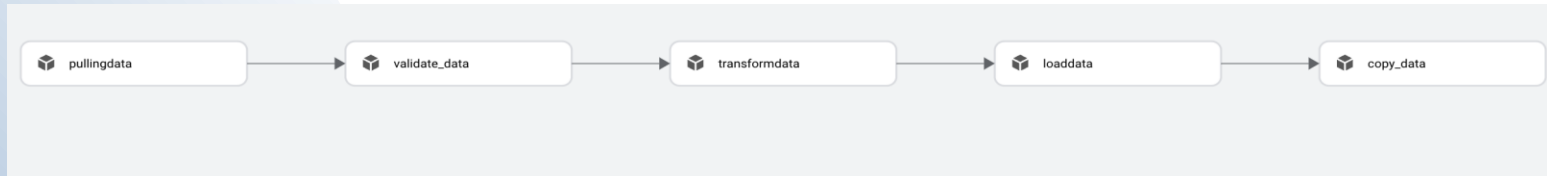
- Datawarehouse
- Tableau

06

# Airflow pipeline

# Airflow

Apache Airflow is an open-source platform to programmatically author, schedule, and monitor workflows. With Airflow, you can automate and schedule complex workflows, such as updating and transforming large datasets in real-time. DAGs can be defined to execute a variety of tasks, and you can easily manage data pipelines. Airflow provides a user-friendly web interface for monitoring and debugging workflows.



---

### **Pulling Data:**

- Extracts data from the openweather API for a specific list of cities.
- Stores the data in **weather\_data\_all**, utilizing Xcom for subsequent task usage.

### **Validate Data:**

- Validates retrieved data, identifying and addressing missing or erroneous entries.
- Processes the data, updating the BigQuery table with the record count and information on any missing records.

### **Transform Data:**

- Structures raw data into a data frame, adjusting column data types as needed.
- Pushes the transformed data into **weather\_df** using Xcom.

### **Load Data:**

- Loads the data into BigQuery, specifying the schema as **data-225-group-project.climate\_dwh.real\_time\_data** and converting to appropriate data types.

### **Scheduling:**

- The pipeline is scheduled to run daily to align with the daily update of the openweather API.
-



07

# DW schema and Implementation

# SCHEMA

## Location Dimension Table:

location_dimension	
location_id	
city_name	
State	
Country_name	
Country_code_ISO3	

## Date Dimension Table:

date_dimension	
date	
Month	
Year	
season	

## Greenhouse Gas Emission Dimension Table:

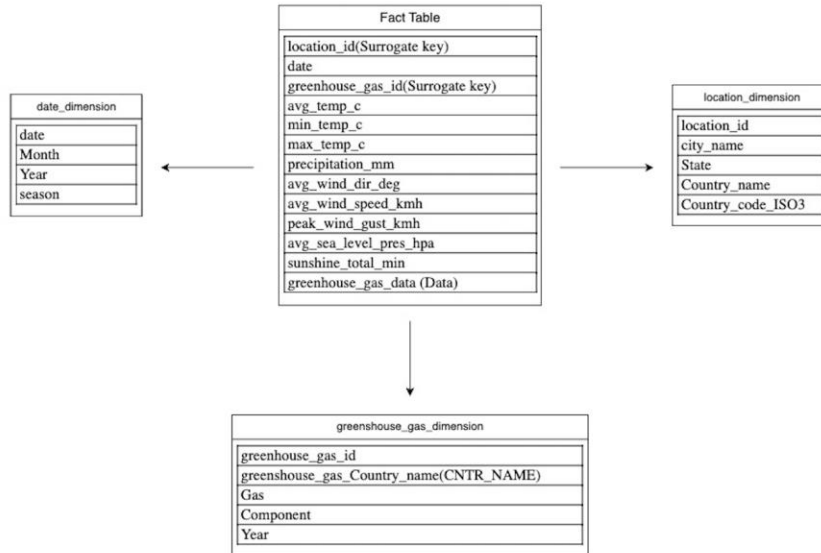
greenhouse_gas_dimension	
greenhouse_gas_id	
greenhouse_gas_Country_name(CNTR_NAME)	
Gas	
Component	
Year	

## Climate Fact Table:

Fact Table
location_id(Surrogate key)
date
greenhouse_gas_id(Surrogate key)
avg_temp_c
min_temp_c
max_temp_c
precipitation_mm
avg_wind_dir_deg
avg_wind_speed_kmh
peak_wind_gust_kmh
avg_sea_level_pres_hpa
sunshine_total_min
greenhouse_gas_data (Data)

## Update Operation on Fact Table:

- An update operation is performed on the fact table to calculate and fill in missing average temperature values.
- The derived average temperature is calculated as the average of minimum and maximum temperatures.

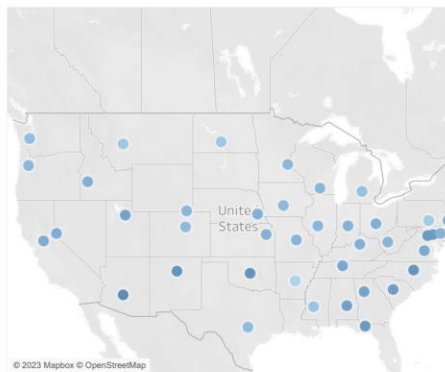


08

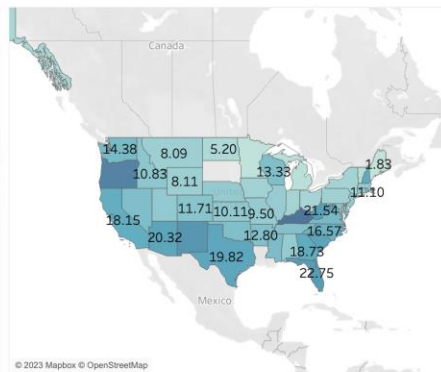
# Visualization

How is the data processed

Temperature across different cities



Maximum Temperatures of each State



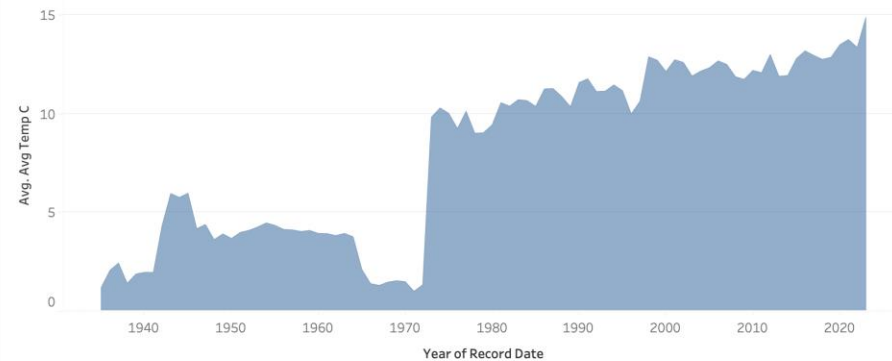
Max. Max Temp C (Rea..

1.83 33.59

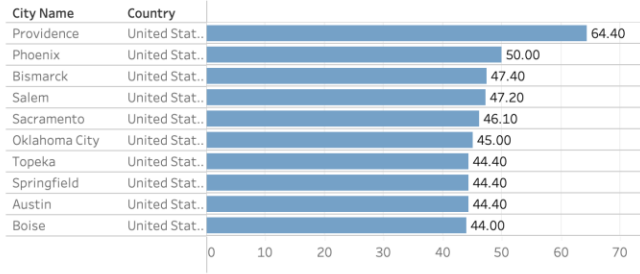
Avg Temp C

35,318 6,707,415

Average Temperature across different years



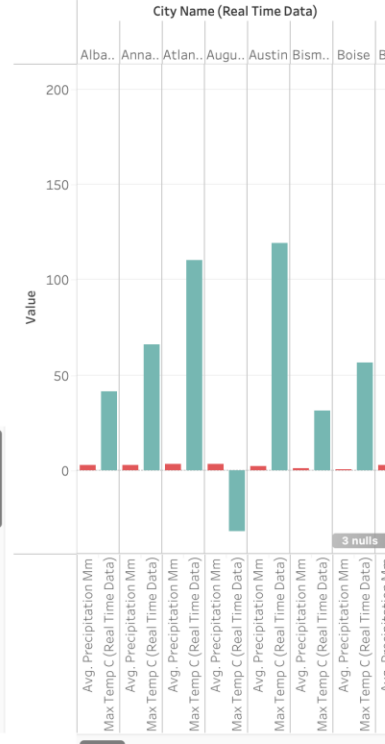
### Top 10 cities with highest temperature



### Monthly Average Temperature Trends by Station:

Stationid	Record Date								
	January	February	March	April	May	June	July	August	Sept
70381	-3.22	-1.60	0.42	4.48	8.79	12.09	13.44	12.82	8.79
72214	11.17	12.72	16.06	19.59	23.61	26.51	27.19	27.15	27.15
72219	6.31	8.15	12.04	16.72	21.17	24.91	26.25	25.86	25.86
72226	8.45	10.64	14.39	18.39	22.59	26.13	27.35	27.09	27.09
72254	10.17	12.27	16.25	20.45	24.26	27.73	29.29	29.48	29.48
72278	12.23	14.17	17.32	21.74	26.53	31.71	34.27	33.31	33.31
72306	4.77	6.30	10.40	15.55	19.77	23.86	25.68	24.91	24.91
72310	7.36	9.23	13.09	17.80	22.13	25.80	27.32	26.63	26.63
72327	3.47	5.50	10.16	15.51	20.25	24.58	26.45	25.78	25.78
72340	4.81	7.10	12.21	17.07	21.58	26.12	28.22	27.76	27.76
72353	2.79	5.29	10.09	15.61	20.21	25.04	27.68	27.20	27.20
72365	1.86	4.68	8.57	13.43	18.57	24.05	25.62	24.31	24.31
72401	3.09	4.59	8.87	14.52	19.20	23.65	25.76	24.86	24.86
72405	2.35	3.69	8.06	13.85	19.10	23.95	26.33	25.42	25.42

### Avg Precipitation of each city by Maximum Temp



09

# Limitations and Future Scope

How is the data processed



---

# Limitations

- Centered to only to all the cities in USA
- The real-time data is not consistent to the archived data

# Future Scope

We hope to improve our analysis and prediction for other countries as well.

---

10

# Conclusion

How is the data processed