

ClimaCity

A Spatial Temporal Visualization Tool of Climate-Demographic Interactions and Built Environments

Presented by:

Jiahe Ling, Galav Minesh Bhatt, Jinxuan Lyu, Yue Liu, Ziqi Meng



COLUMBIA ENGINEERING

The Fu Foundation School
of Engineering and Applied Science



Overview

- Background Information
 - Urban Heat & Climate Change
 - Introduction to Greenhouse Gas Emissions Scenario
 - Existing Visualization Tools
- Data
 - Population
 - Land Use
 - Housing Density
 - City Boundaries
- Building the Visualization Panel
- Temperature Analysis
- Conclusion

Background Information

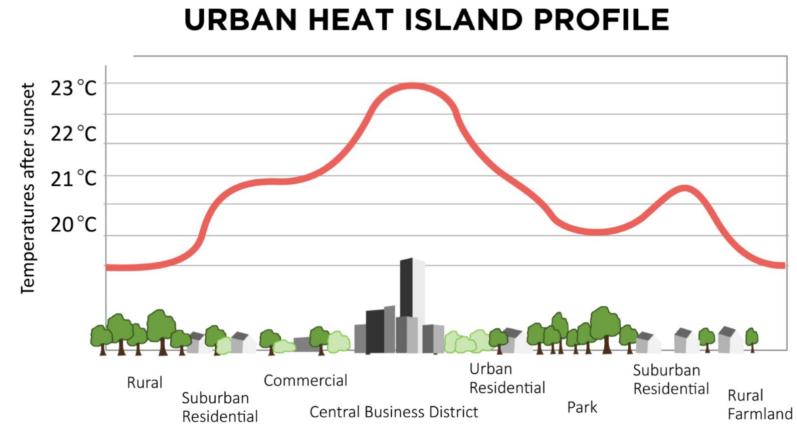
Urban Heat & Climate Change

Urban Heat Island Effect

Cities experience significantly higher temperatures than surrounding rural areas due to extensive concrete surfaces, dark pavement, and limited vegetation.

Why Modeling Matters

By linking temperature data with land-use patterns, housing density, and population demographics, we can generate clearer insights into where heat risks will emerge and how cities can respond proactively.



Background Information

Greenhouse Gas Emissions Scenario

Representative Concentration Pathways (RCPs)

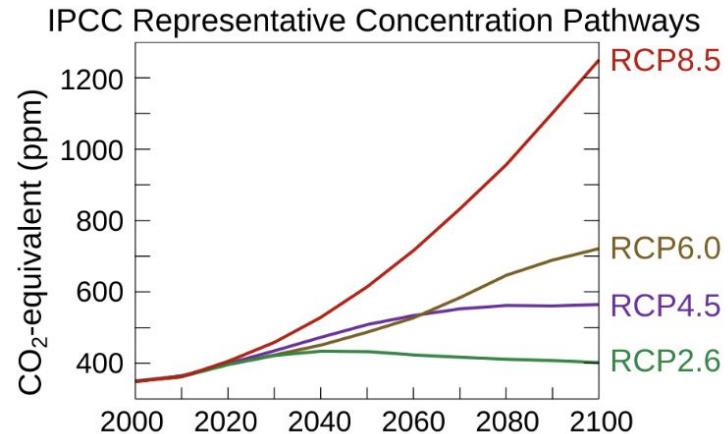
RCP refers to different climate change scenarios that project how greenhouse gas levels may evolve in the future, leading to different degrees of global warming.

RCP 4.5

- Radiative forcing stabilizes at 4.5 W/m² by 2100.
- Results in a moderate level of global warming (roughly 2–3 °C by 2100).

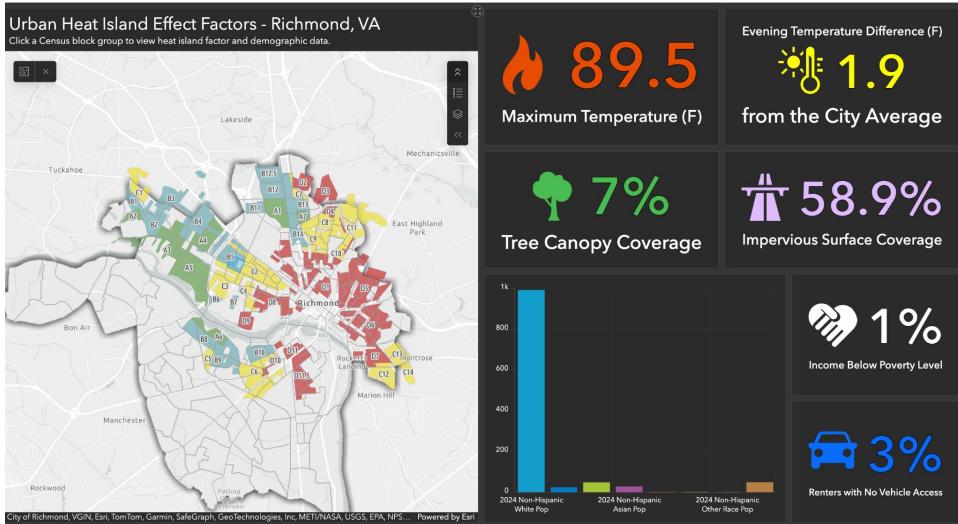
RCP 8.5

- Radiative forcing reaches 8.5 W/m² by 2100.
- Assumes continued growth in fossil-fuel use and minimal climate mitigation.
- The basis for worst-case climate change scenarios.



Background Information

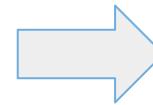
Existing Visualization Tools



Our Project

Datasets

- Temperature
- Population
- Housing density
- Land use



Scenario Framework & Time Span

- RCP 4.5 and RCP 8.5
- Span from 2010 – 2100

Population Data Overview

Data Sources & Methodology

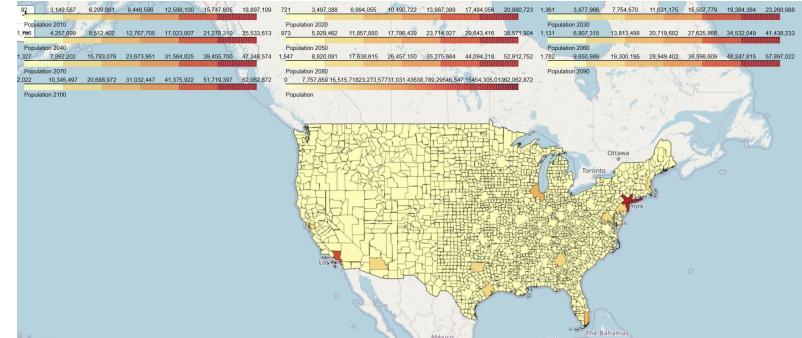
- Base Data: ICLUS population data 2010-2100 every decade
- Spatial Scale: County-level aggregation
- Visualization: Use Choropleth for data visualization

Geographic Coverage

- 48 US States (except Alaska and Hawaii as the data is NA)
- Consistent county boundaries across time periods

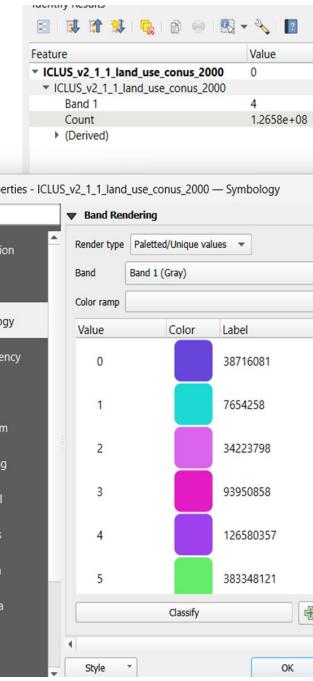
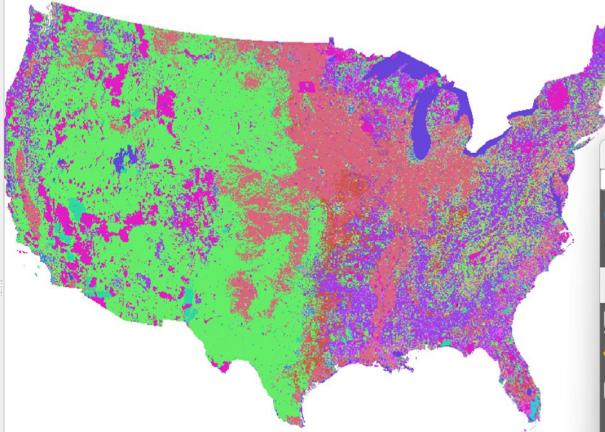
Data Characteristics

- **Format:** The original data format is gbd, convert into GeoJSON before analysis
- **Resolution:** Decadal time steps
- **Variables:** Total population counts



Land Use Data Overview

Data Source: ICLUS Version 2.1.1 United States Land Use Projections Data from 2000 to 2100



Code	Group	Class Name
0	Water	Natural water
1		Reservoirs, canals
2		Wetlands
3		Recreation, conservation
4	Protected	Timber
5		Grazing
6		Pasture
7		Cropland
8	Working/production	Mining, barren land
9		Parks, golf courses
10		Exurban, low density
11		Exurban, high density
12	Developed	Suburban
13		Urban, low density
14		Urban, high density
15		Commercial
16	Industrial	Industrial
17		Institutional
18		Transportation

Land use projection data visualization using QGIS software

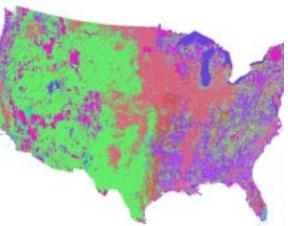
Land use projection color code classification

The provided data consists of raster images with a temporal resolution of 10 years.

Information Credit: <https://www.epa.gov/gcx/iclus-downloads>

Land Use Data Extraction : Raster File to CSV

STEP 1 : INPUT DATA



11 ICLUS Raster
GeoTIFF files from
2000 to 2100
(100+ GB Data)

STEP 2 : DATA PROCESSING



Used the Python **geopandas** library to load the shapefile.
Used the Python **Rasterio** library to load GeoTIFF files.

STEP 3 : OUTPUT DATA



Structured CSV Format
Converted visual map data into quantifiable numbers.

A	B	C	D	E	F	G	H
city_geoid	city_name	state_abbr	year	land_use_code	land_use_name	pixel_count	
1	660620 Richmond	6 CA	2000	0	Water_Natural Water	873	
2	660620 Richmond	6 CA	2000	1	Water_Reservoirs_and_Canals	5	
3	660620 Richmond	6 CA	2000	2	Water_Wetlands	4	
4	660620 Richmond	6 CA	2000	3	Protected_Recreation_Conversation	182	
5	660620 Richmond	6 CA	2000	4	Working_Production_Timber	22	
6	660620 Richmond	6 CA	2000	5	Working_Production_Grazing	257	
7	660620 Richmond	6 CA	2000	9	Developed_Parks_GolfCourses	1335	
8	660620 Richmond	6 CA	2000	10	Developed_Eurban_LowDensity	20	
9	660620 Richmond	6 CA	2000	11	Developed_Eurban_HighDensity	749	
10	660620 Richmond	6 CA	2000	12	Developed_Suburban	874	
11	660620 Richmond	6 CA	2000	13	Developed_Urban_LowDensity	2318	
12	660620 Richmond	6 CA	2000	14	Developed_Urban_HighDensity	278	



US Census
Bureau shapefile
containing all
~32,608 cities
and census
designated
places in the
nation



Processing Engine Spatial Harmonization -
Reprojected all 32000 cities' vector polygons on the specific raster file.
Zonal Extraction -
Processed 32000 shape for land use type pixel count

Final Result : Exact area (pixel count) for 19 different land use types for each city and year.

Housing Density Data Overview

Data Source & Classification :

Source:

ICLUS projected housing density data for conterminous United States, 2010-2000(10-year interval)

Original format:

Gridded raster of housing density classes.

Classifications:

1 = rural (>40 acres/unit) ;

2 = exurban (2–40 acres/unit);

3 = suburban (0.25–2 acres/unit) ;

4 = urban (<0.25 acres/unit) ;

99 = commercial / industrial ;

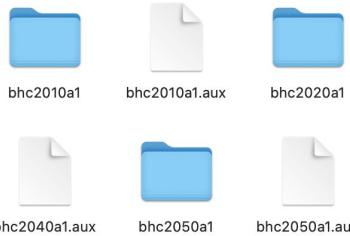
255 = unknown



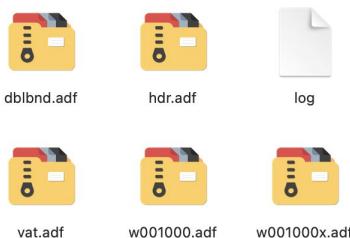
Housing Density Data Overview

How to get the csv version?

1. Import raw data into QGIS



2. Export each interval as GeoTIFF



4. Compute city level statistics, export CSVs

- **Long Table:** important variables(e.g. pixel count, class name...)
- **Summary Table:** more detailed variables(acre for each areas)

3. Load GeoTIFF and City Boundaries shapefiles into python:

- Use rasterio to load housing-density GeoTIFFs (one per year)
- Use geopandas to load the 2023 Census “Places” shapefile (city boundaries) and reproject it to the same CRS as the rasters

	A	B	C	D	E	F	G	H	I	J
1	city_geoid	city_name	state	year	class_cc	class_name	pixel_count	area_m2	area_acres	
2	0660620	Richmond	CA	2010	1	rural (>40 acres/unit)	612	6120000	1512.2849345790519	
3	0660620	Richmond	CA	2010	2	exurban (2–40 acres/	1553	15530000	3837.5465741850776	
4	0660620	Richmond	CA	2010	3	suburban (0.25–2 acr.	1864	18640000	4606.044310547962	
5	0660620	Richmond	CA	2010	4	urban (<0.25 acres/ur	1388	13880000	3429.8226947642547	
6	0660620	Richmond	CA	2010	99	commercial/industrial	254	2540000	627.6477	
7	0660620	Richmond	CA	2010	255	unknown	3118	31180000	7704.745794146215	
8	1206100	Beverly B	FL	2010	2	exurban (2–40 acres/	6	60000	14.82632288802992	
9	1206100	Beverly B	FL	2010	3	suburban (0.25–2 acr.	78	780000	192.74219754438897	
0	1206100	Beverly B	FL	2010	4	urban (<0.25 acres/ur	10	100000	24.710538146716534	
11	1206100	Beverly B	FL	2010	255	unknown	18	180000	44.47896866408976	
2	0477000	Tucson	AZ	2010	1	rural (>40 acres/unit)	3485	34850000	8611.622544130712	
3	0477000	Tucson	AZ	2010	2	exurban (2–40 acres/	7382	73820000	18241.31925906144	
4	0477000	Tucson	AZ	2010	3	suburban (0.25–2 acr.	16543	1.65E+08	40878.64325611316	
5	0477000	Tucson	AZ	2010	4	urban (<0.25 acres/ur	5483	54830000	13548.788065844676	

City Boundaries Data

Data for Plotting

- **Data Source: US Census Bureau**

Census.gov / Census Geographies / Census Mapping Files / [Cartographic Boundary Files - Shapefile](#)

- **Batch Ingestion & Reprojection (56 States)**

Iteratively loads state-level Census shapefiles, and reprojects geometry to **WGS84 (EPSG:4326)** for web-standard compatibility.

- **City Name Standardization (8 Cities)**

Applies a Rule-Based Cleaning Engine to normalize city names for cross-dataset joining (e.g. "St. Paul" vs. "Saint Paul", "Urban Honolulu" vs. "Honolulu")

Place



Building Visualization Panel

Data Preparation:

The system accepts 4 distinct GeoData Frames: Temperature, Land Use, Housing, and Population

Scope Control

Automatically filters by Target State (e.g., "CA") and Climate Scenario (RCP 4.5/8.5), slices data in targeted time period.

Normalization

Standardize temperature scale and corresponding color to ensure consistency, then establish fixed color-to-category assignments for the housing and land use layers to maintain visual continuity.

Spatial Topology

To prioritize city temperature, the system applies a geometric difference operation that masks overlaps. This cuts 'holes' into the environmental data shapes where cities exist, ensuring clear visibility without visual clutter.

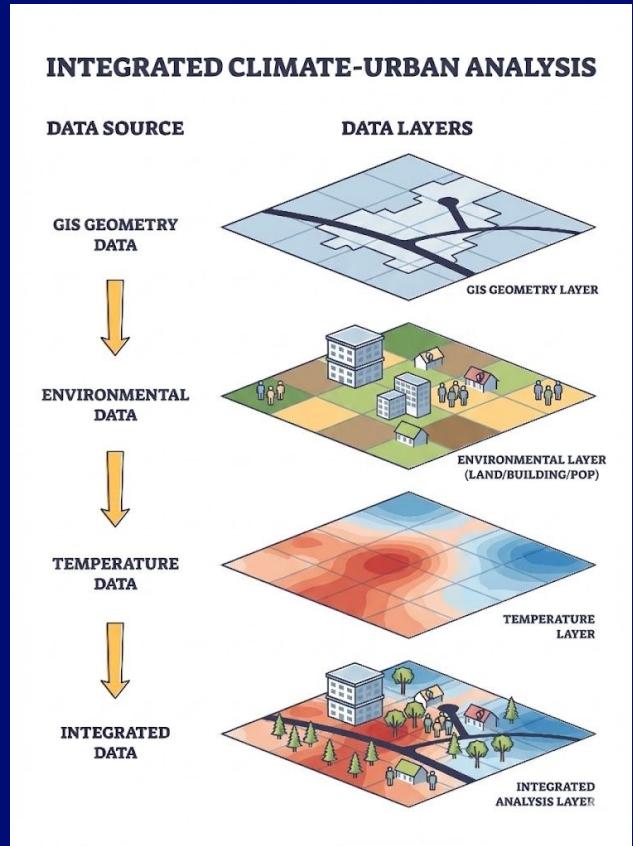
Temporal Alignment

All datasets are aligned to a unified timeline, generating discrete frames for the animation slider.

Building Visualization Panel

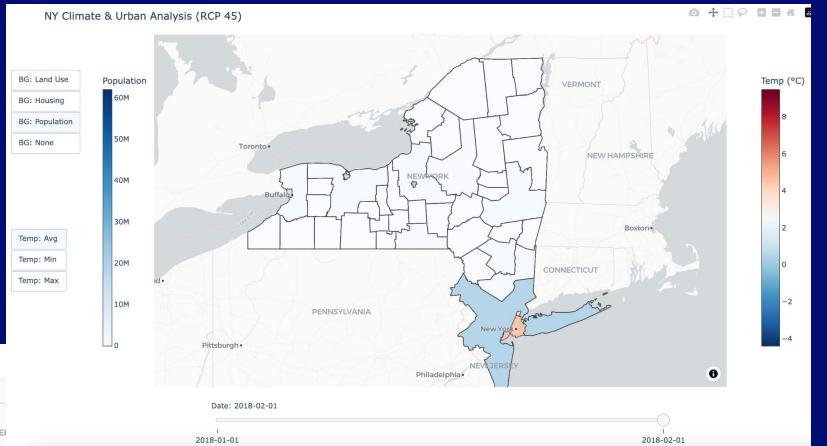
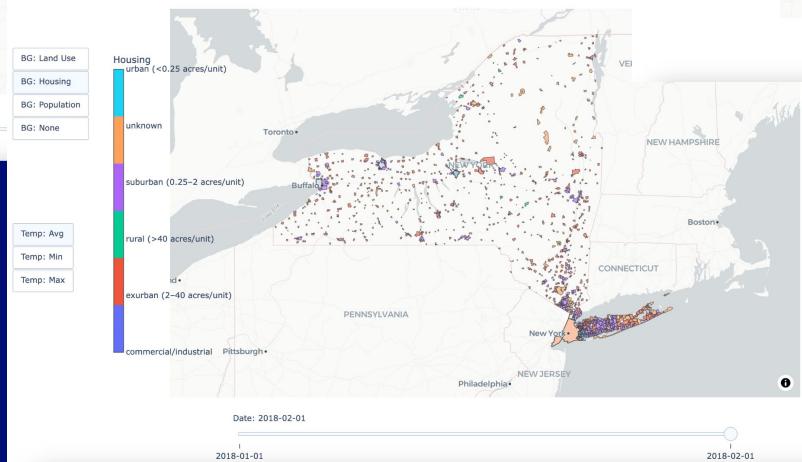
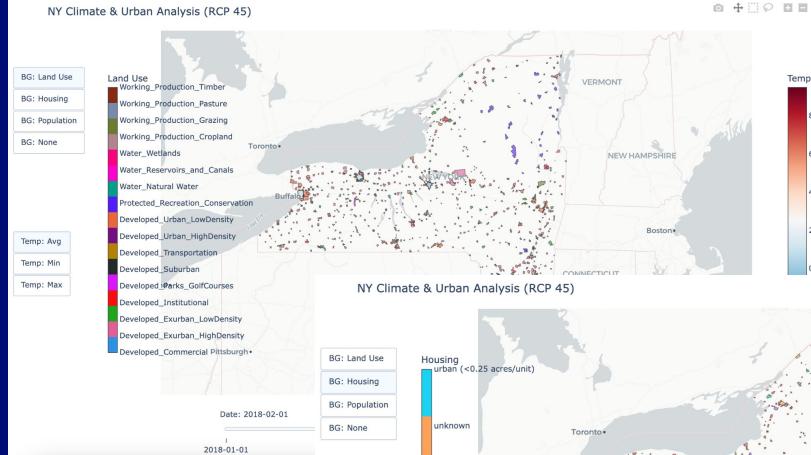
Interactive Rendering Architecture

- **Base Geometry Ingestion:** Establishes the foundational spatial coordinate system using raw GIS shapefiles (Boundaries).
- **Environmental Contextualization:** Maps static urban attributes (Land Use, Building Density, and Population) to provide local context.
- **Climate Signal Injection:** Overlays dynamic temperature data, applying spatial exclusion (masking) to prevent data overlaps and clutter.
- **Holistic Synthesis:** Merges all layers into a single interactive dashboard, directly correlating urban structure with heat exposure.



Building Visualization Panel

Demonstration



Date: 2018-02-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

2018-01-01

2018-02-01

Population Analysis & Insights

Observation: A strong positive correlation exists between population density and local temperatures.

General Pattern: Densely populated urban areas consistently exhibit higher average temperatures than their surrounding rural regions.

Illustrative Example: New York State

New York City (high population density) shows significantly higher temperatures than less populous cities like Buffalo or Rochester.

Conclusion: Urban development and concentrated human activity are key factors in elevating local temperatures.



Correlation between population density and temperature:

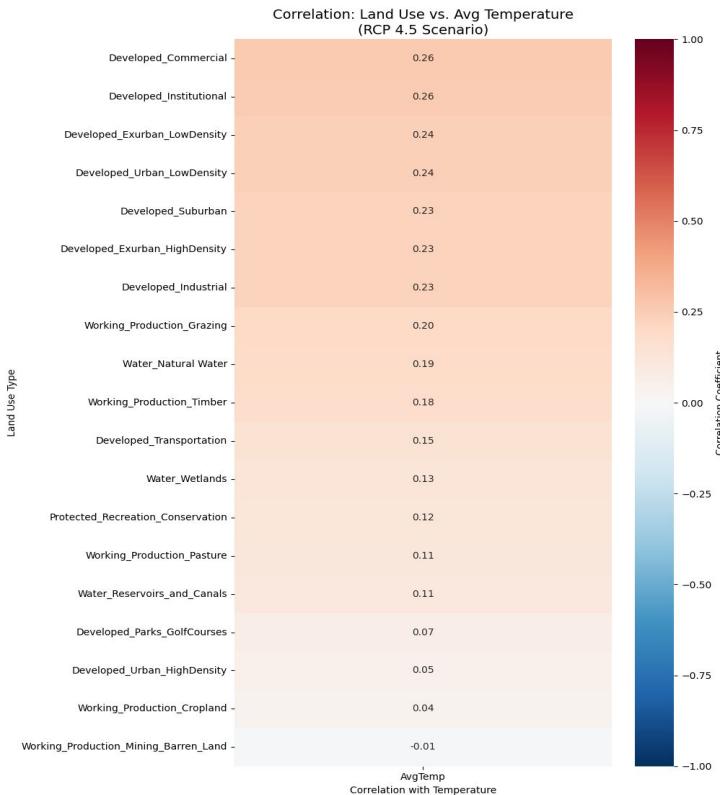
$$\rho = +0.38$$

Regression results:

$$\text{Average Temperature} = \beta_0 + \beta_1 \cdot \text{Population Density} + \varepsilon$$

- $\beta_0 \approx 18.5$ (Estimated baseline temperature)
- $\beta_1 \approx +1.45$ (For a unit increase in population density, the average temperature increases by ~ 1.45 units)
- p-value < 0.001 (The relationship is statistically significant)
- $R^2 \approx 0.144$ (Population density explains about 14.4% of the variation in temperature)

Land Use Projection and Temperature Analysis



Regression Equation:

Avg Temp = 15.87 + (0.00039) X Dev_Commercial + (0.00010) X Dev_Institution + (0.00010) X Dev_UrbanLowDensity + (0.00129) X Work_Pasture + (0.00034) X Work_Timber + (0.00021) X Water_Wetlands + (0.00011) X Prot_Recreation + (0.00010) X Work_Grazing + (0.00006) X Dev_Suburban + (0.00004) X Dev_Urban_HighDensity + (0.00002) X Dev_Exurban_HighDensity - (0.00007) X Water_Natural - (0.00008) X Water_Reservoirs - (0.00008) X Dev_Exurban_LowDensity - (0.00009) X Work_Cropland - (0.00045) X Dev_Industrial - (0.00046) X Dev_Transportation - (0.00061) X Dev_Parks_Golf_Course - (0.00210) X Work_Mining

RESULT

Heating Drivers:

- Commercial & Institutional Zones:** The strongest predictors of heat.
- Low-Density Sprawl:** Adds significant heat load due to road infrastructure.

Cooling Levers:

- Parks & Green Space:** The *only* developed land use that actively lowers temperature.

Heat Map : Land Use Type vs Temperature

Housing Density and Temperature Analysis

Correlation between housing density and temperature: $\rho = -0.15$

Indicate a weak negative relationship: Cities with a higher share of high-density residential land tend to have slightly lower average temperatures, but the magnitude is small.

Regression Results:

- Average Temperature = $\beta_0 + \beta_1 \cdot \text{High Density Share} + \varepsilon$
- $\beta_0 \approx 19.8$, $\beta_1 \approx -2.65$, p-value < 0.001, $R^2 \approx 0.021$

Conclusion:

Many dense cities are located in cooler northern regions, while hot Sunbelt cities tend to be low-density and sprawling.

Potential Improvements in Future

1. Optimizing Visualization Runtime on Large Data
2. Integrate Machine Learning/Smart Analytics into Visualization Panel
3. Uncertainty Quantification: Include Confidence Interval
4. Simulation Mechanism Integration into Visualization
5. Bivariate Mapping (Create a Bivariate Choropleth which combines 2 variables into a single color scale)

THANK YOU!