

# Visualizing the Impact of Climate Change on Housing Prices in California

Kenneth Lui, Michelle Tan, Preston Slane, Christopher Miro, Radhika Cherukat, Nathan Kistler  
CSE 6242 Class Project



Georgia Tech  
College of Computing

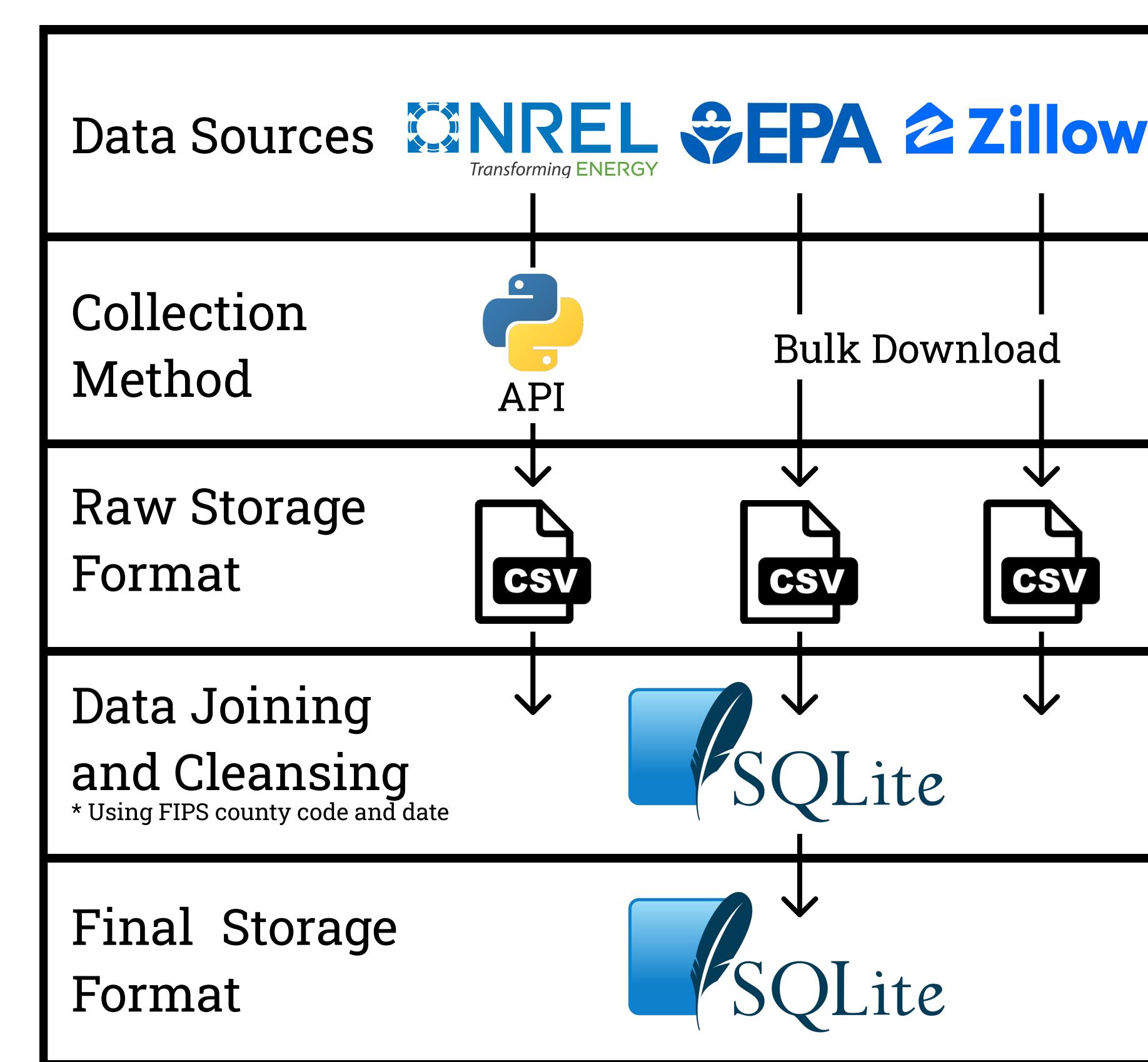
## MOTIVATION

Climate change is causing irreparable damage to our living conditions. Wildfires, floods, hurricanes, and other natural disasters are threatening the infrastructure of the United States' highly populated areas. The frequency and intensity of these climate-related events are worsening causing a trickle down effect on the real estate market through the destruction of peoples' homes and places of business. Real estate investments once thought "safe" are now at risk due to the yearly wildfire season in California or the inevitable summer hurricane in Louisiana.

In this project we will link historical weather patterns with trends in the real estate market in California while utilizing time-series computation methods to predict future trends on a county by county basis.



## DATA GATHERING



Data was collected from open-source providers using APIs or bulk download. Python was used to extract data from APIs. The data was compiled into flat CSV files using the data and county FIPS code as the primary keys.

After data collection the CSV files were loaded into a SQLite database where the data was cleaned and joined to create the final dataset. During the cleansing process 8 counties did not contain sufficient historical data to interpolate historical data and develop times series models (Inyo, Lassen, Tuolumne, Tehama, Yuba, Alpine, Modoc, and Sierra). These counties were excluded from the current project.

The main dataset was stored in a SQLite database to ease distribution.

## APPROACH AND EXPERIMENTS

### Historical Data (2010-2021)



Average Housing Prices



Air Quality Index (AQI)



Transforming ENERGY

Temperature/Precipitation

**ARIMA Model**

Develop optimized model for each attribute and county

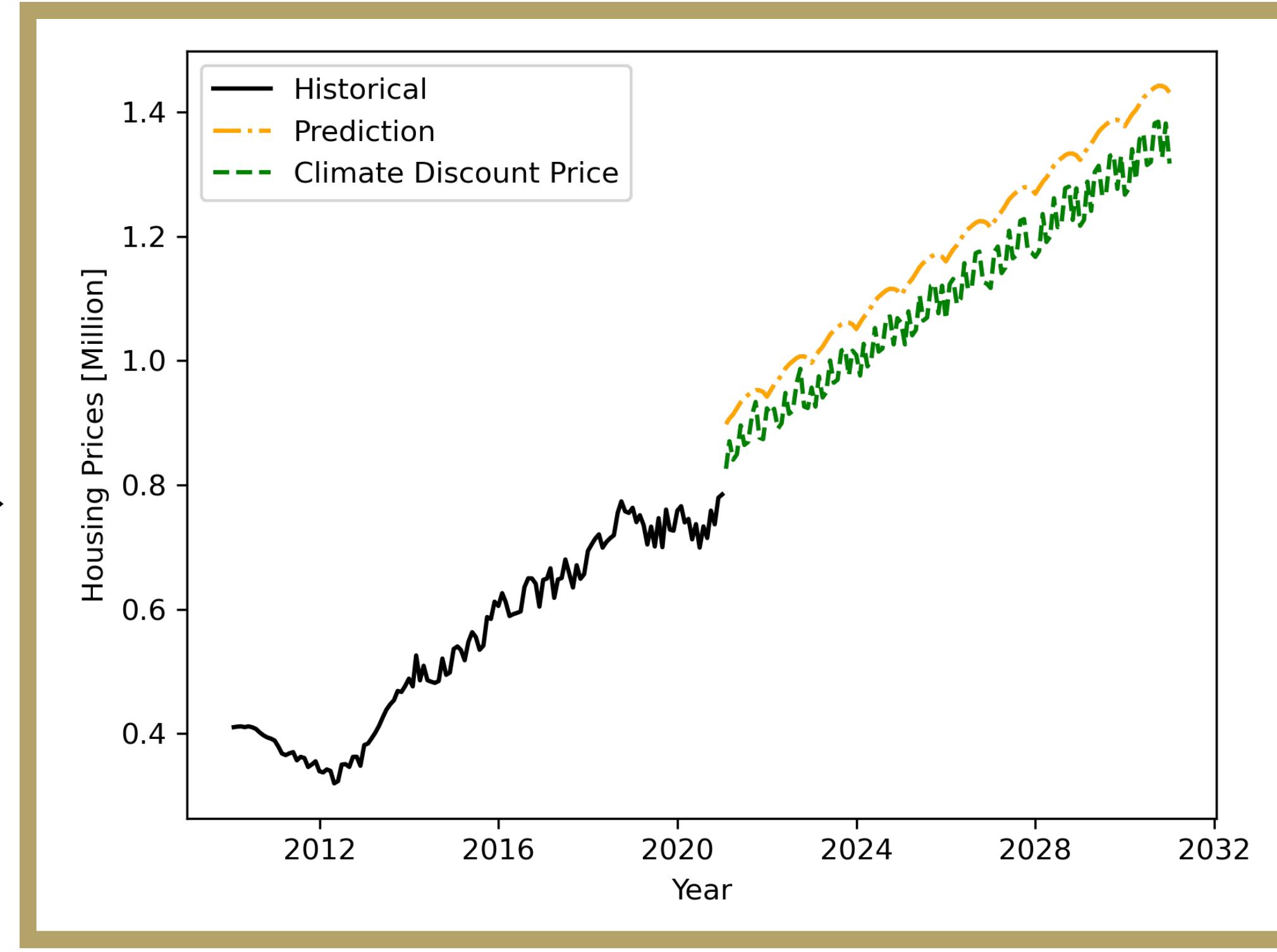
### Future Predictions (2021-2031)

Average Housing Prices

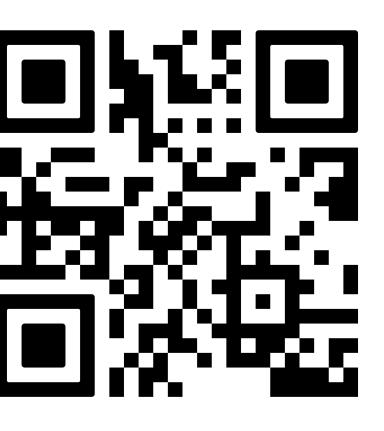
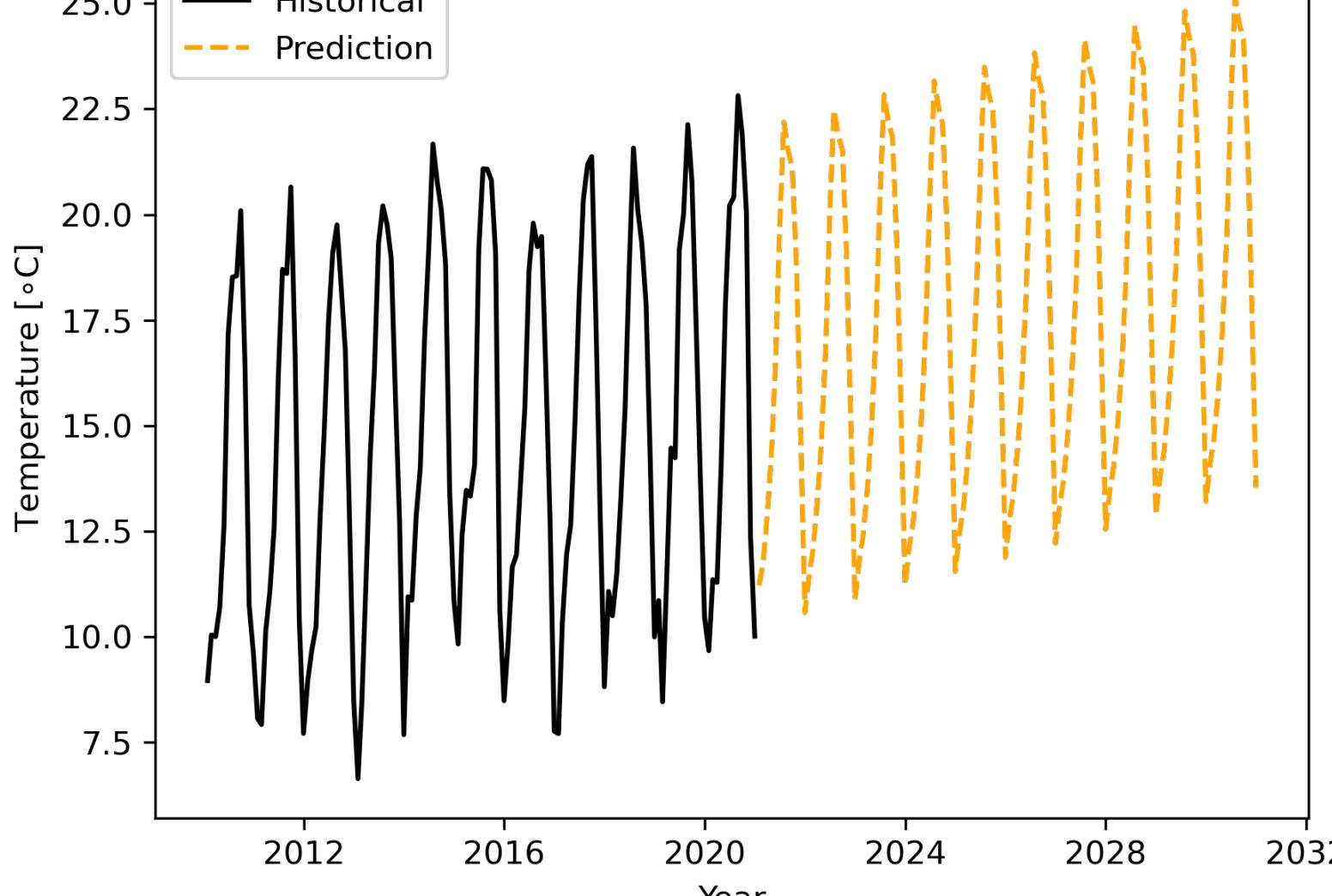
Air Quality Index (AQI)

Temperature/Precipitation

**Scoring Matrix**



### Temperature ARIMA Model



ARIMA models were generated using a stepwise approach (AIC) for each attribute and county to predict future records. An example of the historical and predicted values for Alameda County (Temperature and AQI) are shown above. The data presented is representative of the bulk data where temperature shows strong seasonality and trend while AQI shows less correlation with time. The full results of the ARIMA predictions can be visualized through Tableau using the QR code.

### Risk Weightage

AQI - 40%  
Temp - 30%  
Precipitation - 30%

### Overall Scoring (AQI + Temp + Precipitation)

Score (out of 100%)	Category	% Discount to Housing Prices
>99 and <=100	Great	0%
>97 and <=99	Good	2%
>95 and <=97	Fine	4%
>90 and <=95	Moderate	6%
>85 and <=90	Bad	8%
>80 and <=85	Very Bad	10%
<=80	Extremely bad	15%

### AQI Scoring (40% of total score)

AQI Level	Monthly* Score
AQI <= 50	1
AQI > 50 and <=100	0.98
<=150	0.95
AQI > 150	0.85

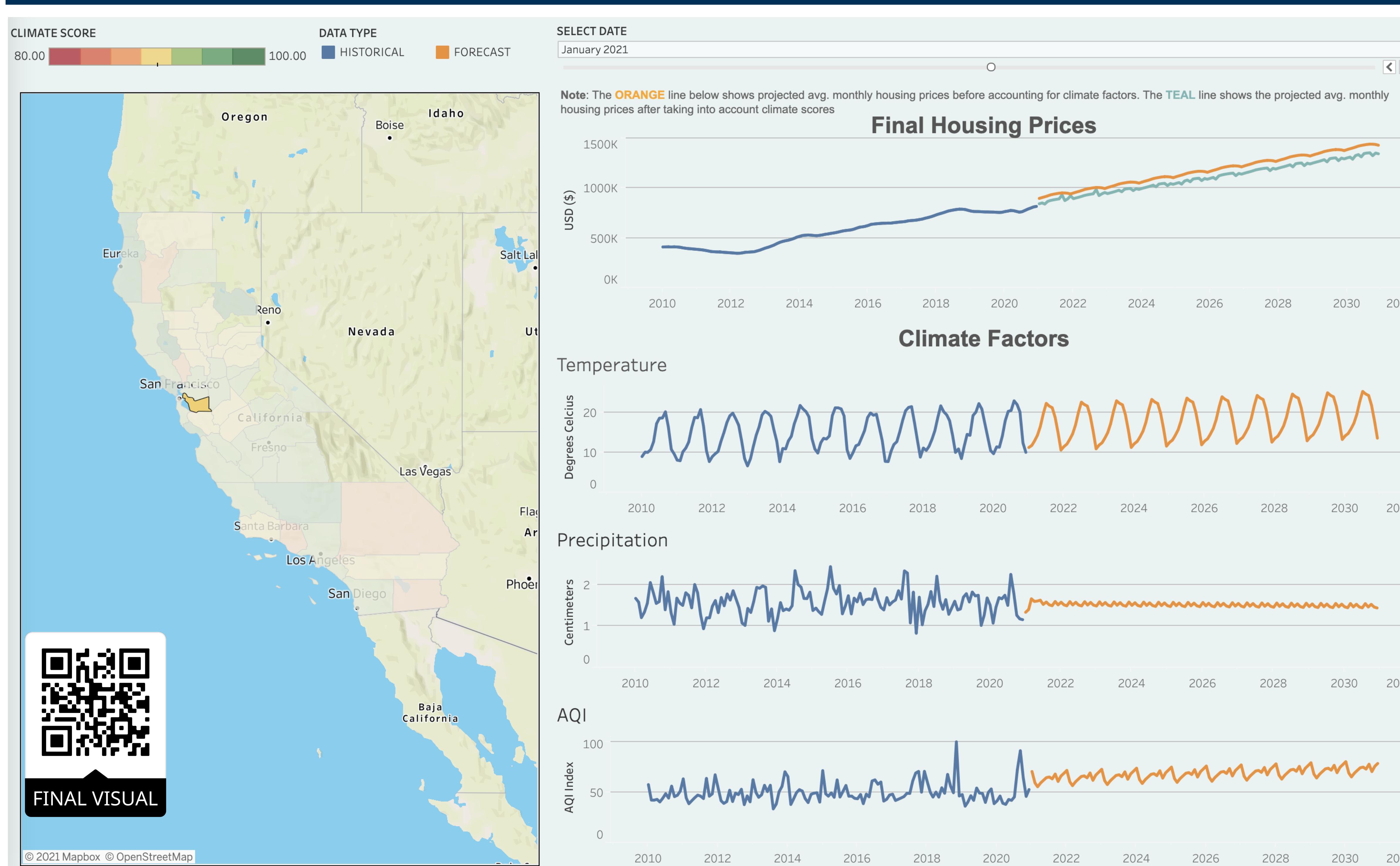
### Temperature Scoring (30% of total score)

Temp Increase (Compared to 2010 baseline)	Monthly Score
<1 C	1
>=1 C and < 1.5 C	0.95
>= 1.5 C and < 2 C	0.85
>= 2 C	0.75

% Change (Compared to 2010 baseline)	Monthly* Score
<10%	1
>= 10% and < 25%	0.95
>= 25% and < 50%	0.85
>= 50%	0.75

\* For Temperature and Precipitation, the daily metrics are averaged to achieve a monthly average. The average monthly metrics are then compared to the same month in 2010 and then scored based on the above criteria.

## RESULTS



The final visualization is shown in the image on the left. The end product can be seen in its entirety by using the QR code.

The visualization uses a map of California segmented by county to depict the climate score observed or predicted for the time selected by the user. The climate score is calculated using the AQI (air quality index), temperature, and precipitation as shown in the scoring matrix. Detailed data of the historical and forecasted AQI, temperature, precipitation and housing prices are shown in the plots. The final housing price plot includes both the forecasted housing price and the climate score discounted price.

Initial analysis using the visualization shows a general warming trend amongst most counties in California. This leads to a poorer climate score causing higher discounts on housing prices.

## FUTURE

The goal of the project was met but additional ideas were generated that could better the product in the future:

- Inclusion of additional climate variables (e.g., wildfires, earthquakes, and drought) to improve the scoring algorithm
- Refine the visualization to allow users to drill down to the zip code and neighborhood granularity
- Extend the visualization and prediction model to more states and countries
- Implement a more rigorous machine learning model to determine the discounted housing rate with better precision
- Improve ARIMA models that display inaccurate predictions