Modeling & Forecasting the Urban Real Estate Market in the US

Eliza Sparrow, Robert Anderco, Timothy Kelley, Ahmad Raiyan, Charlie Geier, Zizhou Lu, Taylor Shinal

Research Question

- **How do micro and macro economic factors influence housing prices in the real estate market, and can prices be forecasted by these factors?**
  - Household income
  - Unemployment
  - GDP
  - Population
  - Crime rate
  - Tax
  - CPI
  - Interest rate

Motivation

- The real estate market is a vital foundation to **building wealth for the middle class**
- 2008-2009 housing crisis created great instability that spilled over into areas beyond the housing market
  - Negative wealth effect
  - Decreased demand/consumption of other goods and greatly economic growth
- COVID-19 pandemic has “squeezed” the middle class
  - Through the exponential increase in home prices coupled with the shortage of affordable housing

Motivation

- Forecasting the real estate market allows for:
  - **Policy makers to determine** if there needs to be changes with creating affordable housing or decrease/increase demand through tax credit, changing interest rates, etc
  - **Residents to plan** or time out buying or selling homes given the specific factors in their area

Motivation

- A potential difficulty of this project is confronting data collection challenges
  - We will need to determine how vital the missing city's data is and how worried we should be about missing data

In 2021, 16.7% of GDP was accounted from spending within the housing market (rent, utilities, new home constructions, remodeling, brokers' fees)

General roadmap

- Dive into the cities
- Explore the individual factor characteristics
- Visualize the data collected
- Challenges?
- ARDL summary statistics
- Examples within the data: Case Studies

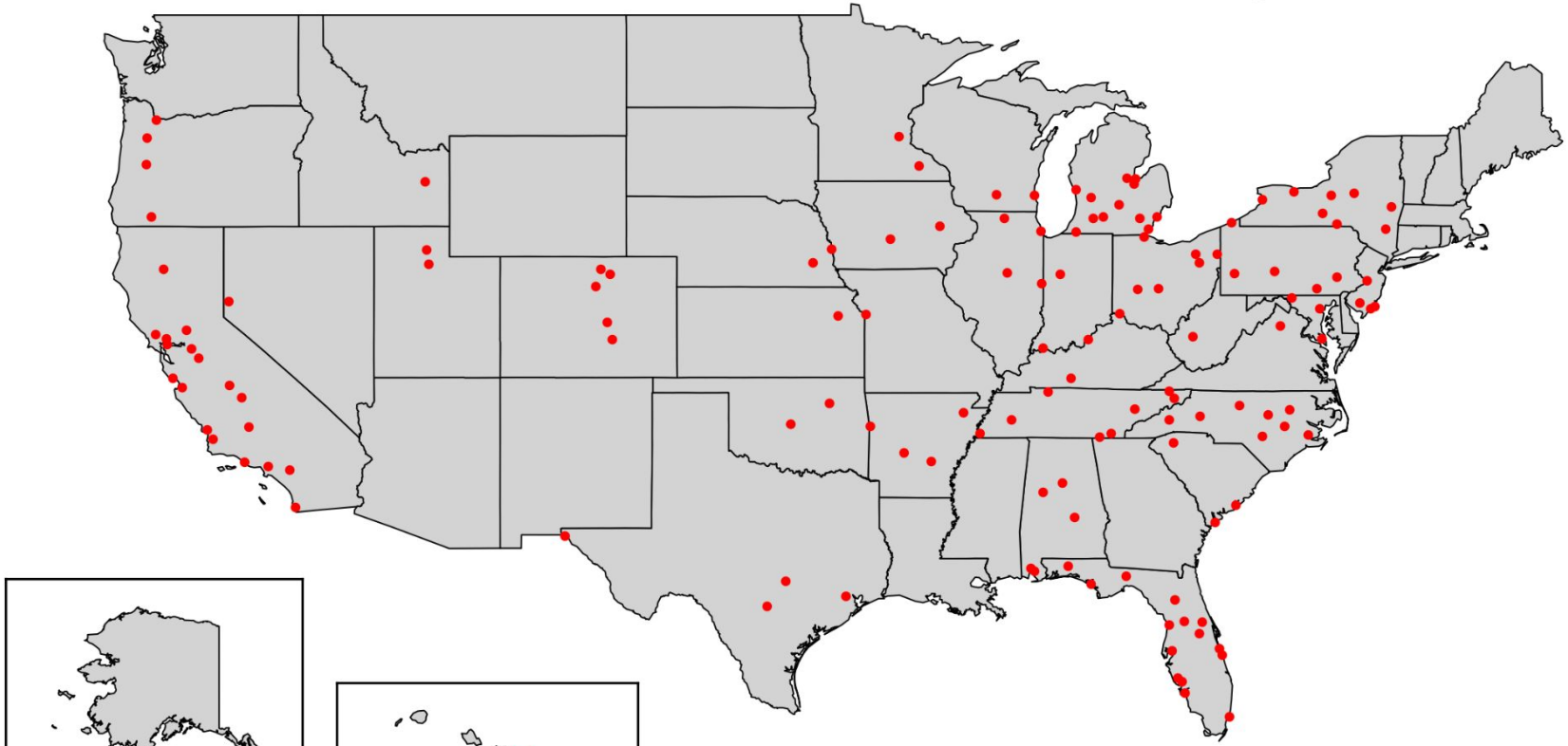
Why these cities

- Capture different population trends in the data. (Increasing, decreasing)
- Gives more options to look into. (Size, geography)
- By combining them, the data is more representative of the country as a whole.

Chosen Cities

- Zillow provides data on 896 U.S. cities
- We dropped cities with missing housing value data
- We then dropped cities for which we were unable to collect all variable data on
- This leaves us with 141 cities to analyze
    - Dropped cities: Varied in characteristics, some incredibly large in various ways (NYC) and some not

Selected Cities for Real Estate Analysis



Data Descriptions

Response Variable (by city; source: Zillow):

- Median Home Values ($) (Monthly, 2000-2023)

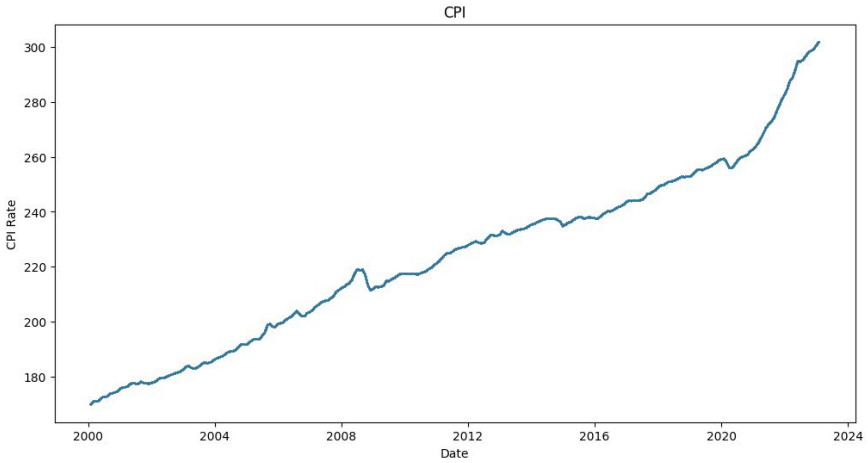
Explanatory Variables (source: FRED) :

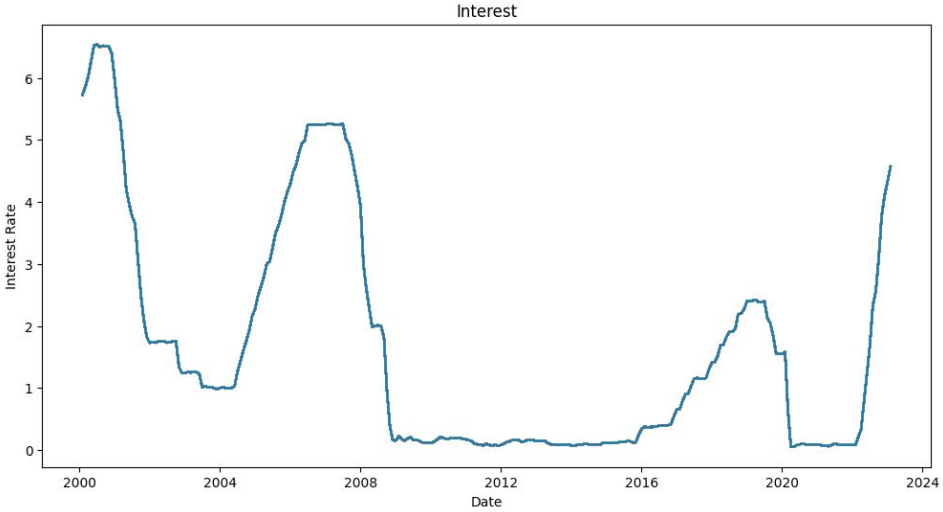
- By Metropolitan Statistical Area
  - Per Capita Personal Income ($) (Annual, 1969-2021)
  - Unemployment Rate (%) (Monthly, 1990-2023)
  - Total Gross Domestic Product ($Millions) (Annual, 2001-2021)
  - Resident Population (Thousands) (Annual, 2000-2021)

Data Descriptions Cont.

- By County
  - Combined Violent and Property Crime (Known Offenses) (Annual, 2004-2021)
- By State
  - Property Taxes ($Millions) (Quarterly, 1994-2021)
- Nationwide
  - Consumer Price Index (Index 1982-1984=100) (Monthly, 1947-2023)
  - Federal Funds Effective Rate (%) (Monthly, 1954-2023)

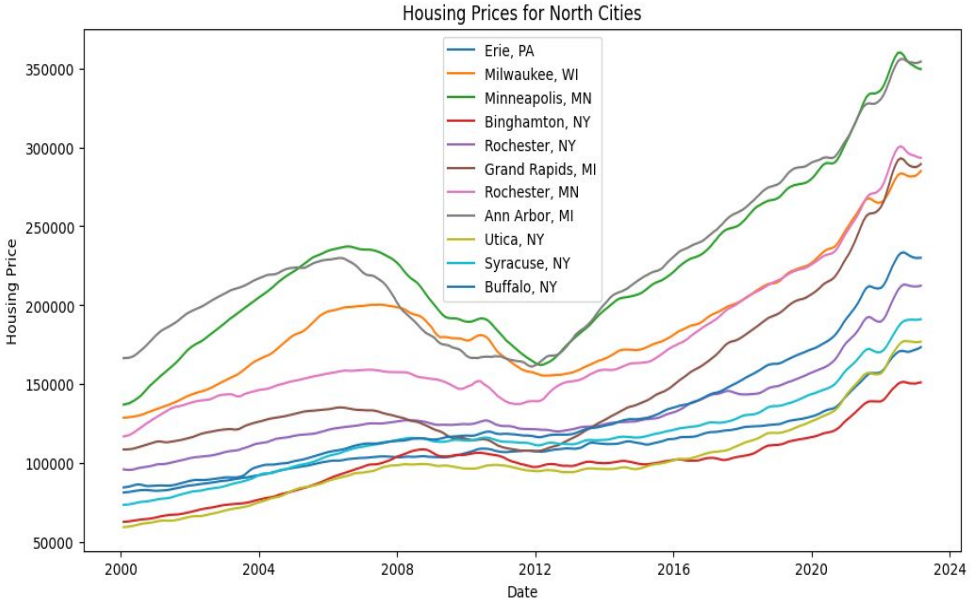
Data visualization

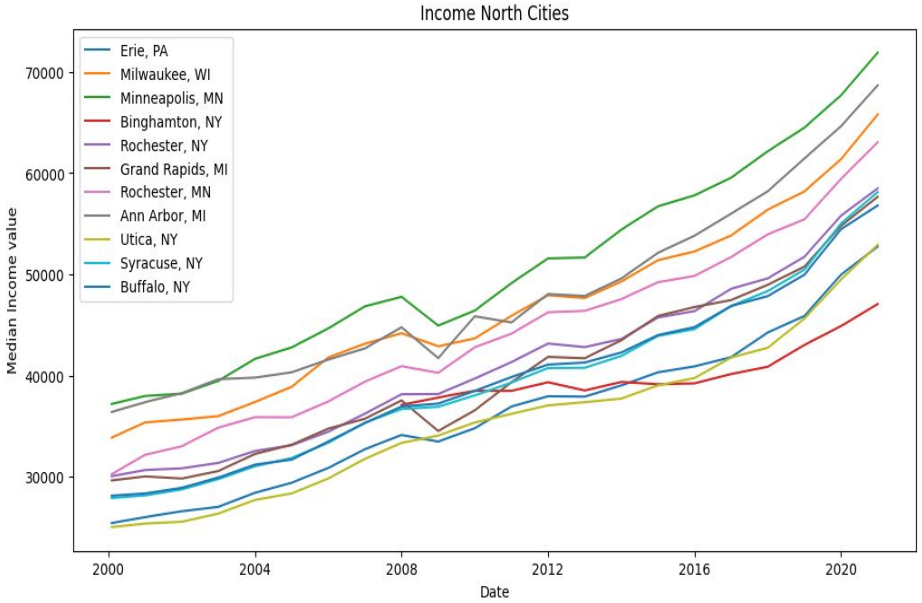




- General US market data
- Can be noticed the monetary policy and the crisis that were happening (post 2000 and 2008)
- How are those reflected in the Housing Price data?

Data visualization North Cities

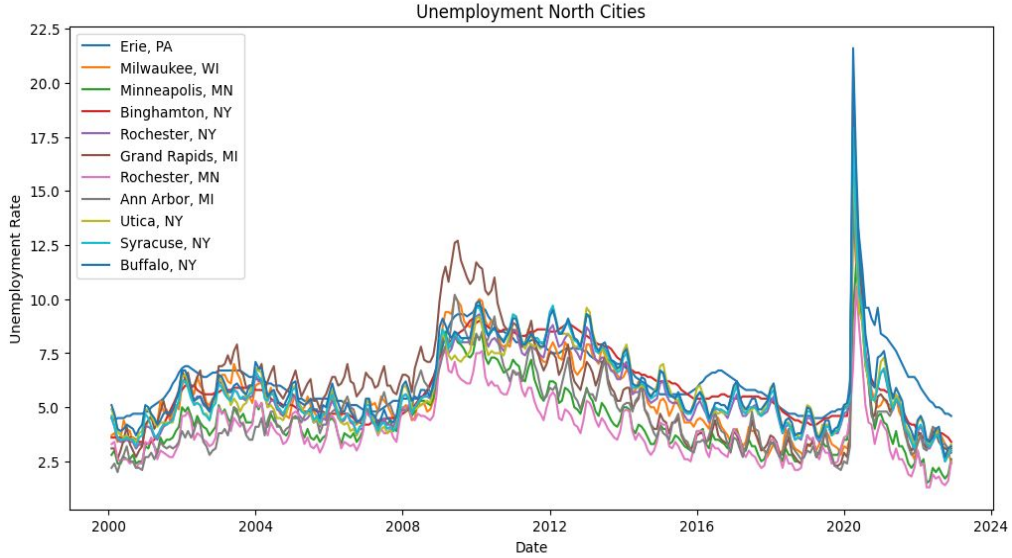


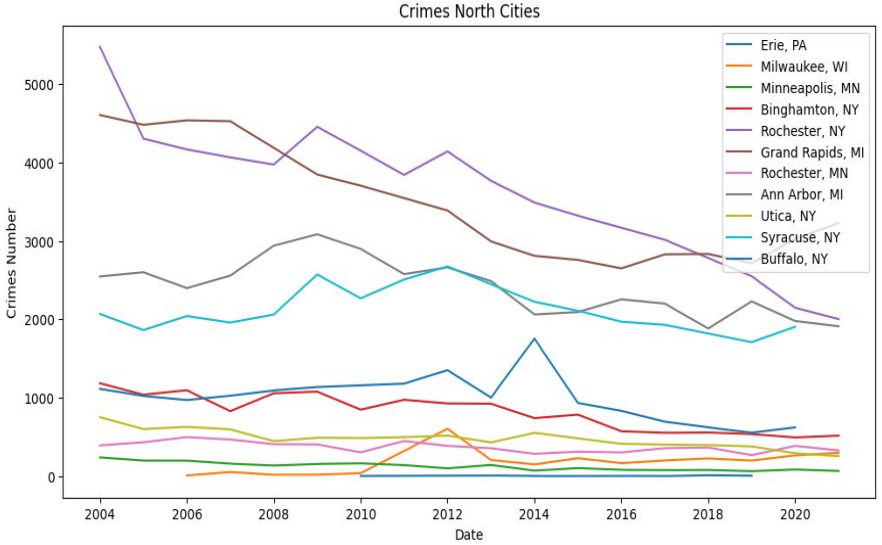


Can be noticed the following:

- Income follow the same rate of growth in almost all the cities
- Housing prices growth rate is not equivalent in all of them

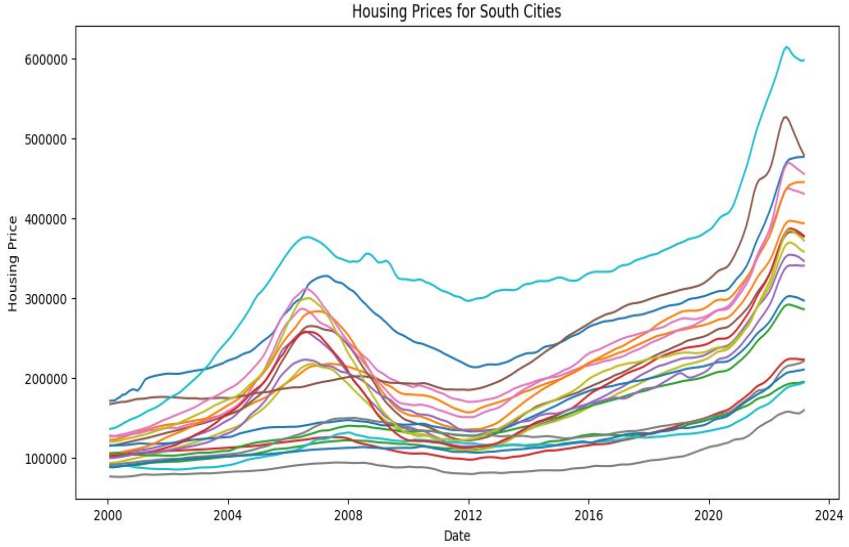
Data visualization North Cities

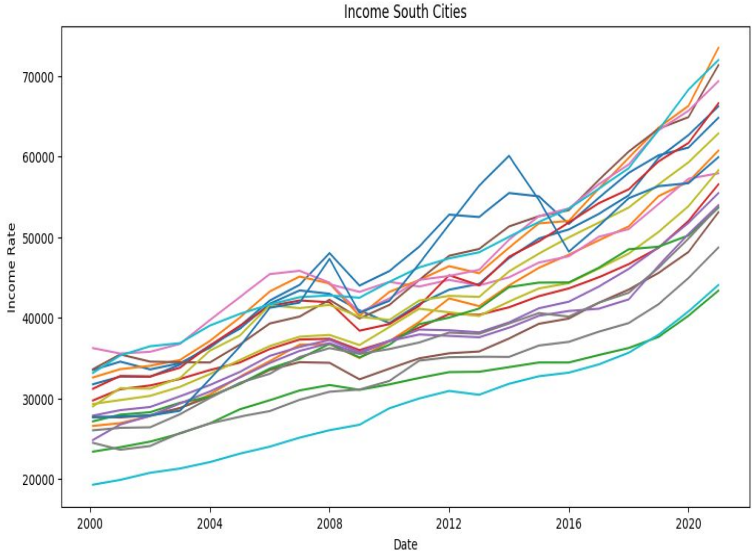




- The unemployment is somehow season that being a general rule
- Crimes tends to lower the number over time

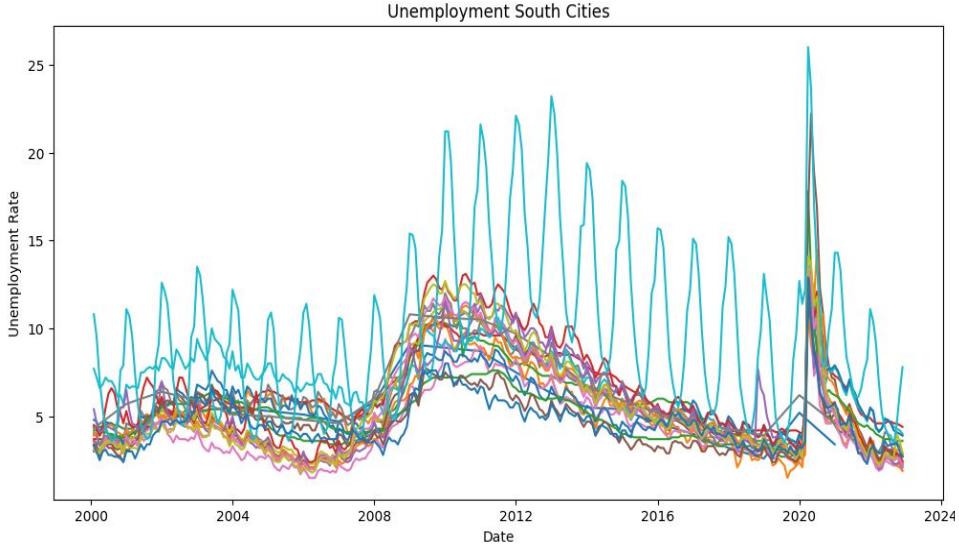
Data visualization South cities

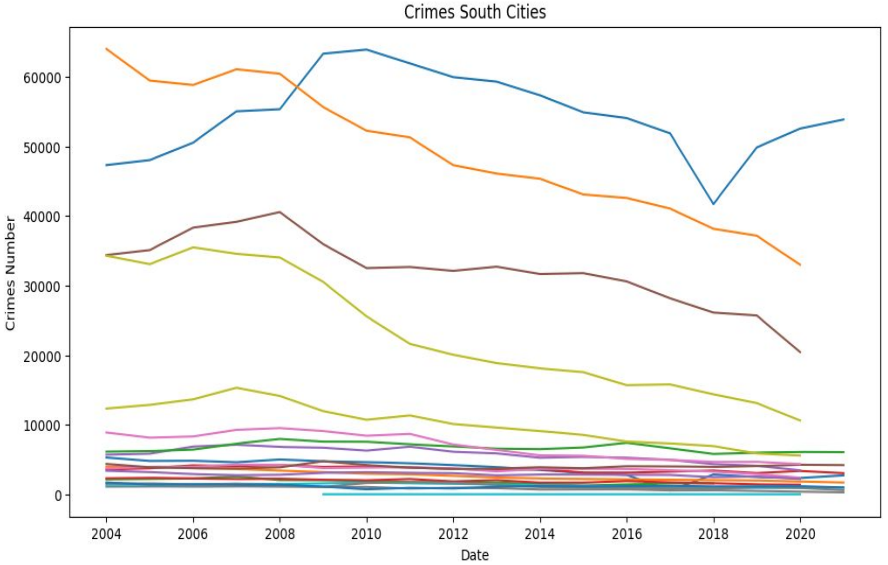




- Lot of cities in the south part all having constant income growth
- Is interesting that the prices are not having an similar grow being lots of irregularities

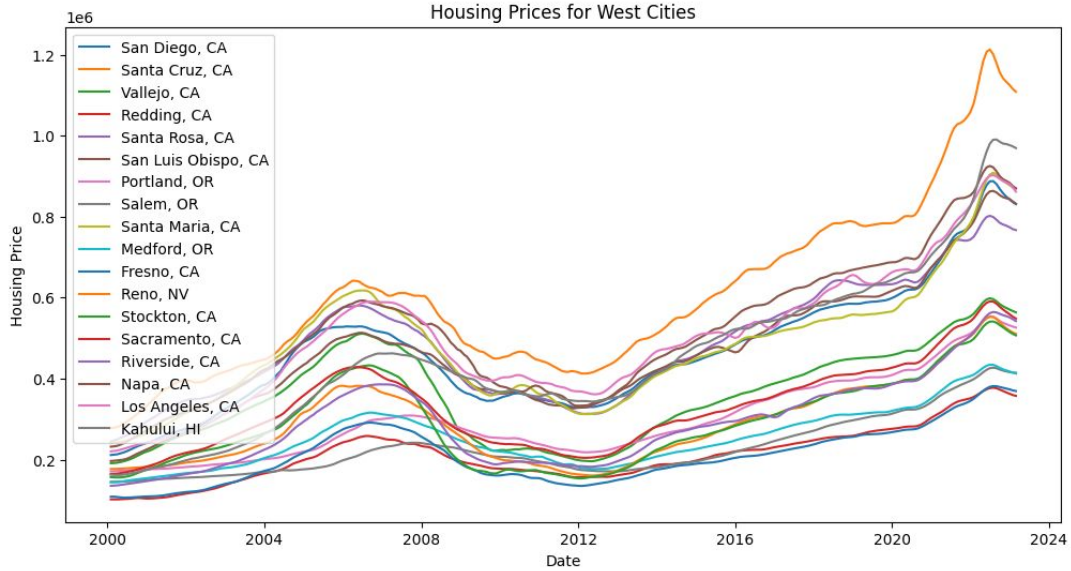
Data visualization South cities

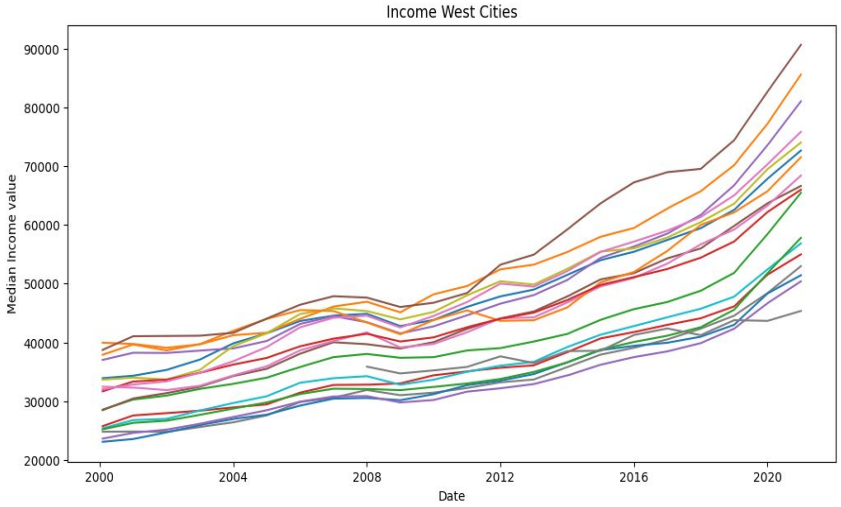




- Ocean City is very season compared to others
- Some crimes rate cities are way higher than the others

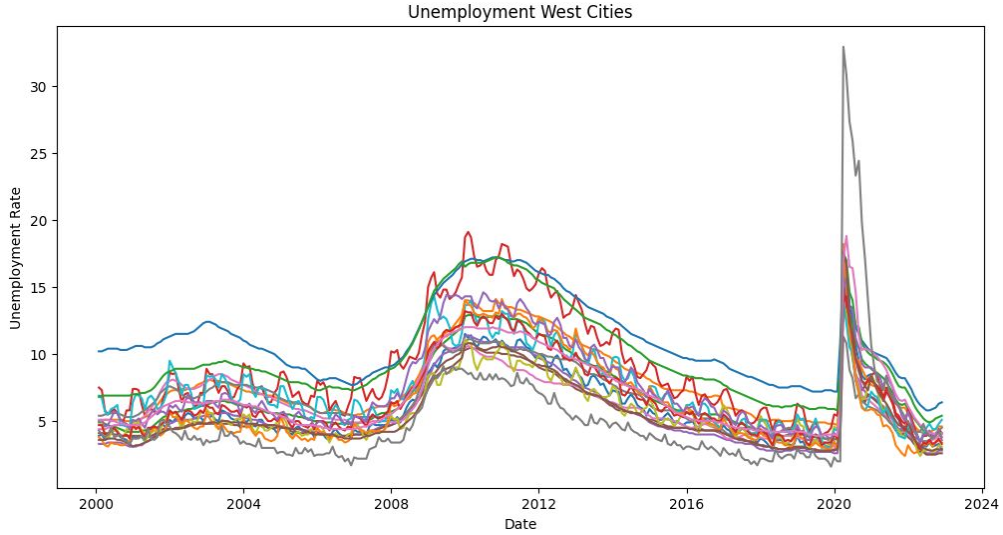
Data visualization West cities

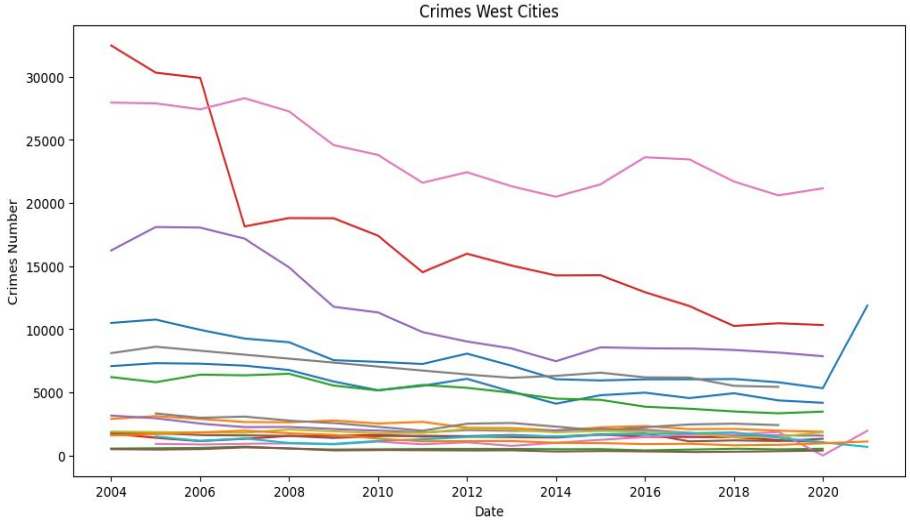




- After 2008 the growth in some cities have changed drastically in terms of housing price

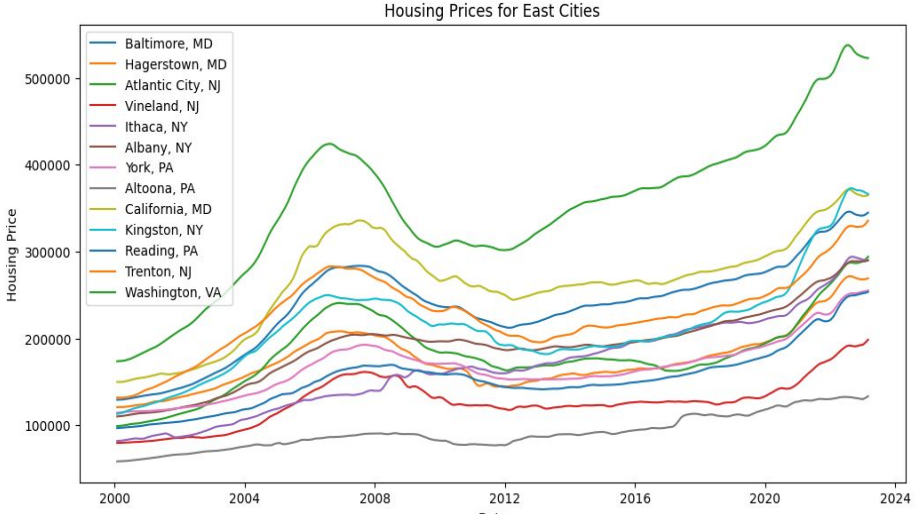
Data visualization West cities

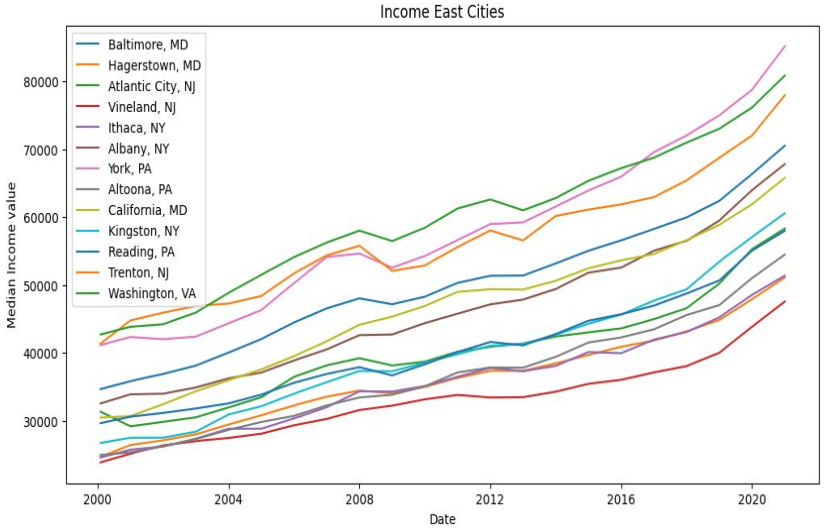


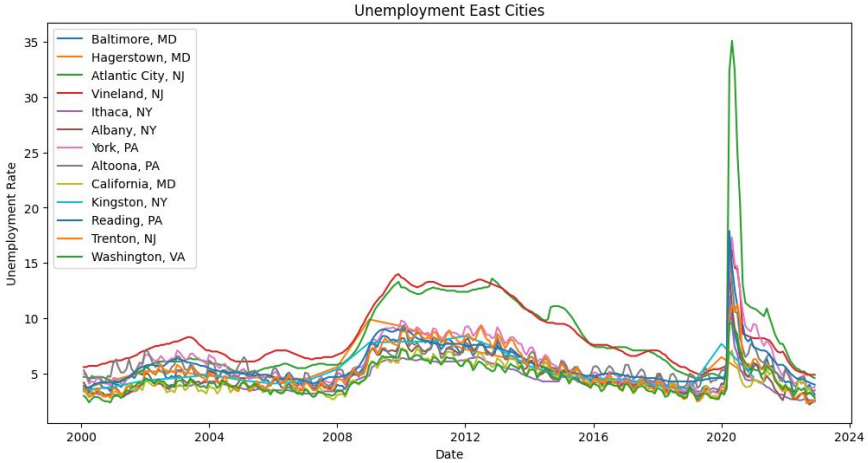


- Some are more seasonal than others in terms of unemployment
- Crimes trend is usually down/constant, but we see a retricement in last years which might be alarming

Data visualization East cities



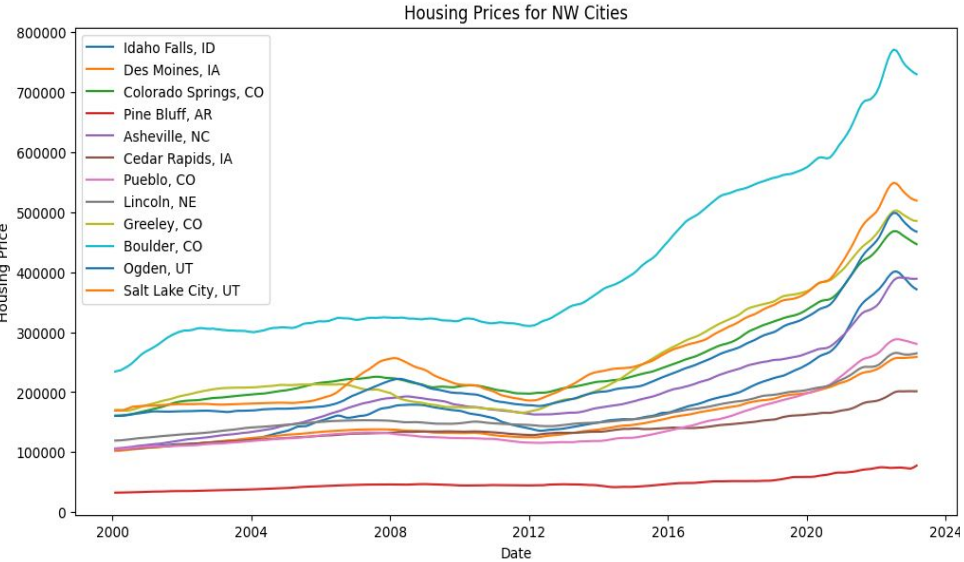


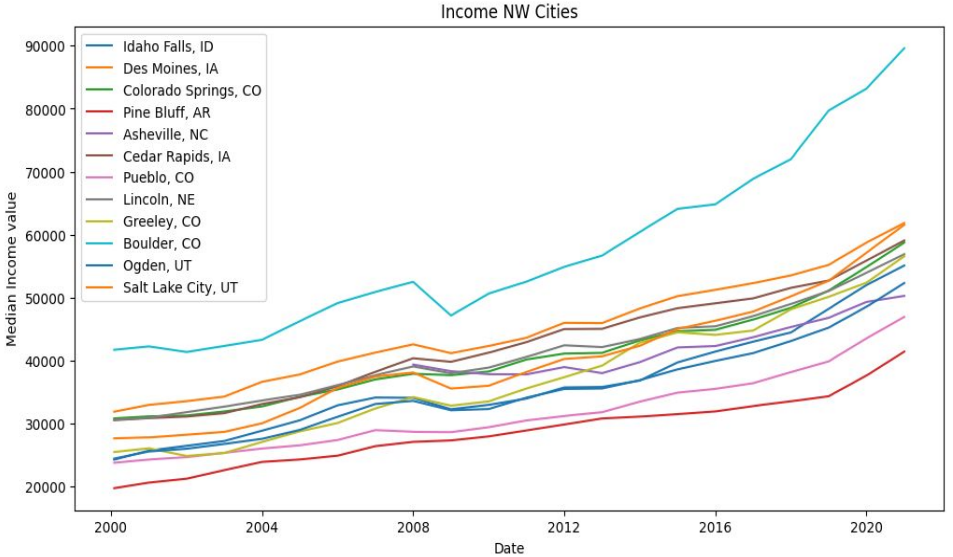


Significant missing thata, s those 3 graphs are the most relevant for that area.

- Housing prices in almost all ranges from less than 100k to over 500k

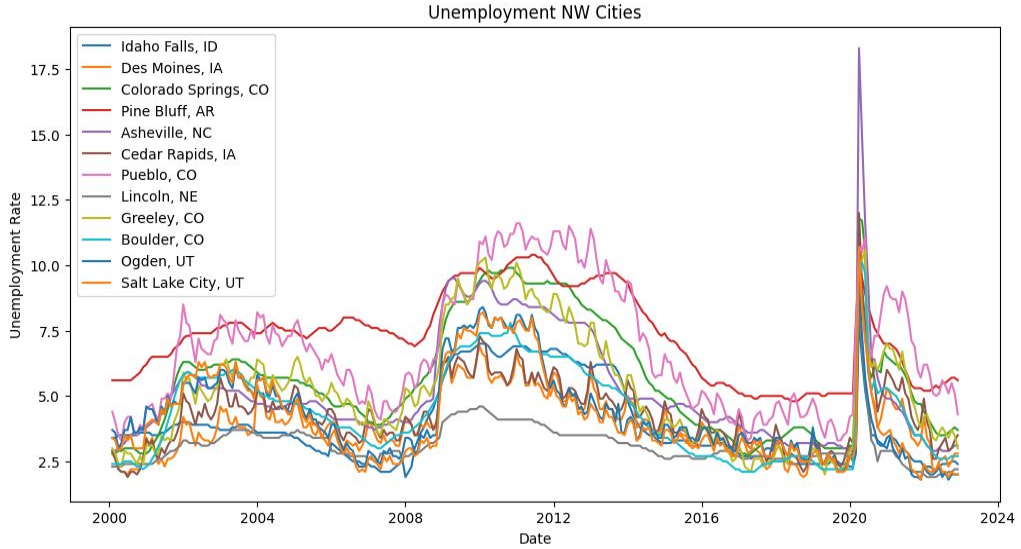
Data visualization NW cities

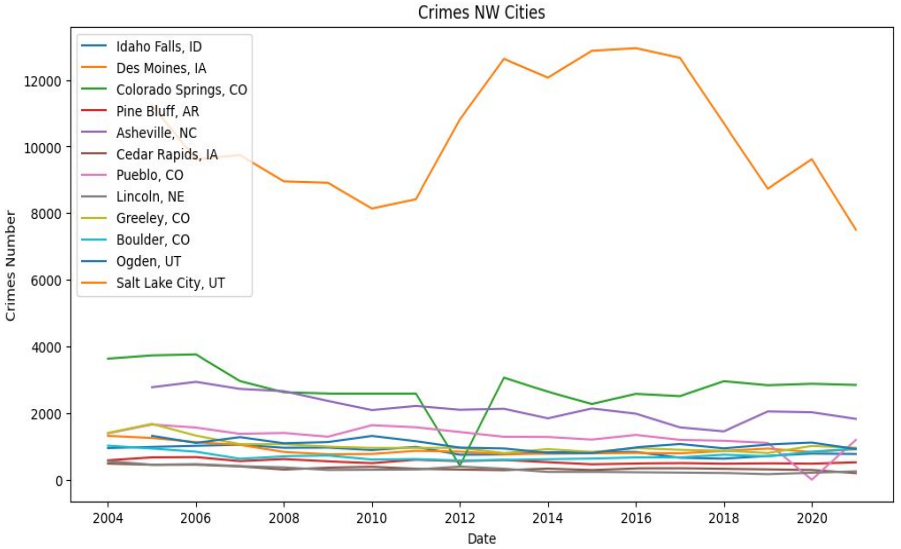




We have Boulder, CO a city that is doing amazing compared to the others from its zone

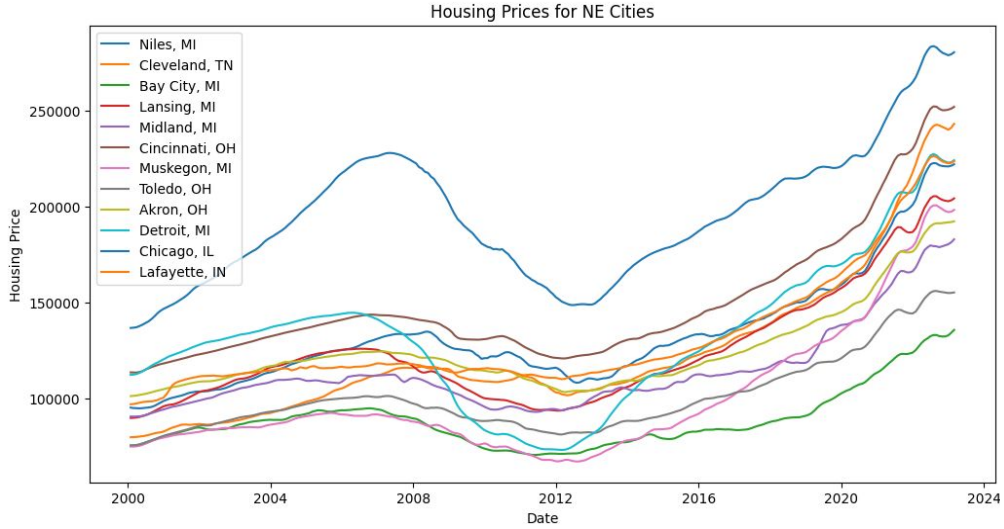
Data visualization NW cities

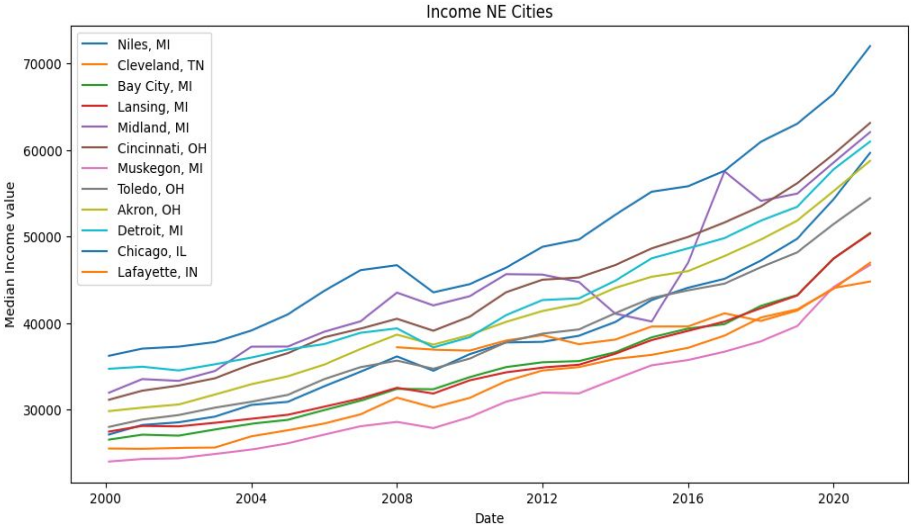




- Boulder CO stays consistent also in unemployment and crimes number
- Unemployment differs a lot from city to city

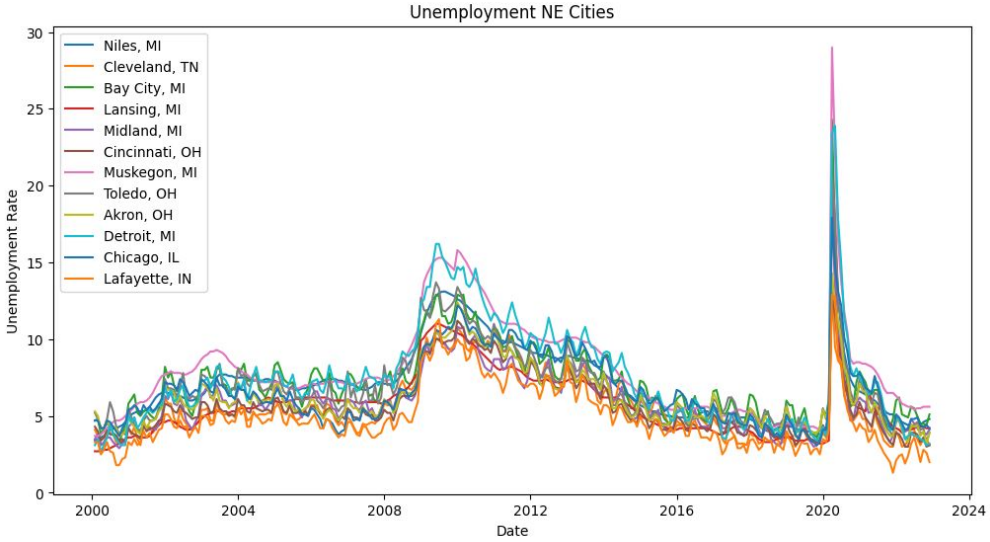
Data visualization NE cities

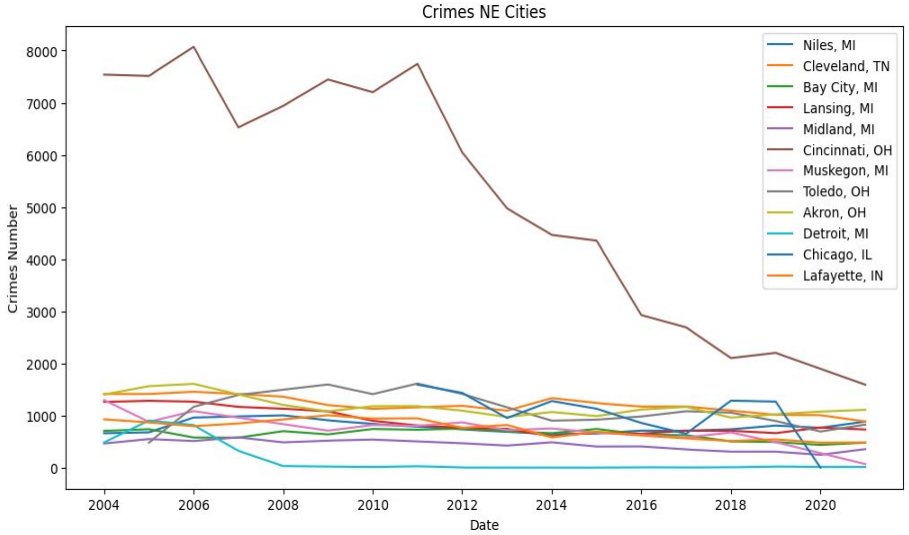




- In 2008 those cities were not affected that bad by the housing crisis

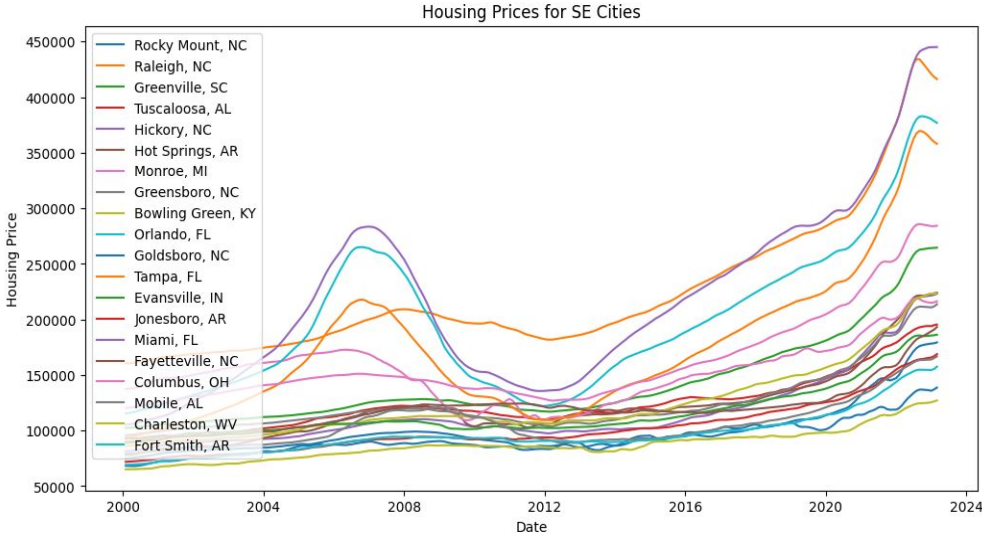
Data visualization NE cities

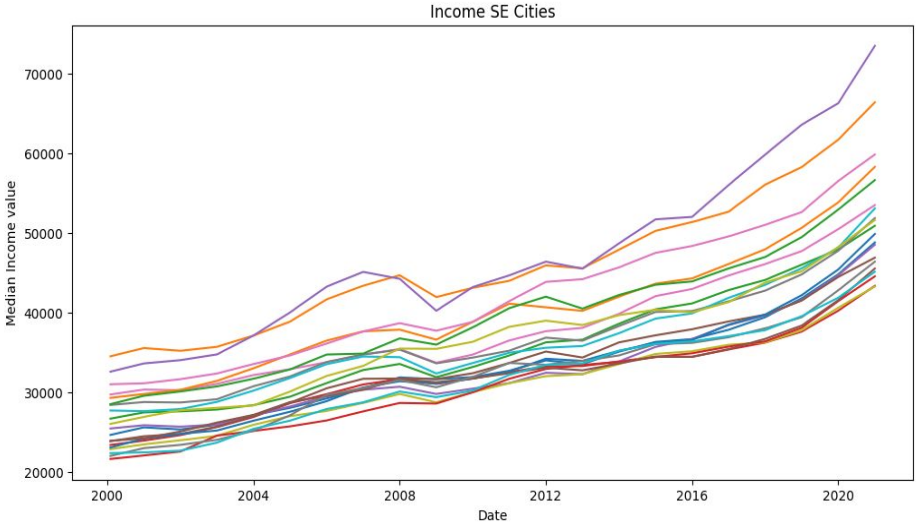




- Unemployment is very similar
- Crimes number is very very low which is interesting

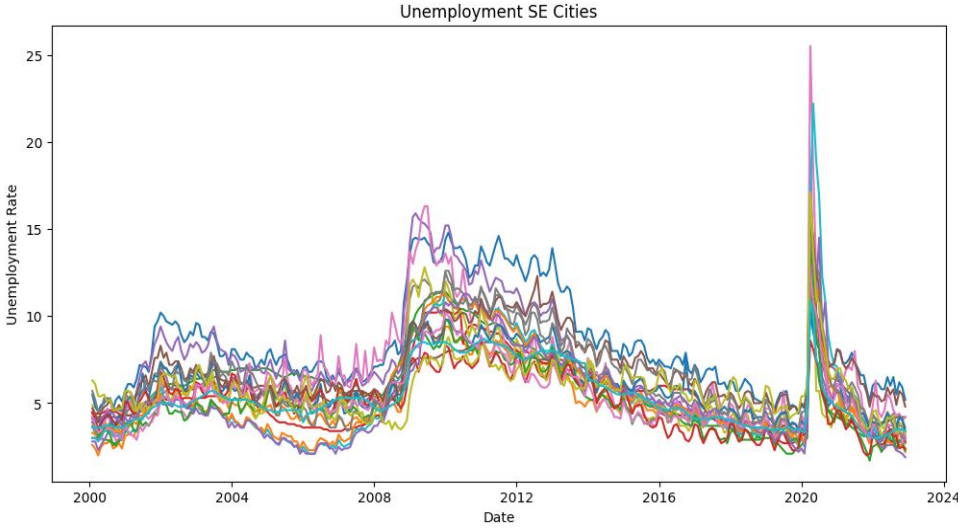
Data visualization SE cities

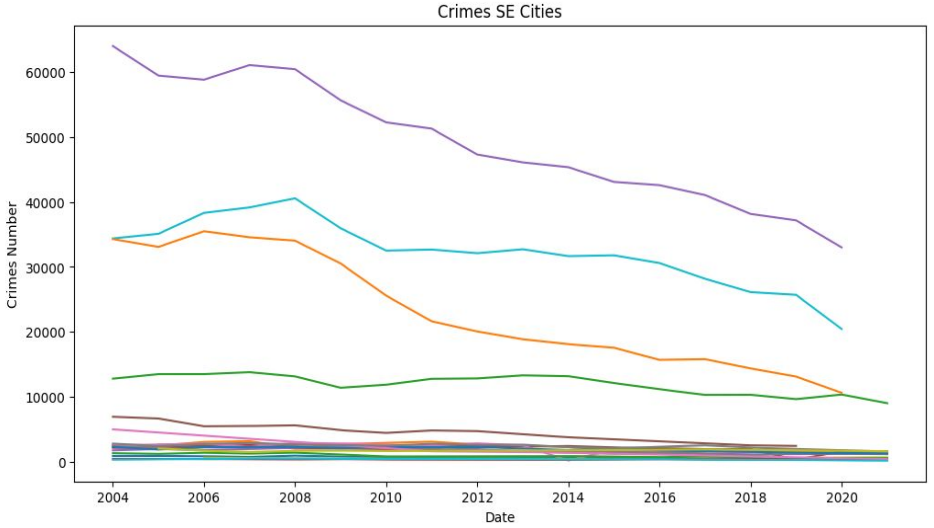




- Salaries amount and growth are similar
- Interesting what we see in house pricing, having only 5 cities that are out of the general line, the rest following slow growth trend (big cities big fluctuations: Miami, Tampa, Orlando, Raleigh)

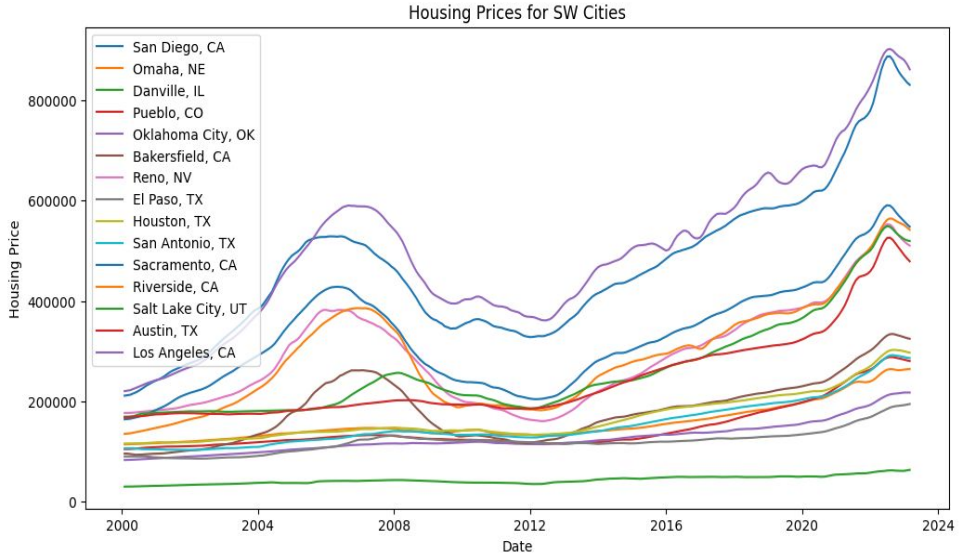
Data visualization SE cities

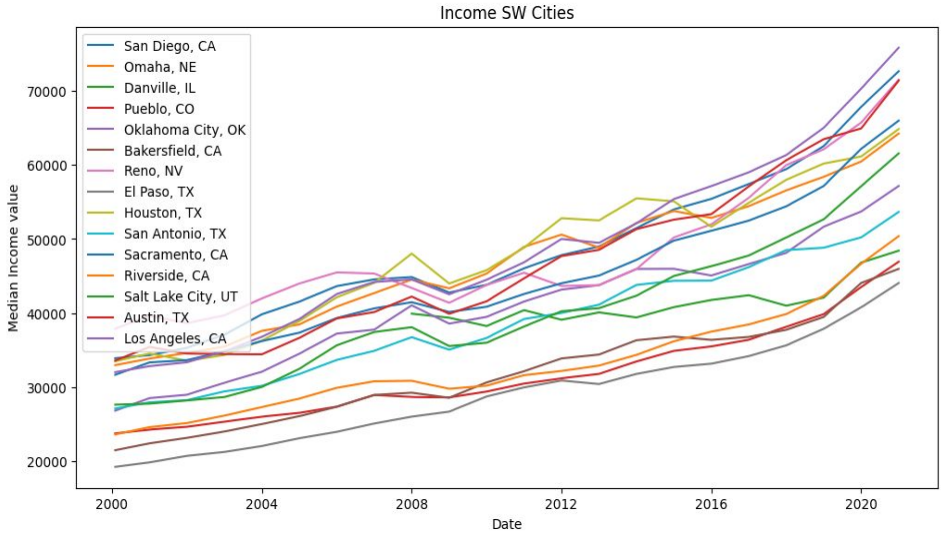




- High crimes number in bigger cities with higher housing prices but lowered over time
- Unemployment is very seasonal especially on ocean-side cities

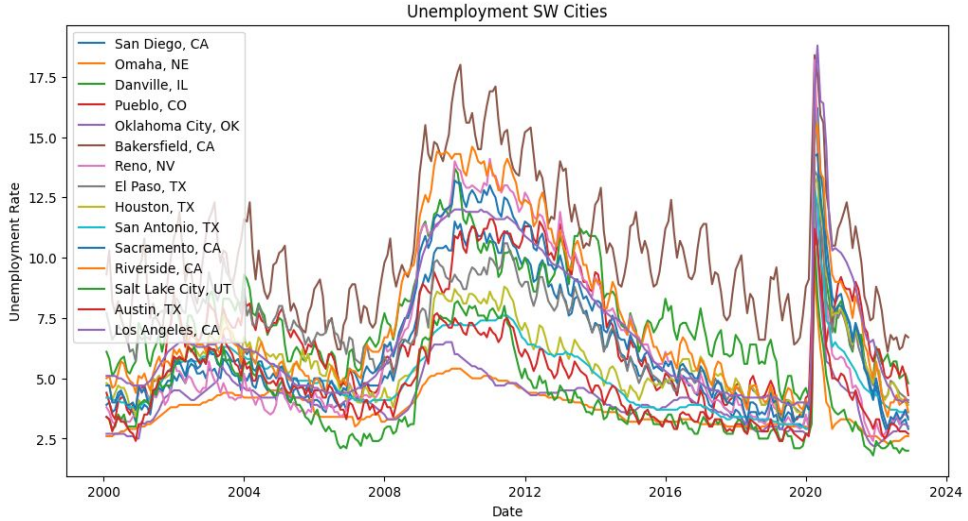
Data visualization SW cities

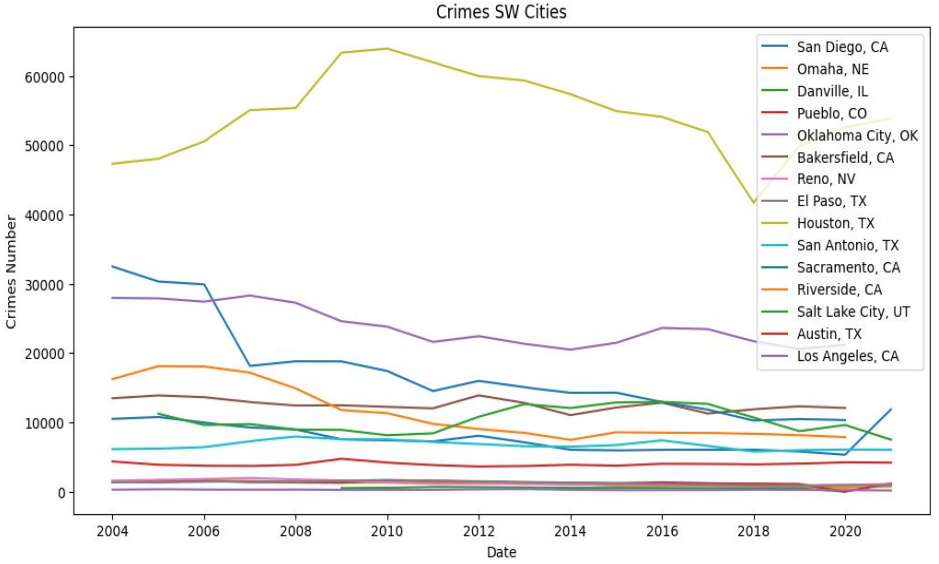




- Growth is consistent even if is some gap between the cities
- No irullegarities in the income

Data visualization SW cities





- Unemployment is the most variant maybe, but we also have to consider that in those cities are also lots on self employment persons winning a good amount

Data visualization conclusions and important takeaways

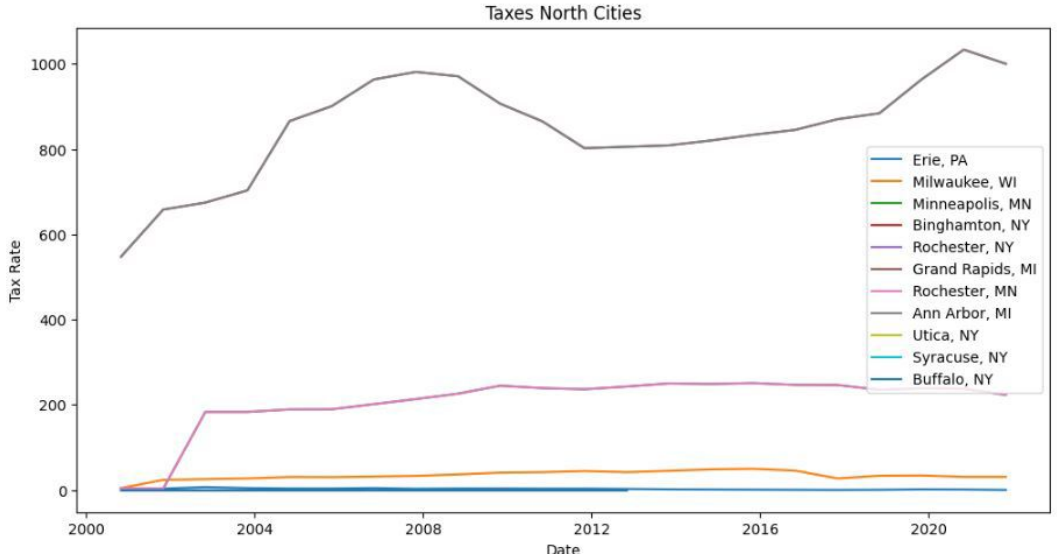
As a general trend in pricing we observed the following rule in all the regions:

- The lower price cities are having a constant growth and are not that affected by big crisis
- Lower risk (crimes) means higher price
- Seasonal cities for trips/holidays are not doing that good, people probably prefer to rent when they go there instead of invest in that place

Data Challenges

- Problems:
  - Different frequencies (monthly, quarterly, yearly)
  - Missing data
  - Inaccurate tax data
- Challenges of Missing Data
  - Can make the analysis and outcomes biased.
  - Analysis less accurate.
  - Unusual behavior.
  - Less statistical power.

Tax rate inaccuracies



City of Buffalo
Adopted Budget 2021-2022
General Fund

	2017-2018 Actual Amount	2018-2019 Actual Amount	2019-2020 Actual Amount	2020-2021 Adopted Budget
01 TAXES	**(151,431,074)**	**(156,711,165)**	**(157,351,261)**	**(157,190,897)**
311100 REAL PROPERTY TAX LEVY	(130,539,085)	(137,520,219)	(139,169,787)	(139,971,280)
312000 OCCUPANCY TAX	(2,644)	(3,217)	(2,318)	(2,000)

https://www.opendatanetwork.... · Traduceți această pagină :

Total Taxable Value Data for Ann Arbor, MI - Property Tax on ...

The total taxable value of Ann Arbor, MI was $6,684,030,058 in 2021. ... Property Tax is the total amount of revenue collected based on a percentage of the ...

Solutions

- Dropping cities from dataset
- Linear interpolation
- Forward fill and backward fill

ARDL (Autoregressive Distributed Lag)

- $y_t = \beta_0 + \beta_1 y_{t-1} + \ldots\ldots + \beta_p y_{t-p} + \alpha_0 x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots\ldots\ldots + \alpha_q x_{t-q} + \varepsilon_t$

- Analyze dynamic relationship with time-series data
- Model lags of both the dependent (housing price) and explanatory (factors) variables
- Used for forecasting

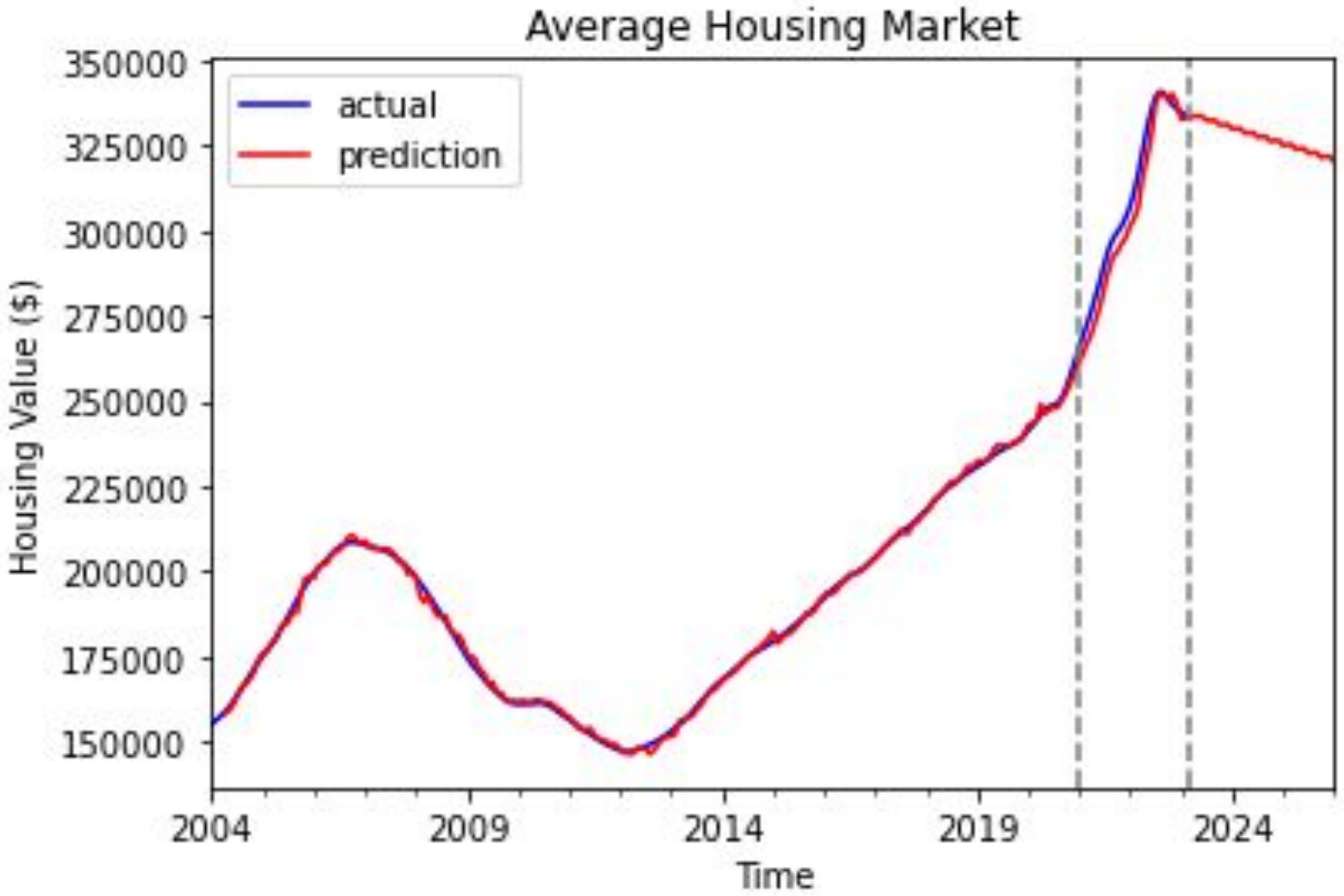
Akaike Information Criteria (AIC)

- Evaluates models
- It helps minimize the amount of variables used in the model
- Penalizes more parameters
- Helps determine if extra parameter is worth it or not

ARDL Model Results

- No. Observations : 28764
- r^2: 0.9928
- RMSE: 9734.54

const	4221.2593
value.L1	1.0201
value.L2	-0.0065
value.L3	-0.0190
value_income.L0	9.3918
value_income.L1	-9.7364
value_income.L2	0.2143
value_income.L3	0.1676
value_unemployment.L0	474.7147
value_unemployment.L1	-630.4867
value_gdp.L0	-0.2321
value_gdp.L1	0.2307
value_population.L0	14.6965
value_population.L1	-14.5992
value_crime.L0	1.2227
value_crime.L1	-1.2140
value_cpi.L0	-1505.8598
value_cpi.L1	1619.3993
value_cpi.L2	-129.9800
value_interest.L0	3558.7019
value_interest.L1	-593.3330
value_interest.L2	-4968.3650
value_interest.L3	1937.5662



In-sample RMSE: $1168.84

Out-of-sample RMSE: $5667.46

Contextualizing the data: Case studies

How we decided the case studies

- Created a code that looped over all 141 cities and organized the cities by growth rate

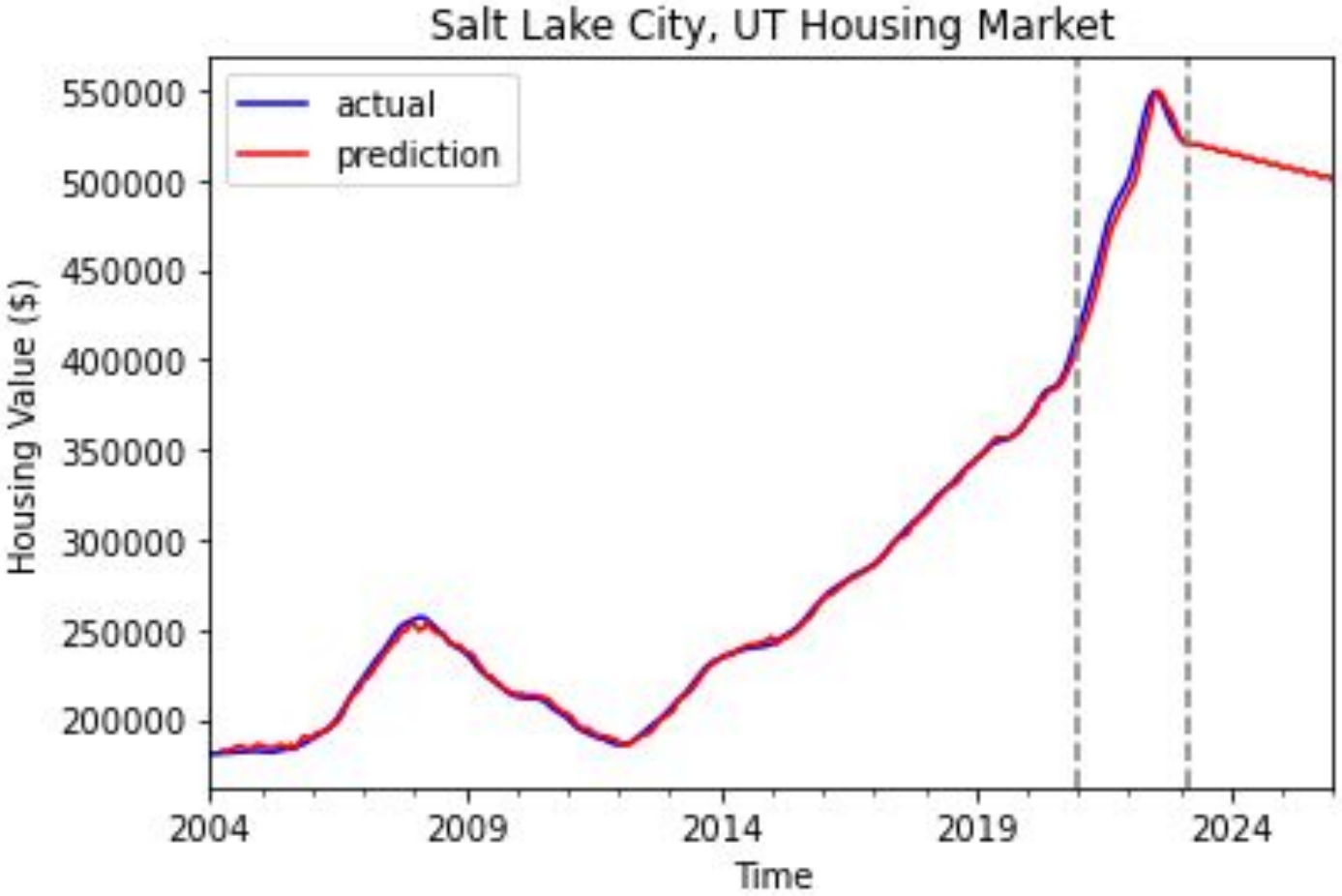
  Growth rate: calculated using housing prices from years

  - Max
  - Min
  - Determined median growth ~1.75

Maximum: a city on fire!

Salt Lake City, UT: 2.99 growth rate

- High amount of immigration to SLC, **low percentage emigrating**
- Limited property available in prime area of the city, creating **high demand** for housing
- Strong local economy, low unemployment
- About half of the population is **Mormon**
- Largest percentage of residents are **25-34 years old**
- Number of births are **higher than national average**, generating a strong population growth



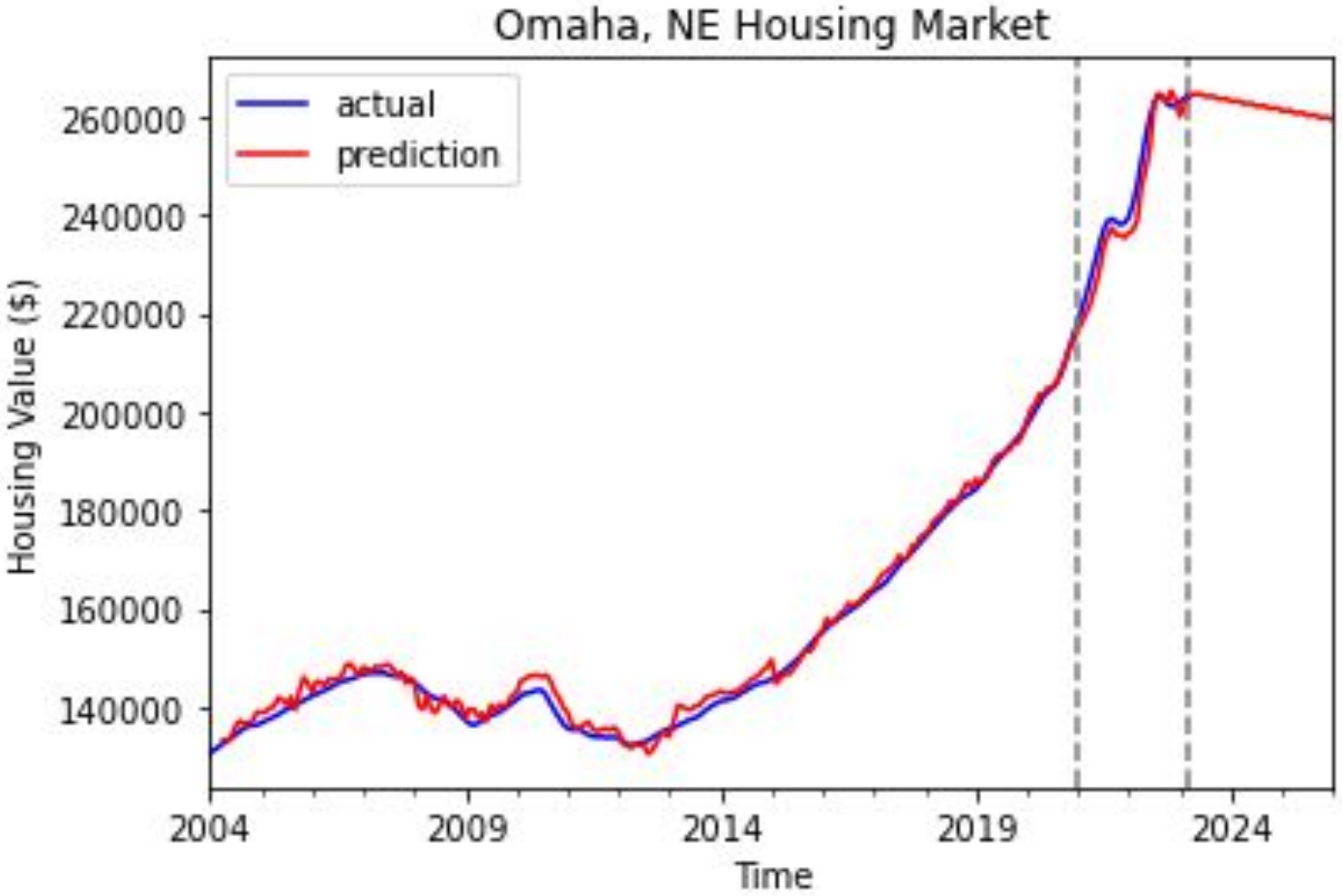
In-sample RMSE: $2075.97

Out-of-sample RMSE: $8872.92

Minimum: a low growth city

Omaha, NE: 0.83 growth rate

- Lots of land/property available, thus **no high demand**/struggle in finding housing
- Population growth **lower than the national average**
- Family-run farming businesses are declining in the area due to much larger companies accruing the benefits of advancing technology and globalization
- **Crowding out farmers** reduces jobs in the industry and they move to another area to work
- **Limited variety of work:** Focused on regional manufacturing, transportation, trade, and service hub



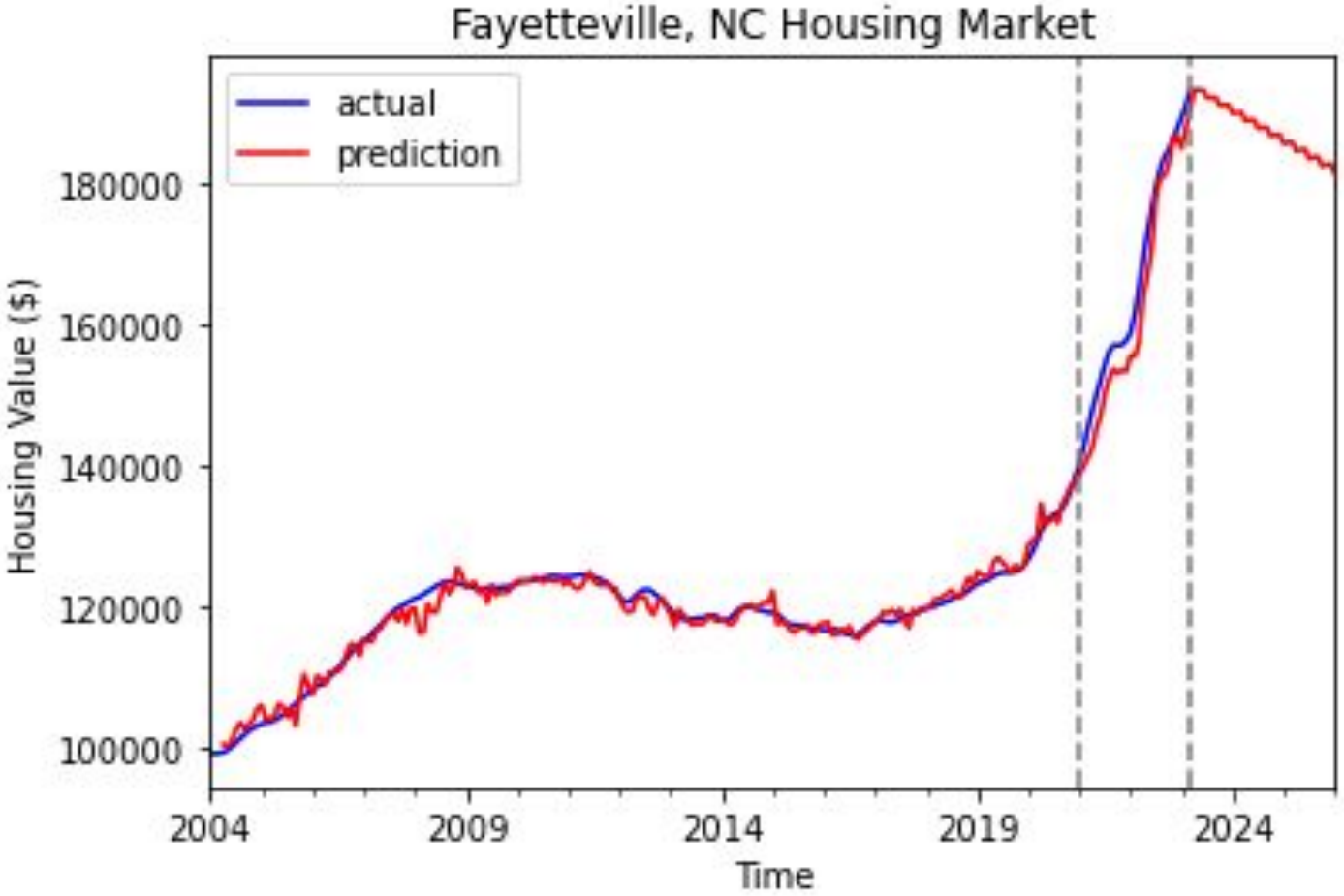
In-sample RMSE: $1802.50

Out-of-sample RMSE: $3863.72

Median: a “standard” city

Fayetteville, NC: 1.77 growth rate

- Known for **Fort Bragg**, a major US army installation
  - Main employer of citizens (14k people)
- Population of **208,501**
  - 6th-largest city in North Carolina
- Cape fear river runs through the city, as does I-95, a key highway
- **Under an hour from Raleigh and Durham**
- Median household income hovers at **$50k**, **30.3 median age**
- Diversity: 41.6% black 35.3% white 12% hispanic
- Includes: historic sites, malls, parks, museums, an arena, multiple colleges, small regional airport



In-sample RMSE: $1279.20

Out-of-sample RMSE: $4217.45

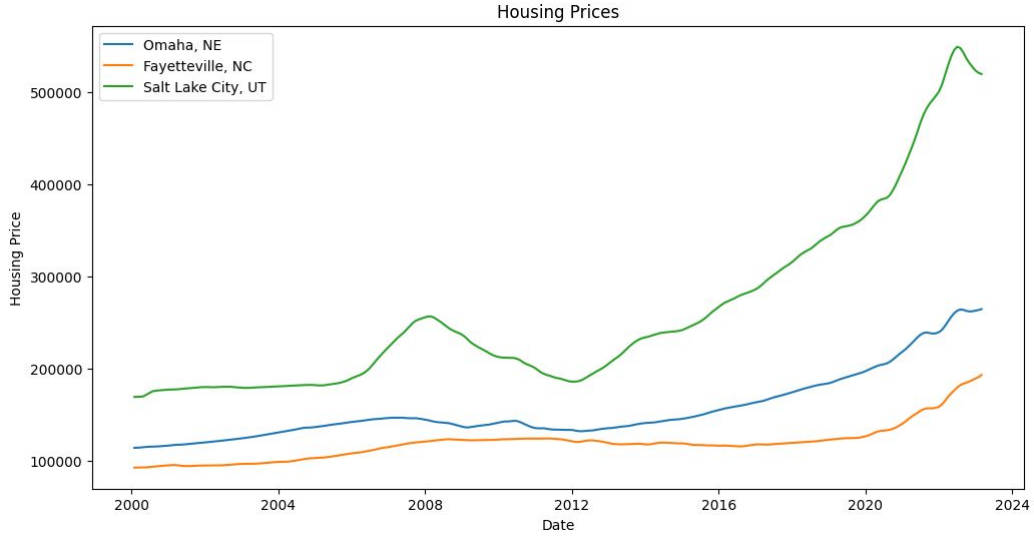
How the case studies compare using key factors

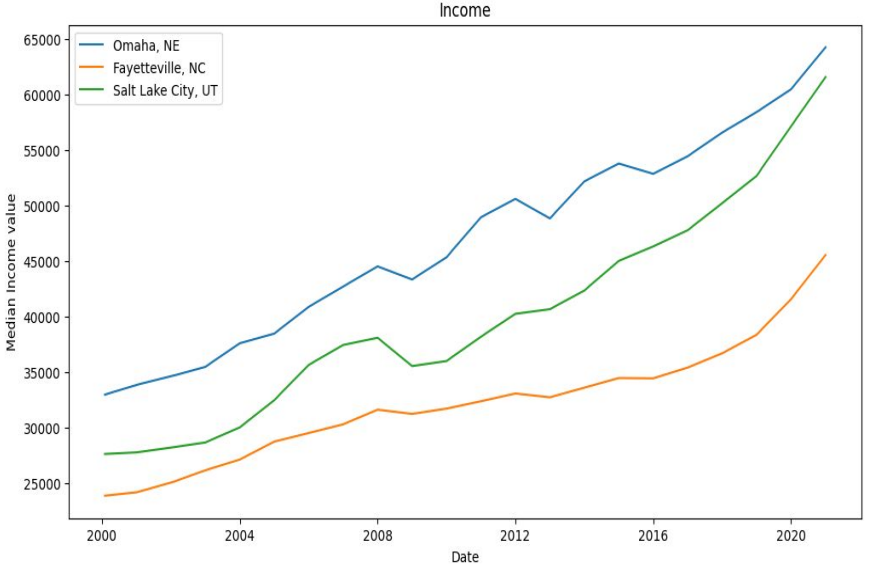
	Fayetteville	Salt Lake City	Omaha
Home value	$119,511.71	$315,634.95	$174,475.30
Income	$36,694	$50,199	$56,556
Unemployment	5.6%	3.1%	3%
GDP	$20,930.74	$98,524.08	$65,880.33

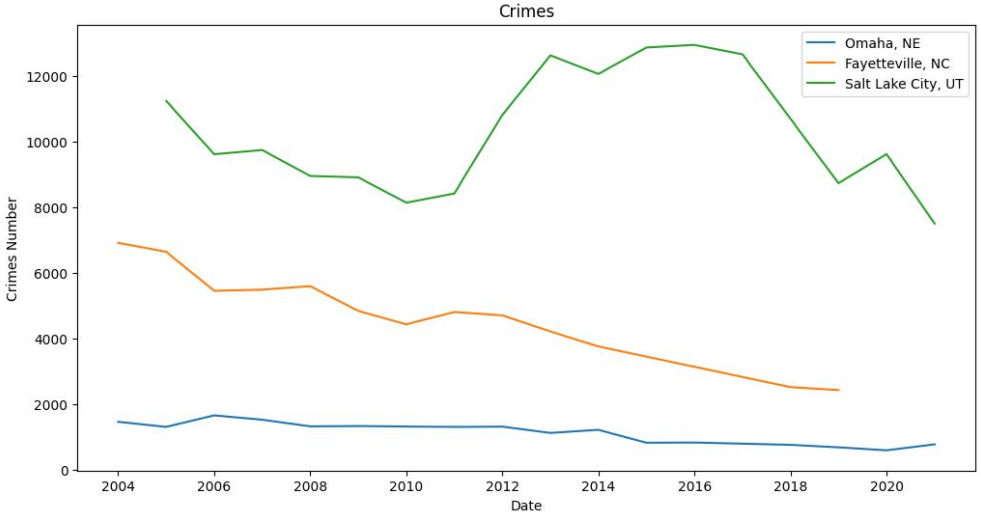
No single factor determines market conditions- context matters!

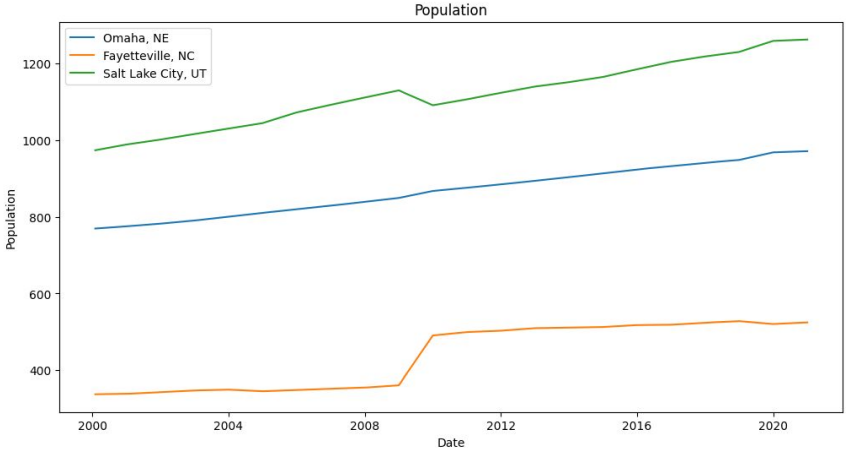
*2018 used for standardization, pre-COVID-19

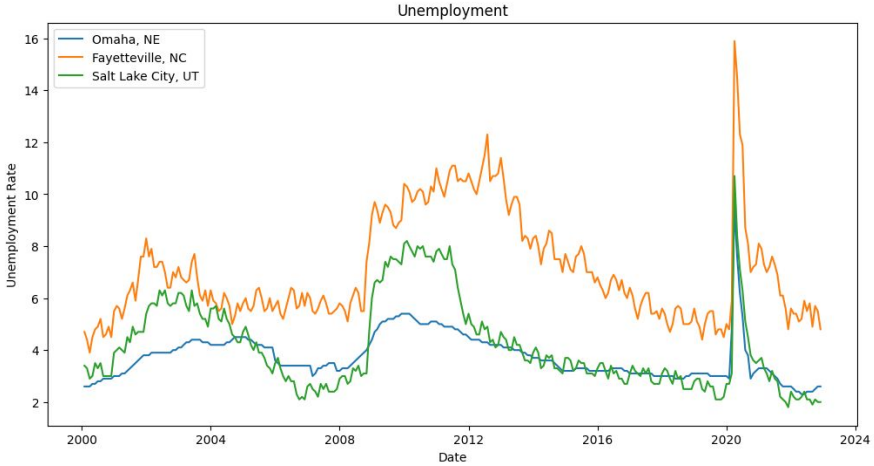
Comparing these 3 cities

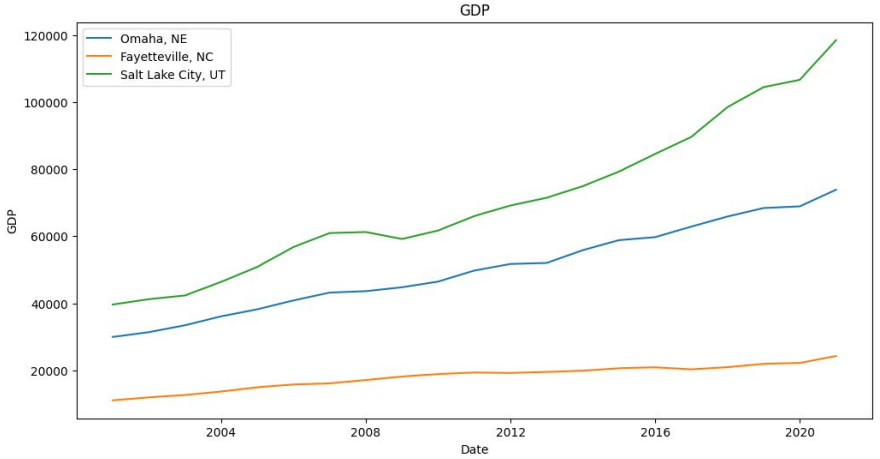





Crimes
Omaha, NE
Fayetteville, NC
Salt Lake City, UT
Crimes Number
Date


Population
Omaha, NE
Fayetteville, NC
Salt Lake City, UT
Population
Date


Unemployment
Omaha, NE
Fayetteville, NC
Salt Lake City, UT
Unemployment Rate
Date


GDP
Omaha, NE
Fayetteville, NC
Salt Lake City, UT
GDP
Date

Data Sources

Zillow Housing Data:
https://files.zillowstatic.com/research/public_csvs/zhvi/Metro_zhvi_uc_sfrcondo_tier_0.33_0.67_sm_sa_month.csv

FRED Database: https://fred.stlouisfed.org/

Simple Maps City Data:
https://simplemaps.com/static/data/us-cities/1.76/basic/simplemaps_uscities_basicv1.76.zip

ArcGIS States Shapefile:
https://services2.arcgis.com/DEoxb4q3EJppiDKC/arcgis/rest/services/States_shapefile/FeatureServer/0/query?outFields=*&where=1%3D1&f=geojson

Stanford Cities Shapefile:
https://stacks.stanford.edu/file/druid:bx729wr3020/data.zip

Questions?