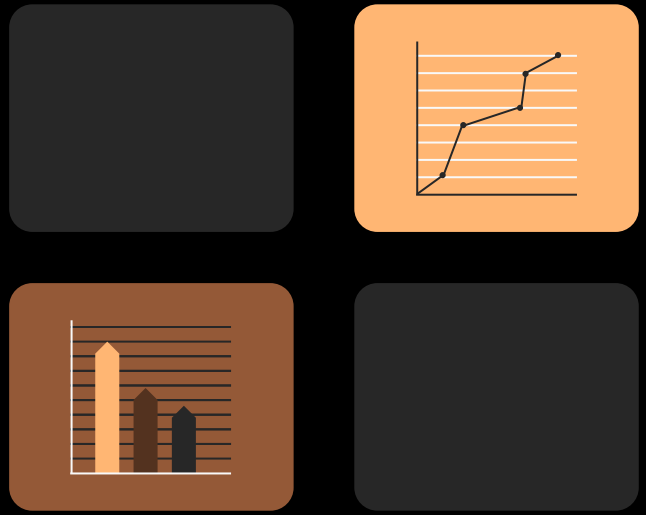
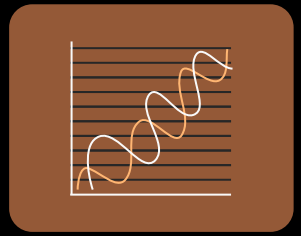


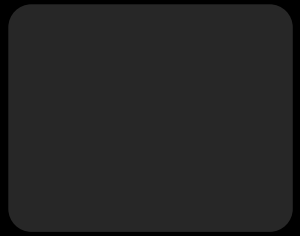
Assessing Ideal U.S. Northeast States for Living

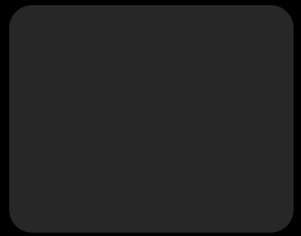
Jabari Coke, Irith Chaturvedi, Rishab Kalra

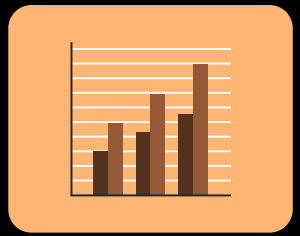
TABLE OF **CONTENTS**

01 *









Introduction

Introduction

01 Why This Project?

The Northeast coast of the United States, comprising states like New York, New Jersey, and Massachusetts, Maine, has long been a location of, cultural, and historical significance. Yet as time progress the natives(minorities) of NYC continue to struggle with affording cost of living with their salaries.

02 Relation To Humanities

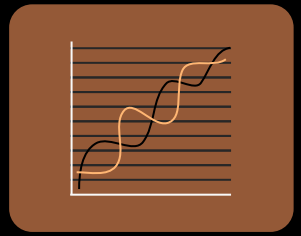
We are not just trying to look at this from an economic perspective. Yet, we aim to understand a cultural and economic perspective.

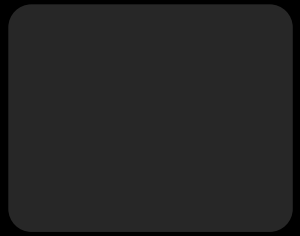
Why **NYC?**

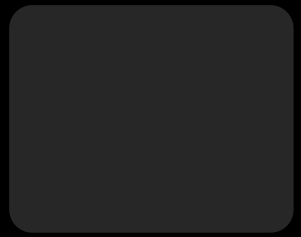
Why do people live in NYC?

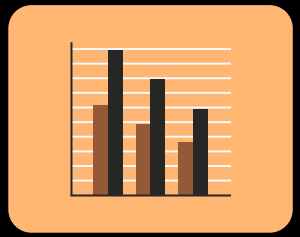
Because of:

- Jobs
- Diverse Cultures
- Great Food









* “65 percent of Latino households struggle to afford living costs, followed by 58 percent of Black households and 51 percent of Asian, Native Hawaiian, or Pacific Islander New Yorkers.”

—Daniel Parra

Why **NOT** NYC?

Why not live in NYC?

Because of:

- High Cost of Living
- Poor standard of living
- High pollution
- High Population Density

Alternative Places in the U.S. Northeast

Why Stay in Northeast:

- Addressing fear of moving from the Northeast due to historical prosperity.
- Explore alternate places near New York with affordability and career opportunities.

Focus on Affordability:

- Emphasis on cities that won't strain the budget.
- Aim to provide options for career development without a hefty cost.

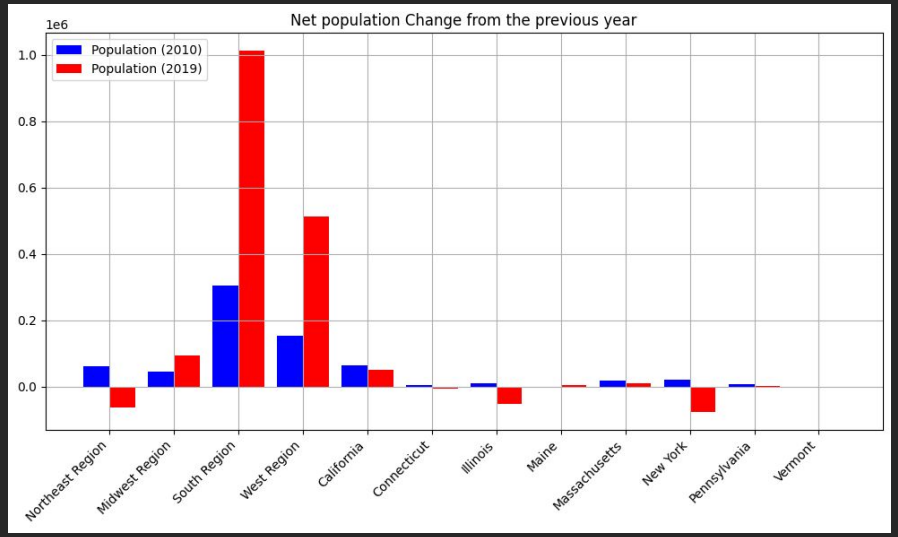
Avoiding Newly Saturated Cities:

- Caution against moving to popular Southern cities like Austin and Dallas.
- Consider sister cities for better career prospects.

Strategic Move from New York:

- Encourage strategic decisions for those leaving New York.
- Importance of exploring nearby cities with growth potential.

Alternative Places in the U.S. Northeast

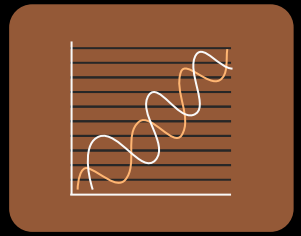


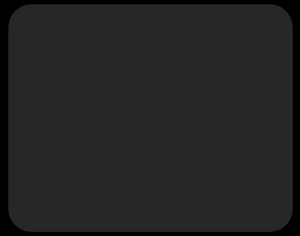
The Y – axis describes the population change in numbers, while the X – axis are the regions being compared.

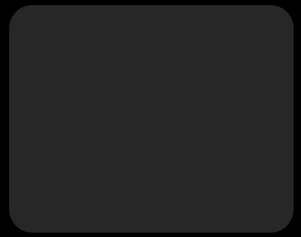
As evident, the South has seen a drastic change in population even from one year to the other. In comparison, the American Northeast has been relatively unchanged in total population.

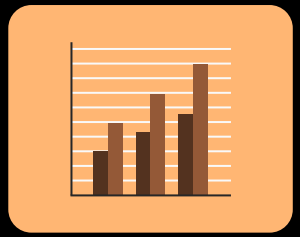
Suggestion – Move to cities in the Northeast when moving from New York.

02*









Methodology

Methodology

Data Retrieval

Datasets were retrieved from Kaggle and Data.gov, a government subsidiary that contains millions of datasets that are regularly updated.. The datasets used for Spatial Analysis containing information about the PCI and the data about the housing market in the Northeast, while the dataset obtained from Data.gov has information on Population changes, international and domestic migration within the US.

Data Cleaning

All Null or empty values were removed from the datasets. Unwanted columns and rows were dropped.
In terms of population analysis, data was cleaned based on two years, 2010 and 2019 to show a definitive change over a decade. Only states and regions directly related to the study were kept in the final dataset for analysis.

Methodology

Data Sorting

Data was sorted and grouped for each state, numerical data points were aggregated on the basis of their mean values. For populations, only data containing information for the years 2010 and 2019 was kept to create a comparative study based on time.

Analysis & Visualizations

Visualizations were created using the Folium Spatial Analysis Library, using the US States as coordinates, and color coding based on the data point values (House Prices and PCI). Population data was analyzed using PANDAS data analysis library, Double bar plots were made to show a comparison between the years.

Datasets and Preprocessing

Per Capita Income

	county	states	pci	household_income	family_income	population	num_of_households
0	New York County	New York	76592	69659	86553	1628706	759460
1	Arlington	Virginia	62018	103208	139244	214861	94454
2	Falls Church City	Virginia	59088	120000	152857	12731	5020
3	Marin	California	56791	90839	117357	254643	102912
4	Santa Clara	California	56248	124055	124055	1927852	640215
...	...	...	...	...	...	...	...
3226	Comerío	Puerto Rico	7047	13523	15799	20778	6075
3227	Western District	American Samoa	6429	24705	24916	31329	5418
3228	Eastern District	American Samoa	6191	23350	24911	23030	3982
3229	Maricao	Puerto Rico	5943	13462	15864	6276	1914
3230	Manu'a District	American Samoa	5441	17614	19226	1143	282

3231 rows × 7 columns

	states	pci	household_income	population	num_of_households
0	Alabama	20481.432836	37778.716418	71631.000000	27443.029851
1	Alaska	29357.655172	62611.344828	24838.482759	8686.172414
2	American Samoa	6911.750000	23292.250000	13879.750000	2422.000000
3	Arizona	20863.933333	42891.266667	431980.200000	158019.266667
4	Arkansas	19443.186667	35861.813333	39111.586667	15062.973333
5	California	27509.327586	56115.827586	651290.086207	216781.000000
6	Colorado	27625.890625	50853.390625	79989.515625	30899.859375
7	Connecticut	36208.500000	70502.875000	447945.125000	169481.125000
8	Delaware	28055.000000	57465.333333	302815.333333	111902.333333
9	Florida	22904.432836	43715.925373	284942.626866	106850.447761
10	Georgia	20093.773585	40365.031447	61700.735849	22126.396226

Datasets and Preprocessing

Population Change

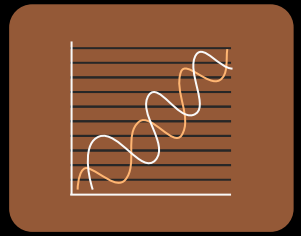
	SUMLEV	REGION	DIVISION	STATE	NAME	CENSUS2010POP	ESTIMATESBASE2010	POPESTIMATE2010	PC
0	10	0	0	0	United States	308745538	308758105	309321666	
1	20	1	0	0	Northeast Region	55317240	55318443	55380134	
2	20	2	0	0	Midwest Region	66927001	66929725	66974416	
3	20	3	0	0	South Region	114555744	114563030	114866680	
4	20	4	0	0	West Region	71945553	71946907	72100436	
5	40	3	6	1	Alabama	4779736	4780125	4785437	
6	40	4	9	2	Alaska	710231	710249	713910	
7	40	4	8	4	Arizona	6392017	6392288	6407172	
8	40	3	7	5	Arkansas	2915918	2916031	2921964	
9	40	4	9	6	California	37253956	37254519	37319502	
10	40	4	8	8	Colorado	5029196	5029319	5047349	
11	40	1	1	9	Connecticut	3574097	3574147	3579114	
12	40	3	5	10	Delaware	897934	897937	899593	
					District of				

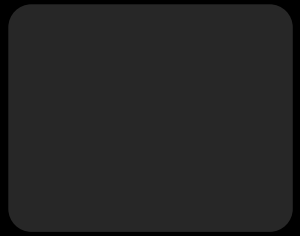
	STATE	NAME	POPESTIMATE2010	POPESTIMATE2019	NPOPCHG_2010	NPOPCHG_2019	INTERNATIONALMIG
1	0	Northeast Region	55380134	55982803	61691	-63817	4
2	0	Midwest Region	66974416	68329004	44691	92376	2
3	0	South Region	114866680	125580448	303650	1011015	6
4	0	West Region	72100436	78347268	153529	512448	3
9	6	California	37319502	39512223	64983	50635	1
11	9	Connecticut	3579114	3565287	4967	-6233	
18	17	Illinois	12840503	12671821	8931	-51250	
24	23	Maine	1327629	1344212	-729	5155	
26	25	Massachusetts	6566307	6892503	18522	9868	
37	36	New York	19399878	19453561	21734	-76790	1
43	42	Pennsylvania	12711160	12801989	8292	1067	
50	50	Vermont	625879	623989	142	-369	

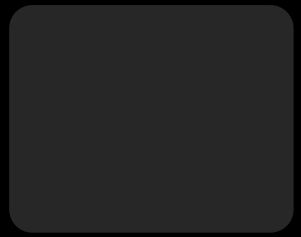
Correlation between chosen variables

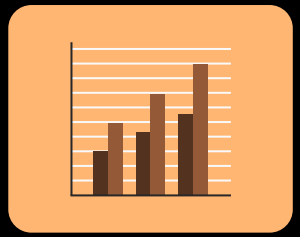
	bed	bath	acre_lot	house_size	price	pci	household_income	population	num_of_households
bed	1.000000	0.007214	0.275547	0.375152	0.369713	-0.097227	-0.162067	0.246835	0.251303
bath	0.007214	1.000000	0.205016	0.522889	0.750192	0.381558	0.395714	0.313867	0.311874
acre_lot	0.275547	0.205016	1.000000	0.587406	0.409986	0.062129	0.068757	0.121635	0.105290
house_size	0.375152	0.522889	0.587406	1.000000	0.509454	0.486238	0.464976	0.429005	0.425434
price	0.369713	0.750192	0.409986	0.509454	1.000000	0.313786	0.309990	0.447593	0.452097
pci	-0.097227	0.381558	0.062129	0.486238	0.313786	1.000000	0.989686	0.668744	0.680355
household_income	-0.162067	0.395714	0.068757	0.464976	0.309990	0.989686	1.000000	0.707060	0.713777
population	0.246835	0.313867	0.121635	0.429005	0.447593	0.668744	0.707060	1.000000	0.999002
num_of_households	0.251303	0.311874	0.105290	0.425434	0.452097	0.680355	0.713777	0.999002	1.000000

03*



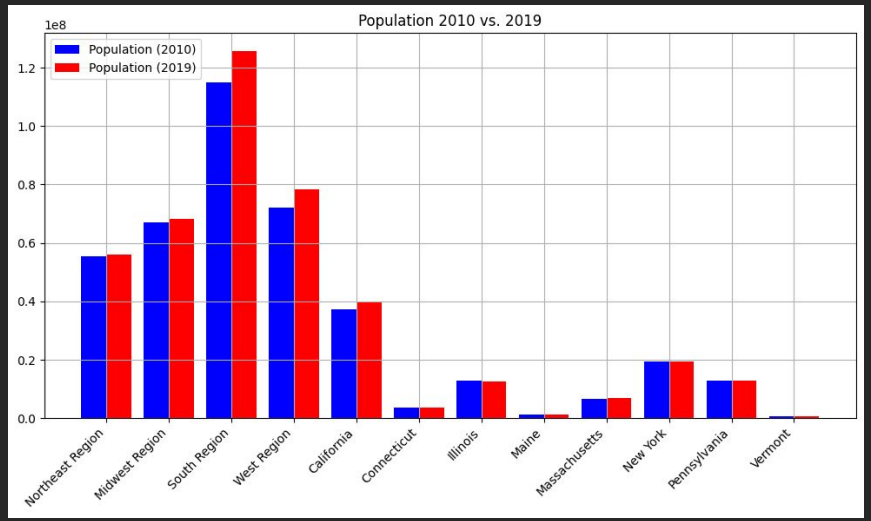






Results

Results



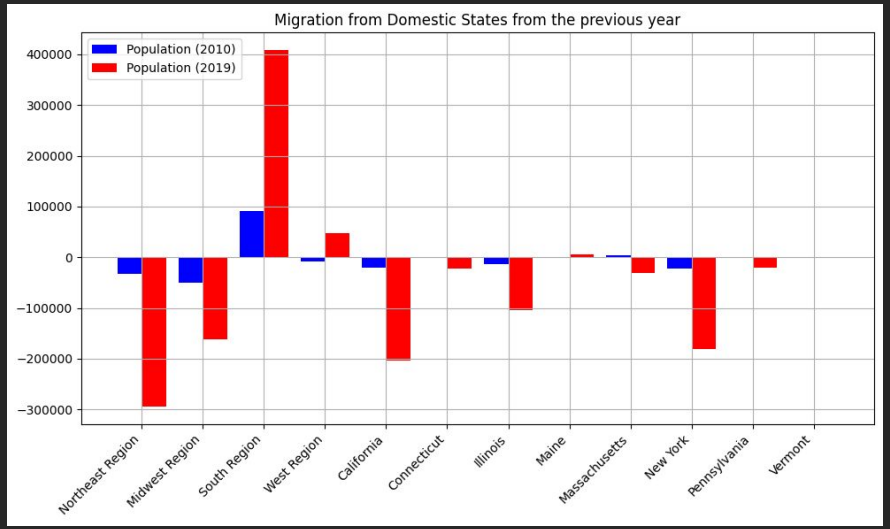
Analysis & Visualizations

Areas where the cost of living is more reasonable have experienced an increase in population between 2010 and 2019

Example: Southern Regions

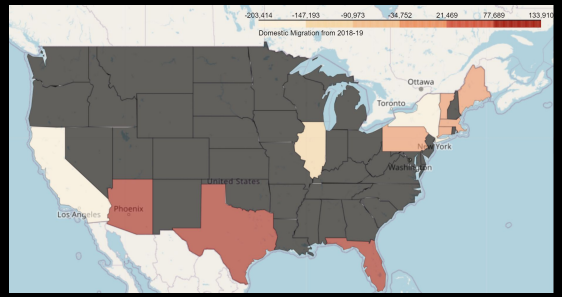
However, most NE states have not seen a drastic increase in population.

Results

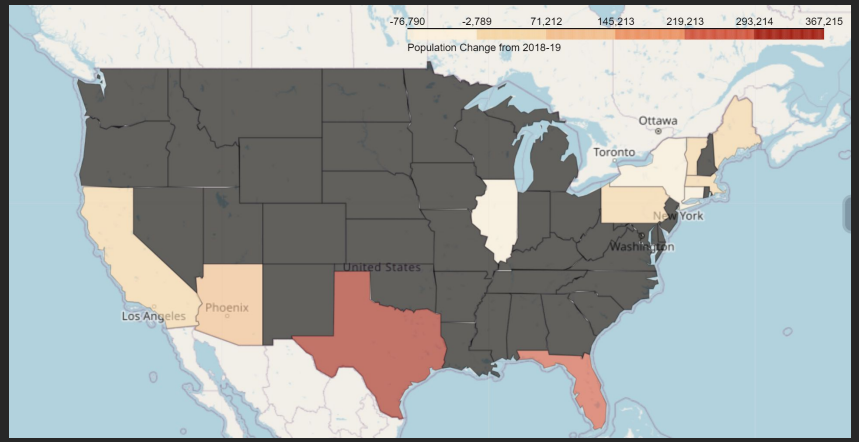


The increase in population for the South is strongly correlated to migration from domestic states

As Expected, NY sees a huge drop in domestic migration rates.

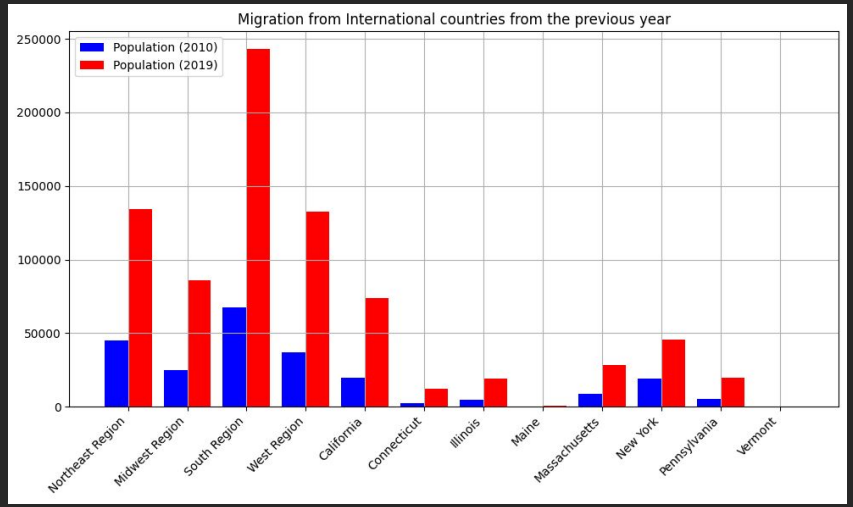


Results



Areas that are more expensive to live in have a drop in population between 2010 and 2019. Population changes from previous years show a trend of stagnancy in most NE states.

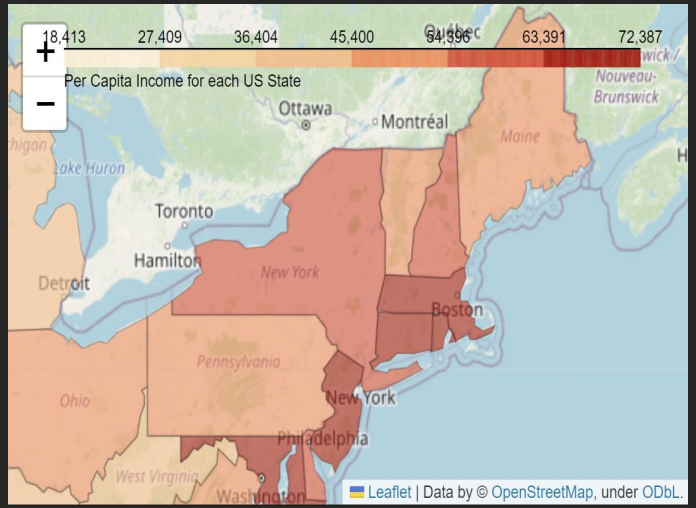
Results

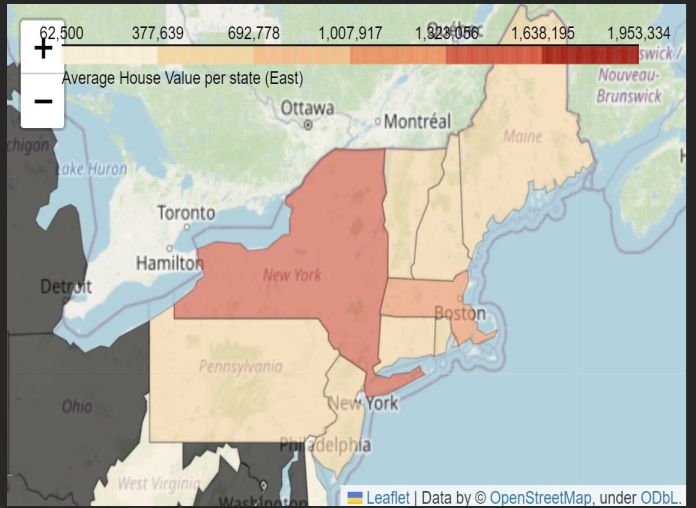


Areas that have a strong job market and reasonable cost of living seem to be favorable for international migrants.

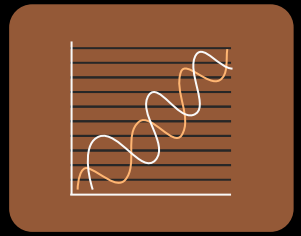
That's why the Southern & Northeastern regions are more popular

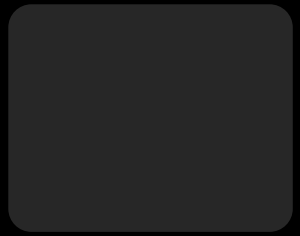
Results

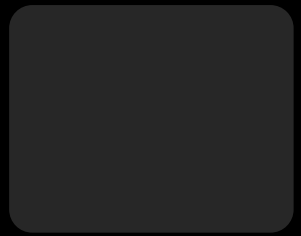

18,413
27,409
36,404
45,400
54,396
63,391
72,387
Per Capita Income for each US State
Ottawa
Montréal
Toronto
Hamilton
Detroit
New York
Boston
Pennsylvania
Ohio
New York
Philadelphia
West Virginia
Maine
Nouveau-Brunswick
Leaflet | Data by © OpenStreetMap, under ODbL.

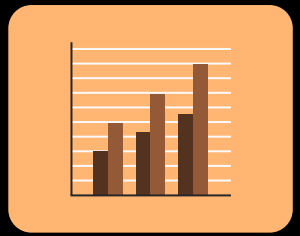

62,500
377,639
692,778
1,007,917
1,323,056
1,638,195
1,953,334
Average House Value per state (East)
Ottawa
Montréal
Toronto
Hamilton
Detroit
New York
Boston
Pennsylvania
Ohio
New York
Philadelphia
West Virginia
Maine
Nouveau-Brunswick
Leaflet | Data by © OpenStreetMap, under ODbL.

04*









Conclusion

Conclusion

Open Availability – Job Market

The presence of the US Northeast as an economic powerhouse can never go unnoticed. As clearly seen, states in the Northeast have seen a relatively unchanged population while the job market there continues to grow,

Open Availability – Housing Market

Housing in every state other than New York is relative to the US average. The states thus show a great blend of affordable housing with massive employment opportunities and a great standard of life.

Alternative States and Cities

We now look at two cities we think work best for people deciding to move to the Northeast and out of New York. Providing them a metropolitain vibe with a great standard of living. As population too has a great correlation with per-capita-income, we will suggest cities that have a good population.

Pennsylvania

Pittsburgh

- Mean House value ~ 160K
- Black or African American (Non-Hispanic) (22.7%), Asian (Non-Hispanic) (5.55%), Two+ (Non-Hispanic) (3.96%), and White (Hispanic) (1.61%)
- Mean Salary for a white-collar job ~ 64K

ML Model

Predicting best neighborhood to live in based on your budget

The model is trained on the best and affordable Neighborhoods to live in Pittsburgh, It can predict the ideal place to live in for your given budget

```
X_train_reshaped = np.array(X_train).reshape(-1, 1)
y_train_reshaped = np.array(y_train).reshape(-1, 1)

# Create a pipeline with StandardScaler and Logistic Regression
pipeline = Pipeline([
    ('scaler', StandardScaler()),
    ('lr', LogisticRegression())
])

# Define the hyperparameter grid
param_grid = {'lr__C': [0.001, 0.01, 0.1, 1, 10, 100, 1000]}

# Create the GridSearchCV object
lr_grid = GridSearchCV(pipeline, param_grid=param_grid, cv=5)

# Fit the GridSearchCV object to the data
lr_grid.fit(X_train_reshaped, y_train_reshaped)

# Print the best parameter and score
print("Best parameter: ", lr_grid.best_params_)
print("Best score: ", lr_grid.best_score_)
```

```
lr_grid.predict([[400000]])
✓ 0.0s
array(['Shadyside'], dtype=object)
```

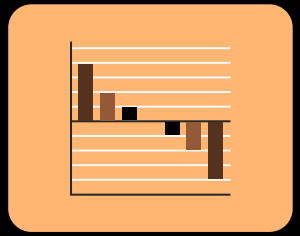
Connecticut

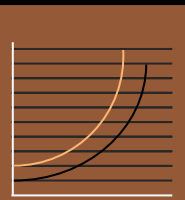
Hartford

- Mean House value ~ 190K
- Black or African American (Non-Hispanic) (34.1%), Other (Hispanic) (20.1%), White (Non-Hispanic) (14.7%), White (Hispanic) (13.1%), and Two+ (Hispanic) (9.41%).
- Mean Salary for a white-collar job ~ 73K

THANKS! *

ANY QUESTIONS?





*

BIBLIOGRAPHY

Sources:

- https://citylimits.org/2023/04/26/more-than-half-of-immigrant-new-yorkers-cant-afford-basic-needs-report-finds/#:~:text=The%20report%2C%20which%20looked%20at,or%20Pacific%20Islander%20New%20Yorkers.
- https://www.migrationpolicy.org/article/caribbean-immigrants-united-states-2017
- https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset
- https://www.kaggle.com/datasets/kabhishm/united-states-counties-by-per-capita-income
- https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html
- https://www.istockphoto.com/photos/
- https://datausa.io/profile/geo/pittsburgh-pa/#:~:text=The%205%20largest%20ethnic%20groups,(Hispanic)%20(1.61%25).
- https://datausa.io/profile/geo/hartford-ct#:~:text=The%205%20largest%20ethnic%20groups,(Hispanic)%20(9.41%25).
- https://www.redfin.com/neighborhood/156434/PA/Pittsburgh/Shadyside/housing-market
- https://catalog.data.gov/dataset/property-data-with-geographic-identifiers/resource/ed31b5da-6c4c-48ba-bcf8-718777778ba9