Investigating on the relationship of House prices and Covid 19 confirmed cases in Polk, Florida

DATA512 Project
Matthew Chau

Motivation

- Confused by the decreasing rent of my apartment and some of my friend's apartments during COVID

- Seeking to find the relationship between the economy and the COVID 19 cases, house prices could be a good indicator

Problem Statement

- **Research Question:**

- Does the pandemic have an impact on the housing prices of Polk, Florida? If so, how was the local housing prices influenced by the pandemic?

- **Methodology:**

- Linear Regression for house prices

- T-test comparing the slope of house prices before/after Covid

- Cross correlation of House prices and Covid confirmed cases

House Prices and Human Centered Data Science

- **Bias in data collection:**

- Did not consider the attributes of the houses

- Did not consider the newly built houses

- Covid data might be biased due to political reasons (US tend to report a falsely higher number whereas some other countries report a falsely lower number than actual)

- **Requires support from Qualitative research:**

- Correlation vs Causation – quantitative approaches could not justify this is a causational relationship

- Pandemic might not be the only trigger – need to consider impact of other factors
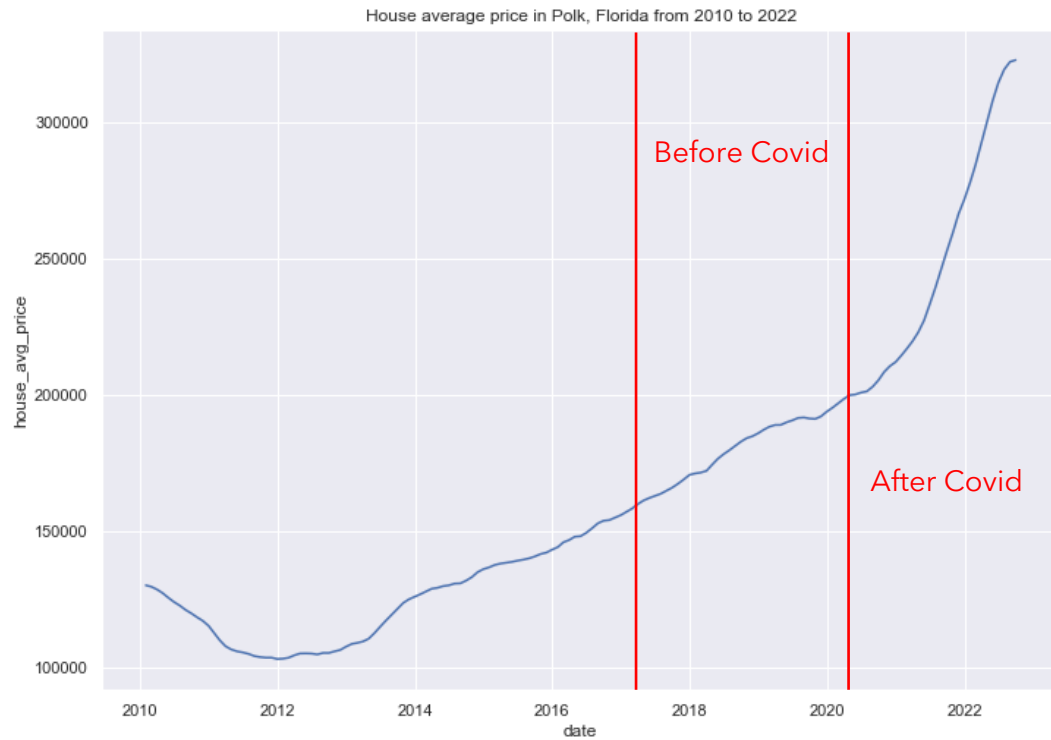
Data

- **Zillow Home Valued Index**

- A smoothed, seasonally adjusted measure of the typical home value and market changes across a given region and housing type.

- Contains monthly averaged house prices across regions up to a neighborhood granularity level. County level is chosen for this project, and house prices of Polk, Florida is selected from entire US data.

	date	house_avg_price
0	2000-01-31	101224.0
1	2000-02-29	101585.0
2	2000-03-31	101858.0
3	2000-04-30	102091.0
4	2000-05-31	102364.0
...	...	...
268	2022-05-31	307742.0
269	2022-06-30	314304.0
270	2022-07-31	319160.0
271	2022-08-31	321991.0
272	2022-09-30	322653.0

273 rows × 2 columns

Quick Glance into Data

Testing Growth difference before and after Covid Outbreak
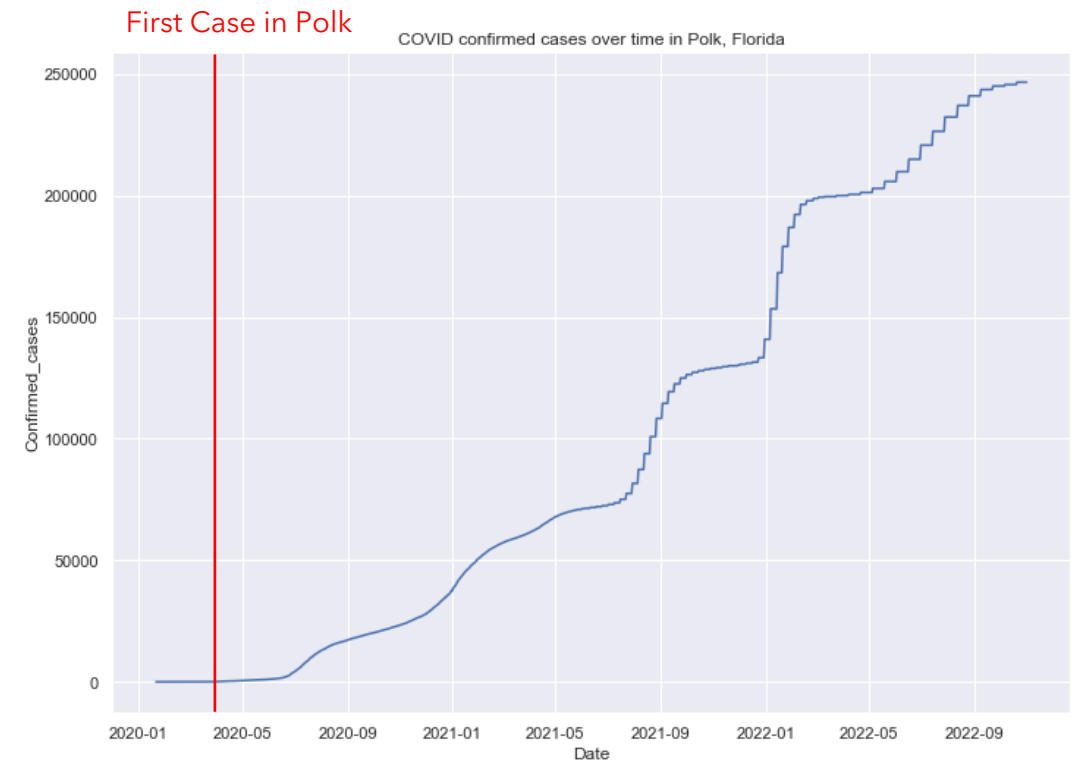
Linear regression fit on house prices before covid

- Fitted Linear Regression on house prices data both before and after Covid Outbreak

Before Covid:

Slope: 33.91

Y-intercept: 168003.41

R2: 0.96

After Covid:

Slope: 151.56

Y-intercept: 177446.25

R2: 0.94

Linear regression fit on house prices after covid

Testing Growth difference before and after Covid Outbreak

Test Statistics

t-Value: 2.66

Degrees of freedom: 57

Probability: 0.0101

$$t = \frac{b_1 - b_2}{\sqrt{s_{b1}^2 + s_{b2}^2}}, df = n_1 + n_2 - 4$$

where b_1 and b_2 are the slopes of lines 1 and 2, s_{b1} and s_{b1} and s_{b2} are the standard errors for lines 1 and 2, and n_1 and n_2 are the sample sizes for lines 1 and 2.

# Finding the correlation between the House prices and Covid Confirm Cases

- **Interpolation**

- Since the house prices are collected on a monthly basis, and the covid confirmed cases is daily data, we need to make sure they are on the same scale.

- The Interpolation method is used instead of linear regression is because it would result in a lower error. Day 15's housing prices is very likely to lie between day 1's price and day 30's price since the change in house prices are very gradual.

Interpolation for House prices
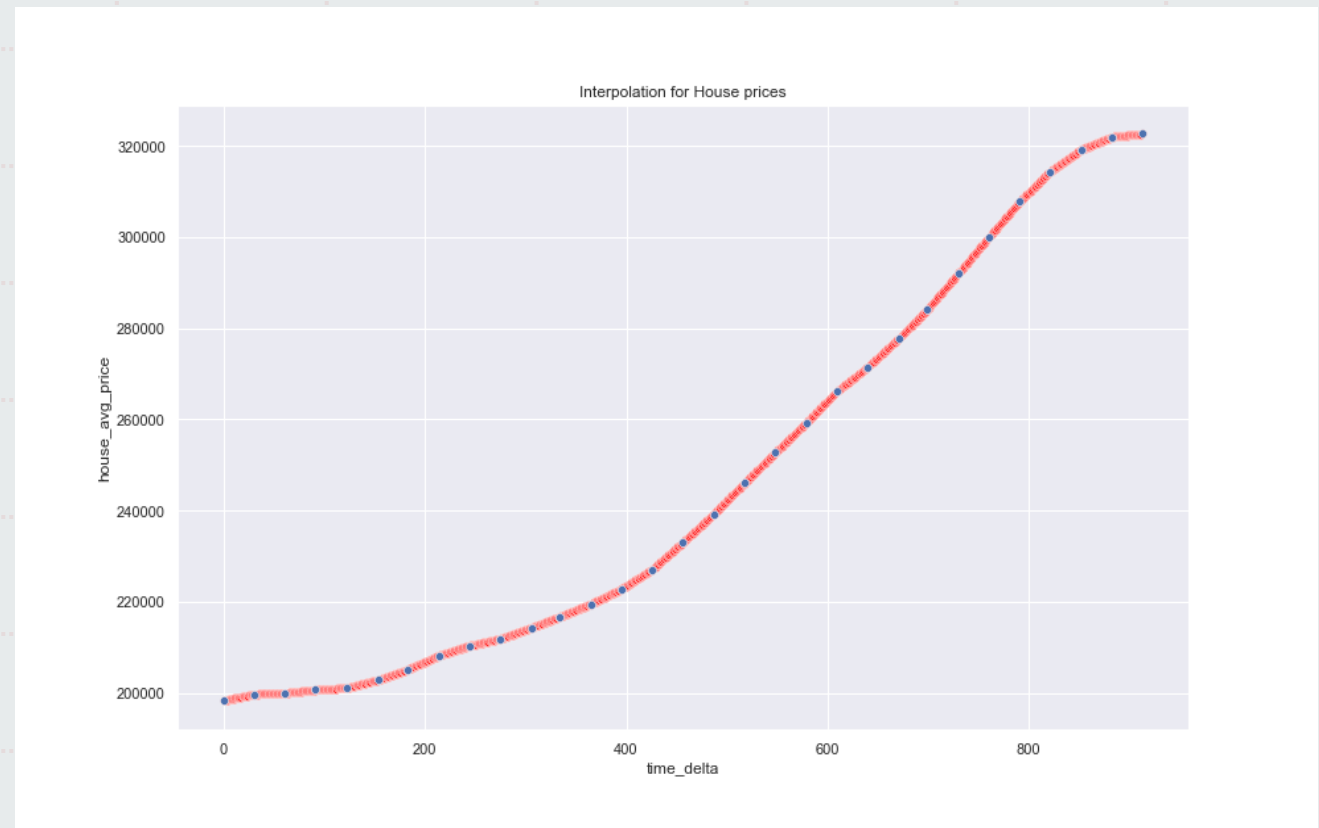
Interpolation and Correlation

- **Interpolation**

- The housing prices of Polk and the Covid Confirmed Cases achieves the highest point at lag 0, where the coefficient almost got to 1.0

- The Correlation decreases as the lag increases, and appears to be almost perfectly linear

cross correlation between house prices and polk confirm cases
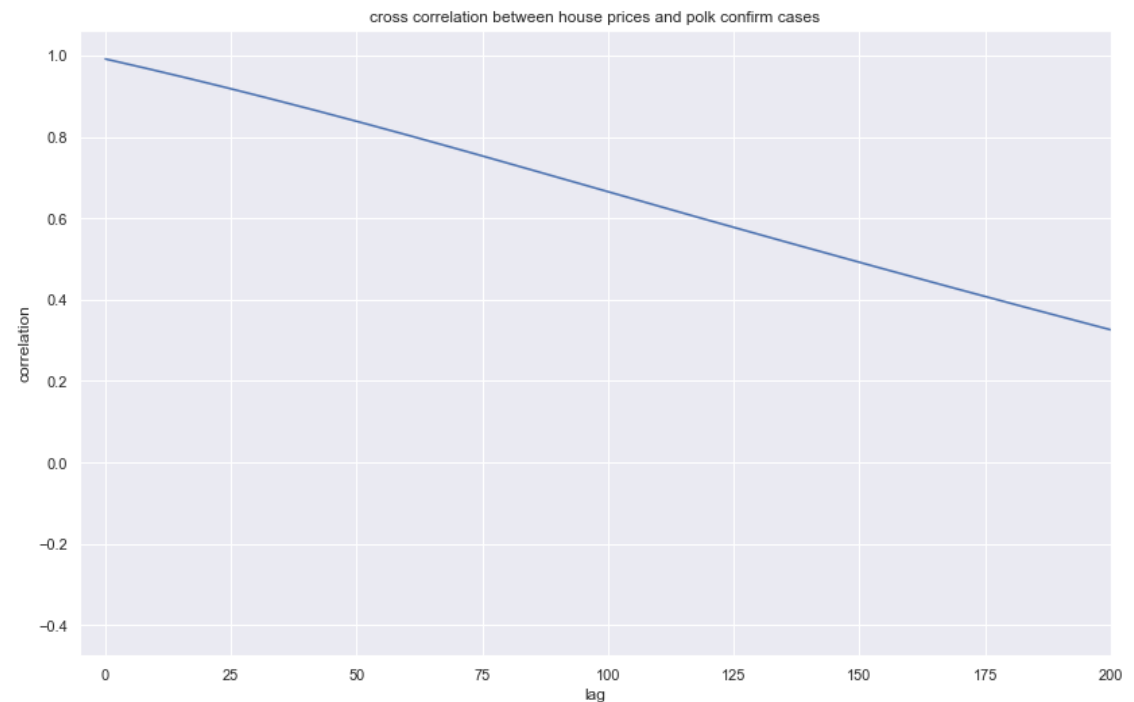
Conclusion and Significance of the Results

• Conclusion

- The increase rate of housing prices in Polk does have a statistically significant difference between Covid and after Covid.

- The change in housing prices have a very strong correlation with the Covid Confirmed cases change in Polk, Florida, and the correlation decreases as the time lag decreases.

• Significance

- The results suggest that there is a high chance that COVID 19, although might not be the primary one, could be a trigger of the surge in the house prices. If further analysis show that this is the main cause, it could imply that similar pandemics could have similar impacts on the house prices as well.

- This brings insights to investors and house owners in Polk, since the strong correlation could imply that there's a high possibility that when confirmed cases increase rate decreases, the house prices would have a lower increase rate or even decrease.

Future Improvements

Apply the research to US scale

Qualitative analysis on Covid, house prices and potential other infleucing factors

Investigate on the derivatives/second derivatives of the two variables (rate of change)