Challenge 3 - Data Incubator

Martin Gascon

11/01/2015

Introduction

Motivation and Goals

My idea was to compare the housing market with the stock market. If there is a correlation between them, I would like to know which index predicts better the Housing price and which cities are more correlated to those indexes.

Some of the various questions that I would like to answer include:

- Does the Stock Market Affect the Housing Market?
- If there is any effect, is this the same for different cities?
- Which index is more connected to the Housing Market?

About the Data

For this project, I've downloaded data from Zillow (www.zillow.com) which provides the monthly mean value per square feet of all homes in U.S from April 1996 to April 2015. This data includes values per month organized by zipcodes for every major city in the country. I've also downloaded data about the stock market index (NASDAQ and DOW JONES INDUSTRIAL AVERAGE) for the same time range from the yahoo finances site.

- http://finance.yahoo.com/q/hp?a=03&b=1&c=2006&d=03&e=1&f=2015&g=m&s=NYA%2C+&ql=1

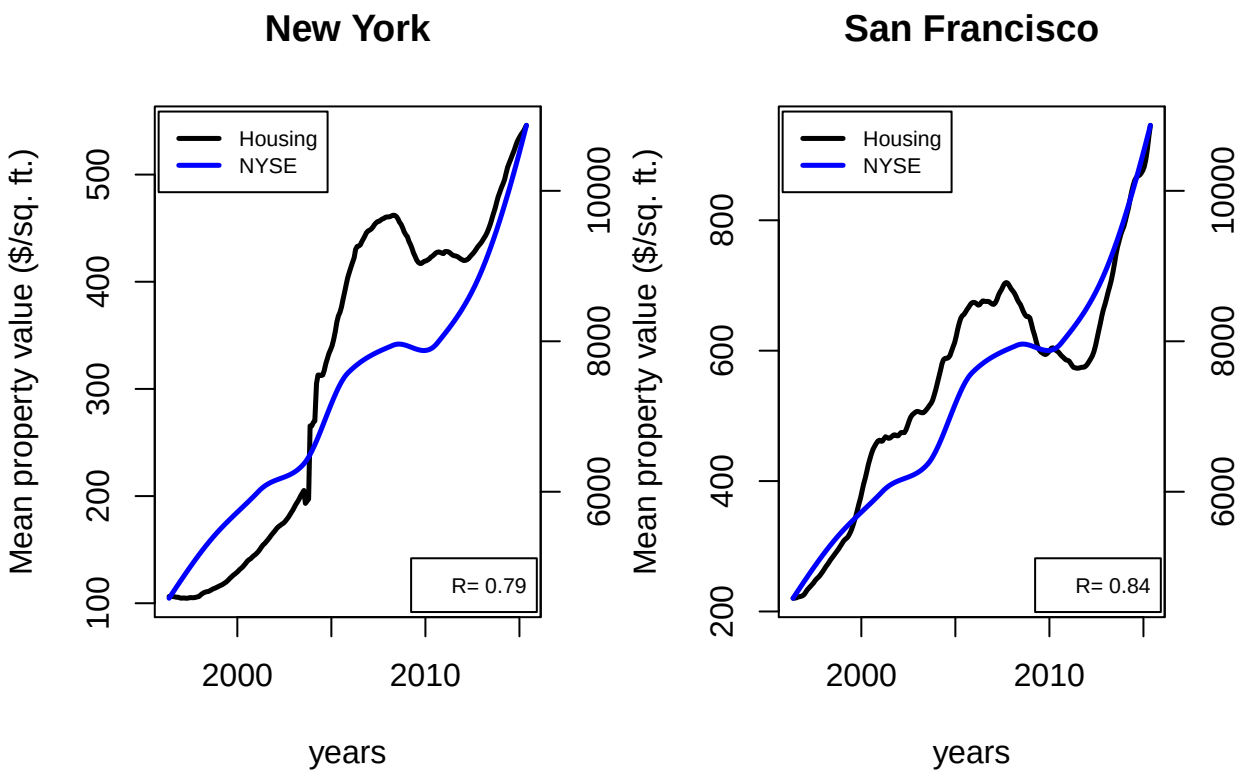
Analysis

Exploratory analysis

The exploratory analysis using R language reveals that we have around 3672 cities and 229 variables which correspond to monthly prices. I averaged the prices for each city (all zipcodes) and selected only complete cases. I also selected the close value for the market index.

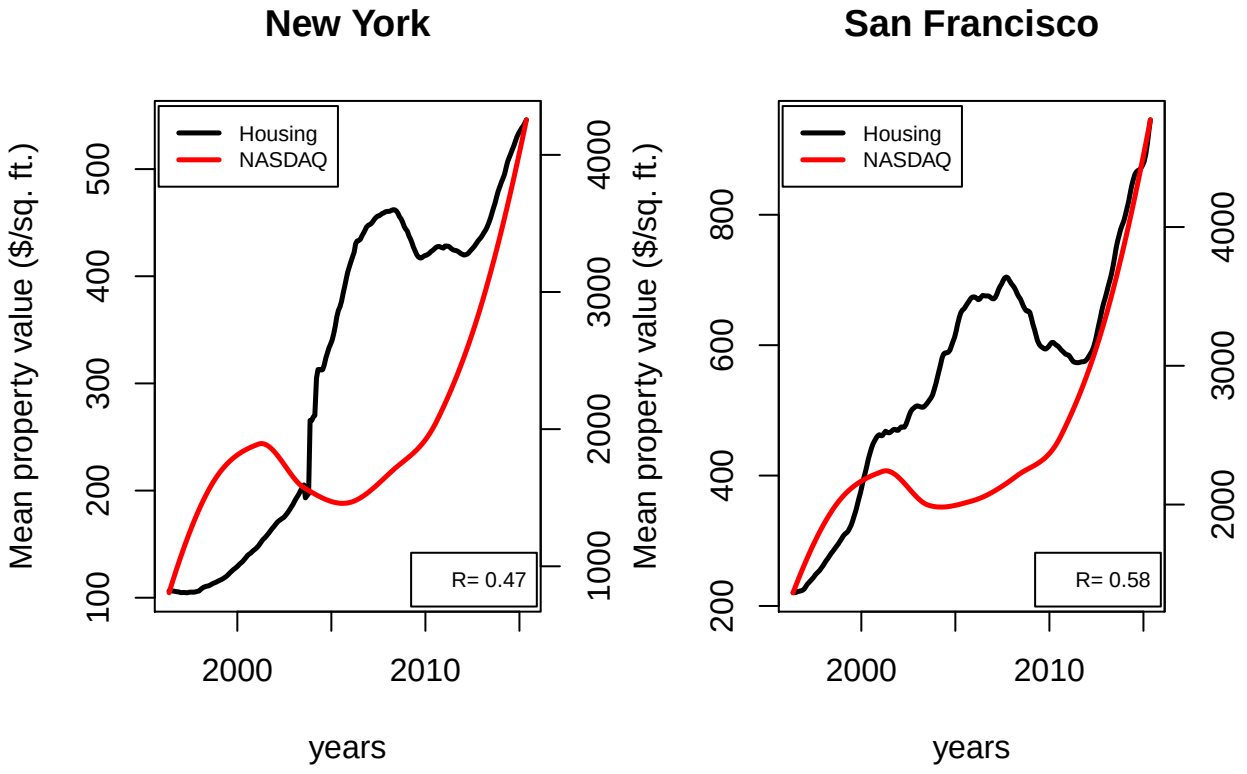
Data Processing

After loading the data for Housing values and the Dow Jones Industrial Average (DJIA), I chose two main cities (New York and San Francisco) to study the correlation between these two datasets.



We observe that the housing and stock markets are clearly interconnected but not exactly in the same way for different cities. The Correlation for SF is higher than the one from NY (0.84 vs 0.79) even though Wall Street is in NY.

Let's do the same comparison with the Nasdaq index.



It was also surprising that NASDAQ, which has more companies from the technology sector, has a lower correlation to the housing market compared to the Dow Jones, despite the fact that the most important companies are close to San Francisco. However, we can see a relative higher correlation for San Francisco than New York.

If we consider all the cities now, we can answer a few questions:

a) which cities are the 5 most/least expensive ones in the U.S.

```
## [1] "Palo Alto"  "Los Altos"  "Menlo Park" "Malibu"     "Stanford"
```

All in California. 4 of them in Silicon Valley.

b) which cities are the 5 least expensive ones in the U.S.

```
## [1] "Shenandoah"        "Town of Salamanca" "Cahokia"
## [4] "East Saint Louis"  "Youngstown"
```

c) which cities are more correlated to Dow Jones?

```
## Mountain View    Eagle Pass     Cupertino    San Carlos     Sunnyvale
##     0.8840694     0.8834987     0.8795367     0.8784956     0.8770255
```

Curiosly, where the headquarters of Google, Apple and Facebook are.

d) which cities are more correlated to NASDAQ?

```
##     Palo Alto     Hohenwald       Purcell Mountain View     Cupertino
##     0.7236257     0.7134041     0.7011296     0.6991091     0.6922888
```

e) which cities are less correlated to Dow Jones?

```
##          Detroit   Town of Beloit      Center Line          Redford
##       -0.4113483       -0.3460930       -0.3104341       -0.2955093
## Garfield Heights
##       -0.2914177
```

f) which cities are less correlated to NASDAQ?

```
##          Detroit   Town of Beloit      Center Line          Redford
##       -0.4113483       -0.3460930       -0.3104341       -0.2955093
## Garfield Heights
##       -0.2914177
```