

# stat133\_group\_project

*Yi Li, Zhang Qin, Zhaoqin Ye*

## Warning: package 'XML' was built under R version 3.2.4

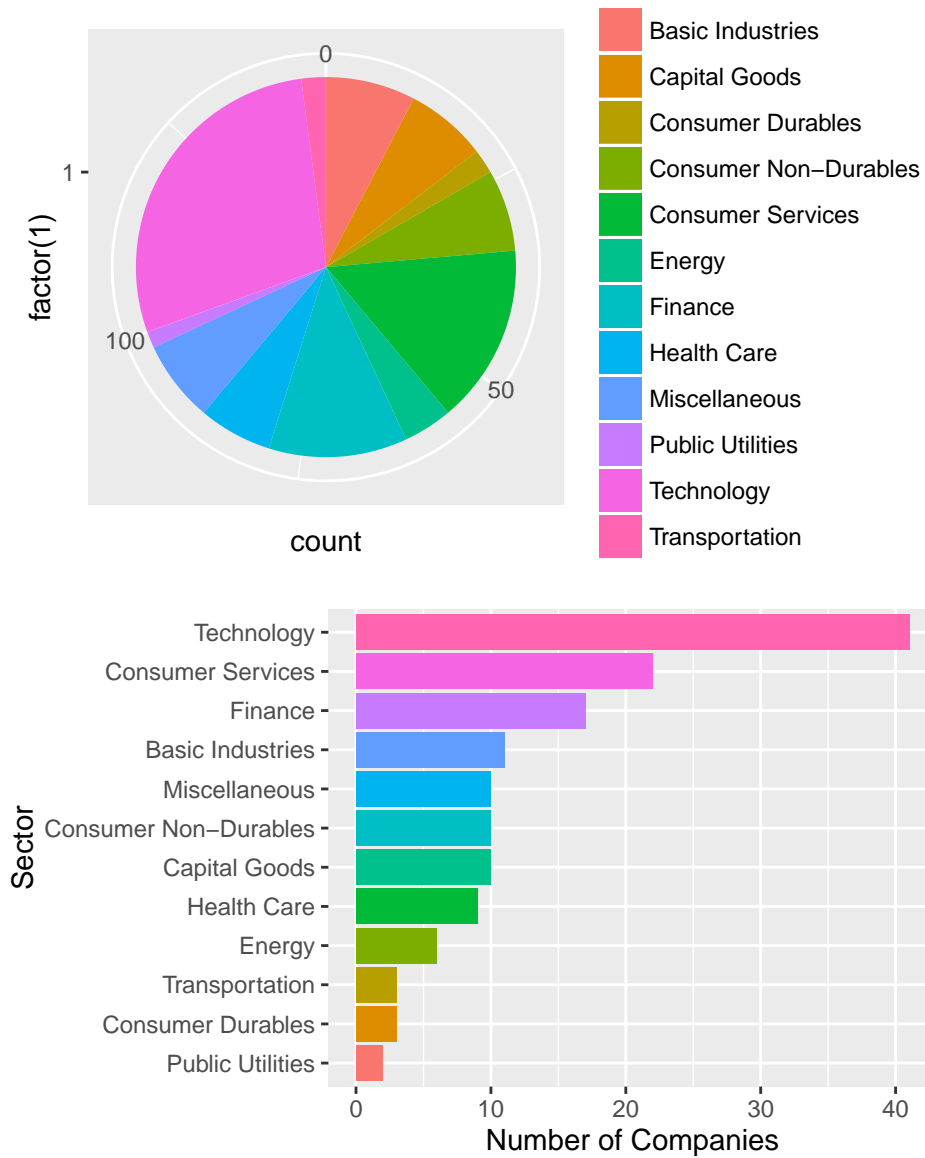
## Analysis of General Information of Chinese Companies in US Stock Market

### 0. Data Preparation

In this section, we get the list of Chinese Company US Stocks `ChinaStockInUS.csv` from NASDAQ official website listing at (<http://www.nasdaq.com/screening/companies-by-region.aspx?region=Asia&country=China>). The website also contains the listing of stocks in major US exchange markets `NASDAQ.csv`, `NYSE.csv`, `AMEX.csv`, downloaded from <http://www.nasdaq.com/screening/company-list.aspx>

### 1. Industry decomposition

| Sector                | Count | Percentage |
|-----------------------|-------|------------|
| Technology            | 41    | 28.47%     |
| Consumer Services     | 22    | 15.28%     |
| Finance               | 17    | 11.81%     |
| Basic Industries      | 11    | 7.64%      |
| Capital Goods         | 10    | 6.94%      |
| Consumer Non-Durables | 10    | 6.94%      |
| Miscellaneous         | 10    | 6.94%      |
| Health Care           | 9     | 6.25%      |
| Energy                | 6     | 4.17%      |
| Consumer Durables     | 3     | 2.08%      |
| Transportation        | 3     | 2.08%      |
| Public Utilities      | 2     | 1.39%      |



- Most in tech

## 2. Exchange Distribution

| Exchange | Count | Percentage |
|----------|-------|------------|
| NASDAQ   | 91    | 63.19%     |
| NYSE     | 49    | 34.03%     |
| AMEX     | 4     | 2.78%      |

- 2/3 in NASDAQ, 1/3 in NYSE

NASDAQ Sector decomposition:

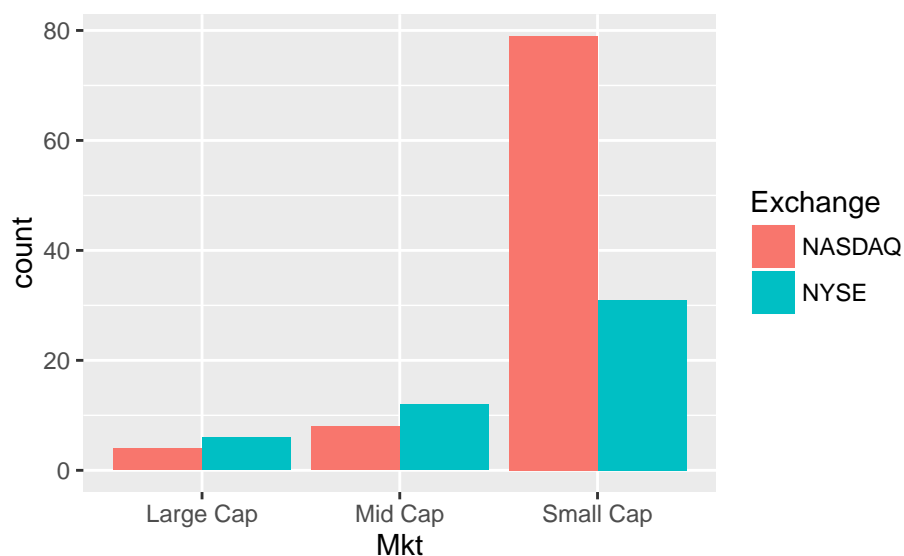
| Sector                | Count | Percentage |
|-----------------------|-------|------------|
| Technology            | 29    | 31.87%     |
| Consumer Services     | 12    | 13.19%     |
| Finance               | 12    | 13.19%     |
| Capital Goods         | 8     | 8.79%      |
| Miscellaneous         | 8     | 8.79%      |
| Basic Industries      | 7     | 7.69%      |
| Consumer Non-Durables | 6     | 6.59%      |
| Health Care           | 5     | 5.49%      |
| Consumer Durables     | 2     | 2.2%       |
| Energy                | 2     | 2.2%       |

NYSE Sector Decomposition:

| Sector                | Count | Percentage |
|-----------------------|-------|------------|
| Technology            | 12    | 24.49%     |
| Consumer Services     | 10    | 20.41%     |
| Finance               | 5     | 10.2%      |
| Basic Industries      | 4     | 8.16%      |
| Energy                | 4     | 8.16%      |
| Consumer Non-Durables | 3     | 6.12%      |
| Transportation        | 3     | 6.12%      |
| Capital Goods         | 2     | 4.08%      |
| Health Care           | 2     | 4.08%      |
| Miscellaneous         | 2     | 4.08%      |
| Public Utilities      | 2     | 4.08%      |

We can see Technology company dominates both NASDAQ and NYSE, and second ranking is both customer service, third Finance...

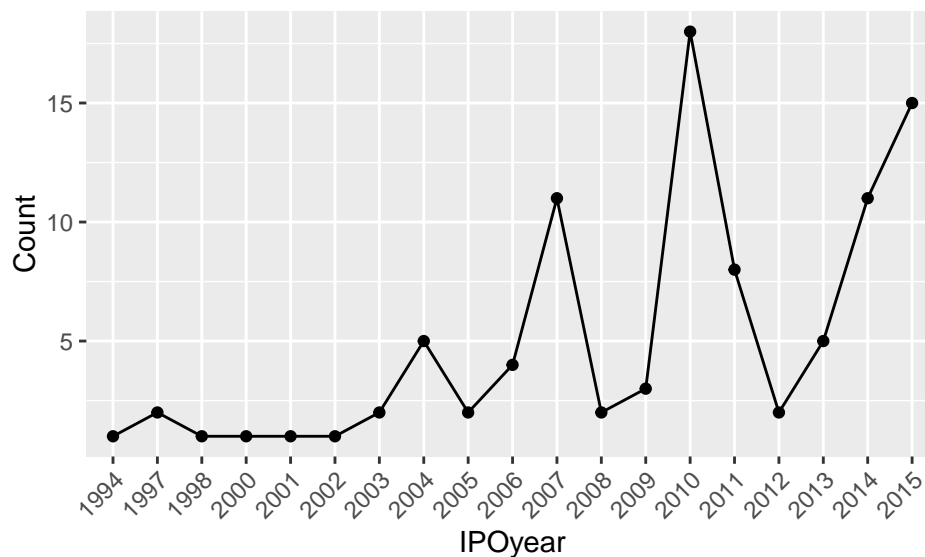
### 3. Market Cap



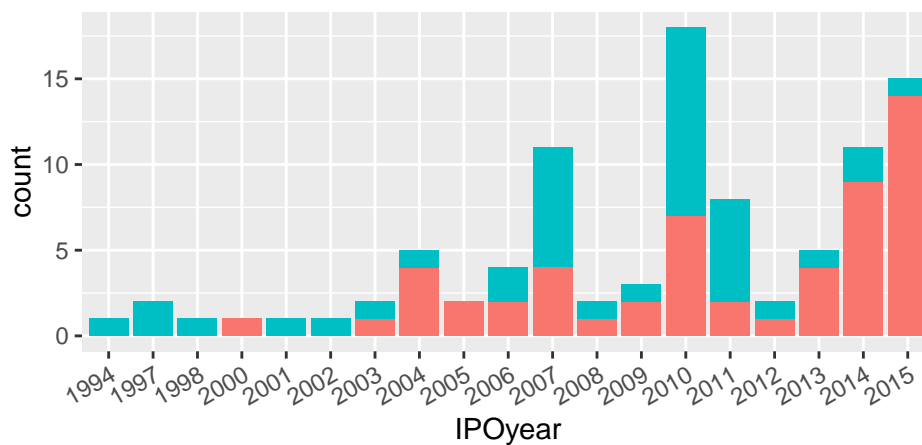
- Average Market cap: \$ 5.47B

- Mostly small cap in NASDAQ
- (The definitions in US and in China are actually different so IDK if this result is meaningful)

#### 4. IPO year



Exchange ■ NASDAQ ■ NYSE



- Peak in 2010
- Upward trend rn
- Used to dominate by NYSE, more NASDAQ in recent years.

For the following analysis, we would only concentrate on the Tech sector as a representative of all other sectors. #Analysis of Chinese Companies' US Stock Performance against US Stock Market

## Data preparation

We download the full history of China technology company US stocks at Yahoo Finance. We also find lists of technology ETFs holding only China tech company US stocks at <http://www.etf.com/sections/features/22712-3-etfs-for-hot-chinese-tech-sector.html?nopaging=1>, and top US Tech ETFs at <http://etfdb.com/etfdb-category/technology-equities/#overview&page=1>, obtaining the history in the same way.

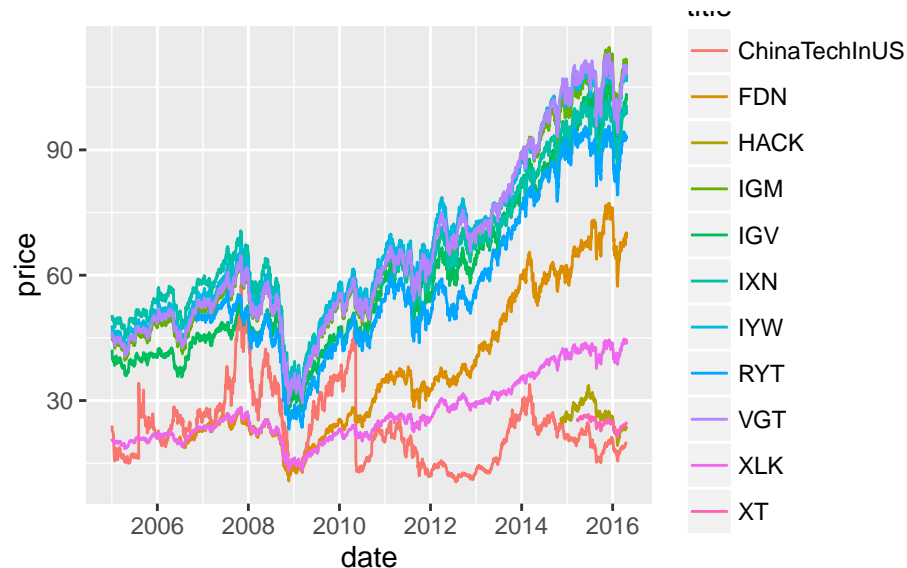
## Plots

Because all history data are obtained in the same format, we can append the all stocks history of China tech Company US stock performance over 2005 to present, group by date and get their average

```
ChinaTechInUS <- getStocksAverageTimeSeries(paste(dataPath, "ChinaTechInUS/", sep=""), "ChinaTechInUS")
```

plot against US top 10 ETF:

```
timeSeriesList <- list(ChinaTechInUS)
USETFPath <- paste(dataPath, "USTechETF/", sep="")
total <- 1
for (etf in list.files(USETFPath)) {
  total <- total + 1
  newTimeSeries <- getStockTimeSeries(paste(USETFPath, etf, sep=""), gsub("(history_|.csv)", "", etf))
  timeSeriesList[[total]] <- newTimeSeries
}
plotMultipleTimeSeries(timeSeriesList)
```

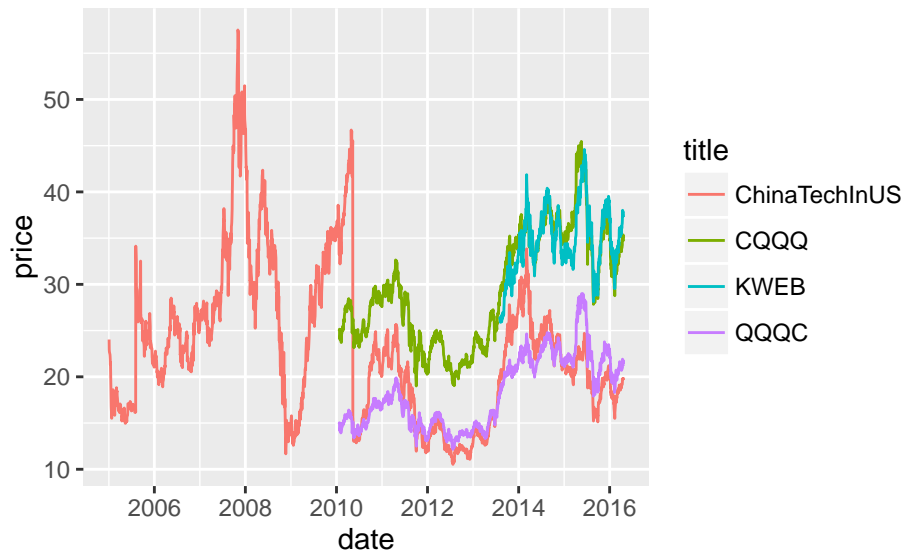


We can see in average all US

tech ETFs outperforms China tech stocks

plot against US ETF holding obly China's US stocks:

```
timeSeriesList <- list(ChinaTechInUS)
USETFPath <- paste(dataPath, "ChinaTechUSETF/", sep="")
total <- 1
for (etf in list.files(USETFPath)) {
  total <- total + 1
  newTimeSeries <- getStockTimeSeries(paste(USETFPath, etf, sep=""), gsub("(history_|.csv)", "", etf))
  timeSeriesList[[total]] <- newTimeSeries
}
plotMultipleTimeSeries(timeSeriesList)
```

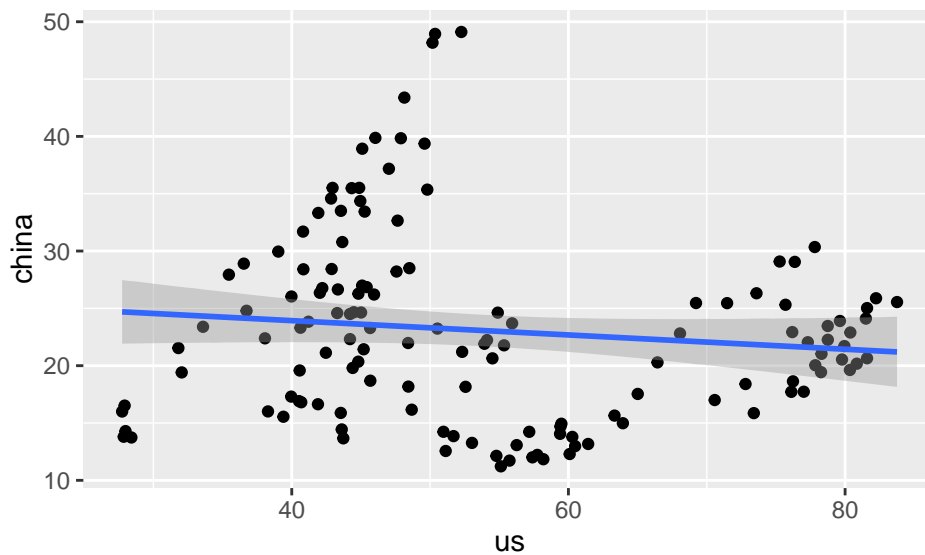


We can see the three ETFs are good representation of general trend of China Company US Stocks average performance, especially QQQC. plot and calculate correlation:

```

USETFAverage <- getStocksAverageTimeSeries(paste(dataPath, "USTechETF/", sep=""), "USETFAverage") %>%
  mutate(year=year(date)+month(date)*0.01) %>%
  group_by(year) %>%
  summarise(price=mean(price))
ChinaTechInUSYearly <- ChinaTechInUS %>%
  mutate(year=year(date)+month(date)*0.01) %>%
  group_by(year) %>%
  summarise(price=mean(price))
USvsChina <- data.frame(us=USETFAverage$price, china=ChinaTechInUSYearly$price)
ggplot(USvsChina, aes(x=us, y=china)) + geom_jitter() + geom_smooth(method="lm")

```



calculated roughly 0.3

Not very strong correlation,

# Analysis of Chinese Companies' US Stock Performance against Chinese Domestic Stock Market

## Data preparation

We find listing of China technology stocks in two subsectors software and hardware at [http://quote.eastmoney.com/center/list.html#28002737\\_0\\_2](http://quote.eastmoney.com/center/list.html#28002737_0_2). Then we obtain history of these stocks in Yahoo Finance in the same way as before.

## Plots

With the same idea to get the average of Software subsector, hardware subsector, and tech sector as a whole:

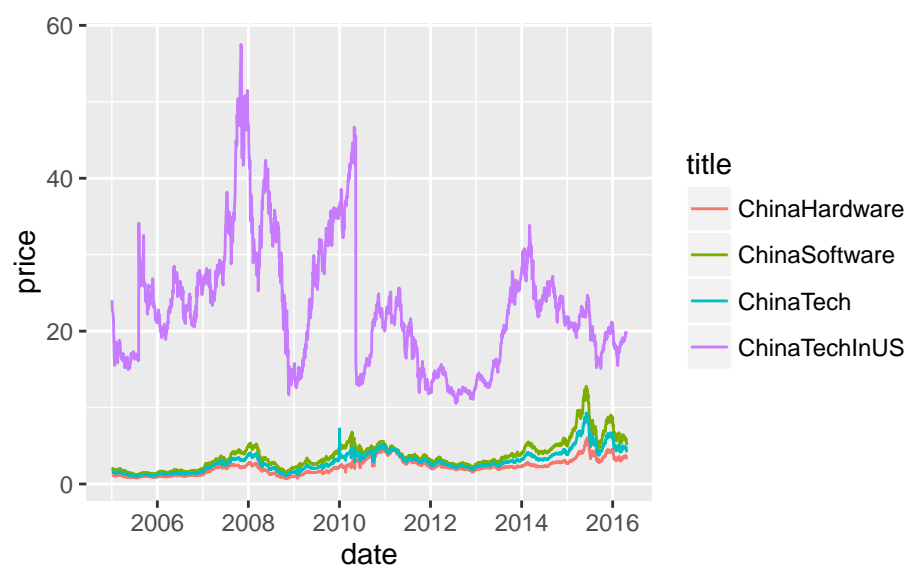
```
ChinaSoftwareRMB <- getStocksAverageTimeSeries(paste(dataPath, "ChinaTech/Software/", sep=""), "ChinaSoft")
ChinaHardwareRMB <- getStocksAverageTimeSeries(paste(dataPath, "ChinaTech/Hardware/", sep=""), "ChinaHard")
ChinaTechRMB <- rbind(ChinaSoftwareRMB, ChinaHardwareRMB) %>%
  group_by(date) %>%
  summarise(price=mean(price))
ChinaTechRMB <- mutate(ChinaTechRMB, title=rep("ChinaTech", nrow(ChinaTechRMB)))
```

Because these data are in RMB, we can convert the price into USD

```
ChinaSoftwareUSD <- mutate(ChinaSoftwareRMB, price=price/6.3)
ChinaHardwareUSD <- mutate(ChinaHardwareRMB, price=price/6.3)
ChinaTechUSD <- mutate(ChinaTechRMB, price=price/6.3)
```

Plot China Company US Stock average against China Domestic performances: \* China stock in USD

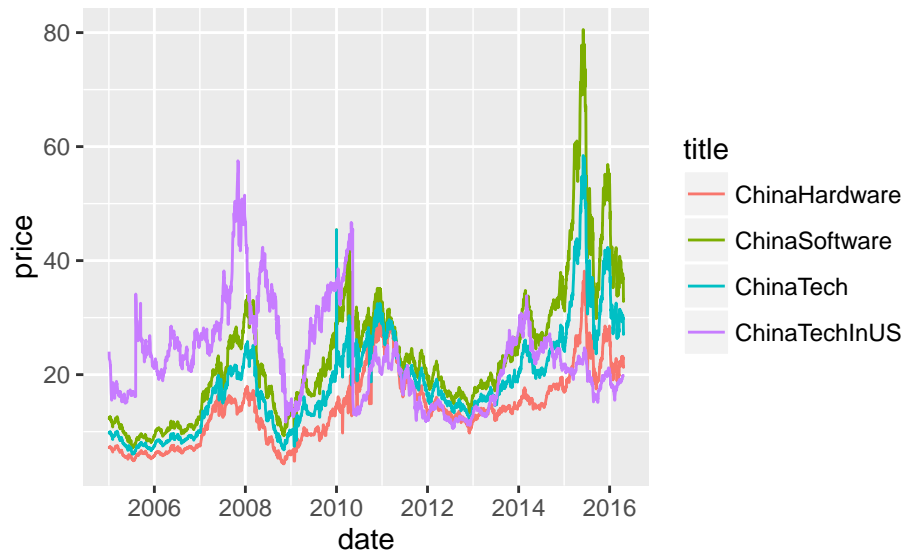
```
timeSeriesList <- list(ChinaTechInUS, ChinaTechUSD, ChinaSoftwareUSD, ChinaHardwareUSD)
plotMultipleTimeSeries(timeSeriesList)
```



We can see China Tech company in US stock market has absolute advantage in total asset value in USD because of the exchange rate

- China stock in RMB

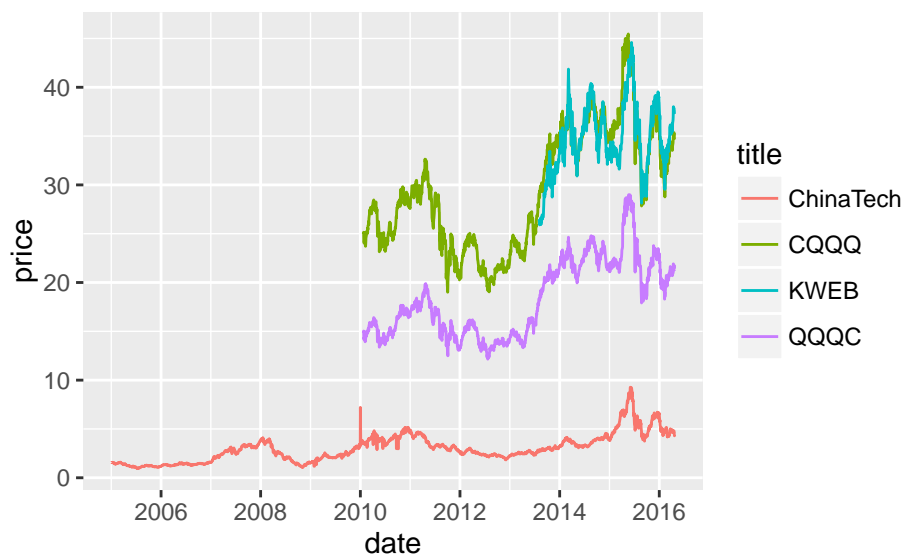
```
timeSeriesList <- list(ChinaTechInUS, ChinaTechRMB, ChinaSoftwareRMB, ChinaHardwareRMB)
plotMultipleTimeSeries(timeSeriesList)
```



We can see comparing the trend actually Domestic and Foreign markets have similar trends before 2015, but during 2015 China domestic stock tends to outperform foreign market because of China's stock boom

Plot China Company US ETF against China Domestic Performance: \* China stock in USD

```
timeSeriesList <- list(ChinaTechUSD)
USETFPath <- paste(dataPath, "ChinaTechUSETF/", sep="")
total <- 1
for (etf in list.files(USETFPath)) {
  total <- total + 1
  newTimeSeries <- getStockTimeSeries(paste(USETFPath, etf, sep=""), gsub("(history_|.csv)", "", etf))
  timeSeriesList[[total]] <- newTimeSeries
}
plotMultipleTimeSeries(timeSeriesList)
```



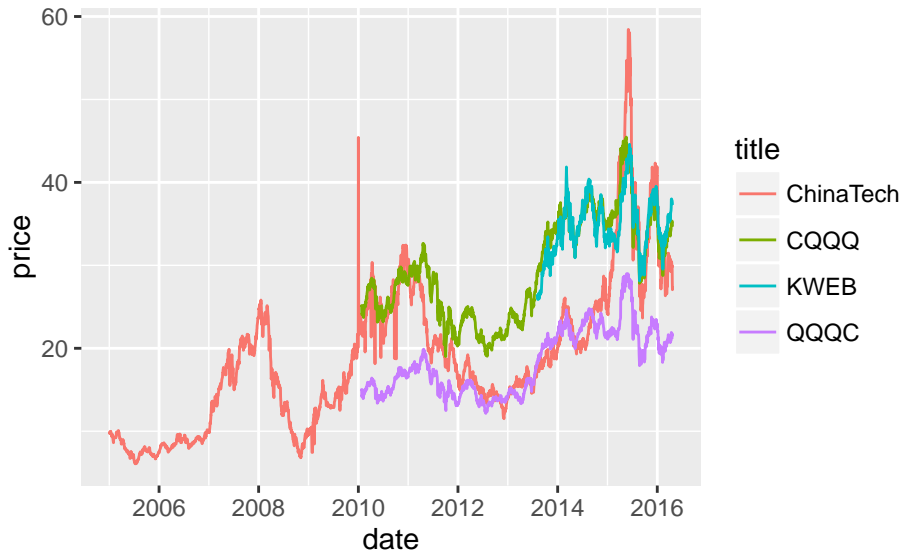
again we can see these foreign



market ETFs have much higher price than domestic average due to the exchange rate

\*China stock in RMB

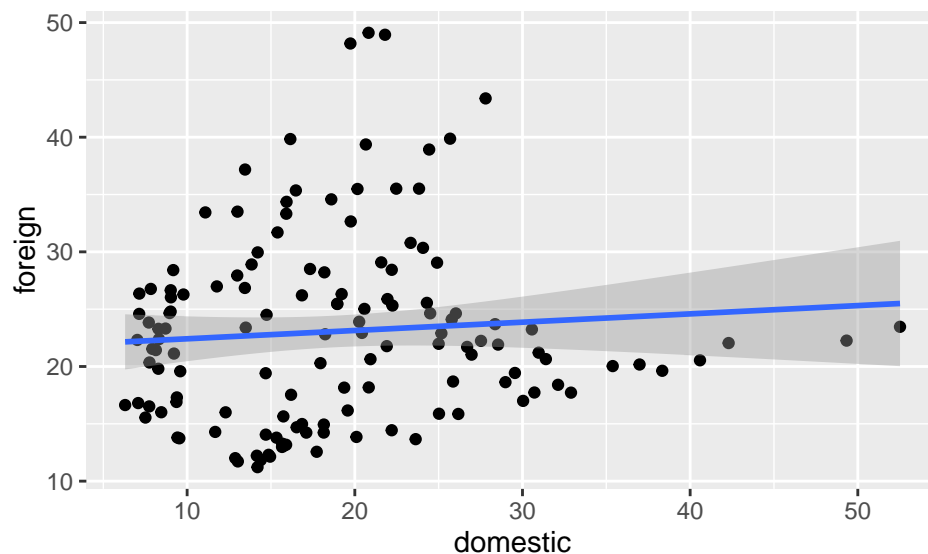
```
timeSeriesList[[1]] <- ChinaTechRMB
plotMultipleTimeSeries(timeSeriesList)
```



However these ETFs follow very similar trend with China domestic average price, implying the China's domestic price is a relatively great driven force for foreign China Company stock price

plot and calculate correlation:

```
ChinaTechDomesticYearly <- ChinaTechRMB %>%
  mutate(year=year(date)+month(date)*0.01) %>%
  group_by(year) %>%
  summarise(price=mean(price))
ForeignVSDomestic <- data.frame(foreign=ChinaTechInUSYearly$price, domestic=ChinaTechDomesticYearly$price)
ggplot(ForeignVSDomestic, aes(x=domestic, y=foreign)) + geom_jitter() + geom_smooth(method="lm")
```



Not very strong correlation