

# What industry do you want to work for?

trying to find out which industry is growing,  
based on PR News

Oct 2018

Jiwon, Cha

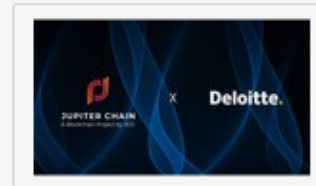
# PR(Public Relations)

- Using the news or business press to carry positive stories about your company or your products; cultivating a good relationship with local press representatives.

18 SEP, 2018, 03:00 BST

## **Jupiter Chain and Deloitte collaborate to deliver secure blockchain-driven data exchange platform for the Southeast Asia market**

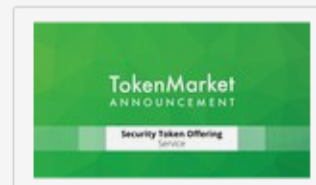
Singapore-based Jupiter Chain and Deloitte announced today a strategic collaboration to implement in Southeast Asia an innovative data exchange...



17 SEP, 2018, 14:35 BST

## **TokenMarket Opens New Security Token Offering Service**

TokenMarket, Gibraltar and London based token sale advisory firm, is pleased to announce the launch of its new Security Token Offering (STO)...



# Basic Idea

- The more, The better
- Growing industries publishes more PR News.



More PR News



Less PR News

# Web scraping

- <http://www.prnewswire.co.uk>

Events Resources Blog Journalists +44 (0)20 7454 5110 Log In Sign Up GDPR

**CISION**  
PR Newswire

News Products Contact Search

News in Focus Business & Money Science & Tech Lifestyle & Health Policy & Public Interest People & Culture

Consumer Technology > Alternative Energies Mining Utilities  
Energy > Chemical Mining & Metals Water Utilities  
Environment > Electrical Utilities Oil & Energy  
Heavy Industry & Manufacturing > Gas Oil and Gas Discoveries  
Telecommunications > View All Energy

### Latest

All Telecommunications News

16:45 BST  
**Arvento Will Present Brand New B2C Products at the Citex 2018**  
One of the biggest players in the field of Fleet Telematics and IoT in the Gulf and the CIS region, Arvento Mobile Systems launches 2 new...

14:51 BST  
**Network Tap/Aggregation Enters the Picosecond Era with ExaLINK Fusion HPT**  
Exablaze today announced the ExaLINK Fusion HPT, a highly anticipated addition to the successful ExaLINK Fusion product series. The Fusion HPT...

12:48 BST  
**Millicom Announces the Pricing of \$500 Million Senior Notes**  
Millicom International Cellular S.A. ("Millicom") (Stockholmsbörsen: MIC) announces the pricing of its \$500 million 6.625% senior notes due 2026...

10:22 BST  
news/alternative-energies-list/ Developer Enablement Program

### Most Viewed

All Telecommunications News

03 OCT, 2018, 20:05 BST  
**Abra Simplifies Cryptocurrency Investing With New Crypto Index Accessible To Everyone, Everywhere**

18 SEP, 2018, 01:30 BST  
**Pareteum Expands to Asia**

02 OCT, 2018, 16:00 BST  
**Swedish Game Studio Bublar Launches First Mobile Sports Game to Leverage Apple's Motion Tracking Technology**

25 SEP, 2018, 14:00 BST  
**Grubhub Announces Acquisition of Tapingo**

26 SEP, 2018, 12:48 BST

177 Industries

95459 news

From April to October – 2018

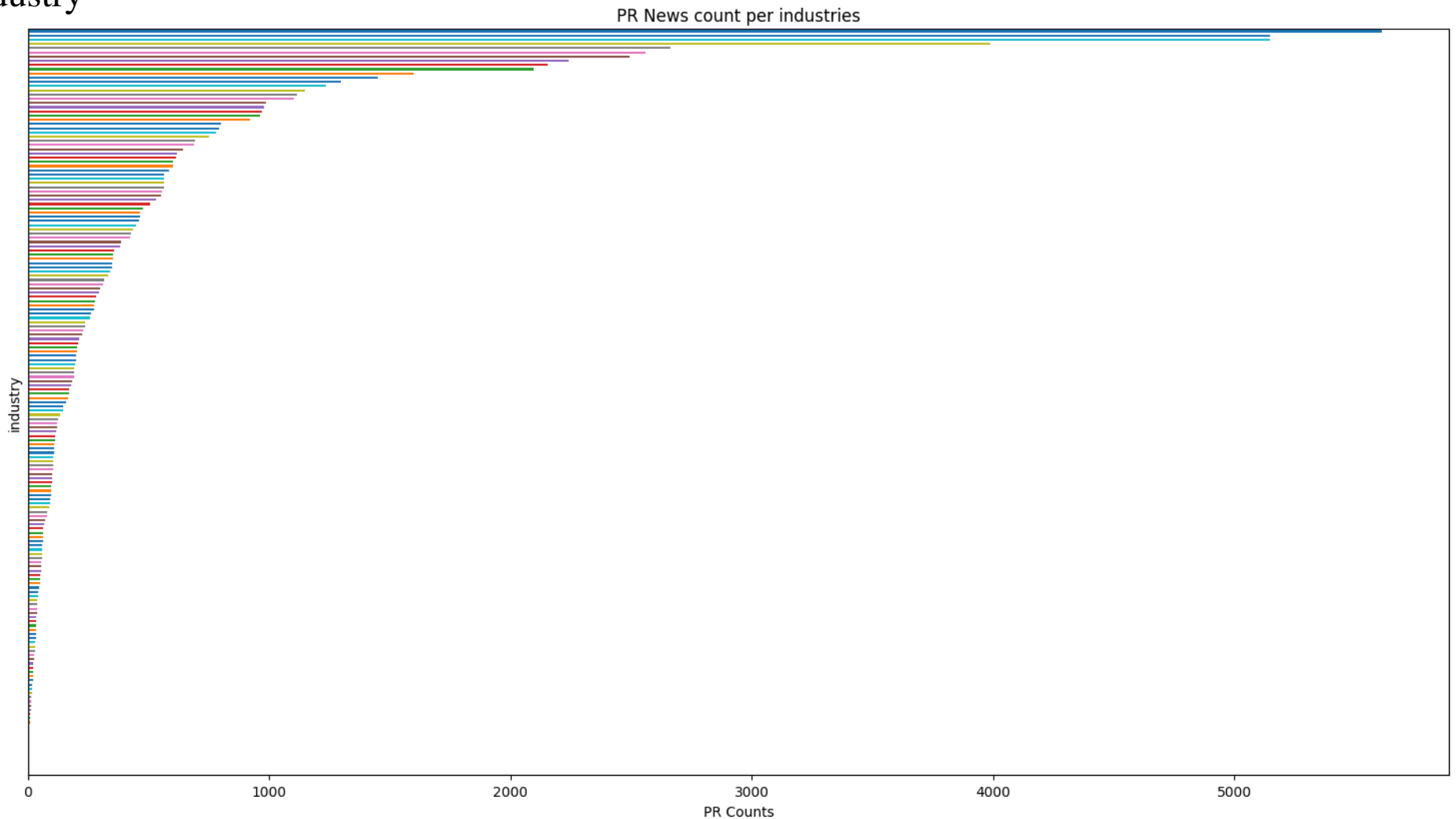
34Mb (News Title only)

Industry	News Title	Release Date
Blockchain	Millicom Announces...	10/11/2018 13:00:00
Chemical	...	...
Blockchain	...	...
Mining	...	...

# PR News Count

```
df_group = df.groupby('level3').size().sort_values(ascending=True)
```

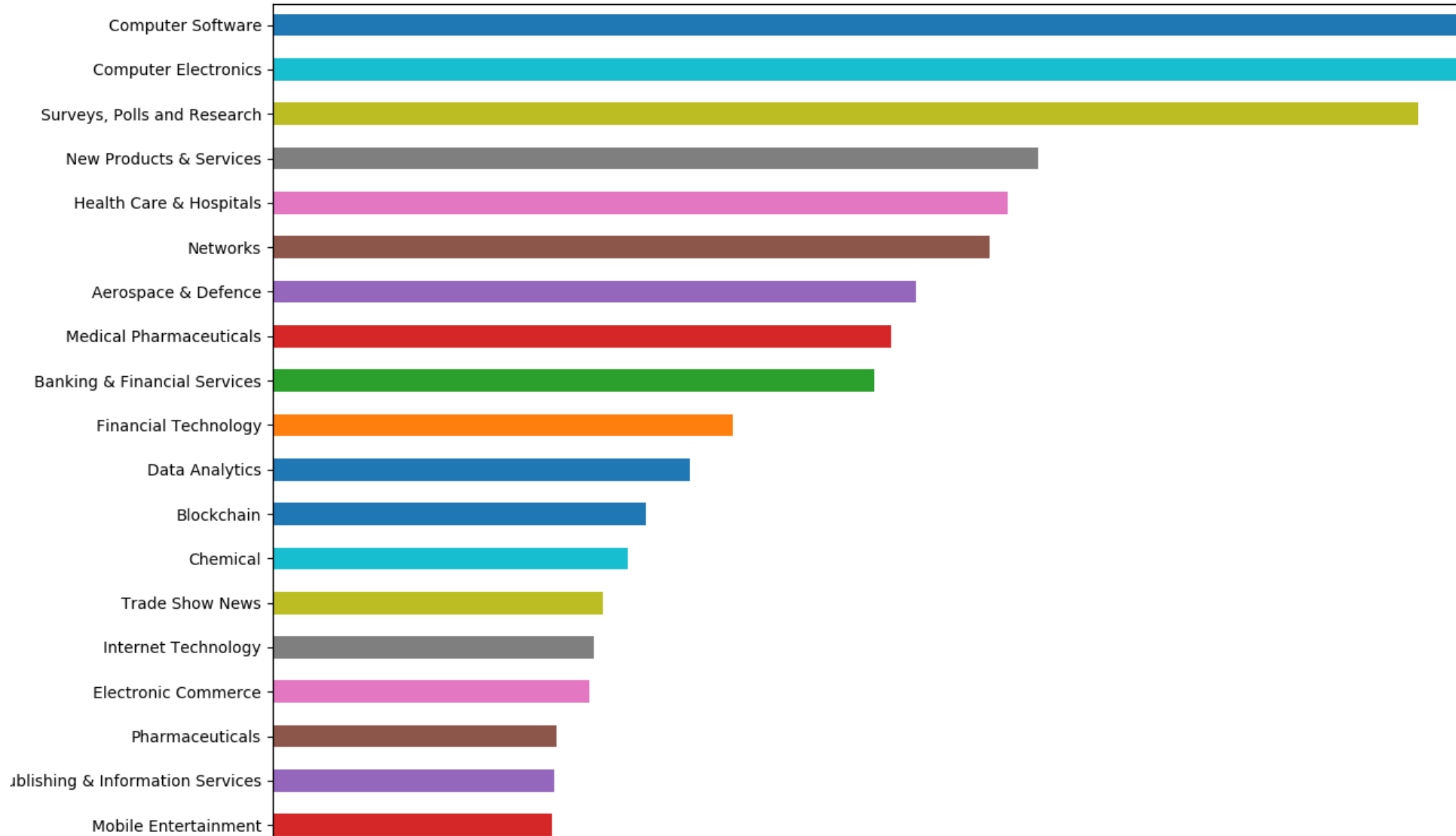
Industry



Number of PR

# Top

PR News count per industries



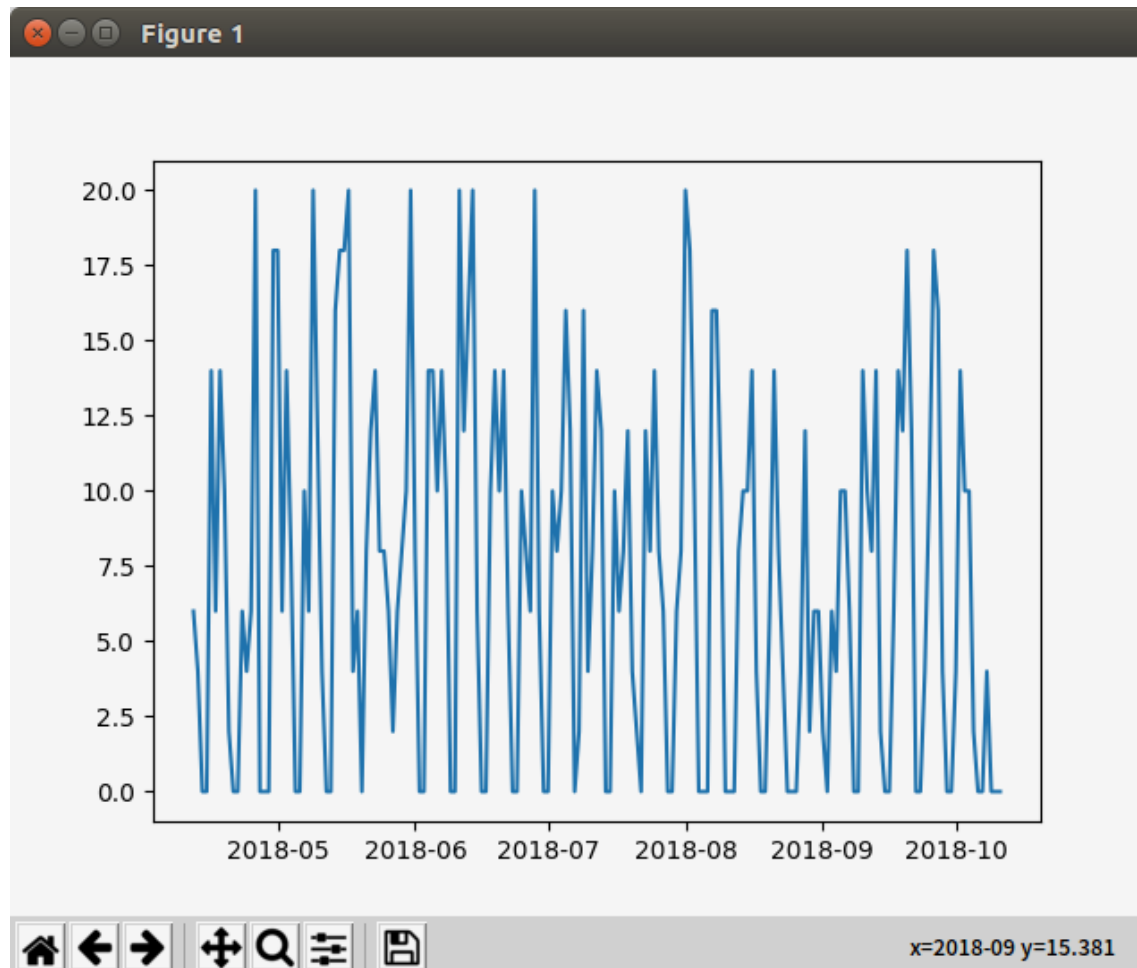
# Bottom

PR News count per industries



# Lets find trend for each industry.

```
time_df = df.groupby(['release_date', 'level3']).size().unstack().fillna(0)  
time_df = time_df.resample('D').sum()
```



Plot for Blockchain Industry



# Nature of PR news...

```
df['day_of_week'] = df['release_date'].dt.weekday_name  
df.groupby('day_of_week').size()
```

"""

day\_of\_week

Friday 9370

Monday 14824

Saturday 590

Sunday 711

Thursday 17890

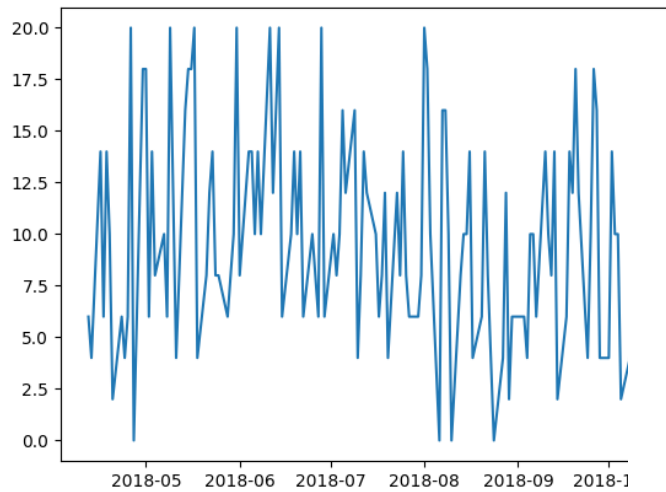
Tuesday 18448

Wednesday 17126

"""

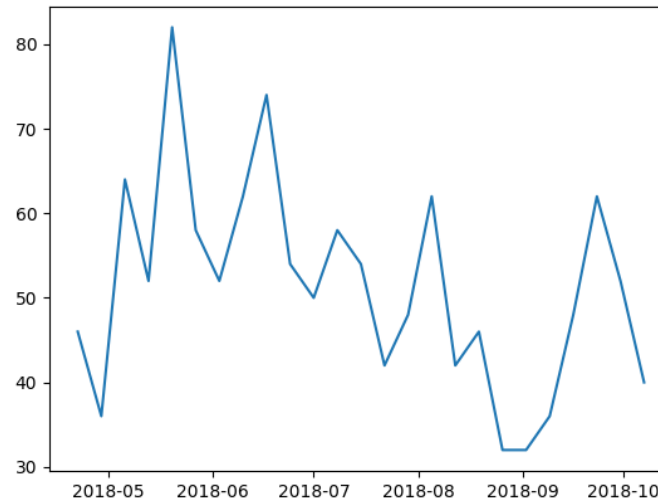
# Can you see the trend?

```
time_df[time_df.index.dayofweek < 5]
```



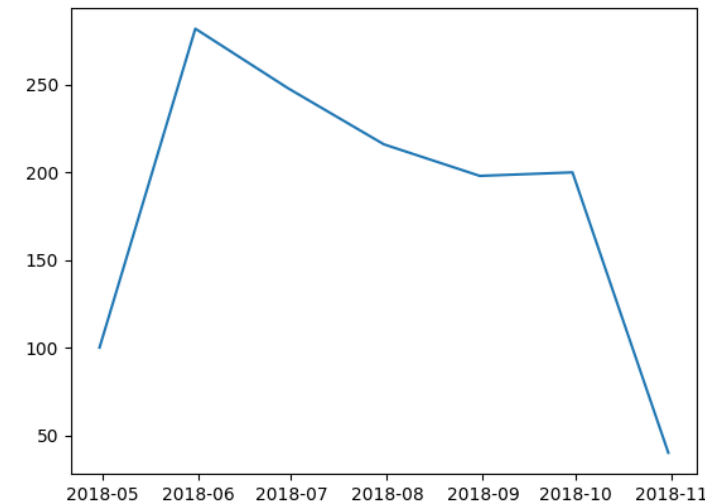
Removed weekends

```
time_df.resample('W').sum()
```



Resampled (weekly)

```
time_df.resample('M').sum()
```

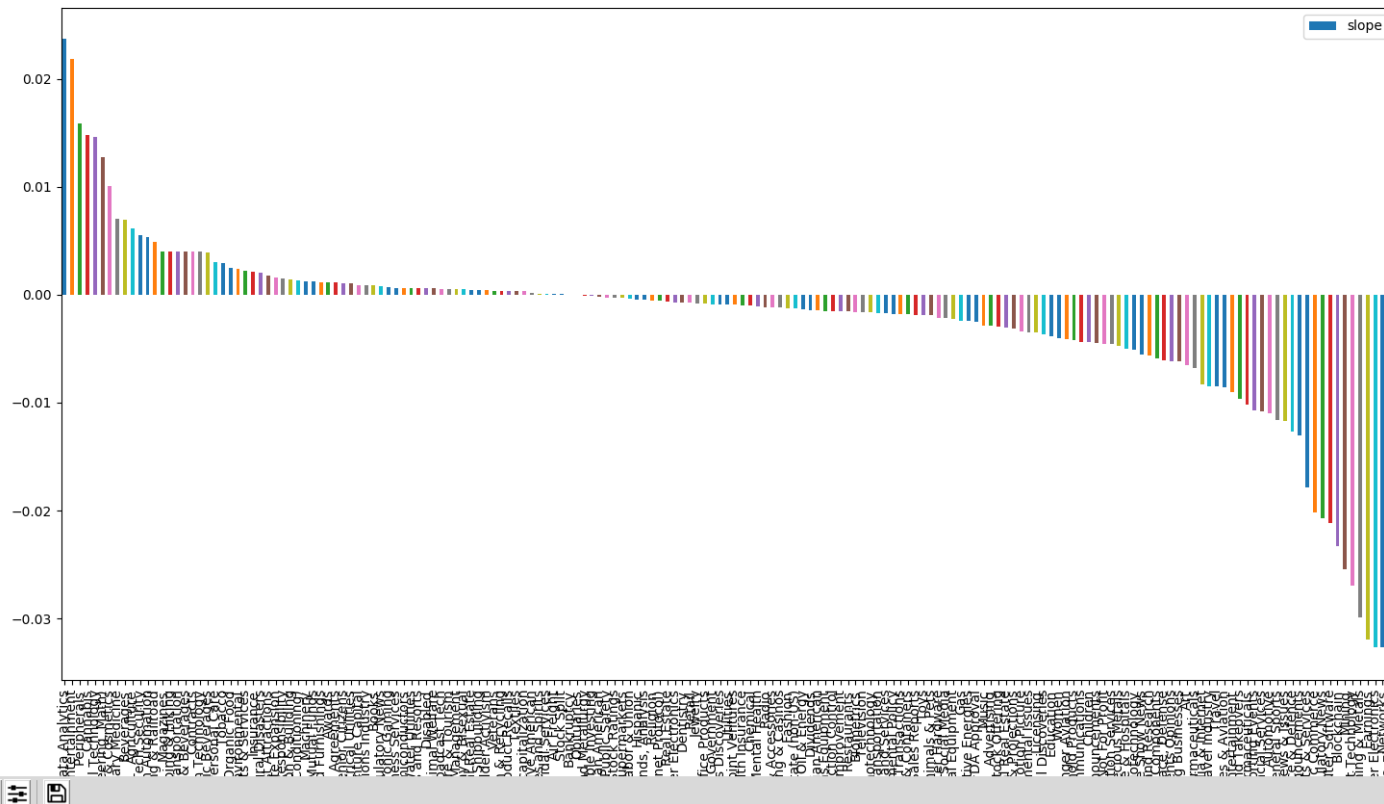


Resampled (Monthly)

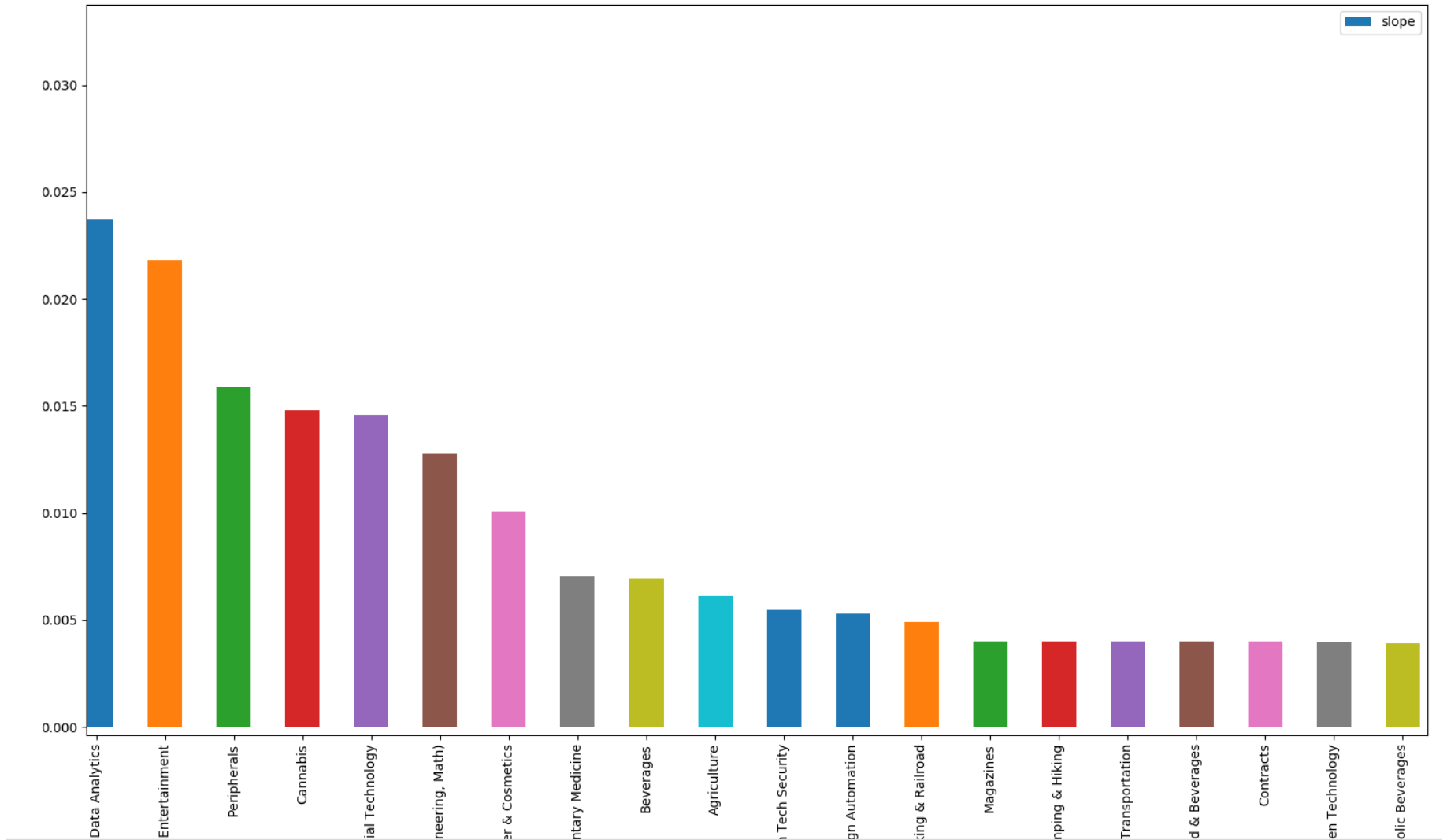
# PR News count trend per industries

```
def trend_line(data, order=1):  
    coeffs = np.polyfit(range(len(data.index)), data, order)  
    slope = coeffs[0]  
    return float(slope)  
  
slope_df = pd.DataFrame(time_df.groupby('level3').apply(lambda x: trend_line(x.pr_count)))
```

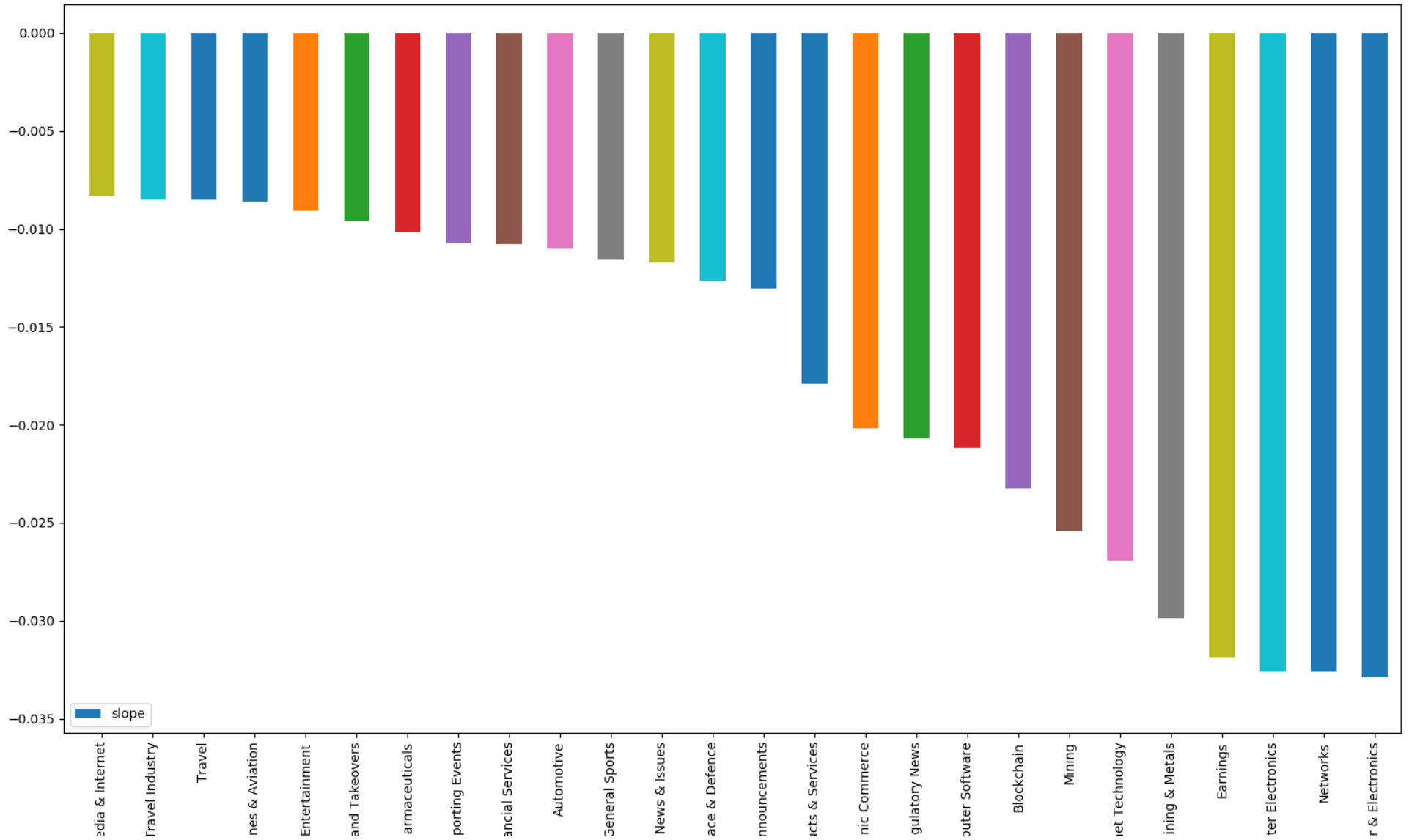
Figure 1



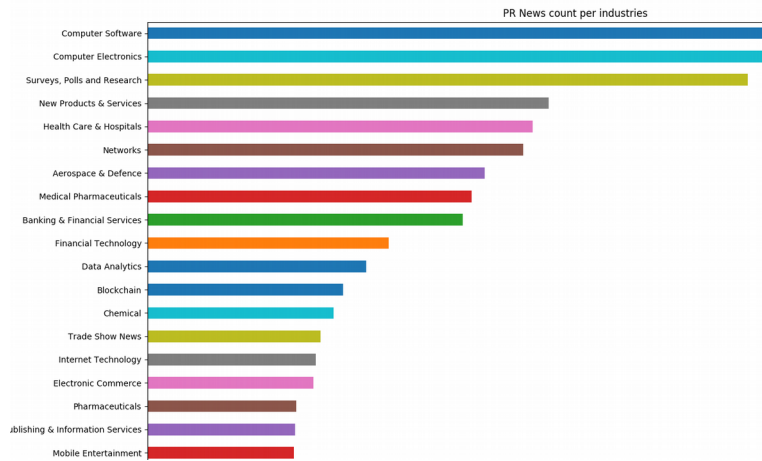
# Top



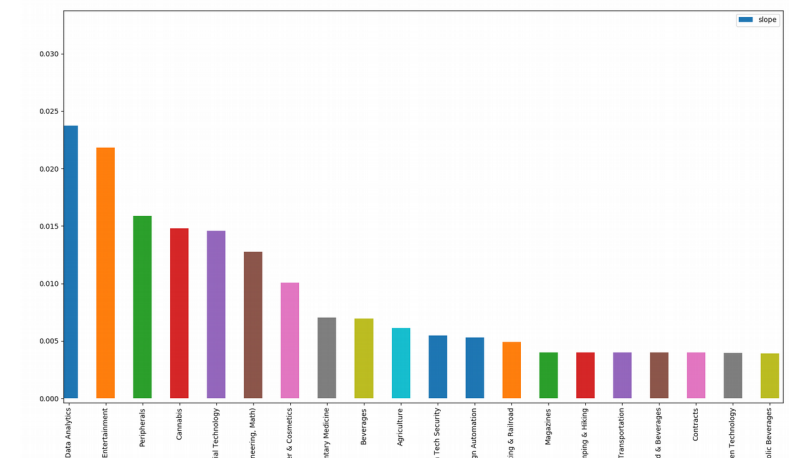
# Bottom



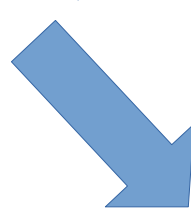
# We need to see 2 variables at the same time



Industry against PR Count



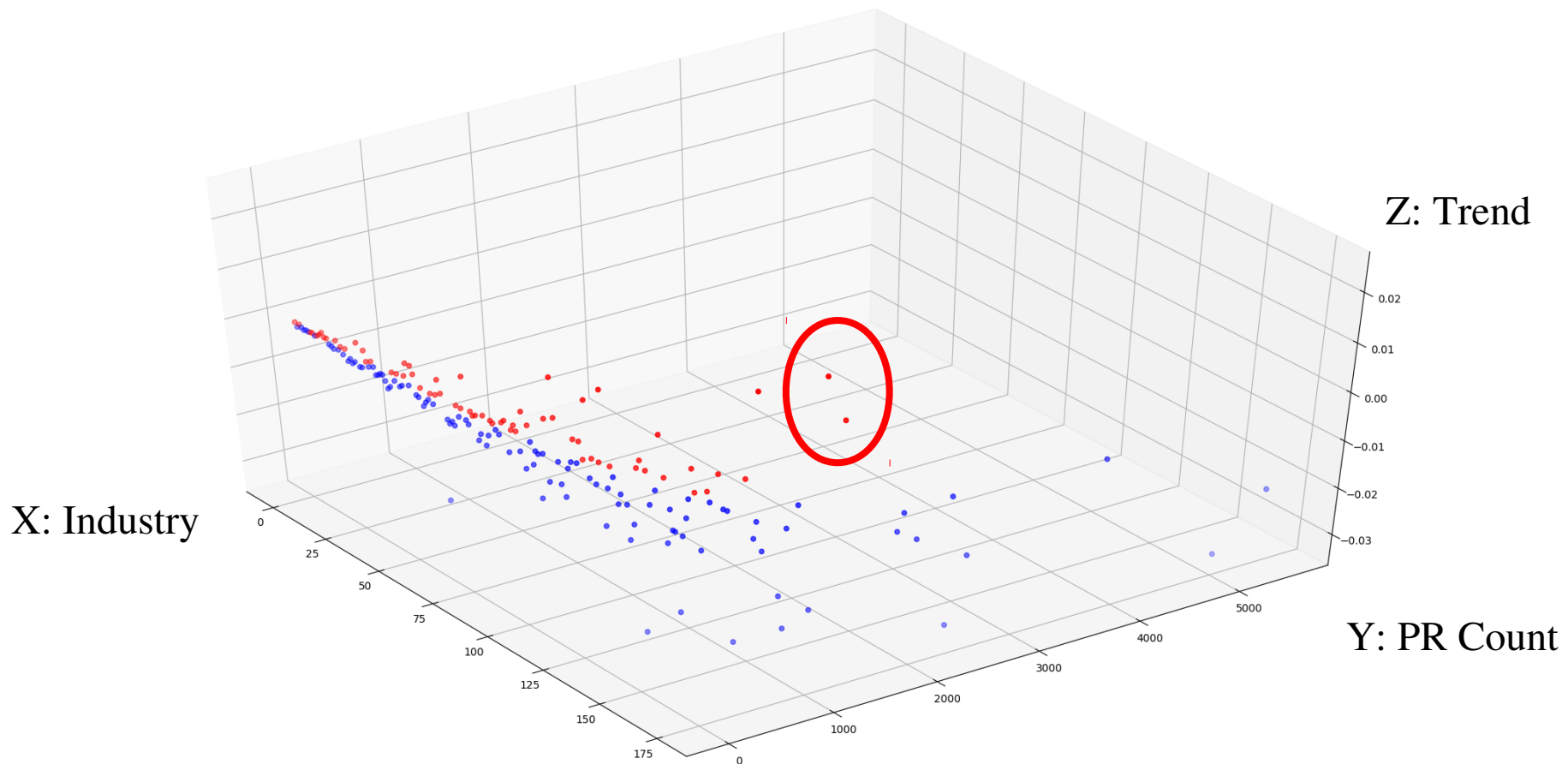
Industry against trend



## 3D Plot

# So, What industries have the most promising future?

- High rank in PR\_Count + Upward slope = good industry



# What industry do you want to work for?

High rank in PR count, Upward trend

- Data Analytics
- Financial Technology



Good!!!

High rank in PR count, but, Downward trend

- Blockchain
- Networks
- Computer Electronics



Watch out!!!



# Future Improvements

- Implementing DB storage for daily batch
- NLP for categorization
-