



CTBC Project Report

Jessie Chen



投資因子分析

QUESTION DEFINITION

- 找出不同環境下，專業投資人的投資因子
- 根據結果、建一投資組合，並觀察追蹤其報酬率

Bill Ackman

PERSHING SQUARE CAPITAL
MANAGEMENT



Datasets



CNBC

- 美股新聞 (2013-now)
- 免費

ROCKET FINANCIAL

- 美股財報
- 資料齊全、免費

DATAROMA

- 機構投資人每季持股變化

Pre-processing

共現圖

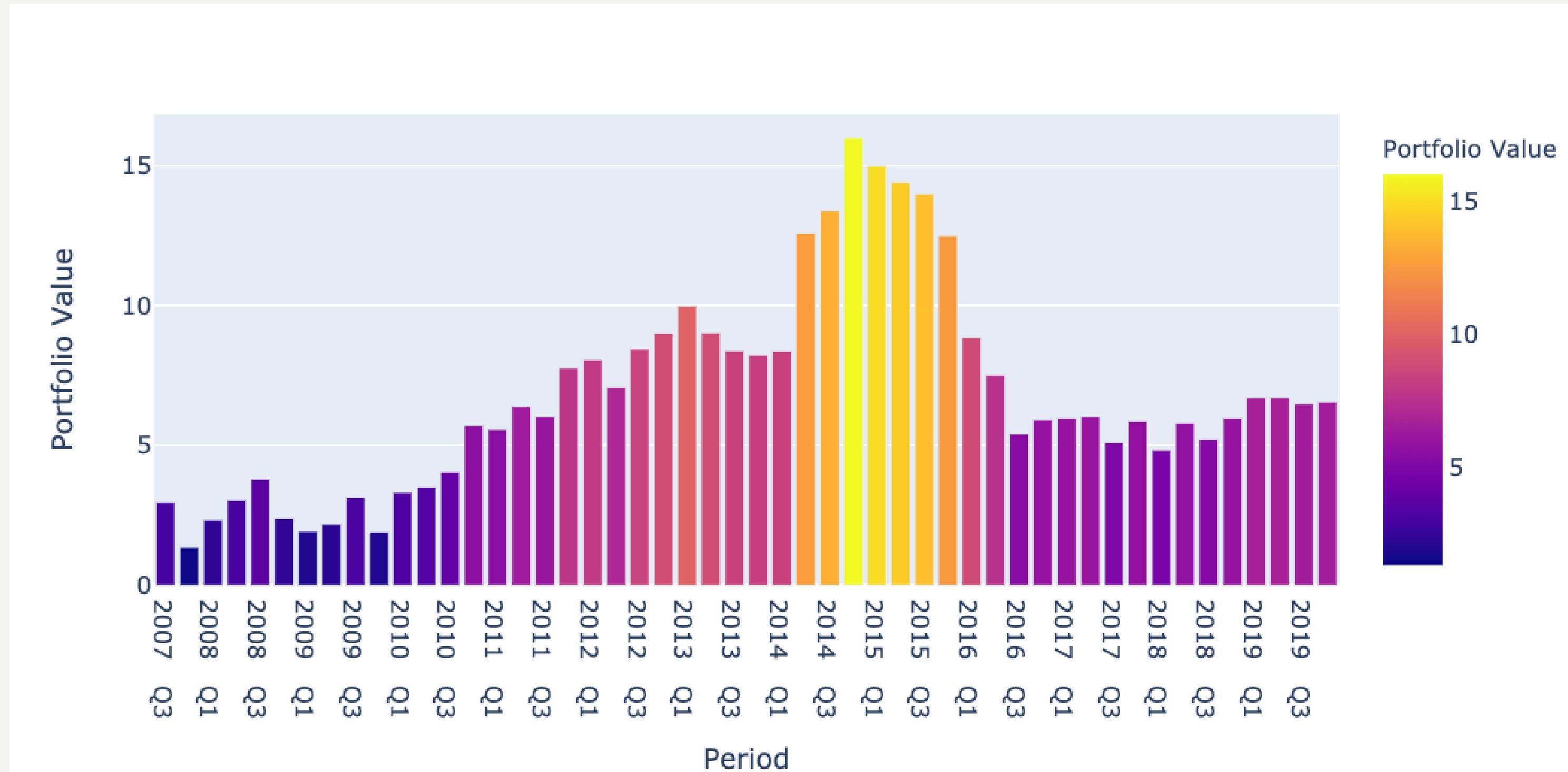
1. 新聞資料
2. 日期切割成年月日
3. 刪除同stop words、標點符號
4. 詞性分類
--> 計算關鍵詞出現次數

投資組合迴歸

1. 財報留下需要的欄位
2. 計算個股投資因子
3. 計算投組投資因子
--> 找投組投資因子迴歸式



PORTFOLIO VALUE



TOP 20 holding history

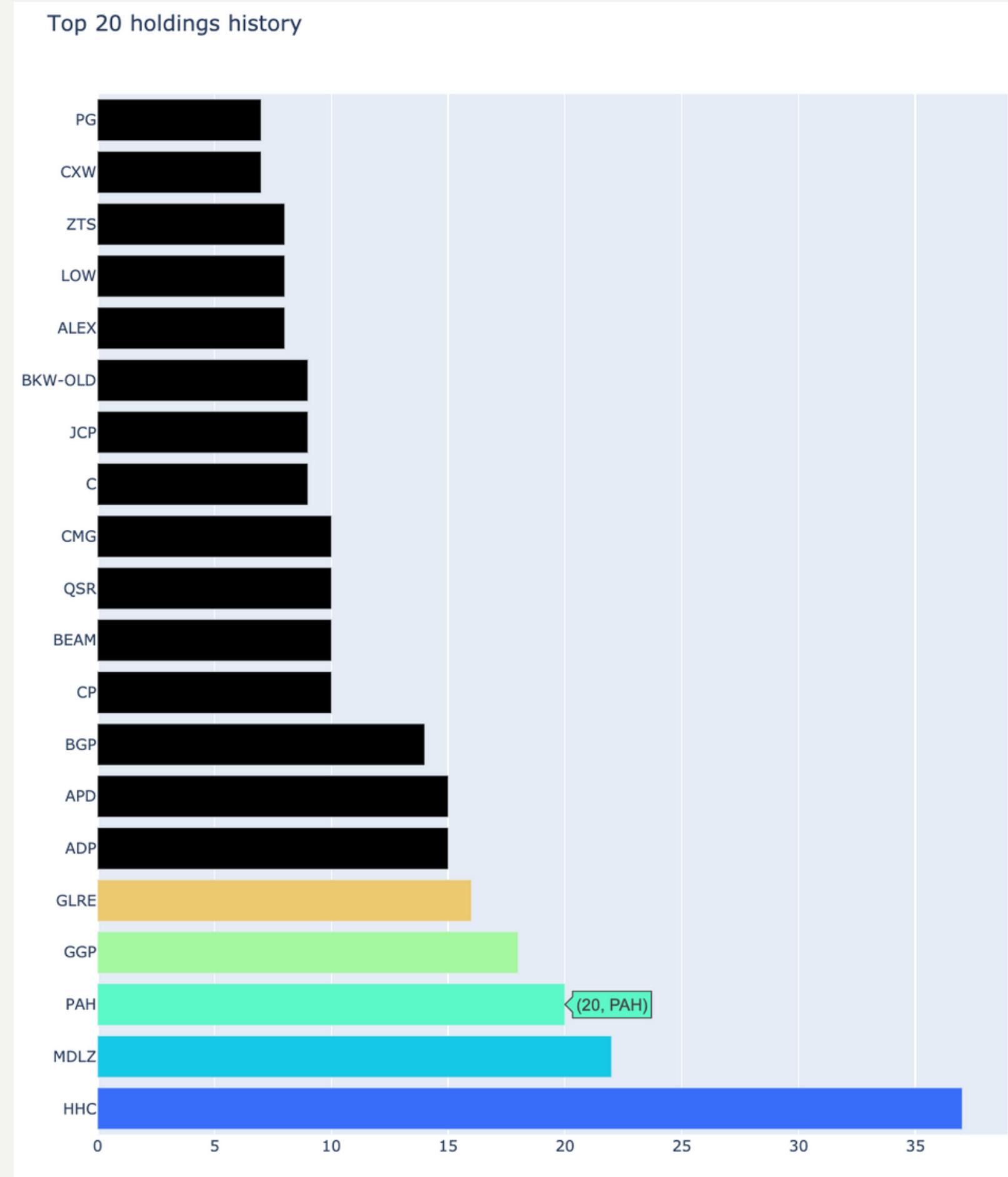
1. HHC

2. MDLZ

3. PAH

4. GGP

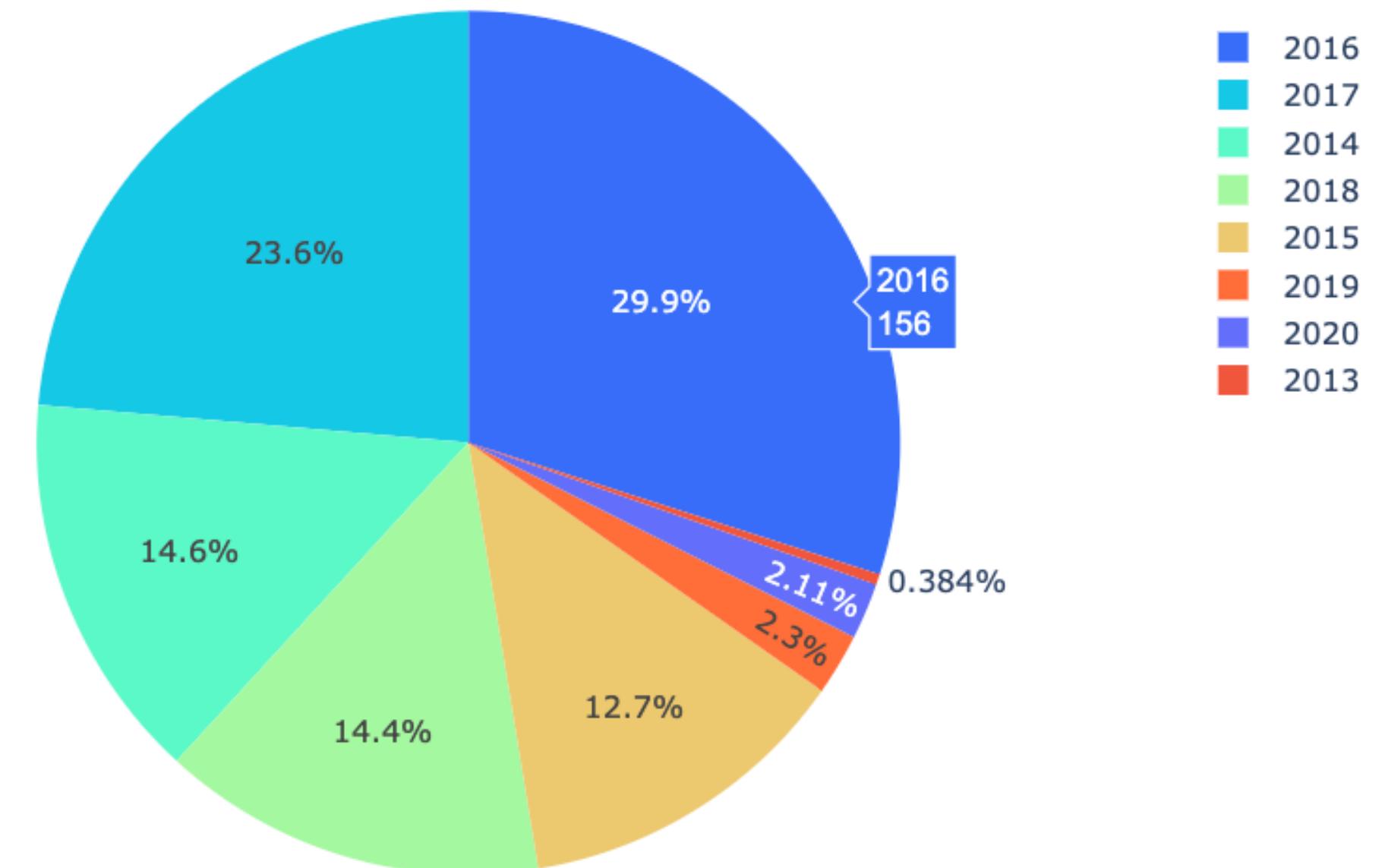
5. GLRE



NEWS

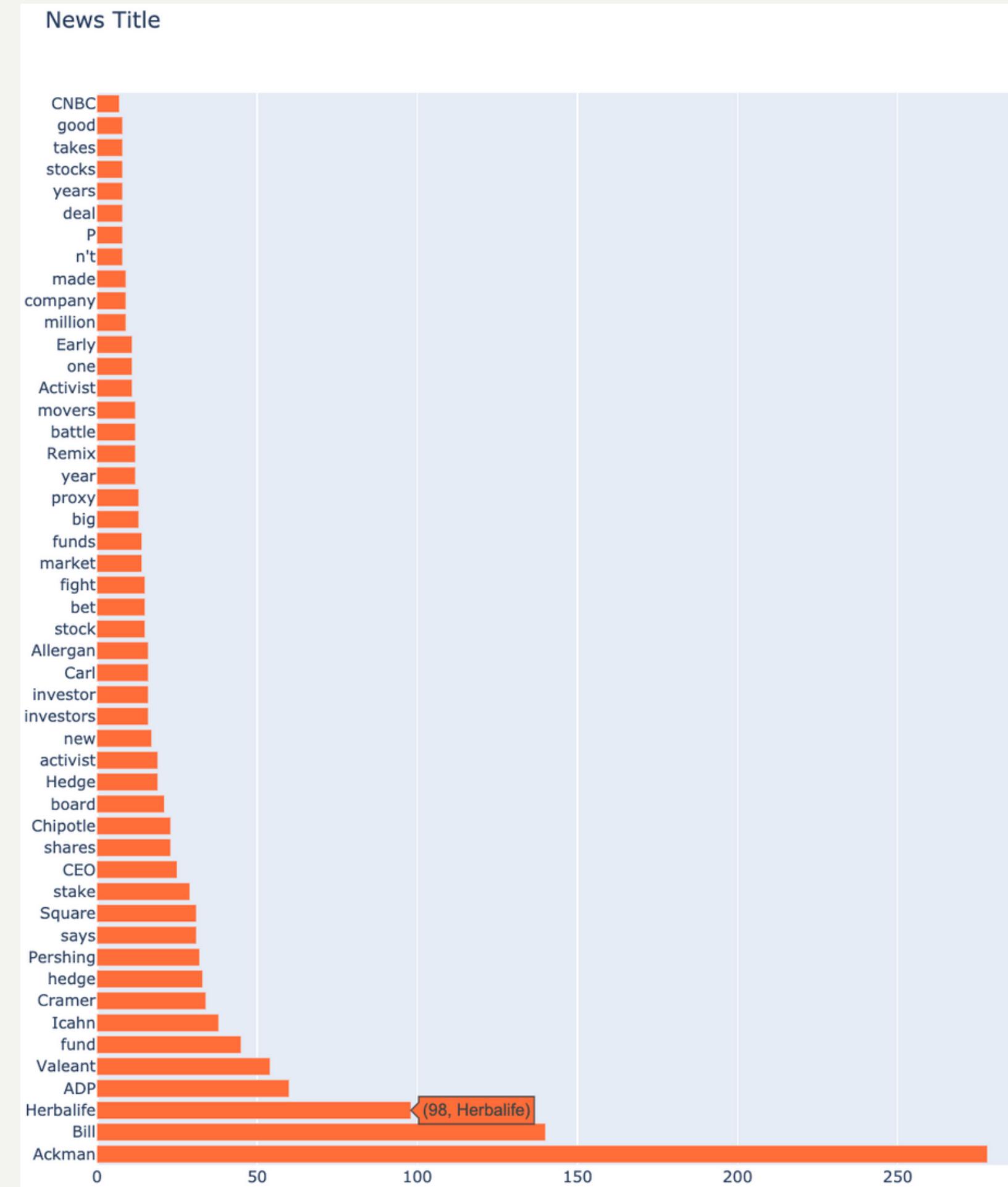
2013-2016

News Amount each year



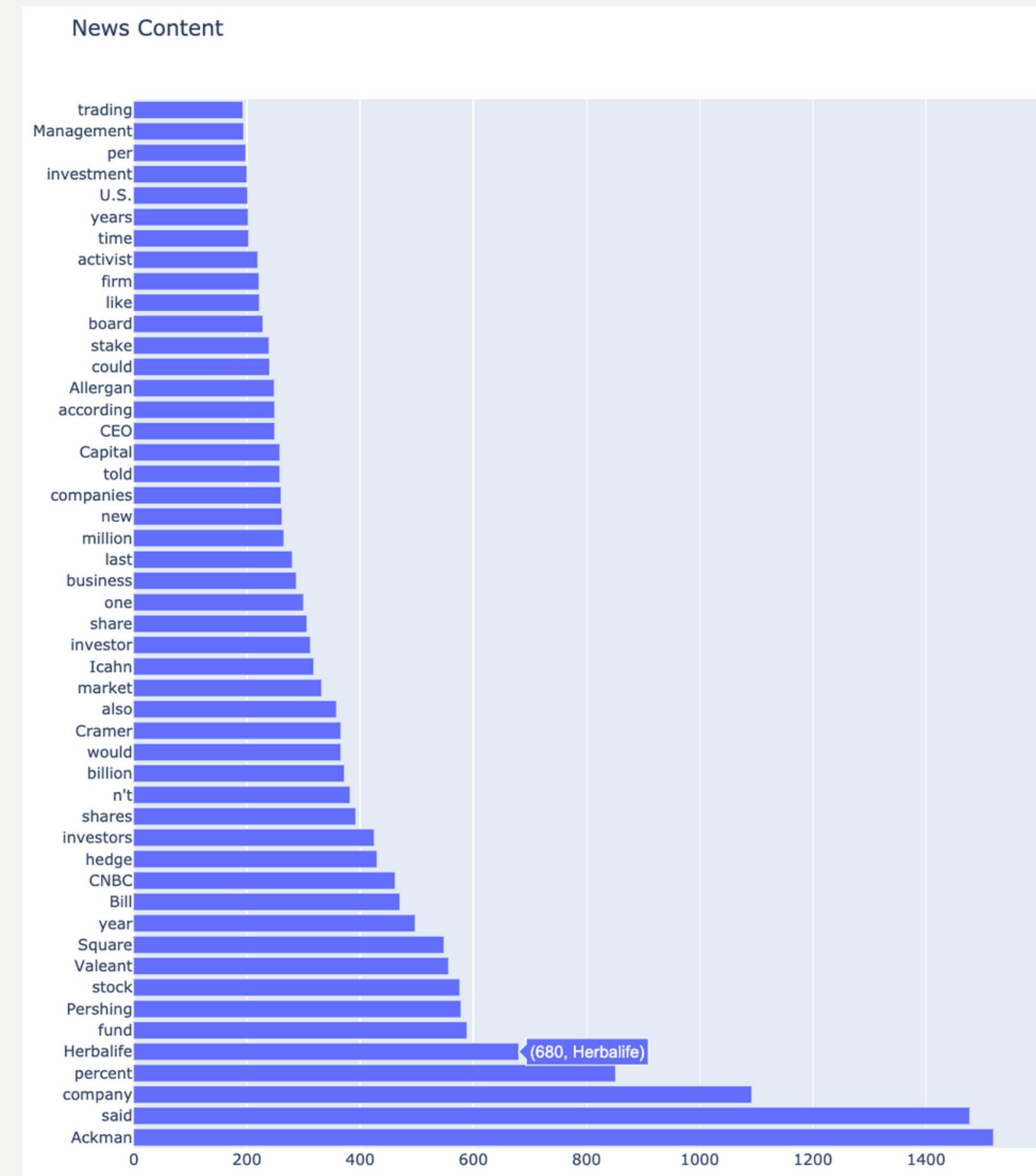
NEWS TITLE

2013-2016



NEWS CONTENT

2013-2016





How to solve problem ?

1. 定義問題
2. 討論、達成共識
3. 查資料
4. 請教教授、專家

OH! WE NEED HELP!

Job Assignment

每週固定討論

1. 每週六晚上18:00
2. 台大社科院研究室

會議結束前總結

1. 總結所有項目目前進行的程度
2. 列出下週須完成的項目

分配工作

按照專業領域分工

問題討論

- 週間進行的過程遇到任何問題：
1. 臨時開會
 2. 互相幫忙

Trello

利用trello追蹤每項
工作進度

中信 第二題 投資因子分析



金融科技 Free

團隊觀看權限



邀請

Butler

顯示選單

方向

1. data來源: 除了yahoo finance，也可以參考quandl，在選擇研究對象的時候要留意一下可取得的data有哪些，以免沒有data可以training
2. 避免確認偏誤: 選擇研究對象或是使用的方法時，可以先初步列舉幾個評估，減少心理上的定錨狀態
3. validation: 要保留一些data來做validation，以確認自己的model是否真的可以解釋該投資名人的投資風格(e.g., r² square for regression)，以此主題為例的話，應該是用過去的data做training，時間較後面的data做validation
4. 時空背景: 分析時需考慮投資名人所在的時空背景，該投資風格都需搭配當時的時空背景一起分析
5. 依據分析出來的投資因子，是否能建立該投資名人風格的porfolio，該porfolio在近期的績效如何? (例：在近期肺炎影響的績效) 或是應該適用怎樣的市場情況?

會議紀錄

完整內容在 [👉 Google Document](#)

04/08 (三)



04/12 (日)



04/18 (六)



4/20 (一)



+ 新增另一張卡片



+ 新增另一張卡片

待討論事項

Q1：放空的話，占比權重怎麼看？

Q2：投資組合的因子回歸式怎麼算
(加權平均?)

Q3：如果我們先算出名人的個股因子，假設發現某個因子在個股皆具有顯著，然而接著再組成屬於我們的投資組合時，再去跑分析時，發現這個因子呈現不顯著，α卻顯著該怎麼辦？

+ 新增另一張卡片

進行中的工作

Bill Ackman

0/6



爬股價歷史報酬



找迴歸計算

爬取歷年投資人新聞

爬取歷年總經新聞

+ 新增另一張卡片



已完成的項目

決定 題目要做的基金投資人

準備想做的知名投資人資料

⌚ 4月12日 ⏳ 1

論文看完，確認題目內容方向

⌚ 4月12日 ⏳

寄email 給mentor

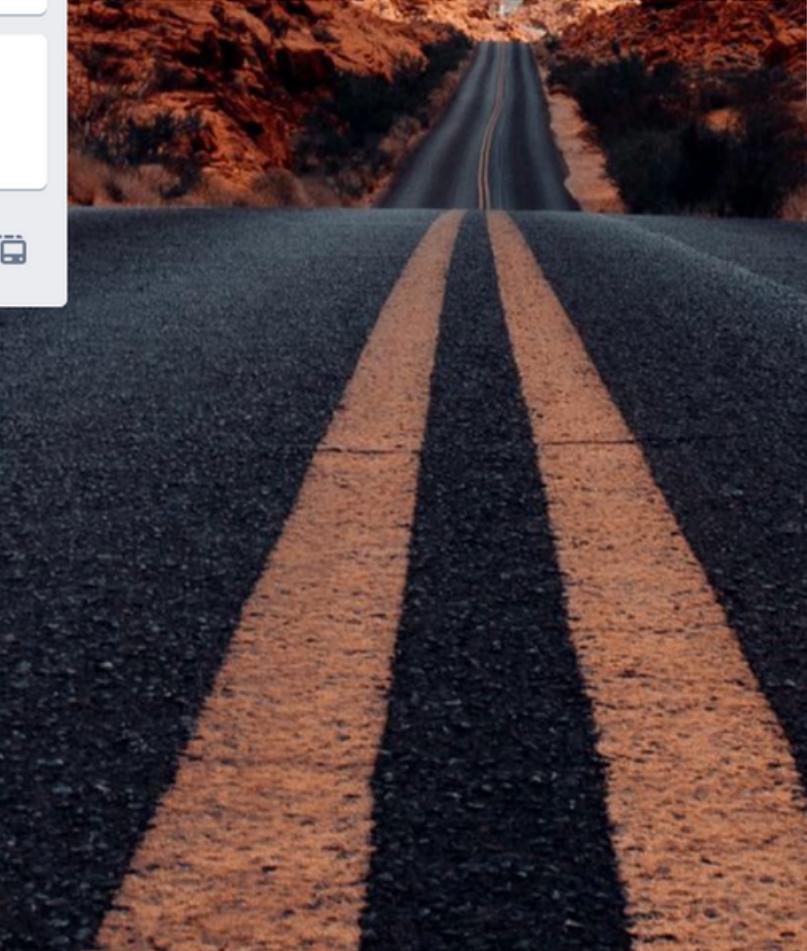


和mentor 更新當前討論進度

提供投資因子之公式等內容



+ 新增另一張卡片



中信第二題投資因子分析



金融科技 Free

團隊觀看權限



邀請

Butler

顯示選單

進行中的工作

Bill Ackman

0/6

BC



爬股價歷史報酬

BC

找迴歸計算

爬取歷年投資人新聞

爬取歷年總經新聞

+ 新增另一張卡片

已完成的項目

決定 題目要做的基金投資人

準備想做的知名投資人資料

🕒 4月12日

≡

1

論文看完，確認題目內容方向

🕒 4月12日

≡

寄email 紙mentor

🕒

目標

找出投資人的「投資因子」，創建「年化報酬高於 6 %」的投資組合

+ 新增另一張卡片

檔案附件

Buffett's Alpha

≡

+ 新增另一張卡片

網站資源

因子

≡

+ 新增另一張卡片

...

...





← 100% 標題 2 Arial 14 B I U A | + | 2 1 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | 編輯 | ^

- 4 / 18 (六)

4 / 18 (六)

大綱

- 一、論文討論
- 二、實踐方式探討
- 三、工作排程
- 四、問題與討論

大綱

- 一、論文討論
- 二、實踐方式探討
- 三、工作排程
- 四、問題與討論

一、論文討論

- 論文討論概要

CAPM	P46、P56-63、P76-P79
Value premium	P5-P8、P13-14、P36-38、P49
Momentum strategy:	P34-47
L4_Investment and profitability anomalies (H)	37-39

$$- r_i = E(r_i) + \beta_{iMP}MP + \beta_{iDEI}DEI + \beta_{iUI}UI + \beta_{iUPR}UPR + \beta_{iUTS}UTS + e_i$$

- MP = monthly growth in industrial production
- DEI = change in expected inflation rate
- UI = unexpected inflation rate
- UPR (credit spread)= excess return of low-grade long-term corporate bonds over long-term government bonds

4/20 (一)

在「會議紀錄」列表中

三 描述

編輯

一、共現圖

- [範例1：客服聯繫找出潛在銷售機會](#)
- [範例2：實際投資行為vs風險屬性評估之落差](#)

二、問題定義

- 希望透過爬取歷年Bill Ackman 的新聞，找出投資人的風格、反應，如何受市場影響，以提供判斷投資因子的初步線索

三、資料來源

- 2007~2020 Bill Ackman 新聞
- 2007~2020 總體經濟新聞

四、方向更新

1. 使用機器學習的方式訓練
 - 總體經濟：以月份為單位分成好與壞（或是更細的類別）
 - 訓練資料：總體經濟環境、投資組合的特徵、投組中個股的特徵或因
 - 結果：給總體狀況、答案是重要的投組或個股因子
 - 建置：再根據這些因子去建置一個相當的投資組合
 - [訓練的資料與結果待討論]
2. 標普五百指數,羅素1000 指數,羅素3000 指數,美國房屋建築業指數,美國消費服務指數,道瓊美國指數,pmi,建業執照（預先）-->判斷總體經濟環境狀況

• 14 commits

• 1 branch

• 0 packages

• 0 releases

• 1 contributor

Branch: master •

New pull request

Create new file Upload files Find file Clone or download

 chenjanice	Add files via upload	Latest commit a85d093 6 minutes ago
 BillAckman.csv	Add files via upload	yesterday
 CNBC_News_BillAckman.ipynb	Add files via upload	yesterday
 HW3_EDA.ipynb	Add files via upload	9 hours ago
 Readme.md	Update Readme.md	9 hours ago
 dataroma_history.csv	Add files via upload	9 hours ago
 fintech_hw3.pdf	Add files via upload	6 minutes ago

 Readme.md

CTBC Project

中信：投資因子分析

- 投資人：Bill Ackman

DATA

- EDA
 - News & Holdings
- 歷年新聞
 - CBNC-Bill Ackman
 - News-crawler
- 歷年持股
 - Dataroma-Bill Ackman
- 歷年財報

Preprocessing

Regression

Model training