Project：高雄發大財

盡量避免用AI這個詞，了解在做甚麼最重要

除了看到故事外，能不能針對問題提出建言?

用資料科學來解決問題

使用資料：<https://www.ptt.cc/bbs/Instant_Food/版>

抓取資料日期：2018/1-2019/05

1. 問題描述+解決的核心關鍵是甚麼?
2. 資料來源說明+如何挑選進行屬性分析?
3. 階段性結果+困難如何解決
4. 分析流程+流程圖
5. 未來走向

🡪為什麼能從這些資料發現是否有成長，去說服大家

1. 高雄發大財?從觀光產業的變化去觀察韓國瑜上任前後的差別

(1).整理出交通()、旅館的總收益(從實際發財的角度來觀察、直條圖)、旅館住宿旅客的組成(觀察國外旅客數的變化、多個圓餅圖)、觀光遊客人數(總計：折線圖、個別：點點散布圖、可以再跟往年的同季做比較、或是同時間的地區做比較)

(2).分別針對以上圖表觀察上任前後的數據波動跟差別，是否真的有發大財(撇除掉特殊節慶)?

(3).挑出特別的時段或月份，比對爬文出的討論變化與熱度(那段時間是否有什麼特殊的事件發生?)

1. <副主題>從新聞或ptt等與論研究對於”高雄”、”韓國瑜”的討論變化與討論熱度(推文與留言的數量)

(1).從前五大新聞電子報去爬出對於高雄、韓國瑜報導量跟提到次數的比較

(2).ptt貼文留言的數量跟熱度

(3).ptt貼文高頻率的詞頻

(4).ptt特定詞的出現頻率

情緒分析semtiment analysis

高雄交通資料(2018/1-2019/5)：

(火車) https://data.gov.tw/dataset/8792

(高鐵) http://m.thsrc.com.tw/tw/Article/ArticleContent/117f6de2-ed8b-403a-ab4a-820d123e37bf

高雄旅館資料(2018/1-2019/3)：https://admin.taiwan.net.tw/FileUploadCategoryListC003330.aspx?Pindex=2&CategoryID=0dcf358f-f875-452d-8d14-2b715d02ab1a&appname=FileUploadCategoryListC003330

高雄觀光遊憩區遊客人數統計(2018/1-2019/5)：

<https://khh.travel/Article.aspx?a=7560&l=1>