# Market Analysis of Mobile

Chao Wen Tsai

2025-03-23

*Open data from Kaggle Mobiles Dataset (2025)*

一、專案概述

在現今數位化時代，手機市場競爭激烈，產品種類繁多。為了深入了解手機市場格局、產品特性以及消費者需求，對手機相關數據進行全面而深入的分析顯得格外重要。

本專案旨在透過對包含多品牌、多型號手機資訊的資料集進行分析，挖掘有價值的信息，為手機製造商、銷售商以及消費者提供決策依據。

加載必要的 **R** 包並讀取數據集。

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.4
## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
# 讀取數據集
df <- read_csv('C:/Users/chao/Desktop/Data Analysis/Datasets/Mobiles Dataset (2025).csv')
```

```
## Rows: 930 Columns: 15
## -- Column specification -----------------------------------------------------
## Delimiter: ","
## chr (14): Company Name, Model Name, Mobile Weight, RAM, Front Camera, Back C...
## dbl  (1): Launched Year
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

查看數據基本資料

```
print('數據基本資料：')
```

```
## [1] "數據基本資料："
```

```
glimpse(df)
```

```
## Rows: 930
## Columns: 15
## $ `Company Name`           <chr> "Apple", "Apple", "Apple", "Apple", "Apple~
## $ `Model Name`             <chr> "iPhone 16 128GB", "iPhone 16 256GB", "iPh~
## $ `Mobile Weight`          <chr> "174g", "174g", "174g", "203g", "203g", "2~
## $ RAM                      <chr> "6GB", "6GB", "6GB", "6GB", "6GB", "6GB", ~
## $ `Front Camera`           <chr> "12MP", "12MP", "12MP", "12MP", "12MP", "1~
## $ `Back Camera`            <chr> "48MP", "48MP", "48MP", "48MP", "48MP", "4~
## $ Processor                <chr> "A17 Bionic", "A17 Bionic", "A17 Bionic", ~
## $ `Battery Capacity`       <chr> "3,600mAh", "3,600mAh", "3,600mAh", "4,200~
## $ `Screen Size`            <chr> "6.1 inches", "6.1 inches", "6.1 inches", ~
## $ `Launched Price (Pakistan)` <chr> "PKR 224,999", "PKR 234,999", "PKR 244,999~
## $ `Launched Price (India)` <chr> "INR 79,999", "INR 84,999", "INR 89,999", ~
## $ `Launched Price (China)` <chr> "CNY 5,799", "CNY 6,099", "CNY 6,499", "CN~
## $ `Launched Price (USA)`   <chr> "USD 799", "USD 849", "USD 899", "USD 899"~
## $ `Launched Price (Dubai)` <chr> "AED 2,799", "AED 2,999", "AED 3,199", "AE~
## $ `Launched Year`          <dbl> 2024, 2024, 2024, 2024, 2024, 2024, 2024, ~
```

查看數據集行數和列數

```
library(dplyr)
  # 用 dplyr 包中的 print() 函数的替代方法，比如 print.data.frame()
  # 也可用 tibble package tbl_df 函數處理
rows <- nrow(df)
columns <- ncol(df)

if (rows < 100 && columns < 20) {
  # 短表數據（行數少於 100 且列數少於 20）查看全量數據資料
  print('數據全部內容訊息：')
  print.data.frame(df, na = 'nan')
} else {
  # 長表數據查看數據前幾行訊息
  print('數據前幾行內容訊息：')
  print.data.frame(head(df), na = 'nan')
}
```

```
## [1] "數據前幾行內容訊息："
##   Company Name        Model Name Mobile Weight RAM Front Camera Back Camera
## 1        Apple   iPhone 16 128GB          174g 6GB         12MP        48MP
## 2        Apple   iPhone 16 256GB          174g 6GB         12MP        48MP
## 3        Apple   iPhone 16 512GB          174g 6GB         12MP        48MP
```

```
## 4          Apple iPhone 16 Plus 128GB          203g 6GB          12MP          48MP
## 5          Apple iPhone 16 Plus 256GB          203g 6GB          12MP          48MP
## 6          Apple iPhone 16 Plus 512GB          203g 6GB          12MP          48MP
##      Processor Battery Capacity Screen Size Launched Price (Pakistan)
## 1 A17 Bionic        3,600mAh  6.1 inches              PKR 224,999
## 2 A17 Bionic        3,600mAh  6.1 inches              PKR 234,999
## 3 A17 Bionic        3,600mAh  6.1 inches              PKR 244,999
## 4 A17 Bionic        4,200mAh  6.7 inches              PKR 249,999
## 5 A17 Bionic        4,200mAh  6.7 inches              PKR 259,999
## 6 A17 Bionic        4,200mAh  6.7 inches              PKR 274,999
##   Launched Price (India) Launched Price (China) Launched Price (USA)
## 1           INR 79,999            CNY 5,799            USD 799
## 2           INR 84,999            CNY 6,099            USD 849
## 3           INR 89,999            CNY 6,499            USD 899
## 4           INR 89,999            CNY 6,199            USD 899
## 5           INR 94,999            CNY 6,499            USD 949
## 6          INR 104,999            CNY 6,999            USD 999
##   Launched Price (Dubai) Launched Year
## 1          AED 2,799          2024
## 2          AED 2,999          2024
## 3          AED 3,199          2024
## 4          AED 3,199          2024
## 5          AED 3,399          2024
## 6          AED 3,599          2024
```
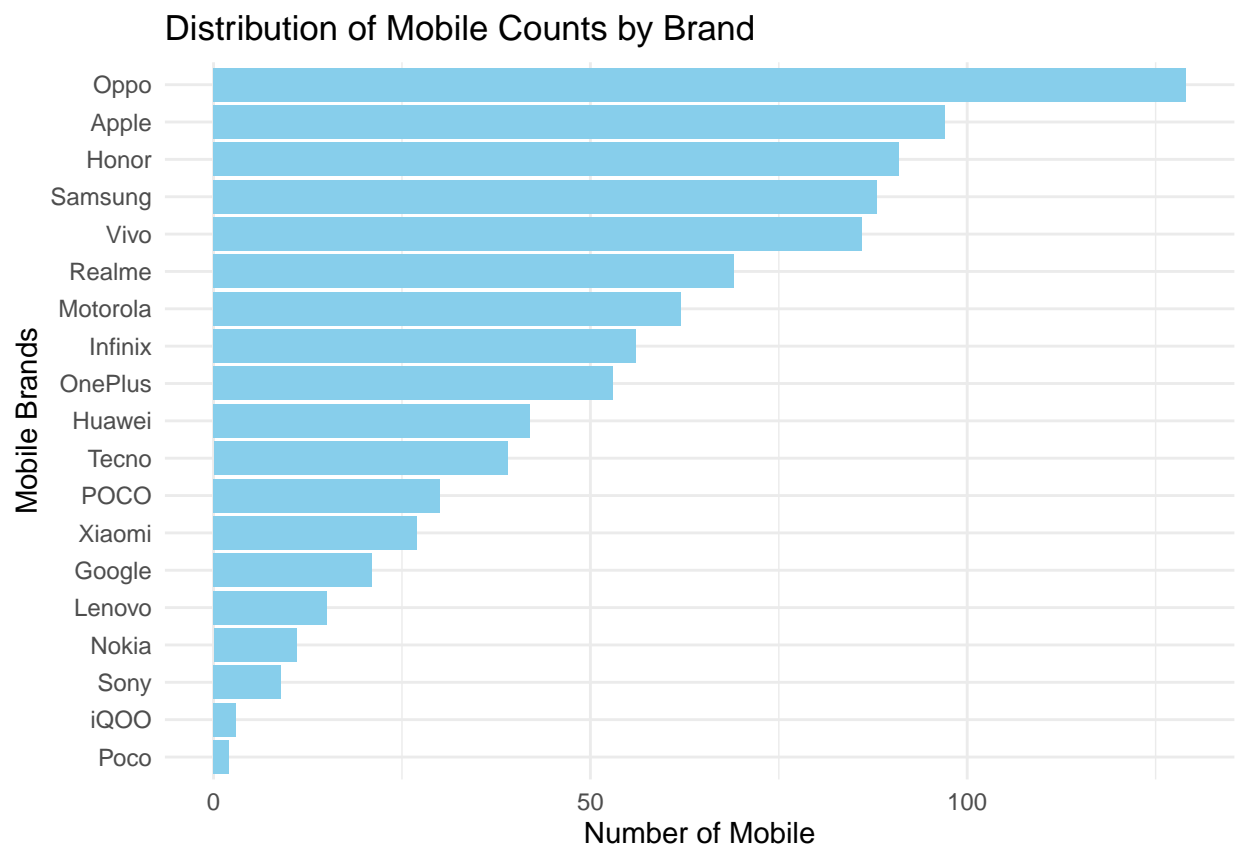
## 二、不同品牌手機數量分布

統計不同品牌手機數量

```r
company_counts <- df %>%
  count(`Company Name`) %>%
  arrange(desc(n))
print(company_counts)
```

```
## # A tibble: 19 x 2
##    `Company Name`      n
##    <chr>           <int>
##  1 Oppo              129
##  2 Apple              97
##  3 Honor              91
##  4 Samsung            88
##  5 Vivo               86
##  6 Realme             69
##  7 Motorola           62
##  8 Infinix            56
##  9 OnePlus            53
## 10 Huawei             42
## 11 Tecno              39
## 12 POCO               30
## 13 Xiaomi             27
## 14 Google             21
## 15 Lenovo             15
```

```
## 16 Nokia              11
## 17 Sony               9
## 18 iQOO               3
## 19 Poco               2
```

繪製柱狀圖

```r
ggplot(company_counts, aes(x = reorder(`Company Name`, n), y = n)) +
  geom_bar(stat = 'identity', fill = 'skyblue') +
  coord_flip() +
  labs(x = 'Mobile Brands', y = 'Number of Mobile', title = 'Distribution of Mobile Counts by Brand') +
  theme_minimal() # 簡約主題函數，移除不必要的裝飾元素，保留基本繪圖元素（去除灰底）
```


Distribution of Mobile Counts by Brand

分析結果:

從柱狀圖中我們可以直觀的看出不同品牌手機數量的差異。數量較多的品牌可能在市場上更具競爭力，擁有更大的市場份額。

三、不同年份手機發布數量

統計不同年份手機發布數量

```
year_counts <- df %>%
  count(`Launched Year`) %>%
  arrange(`Launched Year`)
print(year_counts)
```

```
## # A tibble: 11 x 2
##     `Launched Year`       n
##              <dbl> <int>
##  1            2014     2
##  2            2016     5
##  3            2017     9
##  4            2018    21
##  5            2019    46
##  6            2020   101
##  7            2021   109
##  8            2022   149
##  9            2023   184
## 10            2024   292
## 11            2025    12
```

繪製折線圖

```
ggplot(year_counts, aes(x = `Launched Year`, y = n)) +
  geom_line(group = 1, color ='lightblue') +
  geom_point() +
  labs(x = 'Release Year', y = 'Number of Releases', title = 'Number of Mobile Releases by Year') +
  theme_minimal()
```

## Number of Mobile Releases by Year
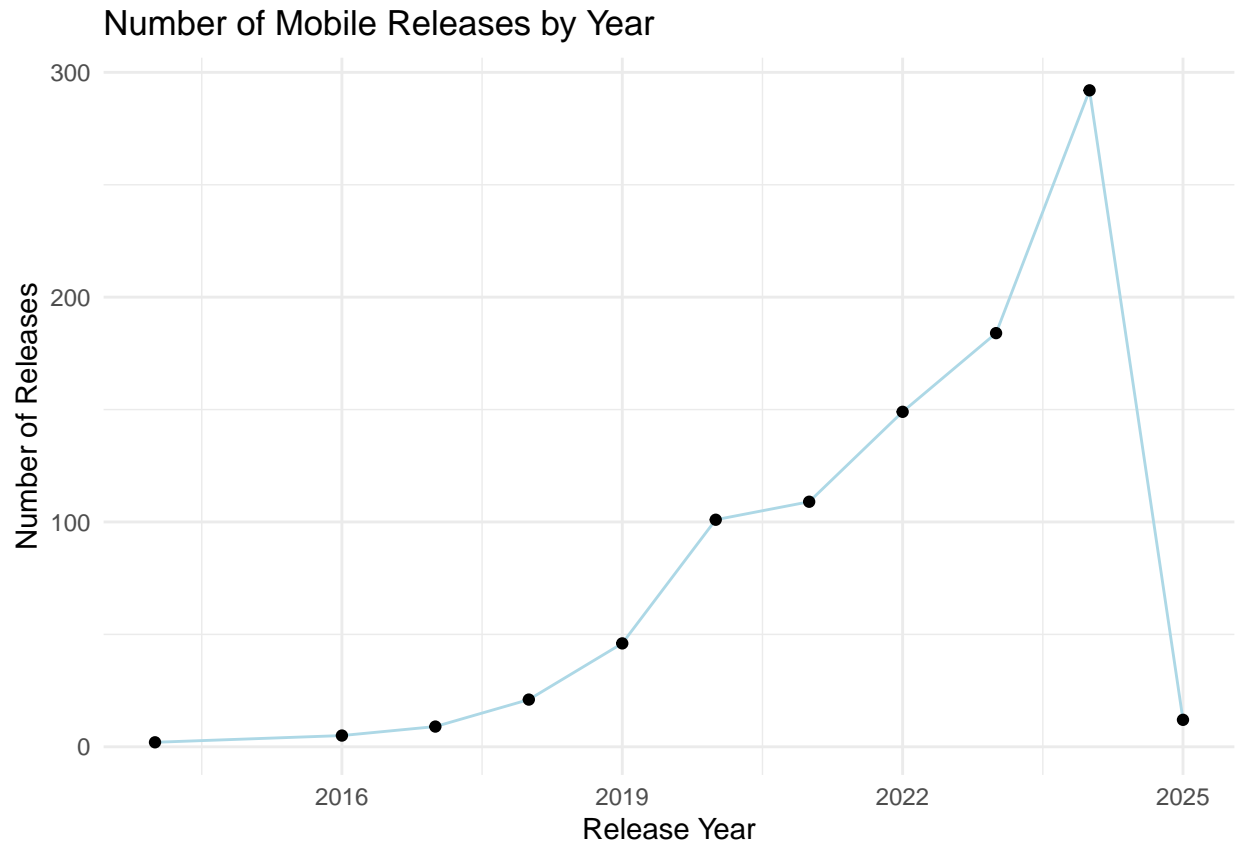


分析結果:

通過折線圖可以觀察到手機發布數量隨時間的變化趨勢。這有助於我們了解手機市場的發展動態，例如是否呈現增長或下降趨勢，是否有週期性變化等。

## 四、手機各性能參數相關性分析

對手機的性能參數列進行數據預處理，提取數值部分並轉換為數值類型。

提取性能參數列

```
df_clean <- df %>%
  select(`Mobile Weight`, `RAM`, `Front Camera`, `Back Camera`, `Processor`, `Battery Capacity`, `Screen
print(df_clean)
```

```
## # A tibble: 930 x 7
##    `Mobile Weight` RAM   `Front Camera` `Back Camera` Processor
##    <chr>           <chr> <chr>          <chr>         <chr>
## 1 174g            6GB   12MP           48MP          A17 Bionic
## 2 174g            6GB   12MP           48MP          A17 Bionic
## 3 174g            6GB   12MP           48MP          A17 Bionic
## 4 203g            6GB   12MP           48MP          A17 Bionic
## 5 203g            6GB   12MP           48MP          A17 Bionic
## 6 203g            6GB   12MP           48MP          A17 Bionic
```

```
##  7 206g          6GB   12MP / 4K      50MP + 12MP   A17 Pro
##  8 206g          8GB   12MP / 4K      50MP + 12MP   A17 Pro
##  9 206g          8GB   12MP / 4K      50MP + 12MP   A17 Pro
## 10 221g          6GB   12MP / 4K      48MP + 12MP   A17 Pro
## # i 920 more rows
## # i 2 more variables: `Battery Capacity` <chr>, `Screen Size` <chr>
```

定義函數提取數值

```
extract_numeric <- function(x) {
  as.numeric(str_extract(x, '\\d+\\.?\\d*'))
# as.numeric 輸入值轉換為數值類型，'\\d+\\.?\\d*'匹配整數、小數。
}
```

對各性能參數列進行數據轉換

```
df_clean <- df_clean %>%
  mutate(across(everything(), extract_numeric))
print(df_clean)
```

```
## # A tibble: 930 x 7
##    `Mobile Weight`  RAM `Front Camera` `Back Camera` Processor
##            <dbl> <dbl>          <dbl>         <dbl>     <dbl>
##  1            174     6             12            48        17
##  2            174     6             12            48        17
##  3            174     6             12            48        17
##  4            203     6             12            48        17
##  5            203     6             12            48        17
##  6            203     6             12            48        17
##  7            206     6             12            50        17
##  8            206     8             12            50        17
##  9            206     8             12            50        17
## 10            221     6             12            48        17
## # i 920 more rows
## # i 2 more variables: `Battery Capacity` <dbl>, `Screen Size` <dbl>
```

計算相關矩陣計算相關性矩陣

```
corr_matrix <- cor(df_clean, use = 'pairwise.complete.obs')
print(corr_matrix)
```

```
##                Mobile Weight        RAM Front Camera Back Camera  Processor
## Mobile Weight     1.00000000 -0.00742740 -0.2832635654 -0.30851338 0.01375491
## RAM              -0.00742740  1.00000000  0.4600931232  0.43963594 0.20182267
## Front Camera     -0.28326357  0.46009312  1.0000000000  0.44932466 0.21287601
## Back Camera      -0.30851338  0.43963594  0.4493246603  1.00000000 0.10558262
## Processor         0.01375491  0.20182267  0.2128760128  0.10558262 1.00000000
```

```
## Battery Capacity     0.18239568  0.08587738 -0.0002920413  0.02639849 0.03395670
## Screen Size          0.97589319  0.04696264 -0.2243110081 -0.26327884 0.03254556
##                     Battery Capacity Screen Size
## Mobile Weight          0.1823956826  0.97589319
## RAM                    0.0858773820  0.04696264
## Front Camera          -0.0002920413 -0.22431101
## Back Camera            0.0263984920 -0.26327884
## Processor              0.0339567012  0.03254556
## Battery Capacity       1.0000000000  0.16952537
## Screen Size            0.1695253719  1.00000000
```
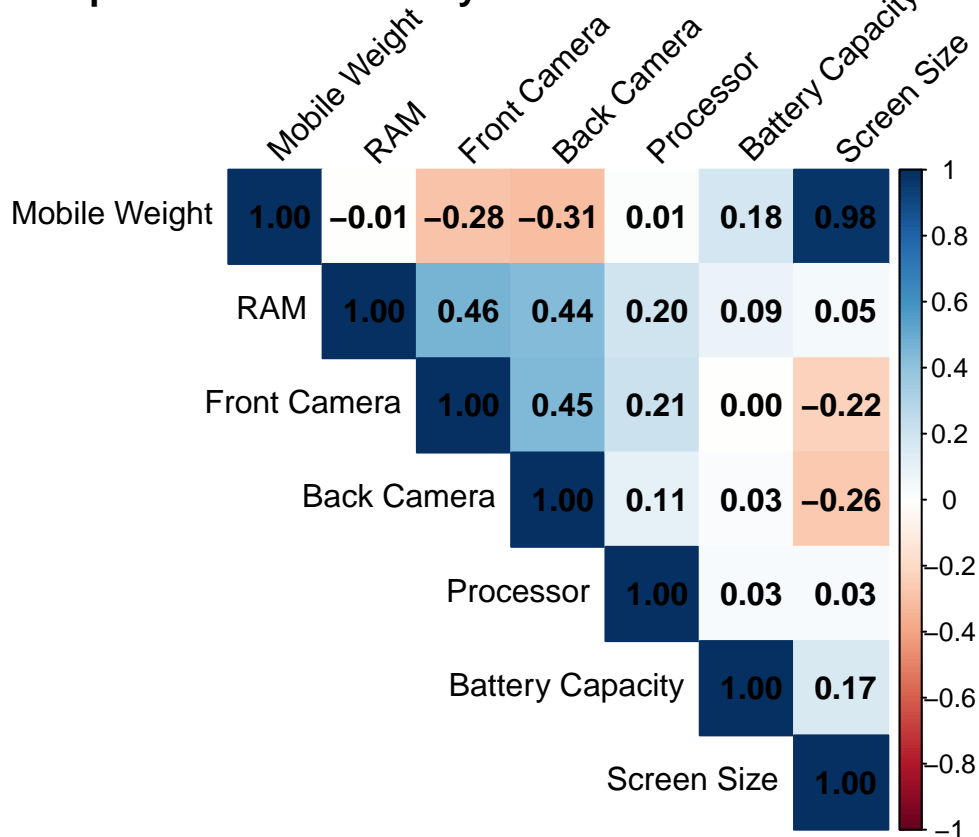
繪製熱力圖

```
library(corrplot)
```

```
## corrplot 0.95 loaded
```

```
# 因 Title 會被切半，故設置圖形的邊距，這裡增加頂部邊距
par(mar = c(5, 4, 8, 2)) #  8 是頂部邊距

corrplot(corr_matrix, method = 'color', type = 'upper',
         addCoef.col = 'black', tl.col = 'black', tl.srt = 45,
         title = 'Heatmap of Correlation Analysis for Mobile Performance Parameters')
```

分析結果:

從相關性熱力圖中，我們可以直觀的看到各性能參數之間的相關強弱。正相關表示兩個參數的變化趨勢一致，負相關表示變化趨勢相反。這有助於我們了解手機性能參數之間的內在關係，例如是否存在某些參數相互制約或協同的情況。

## 五、總結與建議

### 品牌市場狀況

根據不同品牌手機數量的分佈，我們可以推測某些品牌在市場上佔據主導地位，而一些小眾品牌可能需要進一步提升市場競爭力。建議小眾品牌深入研究市場需求，推出更具特色的產品，加強行銷與品牌推廣。

### 發布趨勢

從不同年份手機發布數量的變化趨勢，我們可以預測未來手機市場的發展方向。如果呈現成長趨勢，廠商可以加大研發投入，推出更多新品；如果呈現下降趨勢，則需要謹慎規劃產品發布策略，並專注於產品品質和創新。

### 性能參數關係

根據性能參數的相關性分析，廠商在設計手機時可以考慮參數之間的相互關係，優化產品性能。例如，如果兩個參數正相關且都對使用者體驗很重要，可以同時提升這兩個參數的效能；如果是負相關，則需要在兩者之間進行權衡。