

## 一、原始資料取得

飲料市場銷售量與銷售值資料：

進入「[經濟部工業產銷存動態調查資料庫](#)」選擇產品統計，選擇 5 種飲料類別（果蔬汁飲料、碳酸飲料、運動飲料、咖啡飲料及茶類飲料）的銷售量與銷售值統計項目，期間選擇 1982 年至 2024 年月統計資料，具體設定如下圖。

週期: 月 查詢日期: 西元 1982/1 ~ 2024/9 計算模式: 統計值 製作: 統計表

統計項目 (總數: 10 已選: 2)

- ☐ 生產量
- ☐ 生產值
- ☒ 銷售量
- ☒ 銷售值
- ☐ 內銷量(含間接外銷)
- ☐ 內銷值(含間接外銷)
- ☐ 直接外銷量
- ☐ 直接外銷值
- ☐ 存貨量
- ☐ 存貨值

產品別 (點下面產品名稱可查詢產品定義範圍) (總數: 614 已選: 5)

大行業(收合) 中行業(收合) 細行業(收合) 產品關鍵字搜尋

- (0919)其他酒精飲料製造業
  - ☐ (0919000)其他酒精飲料
- (0920)非酒精飲料製造業
  - ☐ 全選
  - ☒ (0920010)果蔬汁
  - ☒ (0920910)碳酸飲料
  - ☐ (0920920)礦泉水
  - ☒ (0920930)運動飲料
  - ☒ (0920940)咖啡飲料
  - ☒ (0920950)茶類飲料
  - ☐ (0920990)其他非酒精飲料
- (10)菸草製造業

## 二、資料定義

飲料市場銷售量與銷售值資料：

銷售值：飲料每月銷售值（單位：千元新台幣）

銷售量：飲料每月銷售量（單位：千公升）

單位價格：每月銷售飲料之單位價格（單位：千元新台幣/千公升）

銷售量與銷售值可由資料直接得到，單位價格根據銷售量與銷售值兩個變數相除計算得出，名稱定義如下：

### 果蔬汁飲料：

銷售量 (FruitVegetableJuice\_volume)

銷售值 (FruitVegetableJuice\_value)

單位價格 (FruitVegetableJuice\_price)

### 碳酸飲料：

銷售量 (CarbonatedBeverage\_volume)

銷售值 (CarbonatedBeverage\_value)

單位價格 (CarbonatedBeverage\_price)

### 運動飲料：

銷售量 (SportsDrink\_volume)

銷售值 (SportsDrink\_value)

單位價格 (SportsDrink\_price)

### 咖啡飲料：

銷售量 (CoffeeDrink\_volume)

銷售值 (CoffeeDrink\_value)

單位價格 (CoffeeDrink\_price)

### 茶類飲料：

銷售量 (TeaDrink\_volume)

銷售值 (TeaDrink\_value)

單位價格 (TeaDrink\_price)

產品名稱	定義	範圍
果蔬汁飲料	含天然果汁/蔬菜汁或還原果汁/蔬菜汁10%以上，直接供飲用之果汁/蔬菜汁飲料。	稀釋果蔬汁、清淡果蔬汁、發酵果蔬汁、稀釋發酵果蔬汁、清淡發酵果蔬汁、果肉飲料。
碳酸飲料	在除去鹵素飲用水中加壓，添加二氧化碳及果實香料、果汁；或可樂子實葉抽出液；或Saraparilla根抽出液等調味料之碳酸飲料。	汽水、可樂、沙士。
運動飲料	具可調解人體電解質功能之飲料，調整為等張滲透壓，以便自人體腸道迅速吸收，PH值在2.5～3.8之間，電解質濃度(ug/ml)則分別為鈉離子552以下、鎂離子24以下、鉀離子195以下、氯離子639以下、鈣離子60以下、磷酸根離子190以下。	
咖啡飲料	利用咖啡粉或咖啡豆研磨、浸泡、萃取、調理，添加奶精、糖水或調味料之飲料，其咖啡因若超過200ppm則需標示，但不得超過500ppm。	純咖啡飲料、調味咖啡飲料。
茶類飲料	利用茶葉或茶葉梗浸泡、萃取、調理，添加糖水或調味料之飲料，其咖啡因若超過200ppm則需標示，但不得超過500ppm。	烏龍茶、花茶、紅茶、綠茶、調味茶(如檸檬茶)。

### 三、資料處理

在完成資料蒐集後，我們按步驟進行了資料處理，目的是提高數據的可靠性與一致性。首先，我們使用 R 統計軟體對資料讀取後進行清理，包括合併數據、去除重複資料和處理遺漏值等工作。對於部分遺漏值，考量到補值可能帶來的偏誤，我們選擇將無法合理填補的觀察值移除。此外，由於不同飲料類別的統計起始時間不一致，我們將分析的起始時間進行統一，分析時間改由 1991 年 1 月起始，確保資料具有可比性。為了進一步豐富研究變數，我們還根據銷售量與銷售值兩個變數計算並新增了每種飲料的單位價格，為我們後續使用模型估計去衡量價格對消費者需求的影響提供了必要的解釋變數。具體步驟如下：

#### 1. 資料庫位置設定（‘setwd’）

根據不同使用者，設定工作目錄（‘setwd’）到適當的 Dropbox 目錄位置，這樣便能夠從指定路徑讀取數據文件。

#### 2. 讀取資料（‘fread’）

使用 ‘fread’ 函數從指定路徑讀取飲料市場銷售量與銷售值資料的 CSV 資料檔。

#### 3. 銷售數據處理（‘sale\_data’）

- 更改欄位名稱：將第二列作為欄位名稱，並刪除前兩列不需要的數據。
- 年份和月份提取：利用 ‘gsub’ 函數去掉「年」和「月」字，並將年份和月份轉換為數值型。
- 缺失值填補：利用 ‘nafill’ 函數按順序填補年份和月份的缺失值。
- 數據格式轉換：將第三到第十二欄的數據轉換為數值型，並將千分位符號去掉。
- 重新命名：將產品名稱及其對應的數據欄位改為英文名稱。

#### 4. 合併整理資料表

- Full Join：計算並新增單位價格（根據銷售量與銷售值兩個變數相除計算得出）後合併到資料表。
- subset：由於不同飲料類別的統計時間不一致，將分析的時間範圍進行統一，篩選出年份（year）大於等於 1991 且小於 2024 的列。

#### 5. 結果輸出（‘write.csv’）

將合併後的完整數據集保存為 CSV 檔案。