# 📘 今日學習總整理（Jupyter + Pandas 基礎）

## 📌 一、Jupyter Notebook 基礎觀念

|  |  |
| --- | --- |
| 主題 | 說明 |
| Kernel 是什麼 | 是負責「執行你寫的程式」的大腦，每個 Notebook 一個獨立的 Kernel |
| 多個 Kernel | 可以同時開多個 Notebook，各自用不同或相同的 Kernel，彼此獨立不共用變數 |
| Restart Kernel | 重啟會清除所有變數記憶，需重新執行所有 Cell |
| .ipynb\_checkpoints | 是自動儲存的備份資料夾，可忽略或刪除（會自動再生成） |

## 📌 二、Pandas 基礎語法與資料操作

|  |  |
| --- | --- |
| 主題 | 說明 |
| import pandas as pd | 載入 pandas 並使用簡寫 pd |
| pd.DataFrame() | 建立表格型資料 |
| df.shape, df.shape[1] | 回傳資料列與欄數（row, column） |
| df.info() | 查看欄位數、型別、是否有缺值等 |
| df.describe() | 查看數值欄位的統計摘要（平均、標準差、中位數等） |
| df.dtypes | 查看每個欄位的資料型別 |
| dropna() | 刪除有缺值的資料列（可設定 axis, inplace） |
| astype(float) | 將資料從文字（object）轉為數值（float64） |
| pd.to\_numeric(errors='coerce') | 將不能轉成數字的內容變成 NaN，常用來清洗錯誤數據 |

## 📌 三、資料清洗與轉換範例

你成功地對 health\_data 做了以下處理：

• 匯入 .csv 檔案（使用絕對路徑）

• 刪除空值列：dropna()

• 將 "Average\_Pulse" 和 "Max\_Pulse" 從 object 轉為 float

• 使用 describe() 分析健康數據摘要（心跳、卡路里、工作/睡眠時數）

### 🔍 分析結果觀察（來自 describe()）

• 平均心跳 約 102，下限 80，上限 125

• 最大心跳 約 137，範圍 120–150

• 熱量消耗 平均 285，最大 330

• 工作時數 平均 7.4，最少 0，最多 10

• 睡眠時數 很穩定，集中在 7–8 小時之間

## ✅ 建議後續練習方向：

|  |  |
| --- | --- |
| 目標 | 建議做法或範例 |
| 畫圖視覺化 | 使用 matplotlib 或 seaborn 畫折線圖、直方圖、箱型圖等 |
| 數據關聯分析 | df.corr()、相關係數熱圖、散點圖 |
| 計算額外統計指標（如 z-score） | 使用 scipy.stats 或自己計算標準化值 |
| 嘗試建立簡單報告 | 用 markdown Cell 整理輸出結果或轉成 HTML 報告 |