

《大數據分析》課程教學進度安排（共 48 課時）

一、課程結構總覽

課程總時數為 48 課時（每課時 45 分鐘），其中理論課 32 課時，實驗課 16 課時。課程安排為期 6 週，每週五下午 4 節課、週日上午 4 節課，自開學第 3 週起連續上課。

二、每週教學進度安排

第 1 週（開學第 3 週）

- 理論：單元 1（大數據概述）+ 單元 2 前半（NumPy 基礎）
- 實驗 1：pip 與 PyCharm 安裝 + NumPy 創建基本陣列

第 2 週

- 理論：單元 2 後半（通用函數、線性代數）+ 單元 3（Pandas 基礎）
- 實驗 2：考試成績數據統計與排序

第 3 週

- 理論：單元 4（數據存儲）+ 單元 5（數據清洗）+ 單元 6（數據合併與轉換）
- 實驗 3：多源數據讀寫與清洗

第 4 週

- 理論：單元 7（分組與聚合）+ 單元 8（機器學習基礎）
- 實驗 4：sklearn 實作 iris/boston 數據分類與預測

第 5 週

- 理論：單元 9（資料視覺化）+ 案例：單元 10（電力用戶預測）
- 實驗 5：matplotlib/seaborn 視覺化圖表製作
- 實驗 6：用戶行為預測模型建構

第 6 週

- 理論：案例講解：單元 11（文本挖掘）、單元 12（城市通勤分析）
- 實驗 7：文本雲圖與詞頻分析
- 實驗 8：通勤數據熱力圖與區域分析

三、實驗課安排（共 16 課時）

| 實驗編號 | 內容 | 對應單元任務 | 課時 |
|------|----------------|-----------------|----|
| 實驗 1 | NumPy 包安裝與陣列創建 | 1. 2, 2. 1~2. 2 | 2 |

| | | | |
|------|-----------------------------|-----------------|---|
| 實驗 2 | Pandas 排序與統計 | 2. 5, 3. 4~3. 5 | 2 |
| 實驗 3 | CSV 與 Excel 讀寫 + 缺失值處理 | 4. 2, 5. 1~5. 3 | 2 |
| 實驗 4 | Sklearn 建模 (iris/boston) | 8. 1~8. 4 | 2 |
| 實驗 5 | 統計圖表與 Seaborn 熱力圖 | 9. 2~9. 4 | 2 |
| 實驗 6 | 電力行為預測建模 | 10. 2~10. 3 | 2 |
| 實驗 7 | 文本挖掘與詞雲分 析 | 11. 2~11. 4 | 2 |
| 實驗 8 | 城市通勤可視化分 析 | 12. 2~12. 6 | 2 |