**程式設計期末專題 – 迴歸**

**組員: 范睿凱、謝宗穎**

**分析原因:**

我們想知道每個人的各種特徵，包括性別、年齡、地域、是否有小孩、是否有抽菸等因素與保費之間的關聯。並且判斷迴歸模型的準確度。

**程式碼的過程:**

1. **導入標準庫:**導入pandas、線性迴歸模型。
2. **匯入檔案:** 把insurance.csv 檔匯入，把要預測的項目charges，利用drop 刪掉。
3. **處理缺失值:** 依據列(axis = 1)，並且只要有任何一筆資料缺失就將該資料刪除(how = ‘any’)。
4. **減少特徵數**: 再減少不必要的特徵，讓機器學習效率增加。
5. **建立模型**:也就是將數據集拆分為訓練集和測試集(7:3)。
6. **預測結果及準確度:** 將預測的結果及準確度印出來。