

# 2019 ECT 作業七

## (a) 選取以下欄位：（15%）

□ Name、Age、Wage、Skill Moves、Work Rate、Height、Weight、Crossing ~ GKReflexes

□ 總共41個屬性

```
1 import pandas as pd
2 import numpy as np
3
4 #讀取CSV檔案
5 data = pd.read_csv('fifa19.csv')
6 df1 = data[['Name','Age','Wage','Skill Moves','Work Rate','Height','Weight']]
7 df2 = data.iloc[:,54:88]
8 data = df1.join(df2)
```

- i. import 需要的套件
- ii. 讀取 csv 檔案
- iii. 選取所需的 41 個欄位

## (b) 修改[Work Rate]該欄的值（15%）

□ 切開成[High].[Medium].[Low]3欄(有=1，無=0)

□ 完成後，將原本的[Work Rate]欄位刪除

```
1 Medium = data['Work Rate'].str.contains('Medium')
2 High = data['Work Rate'].str.contains('High')
3 Low = data['Work Rate'].str.contains('Low')
4 # 新增 Medium, High, Low 欄位
5 data['Medium'], data['High'], data['Low'] = Medium, High, Low
6 # 將 True 轉為 1, False 轉為 0
7 data[["Medium", "High", "Low"]] *= 1
8 # 刪除 [Work Rate] 欄位
9 data = data.drop(columns="Work Rate")
```

- i. 判斷[Work Rate]內是否存在 Medium、High、Low 之值，將判斷結果分別存入 Medium、High、Low 內。
- ii. 在 dataframe 內新增 Medium、High、Low 三個欄位，並將值存入。
- iii. 將 dataframe 內 Medium、High、Low 三個欄位的值若為 True 轉為 1，為 False 轉為 0。

**(c) 如[Height].[Weight]欄位中有空值(null)，即刪除整筆資料（15%）**

```
1 data = data.dropna(subset=['Height','Weight'])
```

- i. 使用 dropna 函數將[Height].[Weight]欄位中有空值(null)的資料刪除。

**(d) 轉換幣值（15%）**

☐ 修改[Wage]該欄的值

☐ 歐元：台幣→1:35

☐ K→1000

```
1 # 擷取[Wage]欄位的數值
2 data['Wage'] = data['Wage'].str.extract('(\d+)').astype(int)
3 # 轉換幣值 new_Wage = old_Wage * 35 * 1000
4 data['Wage'] = data['Wage']*35*1000
```

- i. 使用 extract 函數擷取[Wage]欄位中的數字。
- ii. 將數值進行幣值轉換(歐元 = \*35、K = \*1000)

**(e) 轉換身高（15%）**

☐ 修改[Height]該欄的值

☐ 將身高單位轉換成公分

☐ 1 feet = 30.48 cm, 1 inch = 2.54 cm

☐ 並把欄位名稱[Height]改成[Height(cm)]

```
1 ### 將 feet 與 inch 分開並轉換成 cm
2 def parse_ht(x):
3     ht_ = x.split(" ")
4     ft_ = float(ht_[0])
5     in_ = float(ht_[1])
6     return (ft_*30.48) + (in_*2.54)
7
8 #使用函數 parse_ht 轉換單位
9 data["Height"] = data["Height"].apply(lambda x:parse_ht(x))
10
11 #使用函數 rename 轉換欄位名稱
12 data.rename(columns={'Height':'Height(cm)'}, inplace=True)
```

- i. 命名函數 parse\_ht，其功能為將 " a' b" 格式之字串，透過 split 函數轉為[a, b]之陣列，a 即是 feet，b 為 inch，並 return 成 cm
- ii. 透過 apply 與 parse\_ht 函數，將[Height]欄位的值轉換成公分。
- iii. 使用函數 rename 轉換欄位名稱成 Height(cm)

## (f) 轉換體重（15%）

- 修改[weight]該欄的值
- 將體重單位轉換成公斤
- 1 lbs = 0.4536 kg
- 並把欄位名稱[Weight]改成[Weight(kg)]

```
1 # 擷取[Weight]欄位的數值
2 data['Weight'] = data['Weight'].str.extract('(\d+)').astype(int)
3
4 # 轉換幣值 new_Weight = old_Weight * 0.4536
5 data['Weight'] = data['Weight']*0.4536
6
7 #使用函數 rename 轉換欄位名稱
8 data.rename(columns={'Weight':'Weight(kg)'}, inplace=True)
```

- i. 使用 extract 函數擷取[Weight]欄位中的數字。
- ii. 將數值進行體重單位轉換(公斤 = \*0.4536)
- iii. 使用函數 rename 轉換欄位名稱成 Weight(kg)

## (g) 匯出CSV（10%）

- 檔名為: 學號datanew.csv

```
1 data.to_csv('105403048datanew.csv')
```

- i. 使用 to\_csv 函數將 dataframe 匯出成 csv 檔。