

運動數據與球員薪資的關係

報告者: 資工系113級洪盛益

問題(Question)

- 球員薪資與戰績是否有關係?
- 球員薪資與特定戰績的相關性?(圖表)
- 球員薪資與數據實際上的意義是否吻合?

資料集(Data)

- 球員薪資表:

<https://www.usatoday.com/sports/mlb/salaries>
格式:html plaintext->python dictionary->txt

- 球員ID表:

<https://www.smartfantasybaseball.com/tools>
格式:mircrosoft xlsx->pandas dataframe->txt

資料集(Data)

- 球員數據:

<https://www.mlb.com>

格式:html plaintext->python dictionary->txt

實作(Implementation)

爬蟲

```
def getContent(url):  
    response=rq.get(url)  
    soup=bs4(response.text, "lxml")  
    return soup
```

```
def getSalary(num):  
    url="https://www.usatoday.com/sports/mlb/salaries/"  
    content=getContent(url)  
    table=content.find_all("tr")[1:num+1]  
    resultLIST=[]  
    for i in table:  
        name=i.find(class_="sp-details-open").text.strip()  
        salary=int(i.find(class_="salary").text.strip()).replace(",","")  
        resultLIST.append({"Name":name, "Salary":salary})  
    return resultLIST
```

實作(Implementation)

寫檔(WriteFile)

```
def readtxt(path):  
    with open(path, "r") as f:  
        result=f.read()  
    return result
```

讀檔(ReadFile)

```
def writetxt(playerLIST):  
    with open('output.txt', "w", newline='\n') as f:  
        for i in playerLIST:  
            f.write(json.dumps(i)+"<my_cutting_tag>")
```

實作(Implementation)

把字串轉成列表

```
def txt_to_list(txt):
    plaintext=txt.split("<my_cutting_tag>")[:-1]
    resultLIST=[]
    for i in plaintext:
        resultLIST.append(eval(i))
    return resultLIST
```

主函式

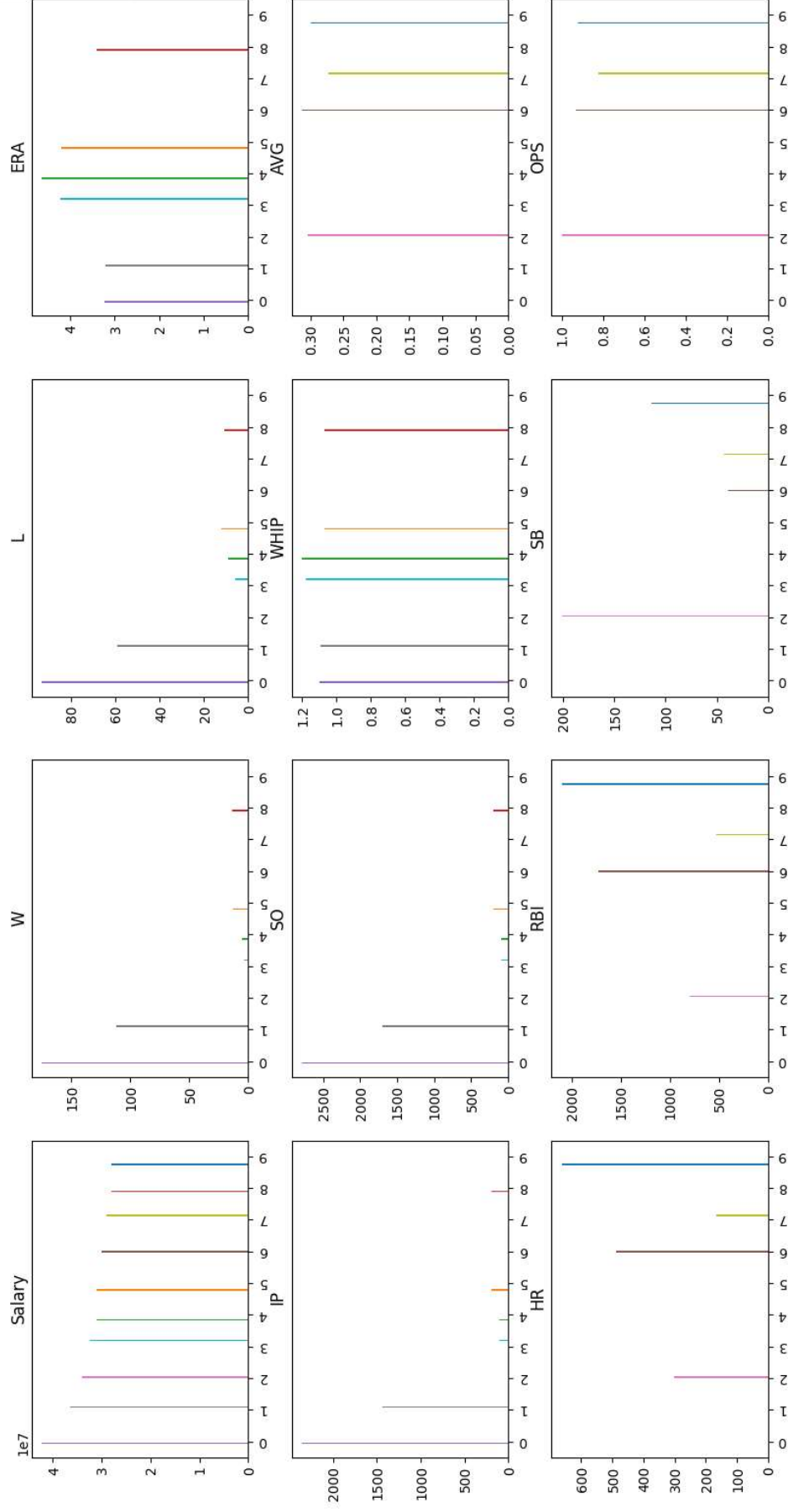
```
if __name__ == "__main__":
    txt=readtxt("./output.txt")
    playerLIST=txt_to_list(txt)
    playerDF=pd.DataFrame(playerLIST)
    print(playerDF)
    makePlot(playerDF)
```

實作(Implementation)

畫表格

```
def makePlot(DF):  
    cnt=0  
    col=DF.columns[1:]  
    labels=DF["Name"].values.tolist()  
    fig, axes = plt.subplots(nrows=(len(col)//4 if len(col)%4  
                             for i in col:  
                        DF.pivot(values=i, columns='Name').plot(kind='bar', ax=  
                            cnt+=1  
    fig.legend(labels)  
    plt.show()
```


結果分析(Analysis)



結果分析(Analysis)

- 球員薪資與戰績是否有關係?
根據圖表顯示，數據與部分戰績有有關係
- 球員薪資與特定戰績的相關性?
正相關:W,L,IP,SO
負相關:HR,RBI
凸起:ERA,WHIP
看不出來:AVG,OPS,SB

結果分析(Analysis)

- 球員薪資與數據實際上的意義是否吻合？
勝場敗場:經驗
IP:投球局數，同上，經驗。
RBI:打點，送了多少位球員回本壘。
ERA:防禦率，越低越好。