

教學方法資料分析

邱鈺婷

科技應用與人力資源學系 116 級

41271124H

一、研究目的

本研究旨在分析不同教學方法（A、B、C）對四個指標（A、B、C、D）數據的影響。我們使用單因子變異數分析（ANOVA）來檢驗不同教學方法是否對數據產生顯著影響。

二、研究方法

使用單因子變異數分析（ANOVA）來檢驗每個欄位（A、B、C、D）的變異數。將教學方法（A、B、C）作為因子，檢查其是否會顯著影響每項欄位的均值。

三、研究內容

產生一個 DataFrame，包含四個指標（A、B、C、D）的數據，以及 E 欄位不同教學方法（A、B、C）的數據，接著使用 ANOVA 進行分析，如圖 1 為產生數值的分析結果。

| | sum_sq | df | F | PR(>F) |
|----------------------------|------------|------|----------|----------|
| E | 1.474842 | 2.0 | 0.697349 | 0.500382 |
| Residual | 102.573892 | 97.0 | NaN | NaN |
| 未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。 | | | | |
| | sum_sq | df | F | PR(>F) |
| E | 0.101981 | 2.0 | 0.053878 | 0.947576 |
| Residual | 91.801719 | 97.0 | NaN | NaN |
| 未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。 | | | | |
| | sum_sq | df | F | PR(>F) |
| E | 1.082597 | 2.0 | 0.564808 | 0.570327 |
| Residual | 92.962480 | 97.0 | NaN | NaN |
| 未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。 | | | | |
| | sum_sq | df | F | PR(>F) |
| E | 7.529061 | 2.0 | 3.854129 | 0.024511 |
| Residual | 94.745012 | 97.0 | NaN | NaN |
| 拒絕原假設：不同教學方法之間的均值存在顯著差異。 | | | | |

圖 1.四個指標的數據結果

1. A 欄位

ANOVA 結果顯示，p-value = 0.500382，大於顯著性水平 0.05，因此不同教學方法之間的均值沒有顯著差異。

2. B 欄位

ANOVA 結果顯示，p-value = 0.947576，大於顯著性水平 0.05，

因此不同教學方法之間的均值沒有顯著差異。

3. C 欄位

ANOVA 結果顯示， $p\text{-value} = 0.570327$ ，大於顯著性水平 0.05，因此不同教學方法之間的均值沒有顯著差異。

4. D 欄位

ANOVA 結果顯示， $p\text{-value} = 0.024511$ ，小於顯著性水平 0.05，因此我們拒絕原假設，並認為不同教學方法之間的均值存在顯著差異。

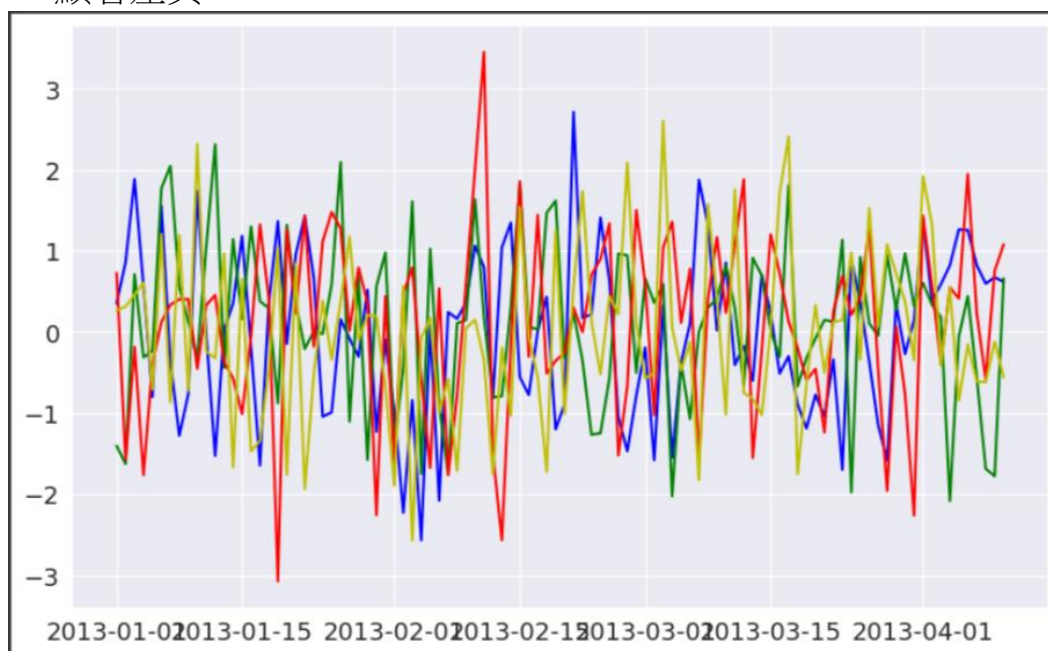


圖 2.四個指標數據的折線圖

根據 ANOVA 結果，A、B、C 三個欄位的數據在不同教學方法之間沒有顯著差異，這表明教學方法對這些欄位的影響較小。然而，D 欄位的數據在不同教學方法之間有顯著差異，這可能表明教學方法對 D 欄位的數據有更強的影響。

四、 結論

本研究顯示，不同教學方法對某些指標（如 D 欄位）具有顯著影響，但對其他指標（如 A、B 和 C 欄位）並未顯示出顯著差異。未來的研究可以進一步探討這些指標的具體性質及其對學習結果的影響。

五、 補充資料

```
import numpy as np
import pandas as pd
```

```

import statsmodels.api as sm
from statsmodels.formula.api import ols

# 建立原始 DataFrame
dates = pd.date_range("20130101", periods=100)
df = pd.DataFrame(np.random.randn(100, 4), index=dates,
columns=list("ABCD"))

# 新增 E 欄位
# 1-33 為 "A", 34-66 為 "B", 67-100 為 "C"
df['E'] = ["A"] * 33 + ["B"] * 33 + ["C"] * 34

# 使用 statsmodels 進行 ANOVA 檢定
# 分析 "A" 欄位的變異數 (可以換成 'B', 'C', 'D' 欄位進行分析)
model1 = ols('A ~ E', data=df).fit()
model2 = ols('B ~ E', data=df).fit()
model3 = ols('C ~ E', data=df).fit()
model4 = ols('D ~ E', data=df).fit()
anova_table1 = sm.stats.anova_lm(model1, typ=2)
anova_table2 = sm.stats.anova_lm(model2, typ=2)
anova_table3 = sm.stats.anova_lm(model3, typ=2)
anova_table4 = sm.stats.anova_lm(model4, typ=2)

# 輸出 ANOVA 表
print(anova_table1)

# 結果解釋
alpha = 0.05
p_value = anova_table1["PR(>F)"][0]
if p_value < alpha:
    print("拒絕原假設：不同教學方法之間的均值存在顯著差異。")
else:
    print("未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。")

# 輸出 ANOVA 表
print(anova_table2)

# 結果解釋

```

```

alpha = 0.05
p_value = anova_table2["PR(>F)"][0]
if p_value < alpha:
    print("拒絕原假設：不同教學方法之間的均值存在顯著差異。")
else:
    print("未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。")

# 輸出 ANOVA 表
print(anova_table3)

# 結果解釋
alpha = 0.05
p_value = anova_table3["PR(>F)"][0]
if p_value < alpha:
    print("拒絕原假設：不同教學方法之間的均值存在顯著差異。")
else:
    print("未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。")

# 輸出 ANOVA 表
print(anova_table4)

# 結果解釋
alpha = 0.05
p_value = anova_table4["PR(>F)"][0]
if p_value < alpha:
    print("拒絕原假設：不同教學方法之間的均值存在顯著差異。")
else:
    print("未能拒絕原假設：不同教學方法之間的均值沒有顯著差異。")

import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

# 繪製折線圖
plt.figure(figsize=(10, 6))
plt.plot(df.index, df['A'], label='A', color='b')
plt.plot(df.index, df['B'], label='B', color='g')

```

```
plt.plot(df.index, df['C'], label='C', color='r')
plt.plot(df.index, df['D'], label='D', color='y')

# 添加標題和標籤
plt.title('A, B, C, D 欄位的折線圖')
plt.xlabel('日期')
plt.ylabel('數值')
plt.legend()

# 顯示圖形
plt.show()
```