

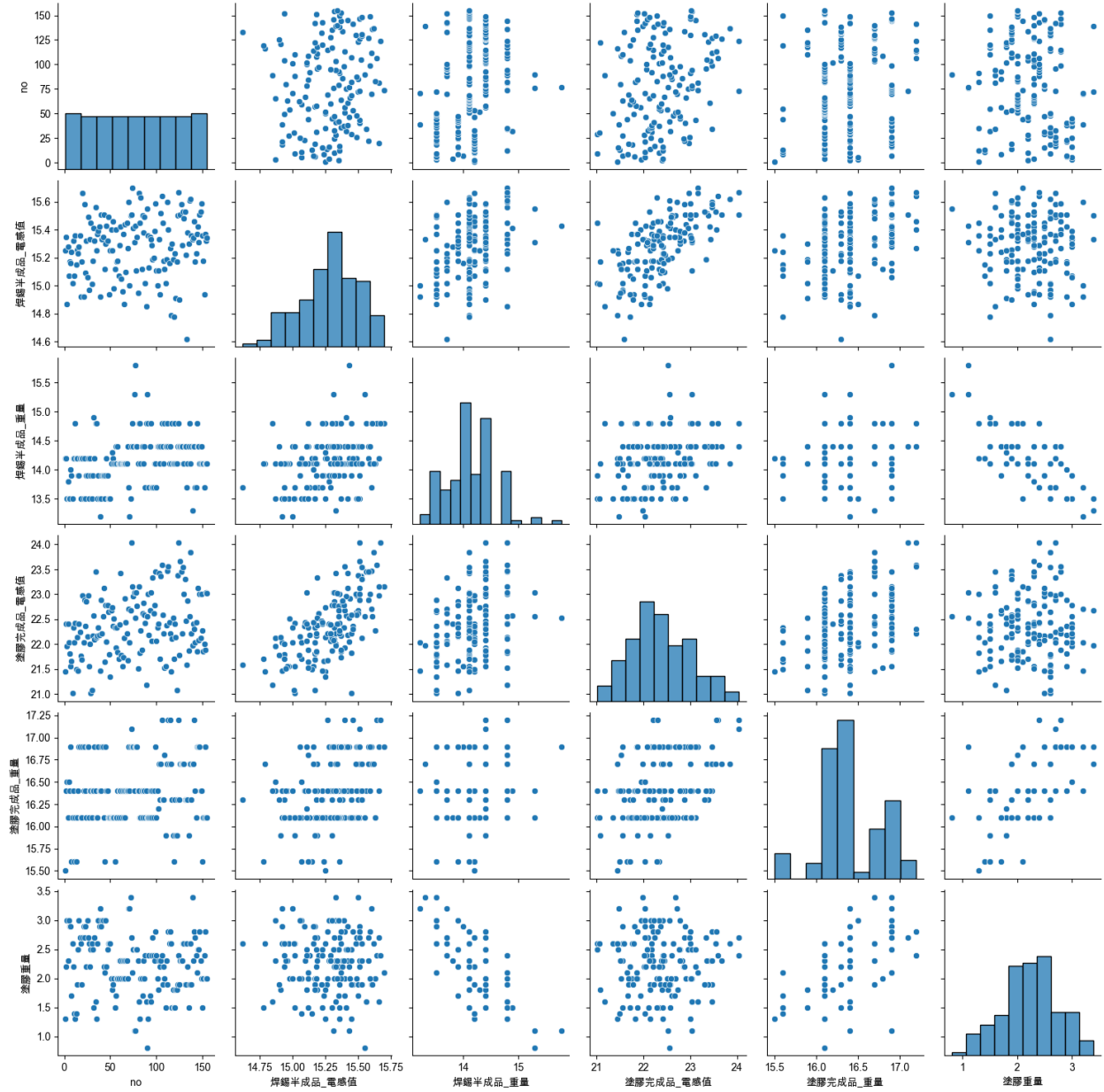
資料初步分析

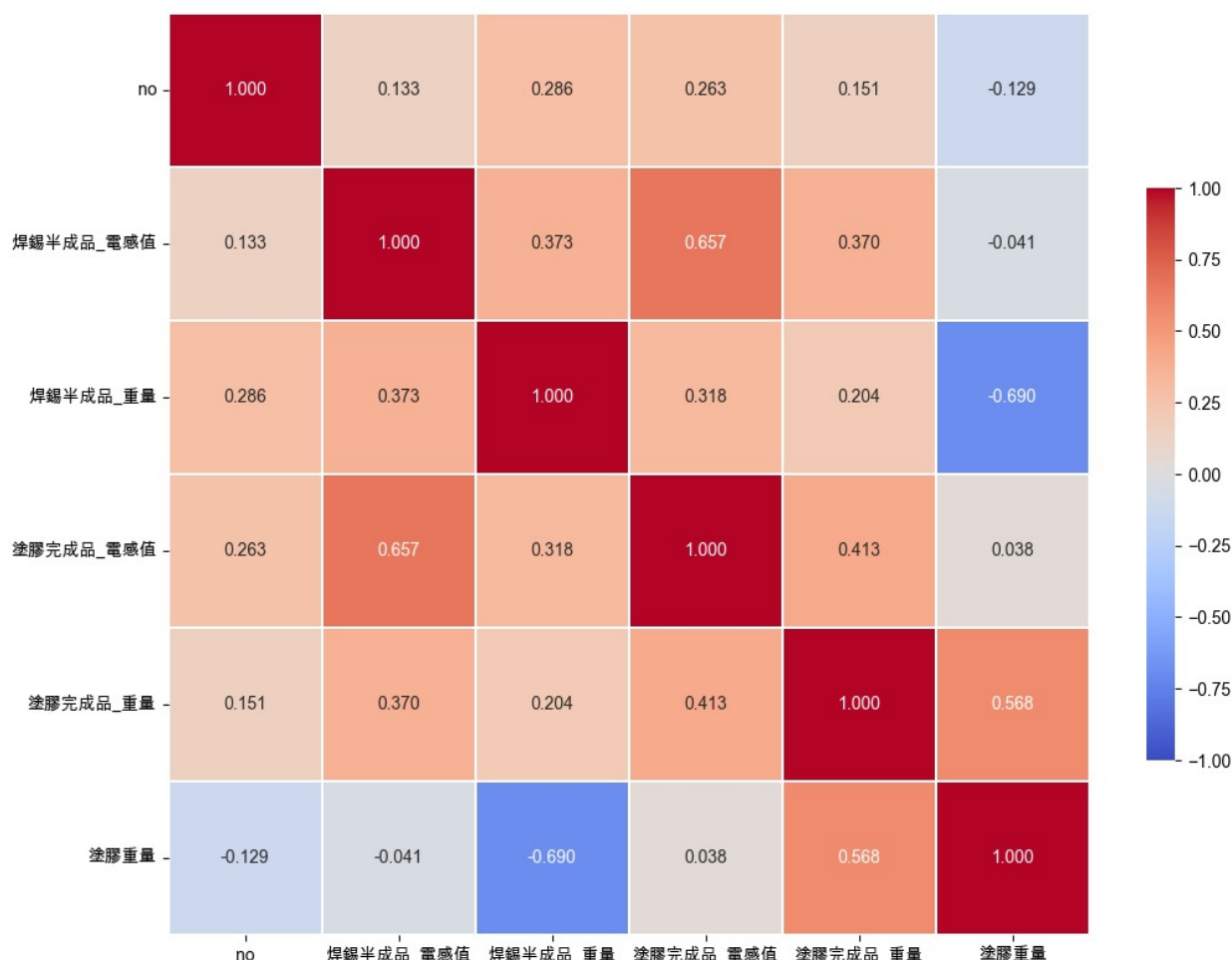
```
# 載入所需套件
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
import statsmodels.api as sm
from scipy.fft import fft
from statsmodels.tsa.stattools import adfuller
from sklearn.preprocessing import PolynomialFeatures

# 中文字體設定
plt.rcParams['font.sans-serif'] = ['Microsoft YaHei']

# 讀取資料為dataframe
df = pd.read_csv('/Users/hanmingcheng/Desktop/dataraw_20241031.csv',
encoding='big5')
plt.rcParams['font.sans-serif'] = ['Arial Unicode Ms']
# 用scatter plot 畫兩兩之間的關係
sns.pairplot(df[df.columns])
plt.show()

# 相關係數矩陣畫熱圖
correlation_matrix = df[df.columns].corr()
plt.figure(figsize=(12, 12))
sns.heatmap(correlation_matrix, annot=True, fmt=".3f",
cmap='coolwarm', square=True, cbar_kws={'shrink': .5}, vmin=-1,
vmax=1, linewidths=0.1, linecolor='white', center=0)
plt.xticks(rotation=0)
plt.yticks(rotation=0)
plt.show()
```





數據線性關係分析

數據分析依據

- 判斷標準：我們將線性相關性絕對值大於 0.5 的情況，視為具有顯著線性關係。
- 工具：以散佈矩陣圖和相關係數熱圖為基礎進行分析。

顯著線性關係（|相關係數| > 0.5）

以下是發現的顯著線性相關性：

- 焊鍋半成品_電感值 與 塗膠完成品_電感值
 - 相關係數：**0.657**
 - 描述：焊鍋半成品的電感值與塗膠完成品的電感值具有中強正相關，可能存在連續加工過程的相關性。
- 焊鍋半成品_重量 與 塗膠重量
 - 相關係數：**-0.690**

- 描述：焊鍋半成品的重量與塗膠重量呈現中強負相關，可能反映了重量在加工中某種負向轉換的特性。
3. 塗膠完成品_重量 與 塗膠重量
- 相關係數：**0.568**
 - 描述：塗膠完成品的重量與塗膠重量呈現中等正相關，可能反映了塗膠過程中重量的累積效應。
-

無顯著線性關係 ($|\text{相關係數}| \leq 0.5$)

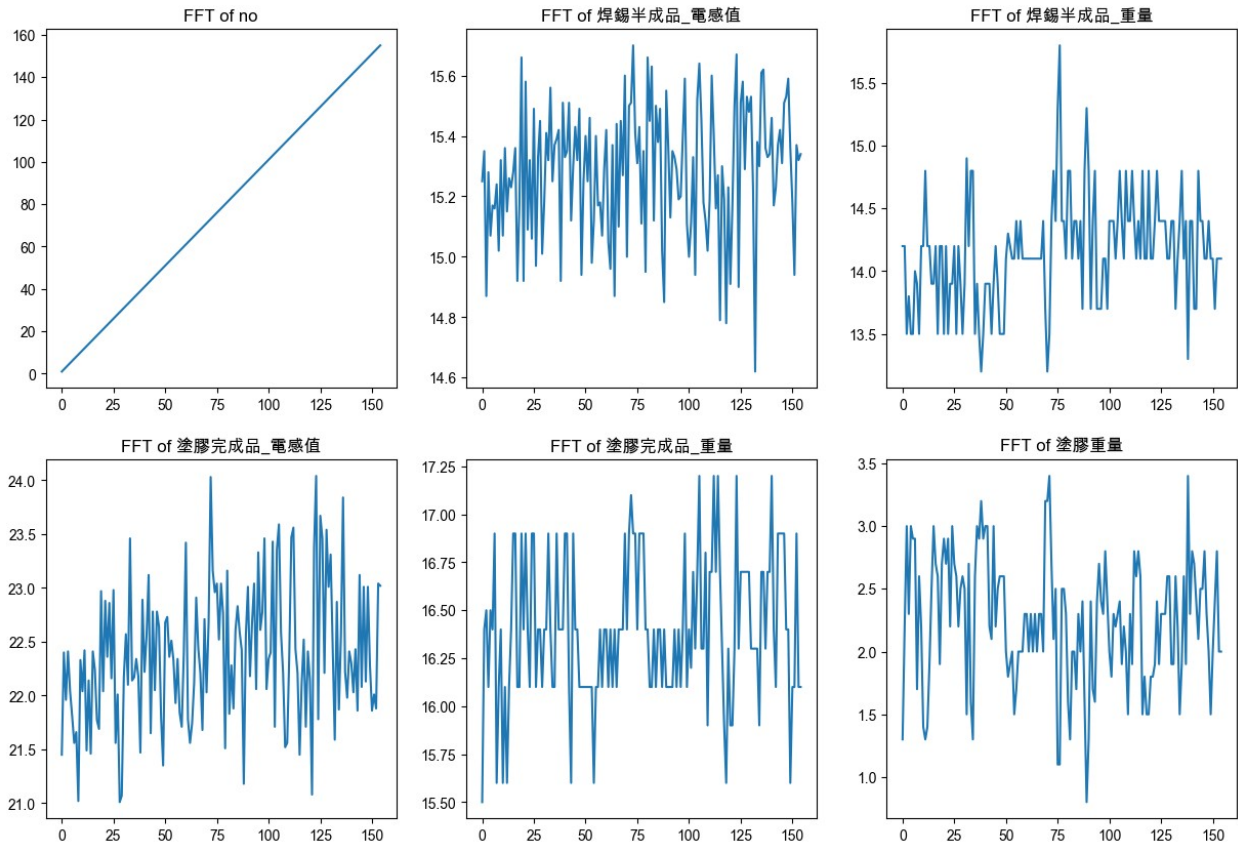
以下變數對之間未發現顯著的線性相關性：

- no 與其他變數
 - 焊鍋半成品_重量 與 焊鍋半成品_電感值
 - 塗膠完成品_重量 與 塗膠完成品_電感值
-

總結

- 具有線性相關的變數組合中，電感值與重量之間的關係顯示出某種可能的加工過程影響。

```
# FFT 畫頻譜圖(時序)
fig, axes = plt.subplots(2, 3, figsize=(15, 10))
for i, col in enumerate(df.columns):
    fft_values = fft(df[[col]])
    axes[i // 3, i % 3].plot(np.abs(fft_values))
    axes[i // 3, i % 3].set_title(f"FFT of {col}")
plt.show()
```



FFT 分析圖解

展示了多個數據集的 FFT 分析結果。以下是對每張圖的觀察與解釋。

FFT 與週期性的關係

FFT（快速傅立葉變換）是一種將數據從時域轉換到頻域的工具，用於檢測數據中的周期性（cycle）。其關係如下：

週期性特徵：

- 如果數據具有明顯的週期性，FFT 分析的頻譜中會顯示對應於該週期的顯著峰值。

頻率與週期關係：

- 頻率 (f) 可通過峰值位置獲得，週期長度為 ($T = 1/f$)。

應用場景：

- FFT 是識別信號中隱含週期性的重要方法，廣泛用於信號處理、設備診斷和數據分析等領域。

1. FFT of no

- 圖解：
 - 圖中呈現出一條線性增長的趨勢，這表明這並非典型的 FFT 頻譜。
 - 可能數據是非週期性的，或僅包含強烈的低頻趨勢。
-

2. FFT of 焊鍋半成品_電感值

- 圖解：
 - 頻譜的波動較為隨機，且沒有明顯的尖峰。
 - 表示數據中沒有顯著的週期性，可能主要是噪聲。
-

3. FFT of 焊鍋半成品_重量

- 圖解：
 - 頻譜中似乎有一些較高的峰值，但不夠明顯。
 - 表明可能存在低頻趨勢或長期變化成分，但週期性特徵不突出。
-

4. FFT of 塗膠完成品_電感值

- 圖解：
 - 頻譜波動隨機且沒有明顯尖峰。
 - 表明數據中沒有顯著的週期性。
-

5. FFT of 塗膠完成品_重量

- 圖解：
 - 頻譜中出現了一些間歇性的較高幅度，但整體沒有顯著尖峰。
 - 需要進一步分析，以確定是否存在潛在的週期性成分。
-

6. FFT of 塗膠重量

- 圖解：
 - 頻譜分佈隨機，且幅值較小。
 - 通常這表明數據中沒有明顯的週期性成分。
-

總結

- 由圖展示的 FFT 頻譜中，多數數據集未表現出明顯的週期性特徵。
 - 對於個別頻譜中可能的尖峰（如焊鍋半成品_重量），可進一步聚焦觀察其頻率和週期特性。
-

ADF 檢定

```
for i, col in enumerate(df.columns):  
    adf_values = adfuller(df[[col]].values.flatten())  
    print(col)  
    print(f'ADF Statistic : {adf_values[0]}')  
    print(f'p-value : {adf_values[1]:.6f}')  
    print(f'Critical Values : {adf_values[4]}')  
    print('')
```

no

ADF Statistic : -3.0984970720361744

p-value : 0.026665

Critical Values : {'1%': -3.474714913481481, '5%': -2.881008708148148,
'10%': -2.5771508444444446}

焊錫半成品_電感值

ADF Statistic : -12.976201619726064

p-value : 0.000000

Critical Values : {'1%': -3.473542528196209, '5%': -2.880497674144038,
'10%': -2.576878053634677}

焊錫半成品_重量

ADF Statistic : -4.93138955630217

p-value : 0.000030

Critical Values : {'1%': -3.474120870218417, '5%': -2.880749791423677,
'10%': -2.5770126333102494}

塗膠完成品_電感值

ADF Statistic : -10.990471159437133

p-value : 0.000000

Critical Values : {'1%': -3.473542528196209, '5%': -2.880497674144038,
'10%': -2.576878053634677}

塗膠完成品_重量

ADF Statistic : -9.53177118672849

p-value : 0.000000

Critical Values : {'1%': -3.473542528196209, '5%': -2.880497674144038,
'10%': -2.576878053634677}

塗膠重量

ADF Statistic : -7.946327923649558

p-value : 0.000000

Critical Values : {'1%': -3.473542528196209, '5%': -2.880497674144038,
'10%': -2.576878053634677}

ADF 檢定與週期性分析

ADF 檢定的意義

ADF (Augmented Dickey-Fuller) 檢定用於測試數據是否具有單位根，目的是判斷數據是否平穩：

1. **ADF Statistic**：檢定統計量，與臨界值比較來做出結論。
2. **p-value**：檢定的 p 值，用於判斷是否拒絕虛無假設（存在單位根）。
3. **Critical Values**：臨界值，對應 1%、5%、10% 的顯著性水平。

指標：

- 如果 $ADF\ Statistic < Critical\ Value$ 或 $p\text{-value} < \text{顯著性水平}$ （如 0.05），則可以拒絕虛無假設，數據是平穩的。
 - 如果數據是非平穩的，可能存在趨勢、週期性或隨機遊走特徵。
-

檢定結果與數據解讀

1. no

- **ADF Statistic:** -3.09
- **p-value:** 0.0266
- 結論：拒絕虛無假設（顯著性水平 5%）。數據接近平穩，但需進一步檢查是否存在趨勢或週期性特徵。

2. 焊鍋半成品_電感值

- **ADF Statistic:** -12.97
- **p-value:** 0.0000
- 結論：數據是平穩的，沒有明顯的隨機遊走特徵，可能存在內在規律（如週期性或短期波動）。

3. 焊鍋半成品_重量

- **ADF Statistic:** -4.93
- **p-value:** 0.00003
- 結論：數據是平穩的，可能具有短期變化的特徵。

4. 塗膠完成品_電感值

- **ADF Statistic:** -10.99

- **p-value:** 0.0000
- 結論：數據平穩，可能存在週期性或特定規律。

5. 塗膠完成品_重量

- **ADF Statistic:** -9.53
 - **p-value:** 0.0000
 - 結論：數據是平穩的，可能存在週期性或趨勢變化。
-

關於週期性的討論

ADF 檢定與週期性的關聯

- ADF 檢定表明數據大多數是平穩的，這意味著它們可能不具有隨機遊走特徵。
- 平穩性並不直接說明數據是否具有週期性，需進一步檢查。

檢查週期性的方法

1. 自相關分析（ACF）：
 - 檢查數據在不同滯後（Lag）間的相關性。
 - 如果自相關圖出現明顯的重複模式（固定間隔的高峰），則表明存在週期性。
 2. 頻譜分析（FFT）：
 - 使用 FFT 分析數據的頻率成分。
 - 如果存在顯著頻率成分（尖峰），則表明數據具有對應的週期。
-

結論

- 數據多為平穩，意味著不存在隨機遊走特徵。
- 某些數據可能存在內在規律或週期性，但需進一步檢查。

迴歸模型建模測試

```
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt

# 載入資料
file_path = '/Users/hanmingcheng/Documents/python_vscode/產學/dataraw_20241031.csv'
data = pd.read_csv(file_path)
```


No. Observations: 155 AIC:
217.6
Df Residuals: 152 BIC:
226.7
Df Model: 2

Covariance Type: nonrobust

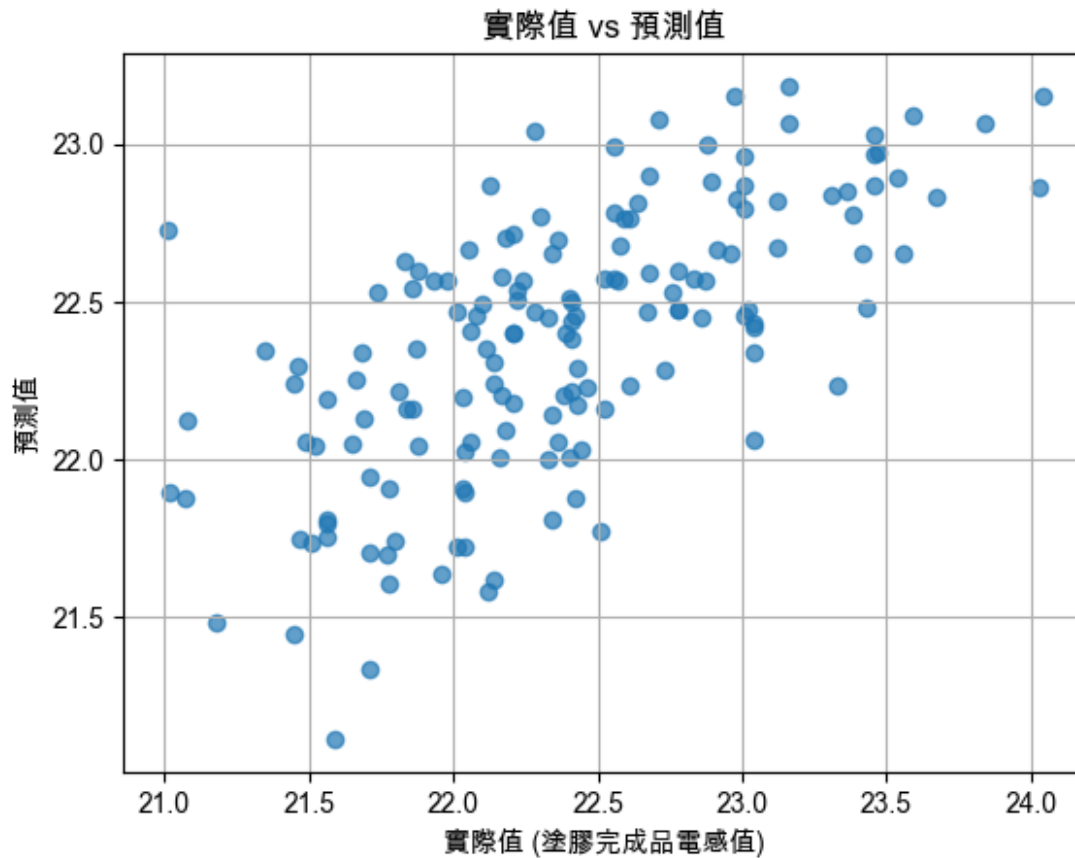
```
=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
const          -7.6238        2.770      -2.753      0.007     -13.096
-2.152
焊錫半成品電感值  1.9514        0.180     10.816      0.000        1.595
2.308
點膠重量        0.0812        0.077        1.061      0.290       -0.070
0.232
=====
=====
```

Omnibus: 1.665 Durbin-Watson:
1.724
Prob(Omnibus): 0.435 Jarque-Bera (JB):
1.247
Skew: -0.184 Prob(JB):
0.536
Kurtosis: 3.241 Cond. No.
1.11e+03

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.11e+03. This might indicate that there are strong multicollinearity or other numerical problems.



```
import statsmodels.api as sm
from statsmodels.stats.stattools import durbin_watson

# 建立交互項
data['交互項'] = data['焊錫半成品電感值'] * data['點膠重量']

# 定義新的解釋變數 (包含交互項)
X_interaction = data[['焊錫半成品電感值', '點膠重量', '交互項']]
X_interaction = sm.add_constant(X_interaction) # 添加常數項

# 重新建構回歸模型
model_interaction = sm.OLS(Y, X_interaction).fit()

# 顯示回歸結果摘要 (包含自由度)
summary = model_interaction.summary2()

# 計算Durbin-Watson 統計量
dw_statistic = durbin_watson(model_interaction.resid)

# 將Durbin-Watson 統計量添加到結果表中
additional_info = pd.DataFrame({
    'Durbin-Watson': [dw_statistic],
    '自由度 (殘差)': [int(model_interaction.df_resid)]
})
```

```

}).T
additional_info.columns = ['值']

# 顯示更新後的回歸結果摘要
print(summary)

# 顯示Durbin-Watson 統計量和自由度的表格
print("\nDurbin-Watson 統計量和自由度:")
print(additional_info)

```

```

Results: Ordinary least squares
=====
Model: OLS Adj. R-squared: 0.425
Dependent Variable: 塗膠完成品電感值 AIC: 219.5830
Date: 2024-11-24 18:46 BIC: 231.7567
No. Observations: 155 Log-Likelihood: -105.79
Df Model: 3 F-statistic: 38.88
Df Residuals: 151 Prob (F-statistic): 1.12e-18
R-squared: 0.436 Scale: 0.23535
-----

```

	Coef.	Std.Err.	t	P> t	[0.025	0.975]
const	-5.7081	13.2294	-0.4315	0.6667	-31.8467	20.4306
焊錫半成品電感值	1.8261	0.8649	2.1114	0.0364	0.1172	3.5350
點膠重量	-0.7508	5.6180	-0.1336	0.8939	-11.8509	10.3493
交互項	0.0544	0.3674	0.1481	0.8825	-0.6714	0.7802

```

-----
Omnibus: 1.686 Durbin-Watson: 1.727
Prob(Omnibus): 0.430 Jarque-Bera (JB): 1.267
Skew: -0.186 Prob(JB): 0.531
Kurtosis: 3.239 Condition No.: 14106
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.41e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Durbin-Watson 統計量和自由度：

值

Durbin-Watson 1.727191

自由度（殘差） 151.000000

```

import statsmodels.api as sm
import statsmodels.formula.api as smf
from statsmodels.stats.anova import anova_lm

```

建交互項的回歸模型，使用公式界面

```
model_interaction = smf.ols('Y ~ Q("焊錫半成品電感值") * Q("點膠重量")',
```

```
data=data).fit()

# 輸出ANOVA 表 (包含自由度)
anova_results = anova_lm(model_interaction, typ=2) # typ=2 表示使用
Type II 誤差平方和

from IPython.display import display

display(anova_results)
```

	Source	sum_sq	df	F	PR(>F)
0	Q("焊錫半成品電感值")	27.356099	1.000000	116.235629	0.000000
1	Q("點膠重量")	0.263447	1.000000	1.119384	0.291741
2	Q("焊錫半成品電感值"):Q("點膠重量")	0.005163	1.000000	0.021937	0.882453
3	Residual	35.537907	151.000000	NaN	NaN

視覺化觀察

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import cm
import seaborn as sns

# Set the style for seaborn
sns.set_style('whitegrid')

# Read the CSV data
file_path = '/Users/hanmingcheng/Documents/python vscode/產
學/dataraw_20241031.csv'
data = pd.read_csv(file_path)

# Define control limits
ucl = 22.99
lcl = 20.95

# 選擇colormap
cmap = plt.cm.viridis

# 使用原始的「點膠重量」作為顏色映射依據
weight_values = data['點膠重量']
weight_min = weight_values.min()
weight_max = weight_values.max()

# Create the plot
plt.figure(figsize=(10, 8))
scatter = plt.scatter(
```

```

data['焊錫半成品電感值'],
data['塗膠完成品電感值'],
c=weight_values,
cmap=cmap,
vmin=weight_min,
vmax=weight_max,
s=80,
edgecolors='black',
linewidth=0.5,
alpha=0.8
)
plt.rcParams['font.sans-serif'] = ['Arial Unicode Ms']
# Add horizontal lines for control limits
plt.axhline(y=lcl, color='red', linestyle='--', linewidth=1.5,
label='LCL')
plt.axhline(y=ucl, color='red', linestyle='--', linewidth=1.5,
label='UCL')

# Add a vertical line at x = 15.1
plt.axvline(x=15.1, color='blue', linestyle='--', linewidth=1.5)

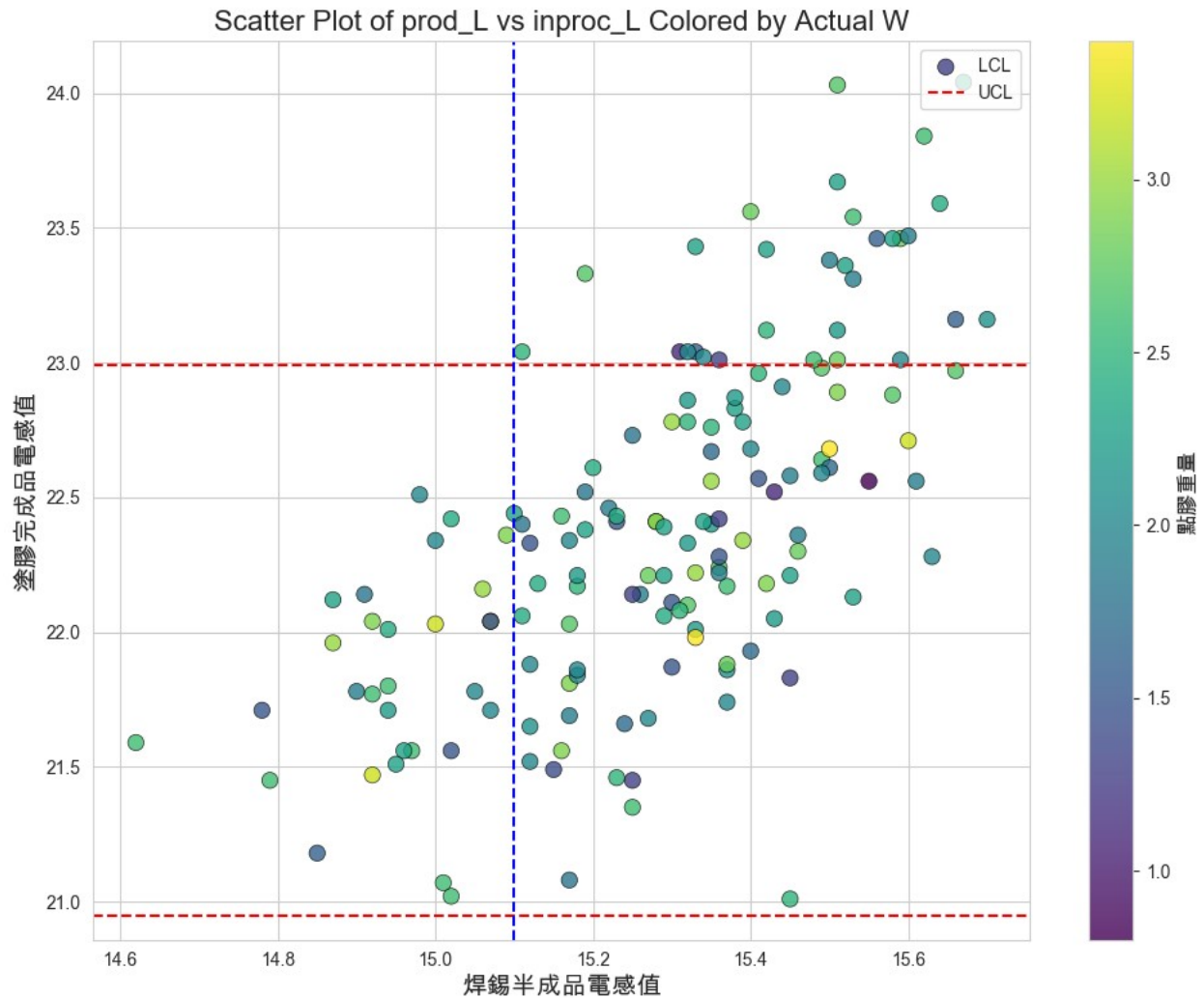
# Add colorbar to indicate actual weight values
cbar = plt.colorbar(scatter)
cbar.set_label('點膠重量', fontsize=12)

# Add labels and title
plt.xlabel('焊錫半成品電感值', fontsize=14)
plt.ylabel('塗膠完成品電感值', fontsize=14)
plt.title('Scatter Plot of prod_L vs inproc_L Colored by Actual W',
fontsize=16)

# Add a legend
plt.legend(['LCL', 'UCL'], loc='upper right')

# Improve layout and display the plot
plt.tight_layout()
plt.show()

```



```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# 設定Seaborn 的風格
sns.set_style('whitegrid')

# 讀取CSV 資料
file_path = '/Users/hanmingcheng/Documents/python vscode/產
學/dataraw_20241031.csv'
data = pd.read_csv(file_path)

# 定義控制界限
ucl = 22.99
lcl = 20.95

# 刪除 '焊錫半成品電感值' 小於15.1 的資料
```



```

filtered_data = data[data['焊錫半成品電感值'] >= 15.1]

# 將處理後的資料存成新的CSV 檔案
output_file_path = '/Users/hanmingcheng/Documents/python vscode/產學/處理 dataraw_20241031.csv'
filtered_data.to_csv(output_file_path, index=False)

# 選擇色彩映射
cmap = plt.cm.viridis

# 創建圖表
plt.figure(figsize=(10, 8))
scatter = plt.scatter(
    filtered_data['焊錫半成品電感值'],
    filtered_data['塗膠完成品電感值'],
    c=filtered_data['點膠重量'],
    cmap=cmap,
    s=80,
    edgecolors='black',
    linewidth=0.5,
    alpha=0.8
)

# 設定中文字體，避免亂碼
plt.rcParams['font.sans-serif'] = ['Arial Unicode MS']

# 添加控制界限的水平線
plt.axhline(y=lcl, color='red', linestyle='--', linewidth=1.5,
            label='LCL')
plt.axhline(y=ucl, color='red', linestyle='--', linewidth=1.5,
            label='UCL')

# 在x=15.1 處添加垂直線
plt.axvline(x=15.1, color='blue', linestyle='--', linewidth=1.5,
            label='x = 15.1')

# 添加坐標軸標籤和標題
plt.xlabel('焊錫半成品電感值', fontsize=14)
plt.ylabel('塗膠完成品電感值', fontsize=14)
plt.title('塗膠完成品電感值與焊錫半成品電感值的散佈圖 (以點膠重量著色)',
          fontsize=16)

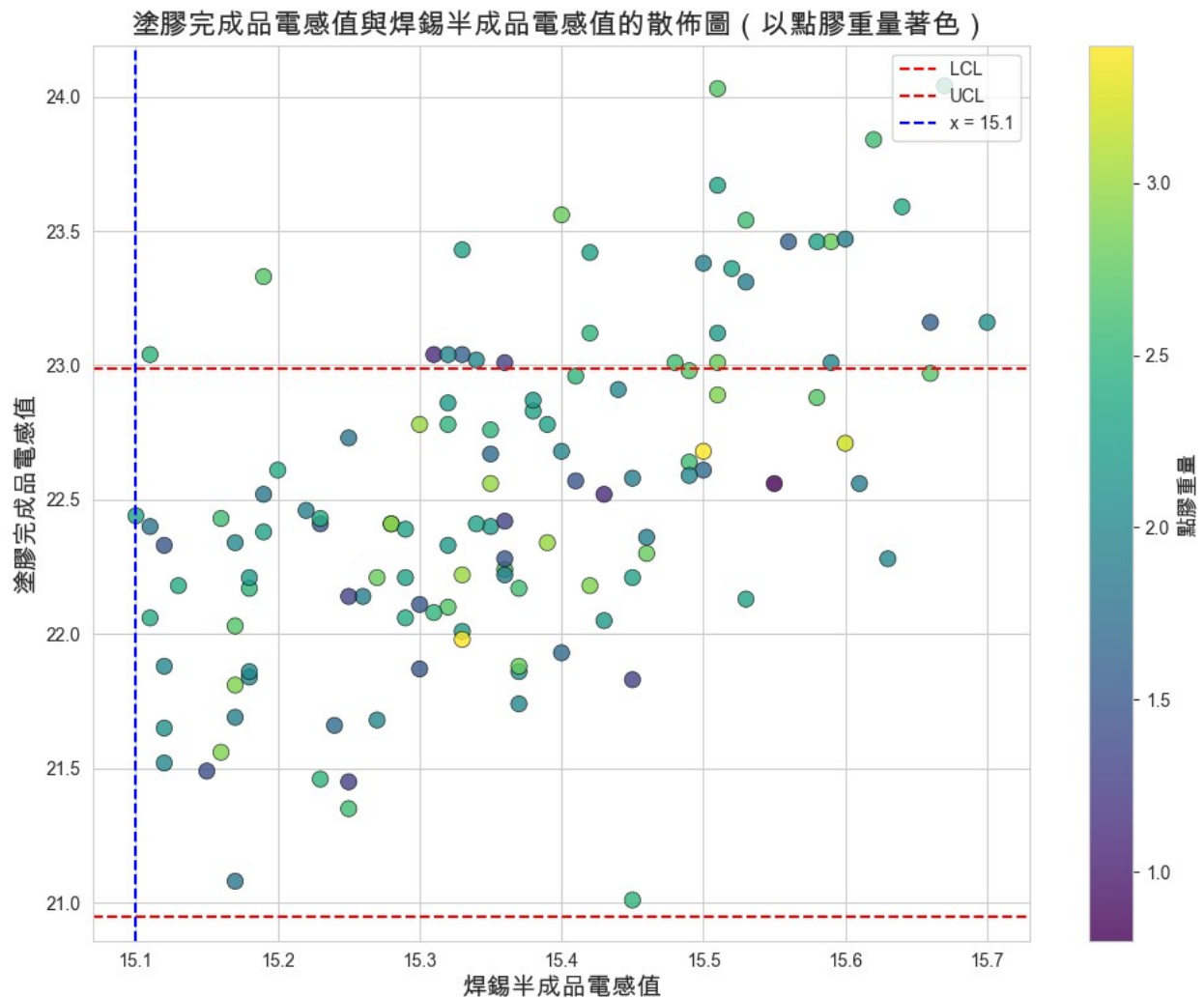
# 添加色彩條以指示 '點膠重量' 的實際值
cbar = plt.colorbar(scatter)
cbar.set_label('點膠重量', fontsize=12)

# 添加圖例
plt.legend(loc='upper right')

# 優化佈局並顯示圖表

```

```
plt.tight_layout()
plt.show()
```



資料清洗後再次迴歸分析

```
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt

# 載入資料
file_path = '/Users/hanmingcheng/Documents/python_vscode/產學/處理
dataraw_20241031.csv'
data = pd.read_csv(file_path)

# 確保數據格式正確
data['焊錫半成品電感值'] = pd.to_numeric(data['焊錫半成品電感值'],
```

```

errors='coerce')
data['點膠重量'] = pd.to_numeric(data['點膠重量'], errors='coerce')
data['塗膠完成品電感值'] = pd.to_numeric(data['塗膠完成品電感值'],
errors='coerce')

# 去除遺失值
data = data.dropna(subset=['焊錫半成品電感值', '點膠重量', '塗膠完成品電感
值'])

# 定義目標變數(Y) 和解釋變數(X)
X = data[['焊錫半成品電感值', '點膠重量']]
Y = data['塗膠完成品電感值']

# 添加常數項(用於回歸截距)
X = sm.add_constant(X)

# 建立回歸模型
model = sm.OLS(Y, X).fit()

# 顯示回歸結果摘要
print(model.summary())

# 預測值計算
data['預測值'] = model.predict(X)

# 繪製實際值與預測值的比較圖
plt.scatter(Y, data['預測值'], alpha=0.7)
plt.xlabel('實際值 (塗膠完成品電感值)')
plt.ylabel('預測值')
plt.title('實際值 vs 預測值')
plt.grid()
plt.show()

```

OLS Regression Results

```

=====
=====
Dep. Variable:          塗膠完成品電感值  R-squared:
0.347
Model:                  OLS      Adj. R-squared:
0.336
Method:                 Least Squares  F-statistic:
32.36
Date:                   Thu, 05 Dec 2024  Prob (F-statistic):
5.30e-12
Time:                   16:55:12    Log-Likelihood:
-89.462
No. Observations:      125    AIC:
184.9
Df Residuals:          122    BIC:

```

```

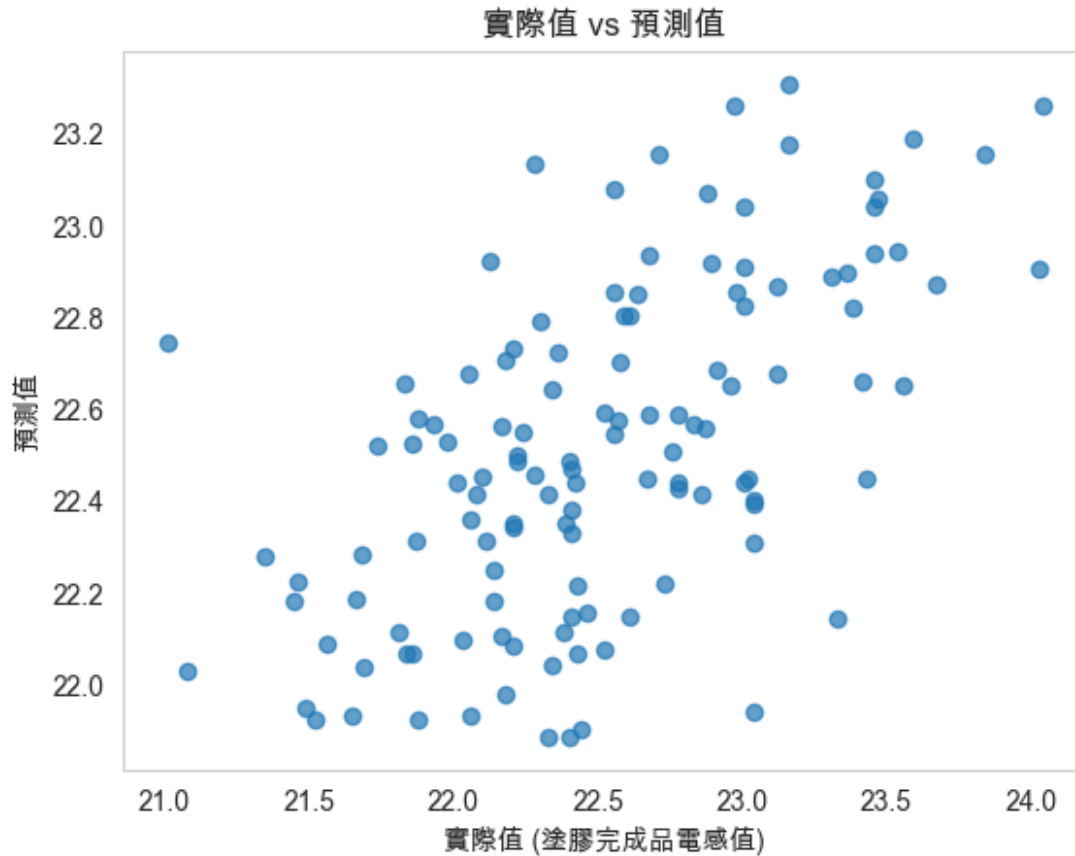
193.4
Df Model:                2

Covariance Type:        nonrobust

=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
const          -14.0695      4.596      -3.061      0.003     -23.167
-4.972
焊錫半成品電感值  2.3705      0.300       7.904      0.000       1.777
2.964
點膠重量        0.0767      0.089       0.860      0.392      -0.100
0.253
=====
=====
Omnibus:                1.199   Durbin-Watson:
1.791
Prob(Omnibus):          0.549   Jarque-Bera (JB):
0.754
Skew:                   -0.126   Prob(JB):
0.686
Kurtosis:               3.285   Cond. No.
1.60e+03
=====
=====

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
[2] The condition number is large, 1.6e+03. This might indicate that
there are
strong multicollinearity or other numerical problems.

```



```
import statsmodels.api as sm
from statsmodels.stats.stattools import durbin_watson

# 建立交互項
data['交互項'] = data['焊錫半成品電感值'] * data['點膠重量']

# 定義新的解釋變數 (包含交互項)
X_interaction = data[['焊錫半成品電感值', '點膠重量', '交互項']]
X_interaction = sm.add_constant(X_interaction) # 添加常數項

# 重新建構回歸模型
model_interaction = sm.OLS(Y, X_interaction).fit()

# 顯示回歸結果摘要 (包含自由度)
summary = model_interaction.summary2()

# 計算Durbin-Watson 統計量
dw_statistic = durbin_watson(model_interaction.resid)

# 將Durbin-Watson 統計量添加到結果表中
additional_info = pd.DataFrame({
    'Durbin-Watson': [dw_statistic],
    '自由度 (殘差)': [int(model_interaction.df_resid)]
})
```

```
}).T
additional_info.columns = ['值']

# 顯示更新後的回歸結果摘要
print(summary)

# 顯示Durbin-Watson 統計量和自由度的表格
print("\nDurbin-Watson 統計量和自由度:")
print(additional_info)
```

```
Results: Ordinary least squares
=====
Model: OLS Adj. R-squared: nan
Dependent Variable: 塗膠完成品電感值 AIC: -107.8769
Date: 2024-12-05 00:09 BIC: -110.4906
No. Observations: 2 Log-Likelihood: 55.938
Df Model: 1 F-statistic: nan
Df Residuals: 0 Prob (F-statistic): nan
R-squared: 1.000 Scale: inf
-----
              Coef.      Std.Err.      t      P>|t|      [0.025      0.975]
-----
焊錫半成品電感值 -1.9608      inf    -0.0000      nan      nan      nan
點膠重量      81.2576      inf     0.0000      nan      nan      nan
交互項      -2.1569      inf    -0.0000      nan      nan      nan
-----
Omnibus: nan Durbin-Watson: 0.005
Prob(Omnibus): nan Jarque-Bera (JB): 0.333
Skew: 0.000 Prob(JB): 0.846
Kurtosis: 1.000 Condition No.: 5333
=====
```

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The input rank is higher than the number of observations.
- [3] The condition number is large, 5.33e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Durbin-Watson 統計量和自由度：

值
Durbin-Watson 0.0053
自由度 (殘差) 0.0000

/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/statsmodels/stats/stattools.py:74: ValueWarning: omni_normtest is not valid with less than 8 observations; 2 samples were given.

warn("omni_normtest is not valid with less than 8 observations; %i " /Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/

```

site-packages/statsmodels/regression/linear_model.py:1718:
RuntimeWarning: divide by zero encountered in double_scalars
    return np.dot(wresid, wresid) / self.df_resid
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/
site-packages/statsmodels/regression/linear_model.py:1796:
RuntimeWarning: divide by zero encountered in divide
    return 1 - (np.divide(self.nobs - self.k_constant, self.df_resid)
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/
site-packages/statsmodels/regression/linear_model.py:1796:
RuntimeWarning: invalid value encountered in double_scalars
    return 1 - (np.divide(self.nobs - self.k_constant, self.df_resid)

```

```

import statsmodels.api as sm
import statsmodels.formula.api as smf
from statsmodels.stats.anova import anova_lm

```

建立交互項的回歸模型，使用公式界面

```

model_interaction = smf.ols('Y ~ Q("焊錫半成品電感值") * Q("點膠重量")',
data=data).fit()

```

輸出ANOVA 表(包含自由度)

```

anova_results = anova_lm(model_interaction, typ=2) # typ=2 表示使用
Type II 誤差平方和

```

```

from IPython.display import display

```

```

display(anova_results)

```

	sum_sq	df	F	PR(>F)
Q("焊錫半成品電感值")	15.681705	1.0	61.976338	1.666144e-12
Q("點膠重量")	0.185573	1.0	0.733410	3.934728e-01
Q("焊錫半成品電感值"):Q("點膠重量")	0.009225	1.0	0.036459	8.488916e-01
Residual	30.616302	121.0	NaN	NaN

```

import pandas as pd
import statsmodels.formula.api as smf
from statsmodels.stats.anova import anova_lm

```

載入資料

```

file_path = '/Users/hanmingcheng/Documents/python_vscode/產學/處理
3dataraw_20241031.csv'
data = pd.read_csv(file_path)

```

確保數據格式正確

```

data['焊錫半成品電感值'] = pd.to_numeric(data['焊錫半成品電感值'],
errors='coerce')
data['點膠重量'] = pd.to_numeric(data['點膠重量'], errors='coerce')
data['塗膠完成品電感值'] = pd.to_numeric(data['塗膠完成品電感值'],

```

```

errors='coerce')

# 去除遺失值
data = data.dropna(subset=['焊錫半成品電感值', '點膠重量', '塗膠完成品電感值'])

# 建交互作用項並使用公式建立模型
formula = '塗膠完成品電感值 ~ 焊錫半成品電感值 * 點膠重量'
model = smf.ols(formula=formula, data=data).fit()

# 顯示回歸結果摘要
print(model.summary())

# 計算ANOVA 表格
anova_results = anova_lm(model, typ=2)

# 顯示ANOVA 表格
print("\nANOVA 表格:")
print(anova_results)

```

OLS Regression Results

```

=====
Dep. Variable:          塗膠完成品電感值  R-squared:
0.134
Model:                  OLS      Adj. R-squared:
0.101
Method:                 Least Squares  F-statistic:
4.061
Date:                  Wed, 04 Dec 2024  Prob (F-statistic):
0.00975
Time:                  23:54:11    Log-Likelihood:
-57.590
No. Observations:      83      AIC:
123.2
Df Residuals:          79      BIC:
132.9
Df Model:              3

Covariance Type:      nonrobust

```

```

=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
Intercept    -38.0230      41.454     -0.917      0.362    -120.536
44.490

```


焊錫半成品電感值	3.9236	2.707	1.449	0.151	-1.465
9.313					
點膠重量	16.9469	19.211	0.882	0.380	-21.292
55.186					
焊錫半成品電感值:點膠重量	-1.0973	1.255	-0.875	0.384	-
3.594	1.400				

```

=====
=====
Omnibus:                2.033    Durbin-Watson:
1.981
Prob(Omnibus):          0.362    Jarque-Bera (JB):
1.405
Skew:                   -0.278    Prob(JB):
0.495
Kurtosis:               3.313    Cond. No.
3.00e+04
=====
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 3e+04. This might indicate that there are strong multicollinearity or other numerical problems.

ANOVA 表格:

	sum_sq	df	F	PR(>F)
焊錫半成品電感值	2.473602	1.0	10.038791	0.002180
點膠重量	0.306412	1.0	1.243533	0.268172
焊錫半成品電感值:點膠重量	0.188490	1.0	0.764962	0.384431
Residual	19.465948	79.0	NaN	NaN