In [2]:

```
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as stats
import pandas as pd
```

Question a

- Is there evidence to support the claim that cotton content affects the mean tensile strength? Use α = 0.05 (Hint: Complete the ANOVA table and then comment).

In [3]:

```
# define matrix containing the values
rows=int(input("Enter number of rows"))
cols=int(input("Enter number of Columns"))
A=np.zeros((rows,cols))

# Enter values

for i in range(rows):
    for j in range(cols):
        A[i,j]=float(input(f"Enter the value of {i}th row and {j}th column"))

print("Matrix : ",A)
```

```
Matrix :  [[ 7.  7. 15. 11.  9.]
 [12. 17. 12. 18. 18.]
 [14. 19. 19. 18. 18.]
 [19. 25. 22. 19. 23.]
 [ 7. 10. 11. 15. 11.]]
```

In [31]:

```
# Calculate the sum of rows
Sum_of_rows=np.sum(A,axis=1)
Sum_of_rows
```

Out[31]:

```
array([ 49.,  77.,  88., 108.,  54.])
```

In [32]:

```
# Calculate the sum of Columns
Sum_of_cols=np.sum(A,axis=0)
Sum_of_cols
```

Out[32]:

```
array([59., 78., 79., 81., 79.])
```

In [33]:

```
Total_sum_of_elements=np.sum(Sum_of_cols,axis=0)
Total_sum_of_elements
```

Out[33]:

```
376.0
```

In [34]:

```
# Calculate Sum of Sqauares total
SST=0
for i in range(rows):
    for j in range(cols):
        SST+=A[i,j]**2

print(SST)
SST-=Total_sum_of_elements**2/(rows*cols)
print("Sum of total :",SST)
```

```
6292.0
Sum of total : 636.96
```

In [35]:

```
# Calculation on treatments
SStreatments=0
print(Sum_of_rows.shape)
for i in range(Sum_of_rows.shape[0]):
    SStreatments+=Sum_of_rows[i]**2
SStreatments/=Sum_of_cols.shape[0]
SStreatments-=Total_sum_of_elements**2/(rows*cols)
print("Sum of Treatments :",SStreatments)
```

```
(5,)
Sum of Treatments : 475.7600000000002
```

In [36]:

```
# Error
SSerror=SST-SStreatments
print("Error",SSerror)
```

```
Error 161.19999999999982
```

In [37]:

```
# Calculation of degrees of Freedom
dft=Sum_of_rows.shape[0]-1
dfw=rows*cols-dft-1
dftotal=rows*cols-1
```

In [38]:

```
# Calculation of Mean squares
mst=(SStreatments/dft)
mse=SSerror/dfw
```

In [39]:

```
# calculate the F statistic and p-value
# F-value
Ftreat=mst/mse
p_val = stats.f.sf(Ftreat, dft, dfw)
```

In [46]:

```
# create the ANOVA table
anova_table = pd.DataFrame({
    'Source of Variation': ['Between Treatments', 'Error (Within Treatments)', 'Tota
l'],
    'Sum of Squares': [SStreatments, SSerror, SST],
    'Degrees of Freedom': [dft, dfw, dftotal],
    'Mean Square': [mst, mse, np.nan],
    'F value': [Ftreat, np.nan, np.nan],
    'p-value': [p_val, np.nan, np.nan]
})

# set the index to the source column
anova_table.set_index('Source of Variation', inplace=True)

# display the ANOVA table
print(anova_table)
```

```
                         Sum of Squares  Degrees of Freedom  Mean Square
\
Source of Variation
Between Treatments               475.76                   4       118.94
Error (Within Treatments)        161.20                  20         8.06
Total                            636.96                  24          NaN

                            F value   p-value
Source of Variation
Between Treatments         14.756824  0.000009
Error (Within Treatments)        NaN       NaN
Total                            NaN       NaN
```

Comment :

- Since the resultant Fo=14.756 is greater than the F value obtained from the F table for alpha=0.05 v1=4 v2=20 is 2.8661. Also the p value is 0. As resultant Fo is > F value from table we reject Null Hypothesis that cotton % doesnot affect tensile strength.
- Thus we accept the alternate Hypothesis and can say that Percentage of cotton in fibre appears to have an effect on tensile strength.

Question b

- Analyze the residuals from this experiment and comment on model adequacy. (The plot must have the x-axis as Residuals and the y-axis as Normal Score, in BLUE colour).

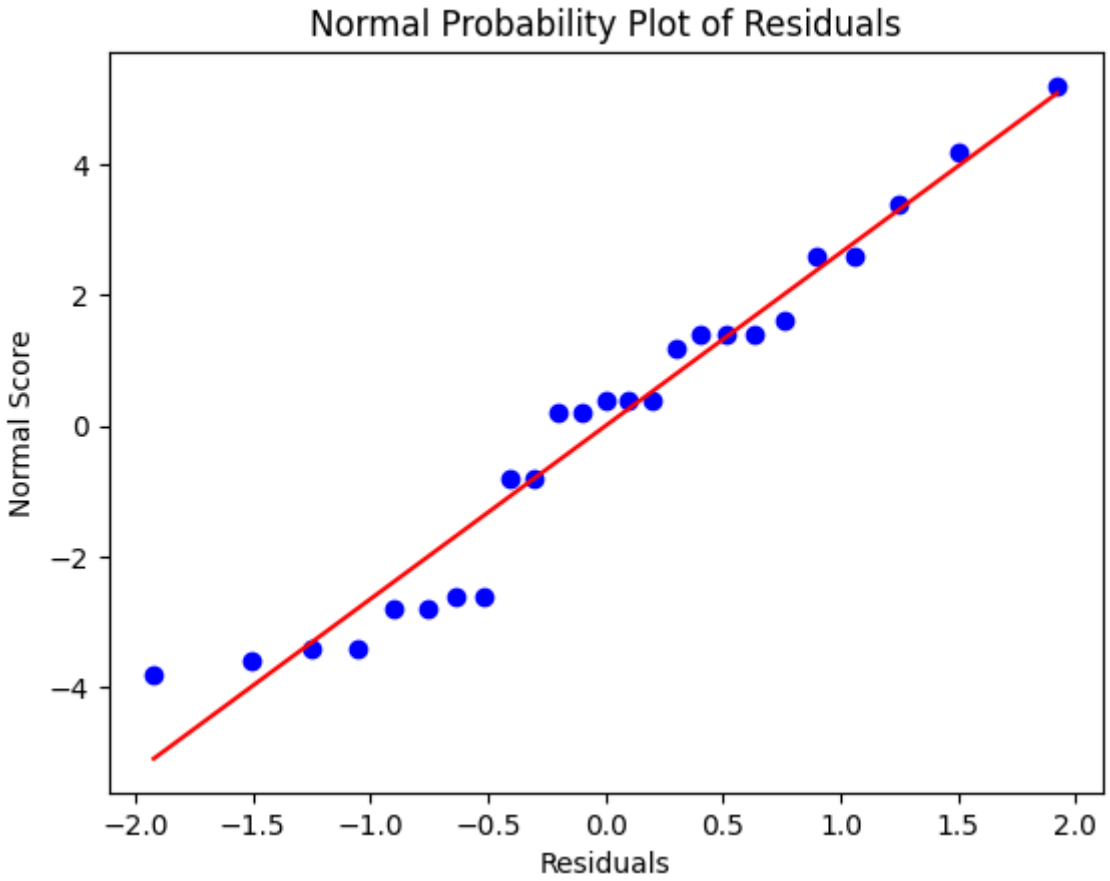
In [48]:

```
# create the residual plot
residuals = []
for i in range(rows):
    for j in range(cols):
        residuals.append(A[i,j]-Sum_of_rows[i]/Sum_of_cols.shape[0])  # Sum_of_rows[i]/
Sum_of_cols.shape[0]
#stats.probplot(residuals, plot=plt)
stats.probplot(residuals, dist="norm", plot=plt)

# Calculate normal scores


# set the title and labels
plt.title('Normal Probability Plot of Residuals')
plt.xlabel('Residuals')
plt.ylabel('Normal Score')

# show the plot
plt.show()
```



Comments :

- Based on the plot, we can see that the residuals appear to be roughly normally distributed, with a few outliers at the extremes. Therefore, the model appears to be adequate for analyzing the effect of cotton content on tensile strength.