

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

## 1. Import data and check top 6 rows.

```
In [2]: data = pd.read_csv("VAS_DATA.csv")
data.head(6)
```

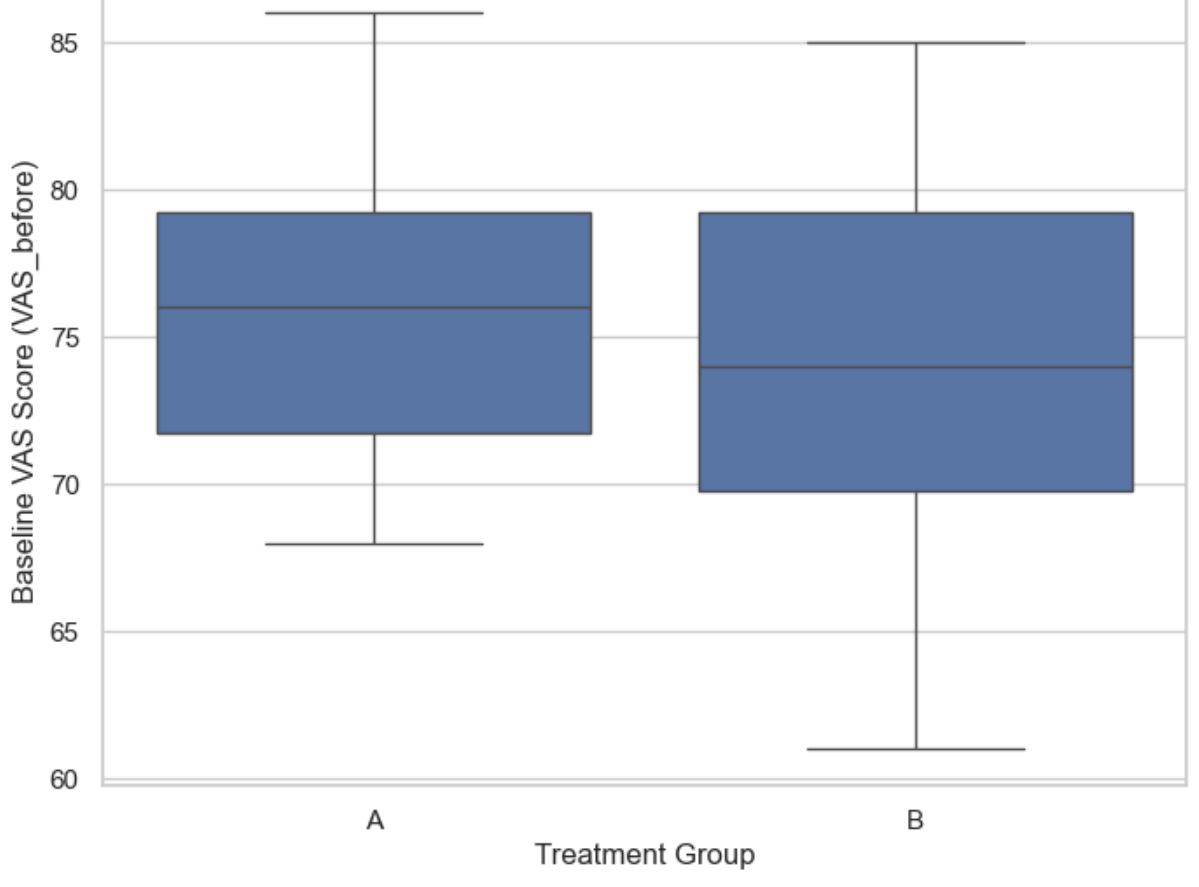
Out[2]:	Group	VAS_before	VAS_after
0	A	86	71
1	A	77	59
2	A	75	44
3	A	83	49
4	A	72	32
5	A	70	42

## 2. Visualize baseline VAS score (VAS\_before) by treatment group.

```
In [3]: # Create a boxplot using Seaborn
plt.figure(figsize=(8, 6))
sns.set(style="whitegrid")
sns.boxplot(x='Group', y='VAS_before', data=data)

# Set labels and title
plt.xlabel("Treatment Group")
plt.ylabel("Baseline VAS Score (VAS_before)")
plt.title("Baseline VAS Score by Treatment Group (Boxplot)")

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



## 3. Obtain measures of central tendency and variation for VAS\_before by treatment group.

```
In [4]: # Calculate and select n, mean, median, and std deviation
result = data.groupby('Group')['VAS_before'].describe()[[

# Rename the columns for clarity
result.columns = ['Count (n)', 'Mean', 'Median', 'Std Dev
result
```

Out[4]:	Count (n)	Mean	Median	Std Deviation
Group				
A	16.0	76.0000	76.0	5.561774
B	16.0	74.3125	74.0	6.680506

## 4. Derive a new variable- Change from baseline after 3 days of treatment

```
In [5]: data['Change'] = data['VAS_before'] - data['VAS_after']
data.head()
```

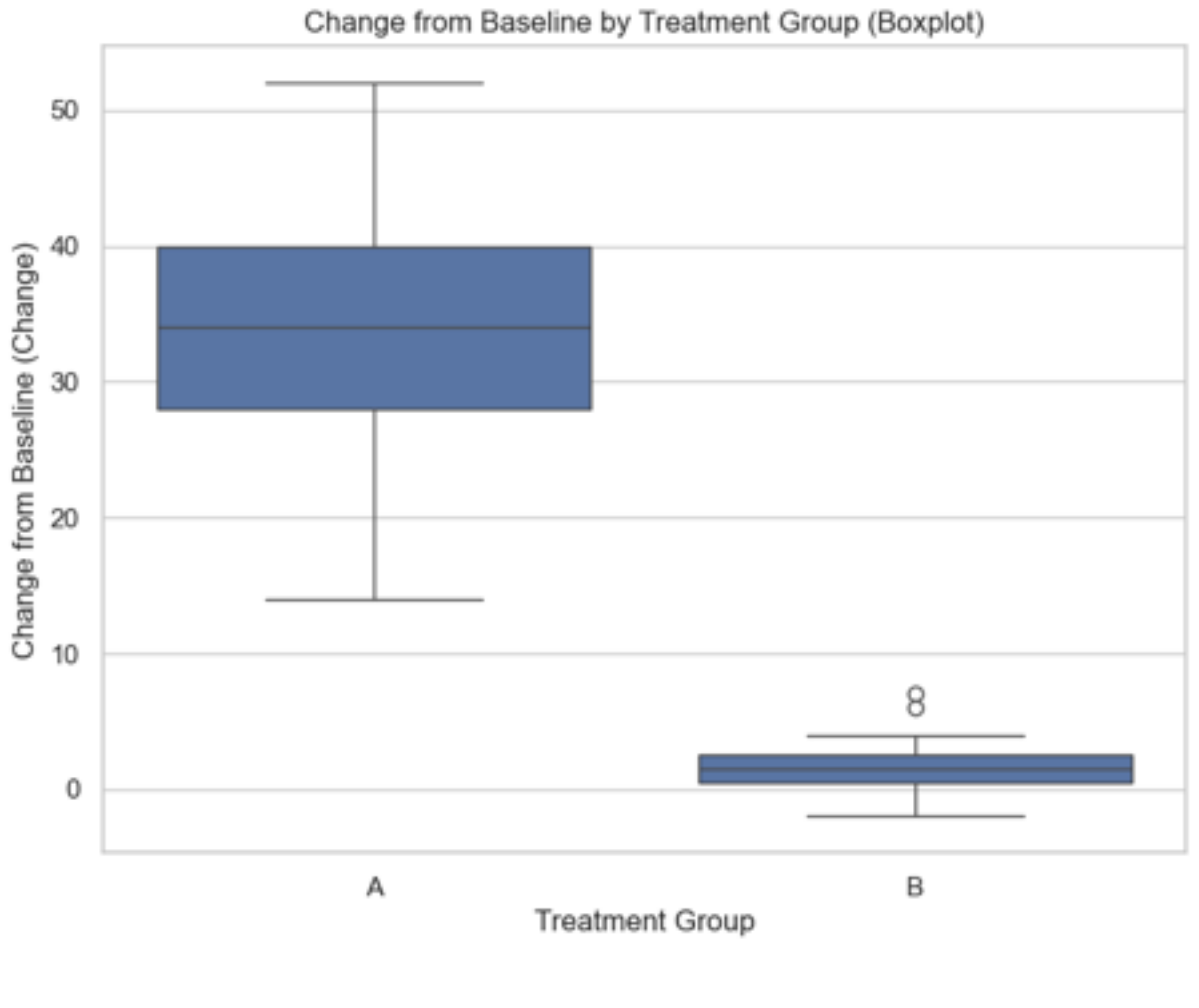
Out[5]:	Group	VAS_before	VAS_after	Change
0	A	86	71	15
1	A	77	59	18
2	A	75	44	31
3	A	83	49	34
4	A	72	32	40

## 5. Visualize the change from baseline by treatment group

```
In [6]: # Create a boxplot using Seaborn
plt.figure(figsize=(8, 6))
sns.set(style="whitegrid")
sns.boxplot(x='Group', y='Change', data=data)

# Set labels and title
plt.xlabel("Treatment Group")
plt.ylabel("Change from Baseline (Change)")
plt.title("Change from Baseline by Treatment Group (Boxplot)")

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



## 6. Derive a new variable indicating 20 points drop in VAS score from baseline

```
In [7]: data['Change_20'] = np.where(data['Change']>20,"Yes","No")
data.head()
```

Out[7]:	Group	VAS_before	VAS_after	Change	Change_20
0	A	86	71	15	No
1	A	77	59	18	No
2	A	75	44	31	Yes
3	A	83	49	34	Yes
4	A	72	32	40	Yes

## 7. Obtain cross table of above indicator variable with treatment group

```
In [8]: cross_table = pd.crosstab(data['Group'], data['Change_20'])
cross_table
```

```
Out[8]:
```

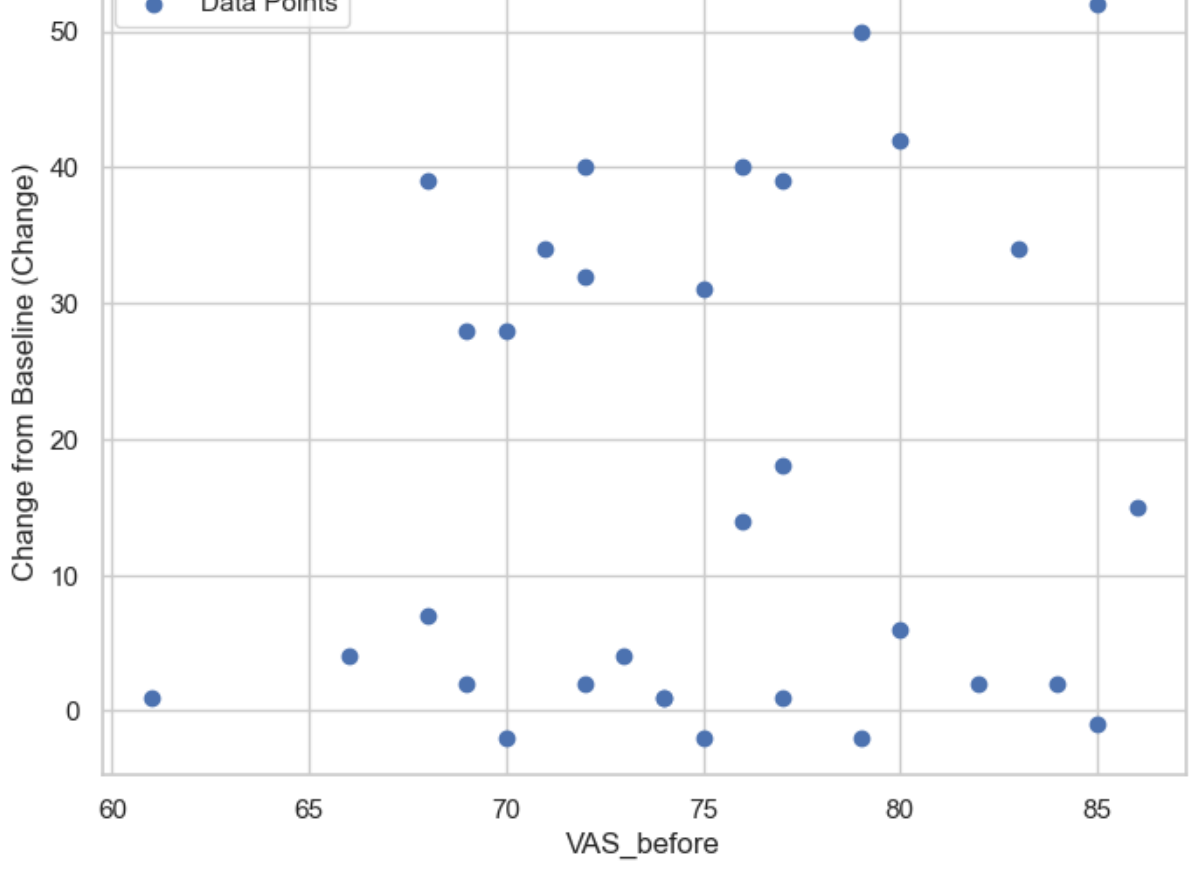
Change_20	No	Yes
Group		
A	3	13
B	16	0

## 8. Visualize the relationship between Change from baseline and baseline score

```
In [9]: # Create a scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(data['VAS_before'], data['Change'], c='b', marker='o')

# Set labels and title
plt.xlabel("VAS_before")
plt.ylabel("Change from Baseline (Change)")
plt.title("Relationship between Change from Baseline and VAS_before Score")

# Show the plot
plt.grid(True)
plt.legend()
plt.show()
```



## 9. Obtain correlation coefficient between Change from baseline and baseline score

```
In [10]: # Calculate the correlation coefficient
correlation_coefficient = round(data['VAS_before'].corr(data['Change']), 2)

# Display the correlation coefficient
print(f"Correlation Coefficient: {correlation_coefficient}")
```

Correlation Coefficient: 0.13