

ent-no-4-statistical-description

October 23, 2023

1 Statistical Description

Experiment no.: 4

Aim: Statistical Description

```
[1]: #Name: Krutika Nemade  
     #Roll no.: 55  
     #Sec: A  
     #Subject: Data Science and Statistics (Lab 1)
```

```
[2]: import pandas as pd
```

```
[3]: import matplotlib.pyplot as plt
```

```
[4]: import seaborn as sns
```

```
[5]: import numpy as np
```

```
[6]: import os
```

```
[7]: os.getcwd()
```

```
[7]: 'C:\\Users\\hp\\Downloads'
```

```
[8]: os.chdir('C:\\Users\\hp\\Desktop')
```

```
[9]: df=pd.read_csv("Salary_dataset.csv")
```

```
[10]: df.head()
```

```
[10]:
```

| | Unnamed: 0 | YearsExperience | Salary |
|---|------------|-----------------|---------|
| 0 | 0 | 1.2 | 39344.0 |
| 1 | 1 | 1.4 | 46206.0 |
| 2 | 2 | 1.6 | 37732.0 |
| 3 | 3 | 2.1 | 43526.0 |
| 4 | 4 | 2.3 | 39892.0 |

```
[11]: df.tail()
```

```
[11]:      Unnamed: 0  YearsExperience  Salary
25          25          9.1  105583.0
26          26          9.6  116970.0
27          27          9.7  112636.0
28          28         10.4  122392.0
29          29         10.6  121873.0
```

```
[12]: df.head(30)
```

```
[12]:      Unnamed: 0  YearsExperience  Salary
0          0          1.2   39344.0
1          1          1.4   46206.0
2          2          1.6   37732.0
3          3          2.1   43526.0
4          4          2.3   39892.0
5          5          3.0   56643.0
6          6          3.1   60151.0
7          7          3.3   54446.0
8          8          3.3   64446.0
9          9          3.8   57190.0
10         10          4.0   63219.0
11         11          4.1   55795.0
12         12          4.1   56958.0
13         13          4.2   57082.0
14         14          4.6   61112.0
15         15          5.0   67939.0
16         16          5.2   66030.0
17         17          5.4   83089.0
18         18          6.0   81364.0
19         19          6.1   93941.0
20         20          6.9   91739.0
21         21          7.2   98274.0
22         22          8.0  101303.0
23         23          8.3  113813.0
24         24          8.8  109432.0
25         25          9.1  105583.0
26         26          9.6  116970.0
27         27          9.7  112636.0
28         28         10.4  122392.0
29         29         10.6  121873.0
```

```
[13]: df.info()
#attribute
```

```
<class 'pandas.core.frame.DataFrame'>
```

```

RangeIndex: 30 entries, 0 to 29
Data columns (total 3 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Unnamed: 0             30 non-null     int64
1   YearsExperience         30 non-null     float64
2   Salary                  30 non-null     float64
dtypes: float64(2), int64(1)
memory usage: 852.0 bytes

```

```
[14]: df.describe()
      #record
```

```
[14]:
```

| | Unnamed: 0 | YearsExperience | Salary |
|-------|------------|-----------------|---------------|
| count | 30.000000 | 30.000000 | 30.000000 |
| mean | 14.500000 | 5.413333 | 76004.000000 |
| std | 8.803408 | 2.837888 | 27414.429785 |
| min | 0.000000 | 1.200000 | 37732.000000 |
| 25% | 7.250000 | 3.300000 | 56721.750000 |
| 50% | 14.500000 | 4.800000 | 65238.000000 |
| 75% | 21.750000 | 7.800000 | 100545.750000 |
| max | 29.000000 | 10.600000 | 122392.000000 |

```
[15]: df.shape
```

```
[15]: (30, 3)
```

```
[16]: df.size
```

```
[16]: 90
```

```
[17]: df.ndim
```

```
[17]: 2
```