

odjeli_analiza

August 1, 2025

```
[1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
```

```
[2]: odjeli_place = pd.read_csv("../..../datasets/odjeli_place.csv")
```

```
[3]: odjeli_place
```

```
[3]:      Unnamed: 0  Department      Salary
0              0  Marketing  52483.570765
1              1  Marketing  49308.678494
2              2  Marketing  53238.442691
3              3  Marketing  57615.149282
4              4  Marketing  48829.233126
..           ...      ...      ...
85            85      Sales  51989.457738
86            86      Sales  60492.412706
87            87      Sales  56972.506658
88            88      Sales  51821.438777
89            89      Sales  58079.604599
```

[90 rows x 3 columns]

```
[4]: odjeli_place.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   Unnamed: 0   90 non-null    int64
 1   Department   90 non-null    object
 2   Salary       90 non-null    float64
dtypes: float64(1), int64(1), object(1)
memory usage: 2.2+ KB
```

```
[5]: odjeli_place.Department.value_counts()
```

```
[5]: Department
Marketing    30
Logistics    30
Sales        30
Name: count, dtype: int64
```

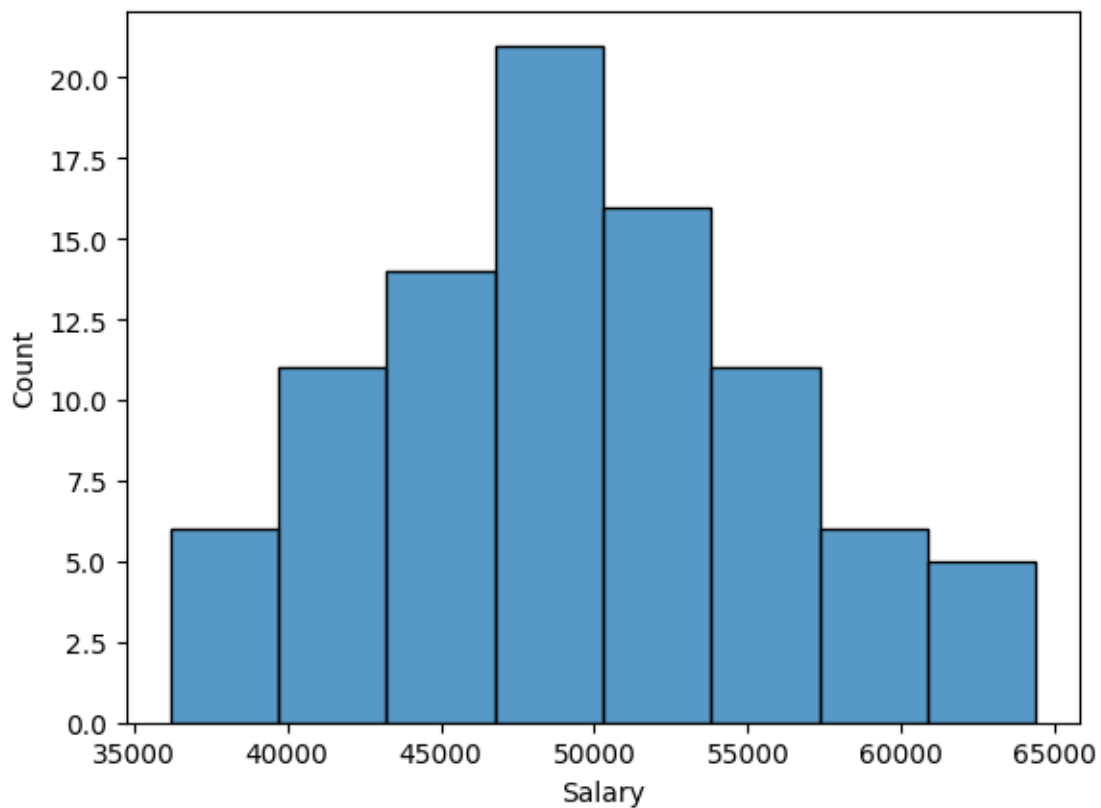
```
[6]: odjeli_place_department = odjeli_place.groupby("Department")
```

```
[7]: odjeli_place_department["Salary"].mean()
```

```
[7]: Department
Logistics    44454.768884
Marketing    49059.265521
Sales        55077.308635
Name: Salary, dtype: float64
```

```
[8]: sns.histplot(odjeli_place.Salary)
```

```
[8]: <Axes: xlabel='Salary', ylabel='Count'>
```



```
[9]: odjeli_place.corr(numeric_only=True)
```

```
[9]:          Unnamed: 0    Salary
      Unnamed: 0      1.00000  0.34113
      Salary          0.34113  1.00000
```

```
[10]: odjeli_place.describe()
```

```
[10]:          Unnamed: 0      Salary
      count    90.000000    90.000000
      mean     44.500000  49530.447680
      std      26.124701   6556.812273
      min       0.000000  36181.484443
      25%      22.250000  44936.698921
      50%      44.500000  48850.216856
      75%      66.750000  53595.289462
      max      89.000000  64387.861935
```

```
[11]: marketing = odjeli_place[odjeli_place['Department'] == 'Marketing']['Salary']
      logistics = odjeli_place[odjeli_place['Department'] == 'Logistics']['Salary']
      sales = odjeli_place[odjeli_place['Department'] == 'Sales']['Salary']
```

```
[12]: F_statistic, p_value = stats.f_oneway(marketing, logistics, sales)
```

```
[13]: F_statistic
```

```
[13]: np.float64(34.87380559925138)
```

```
[14]: p_value
```

```
[14]: np.float64(7.548262426790143e-12)
```