

CS23334-FUNDAMENTALS OF DATA SCIENCE

DEVA

DHARSHINI P (240701107)

14.) ANOVA TEST

Aim:

To perform an ANOVA (Analysis of Variance) test to determine whether there are any statistically significant differences between the means of three or more groups.

Code:

```
import numpy as np
import scipy.stats as stats
```

```
np.random.seed(42)
```

```
n_plants = 25
```

```
growth_A = np.random.normal(loc=10, scale=2, size=n_plants)
growth_B = np.random.normal(loc=12, scale=3, size=n_plants)
growth_C = np.random.normal(loc=15, scale=2.5, size=n_plants)
```

```
all_data = np.concatenate([growth_A, growth_B, growth_C])
```

```
treatment_labels = ['A'] * n_plants + ['B'] * n_plants + ['C'] * n_plants
```

```
from scipy import stats
f_statistic, p_value = stats.f_oneway(growth_A, growth_B, growth_C)
```

```
print("Treatment A Mean Growth:", np.mean(growth_A))
print("Treatment B Mean Growth:", np.mean(growth_B))
print("Treatment C Mean Growth:", np.mean(growth_C))
print()
print(f"F-Statistic: {f_statistic:.4f}")
print(f"P-Value: {p_value:.4f}")
```

```
Treatment A Mean Growth: 9.672983882683818
Treatment B Mean Growth: 11.137680744437432
Treatment C Mean Growth: 15.265234904828972
```

```
F-Statistic: 36.1214
P-Value: 0.0000
```

```
alpha = 0.05
if p_value < alpha:
    print("Reject the null hypothesis: There is a significant difference in mean growth rates among the three treatments.")
else:
    print("Fail to reject the null hypothesis: There is no significant difference in mean growth rates among the three treatments.")
```

Reject the null hypothesis: There is a significant difference in mean growth rates among the three treatments.

```
from statsmodels.stats.multicomp import pairwise_tukeyhsd
if p_value < alpha:
    tukey_results = pairwise_tukeyhsd(all_data, treatment_labels, alpha=0.05)
    print("\nTukey's HSD Post-hoc Test:")
    print(tukey_results)
```

```
Tukey's HSD Post-hoc Test:
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj  lower  upper  reject
-----
    A     B   1.4647 0.0877 -0.1683 3.0977  False
    A     C   5.5923   0.0   3.9593 7.2252   True
    B     C   4.1276   0.0   2.4946 5.7605   True
-----
```

Result:

The ANOVA test was successfully conducted, and the result showed whether there was a significant difference among the group means.