ANOVA

- Analysing the variance (differences) between the variables.
- It is of two types.

1. One way classification – uses only one independent variable

```
# One Way Classification

import scipy.stats as stats
stats.f_oneway(dataset['ssc_p'],dataset['hsc_p'],dataset['degree_p'])
```

```
F_onewayResult(statistic=0.6719700864663097, pvalue=0.5110602818995302)
```

H_0 : There is no similarity between 10^{th}, 12^{th} and degree marks.

H_1: There is a similarity between 10^{th}, 12^{th} and degree marks.

p-value < 0.05 is not satisfied to reject null hypothesis. So,

Accepting H_0 and rejecting H_1.

2. Two way Classification – uses two independent categorical variables and one dependent numerical variable.

	sum_sq	df	F	PR(>F)
C(gender)	2.157904e+10	1.0	3.638124	0.057829
C(status)	1.774039e+08	1.0	0.029909	0.862862
C(gender):C(status)	1.084427e+10	1.0	1.828293	0.177777
Residual	1.251518e+12	211.0	NaN	NaN

Since all your **p-values > 0.05**, you **fail to reject H_0** for all three tests.
That means accepting **the null hypotheses**.

Gender: No significant effect on the dependent variable.
Status: No significant effect on the dependent variable.
Gender × Status: No significant interaction effect.

Two-way ANOVA results indicate that neither gender ($p = 0.0578$), status ($p = 0.8629$), nor their interaction ($p = 0.1777$) have a statistically significant effect on the dependent variable at the 0.05 significance level.