**Correlation Between Attrition & other Columns**

1. **There is no significant relation between the monthly income & attrition level in the XYZ company**.

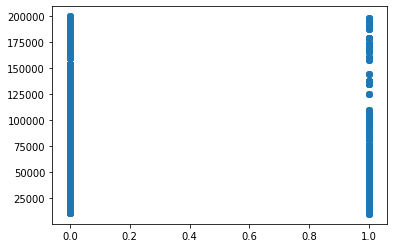
stats,p= pearsonr(dataset.Attrition, dataset.MonthlyIncome)

print(stats,p)

-0.031176281698114025 0.0384274849060192

plt.scatter(dataset.Attrition, dataset.MonthlyIncome)

Out[57]: <matplotlib.collections.PathCollection at 0x16c8bc596c8>



P < 0.05 🡺 P value is less than 0.05 hence H0 can be rejected.

*Therefore it could be concluded that there is a significant relation between ‘Attrition’ and ‘MonthlyIncome’ and the hypothesis stands rejected. Monthly income does have a significant impact on attrition levels.*

1. **There is no realtion between percentage hike in salary vis-a-vis Attrition levels.**

stats,p = pearsonr(dataset.Attrition, dataset.PercentSalaryHike)

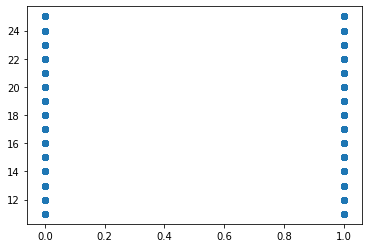
print(stats,p)

0.03253259489105223 0.030743386433369824

P < 0.05 🡺 P value is less than 0.05 hence H0 is rejected.

plt.scatter(dataset.Attrition, dataset.PercentSalaryHike)

Out[65]: <matplotlib.collections.PathCollection at 0x16c8bcc3d08>



*Therefore it could be concluded that percentage salary hike has a significant impact on attrition levels.*

1. **There is no impact on attrition of the number of years of each employee at the company.**

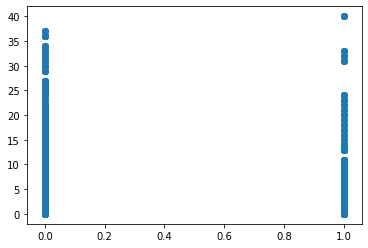
stats,p = pearsonr(dataset.Attrition, dataset.YearsAtCompany)

print(stats,p)

-0.13439221398997386 3.163883122493571e-19

plt.scatter(dataset.Attrition, dataset.YearsAtCompany)

Out[69]: <matplotlib.collections.PathCollection at 0x16c8bd96b88>



*The corresponding p value < 0.05. Hence it could be concluded the hypothesis is rejected.*

1. **There is no significant relation between Attrition and the numbers of years an employee has spent since last promotion.**

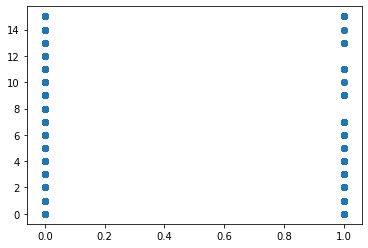
stats,p = pearsonr(dataset.Attrition, dataset.YearsSinceLastPromotion)

print(stats,p)

-0.03301877514258329 0.02833033618939086

plt.scatter(dataset.Attrition, dataset.YearsSinceLastPromotion)

Out[72]: <matplotlib.collections.PathCollection at 0x16c8be886c8>



*Since p value < 0.05, it could be concluded that the H0 hypothesis can be rejected and the alternate hypothesis H1 to be accepted.*

1. **The years spent with the Current Manager do not have any impact on Attrition.**

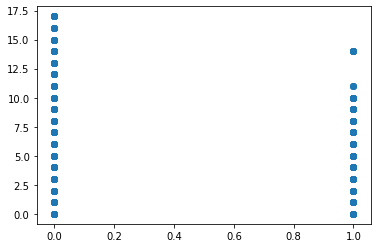
stats,p = pearsonr(dataset.Attrition, dataset.YearsWithCurrManager)

print(stats,p)

-0.1561993159016244 1.7339322652951965e-25

plt.scatter(dataset.Attrition, dataset.YearsWithCurrManager)

Out[75]: <matplotlib.collections.PathCollection at 0x16c8bee7388>



*Since the p value < 0.05, hence it could be concluded that the number of years spent with a manager can have a significant impact of attrition levels.*

1. **Attrition has no impact on the -- distance to home from workplace, aspect of the employee.**

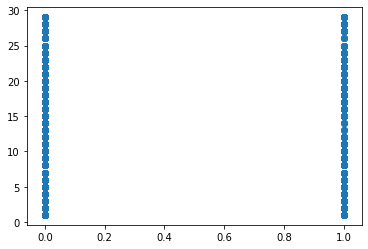
stats,p = pearsonr(dataset.Attrition, dataset.DistanceFromHome)

print(stats,p)

-0.009730141010179435 0.5182860428049617

plt.scatter(dataset.Attrition, dataset.DistanceFromHome)

Out[78]: <matplotlib.collections.PathCollection at 0x16c8bf4f6c8>



*The p value > 0.05, hence it could be concluded that the DistanceFromHome variable has an no impact on the Attrition levels. Therefore hypothesis proves correct.*

1. Education has no great impact on Attrition in the company.

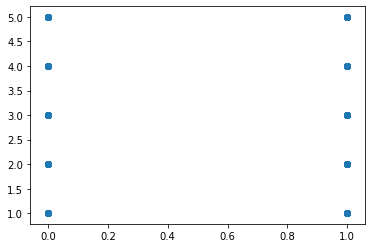
stats,p = pearsonr(dataset.Attrition, dataset.Education)

print(stats,p)

-0.015111167710968753 0.3157293177118575

plt.scatter(dataset.Attrition, dataset.Education)

Out[81]: <matplotlib.collections.PathCollection at 0x16c8bfb9308>



*P value > 0.05, hence it could be concluded that the above hypothesis is correct and Education has no bearing on the attrition levels.*