HR Project

Chi Square Test

After this, we applied **chi square test** on the data, from which following things were noticed:

- No employee with years **more than 6** in company left the company,
- Smaller percentage of employees with high salary left as compare to low and medium,
- Smaller percentage of employees with promotion left as compare to others,
- Smaller percentage of employees with work accidents left.

```
     years spent    left   stayed
0               2.0   53.0  3191.0
1               3.0 1586.0  4857.0
2               4.0  890.0  1667.0
3               5.0  833.0   640.0
4               6.0  209.0   509.0
5               7.0    0.0   188.0
6               8.0    0.0   162.0
7               9.0    0.0   214.0
[[ 772.33975598 2471.66024402]
 [1533.96579772 4909.03420228]
 [ 608.7770518  1948.2229482 ]
 [ 350.69557971 1122.30442029]
 [ 170.94326288  547.05673712]
 [  44.7595173   143.2404827 ]
 [  38.56937129  123.43062871]
 [  50.94966331  163.05033669]]
0.0
```

(a)

```
       left   stayed
0   2172.0  5144.0
1   1317.0  5129.0
2     82.0  1155.0
[[1741.81185412 5574.18814588]
 [1534.68004534 4911.31995466]
 [ 294.50810054  942.49189946]]
1.652086749296005e-83
```

(b)

```
     left    stayed
0  3552.0  11128.0
1    19.0    300.0
[[ 3495.05167011 11184.94832989]
 [   75.94832989   243.05167011]]
6.34415545791872e-14
```

(c)

```
     left   stayed
0  3402.0  9428.0
1   169.0  2000.0
[[3054.59897326 9775.40102674]
 [ 516.40102674 1652.59897326]]
9.55823958002199e-80
```

(d)

(a) shows stats of no. of years, (b) shows stats of salary, (c) shows stats of promotion and (d) stats of work accidents.

Also all of the attributes mentioned above also had **very small p value**, which **shows association** with attribute '**left**' as shown in the figures above.

From the above discussion, we can deduce that more than **90%** of employees previously involved in **accidents stayed** and also more than **90%** of the employees with a **high salary stayed**.
And more than **90%** of the employees **who left** were working in the company between **4 to 5 years**.

More Results:

Below are results of chi square test for attributes departments and no of projects the employee has worked on.
From the chi test of 'departments', it can be seen that there is association between the employee department and their decision to leave. But the no. of people who stayed is always greater than who left, with a slight change in the ratio.

```
     left   stayed
0  1014.0  3126.0
1   204.0   563.0
2   215.0   524.0
3   697.0  2023.0
4   555.0  1674.0
5    91.0   539.0
6   273.0   954.0
7   198.0   704.0
8   203.0   655.0
9   121.0   666.0
[[ 985.66171078 3154.33828922]
 [ 182.60930729  584.39069271]
 [ 175.9429962   563.0570038 ]
 [ 647.58450563 2072.41549437]
 [ 530.68597907 1698.31402093]
 [ 149.99199947  480.00800053]
 [ 292.12727515  934.87272485]
 [ 214.75045003  687.24954997]
 [ 204.27481832  653.72518168]
 [ 187.37095806  599.62904194]]
7.042130463822518e-15
```

And from the chi test of attribute 'no of projects', we can see there is **strong association** between no of proj and the decision whether to leave or stay. And it is also clear that all the employees with 'no of project' equal to **7 left the company**. And the majority of employees with 'no of projects' equal to **3,4 and 5 stayed**.

```
   no of proj    left  stayed
0         2.0  1567.0   821.0
1         3.0    72.0  3983.0
2         4.0   409.0  3956.0
3         5.0   612.0  2149.0
4         6.0   655.0   519.0
5         7.0   256.0     0.0
0.0
[[ 568.54110274 1819.45889726]
 [ 965.42469498 3089.57530502]
 [1039.23028202 3325.76971798]
 [ 657.34588973 2103.65411027]
 [ 279.50890059  894.49109941]
 [  60.94912994  195.05087006]]
```

Other techniques like **t-test** and **pca** were also tried, but it could not produce any insightful or desired results.

Below are the answers of some key questions:

a. **What factors cause the employees to leave?**
Ans: From above tables and assessment, we can say that employees with low salary, no promotions and relatively new in the company left.

b. **Which factor has most influence on employee retention?**
Ans: Employees with high salary, promotion, less no of projects (3 and 4) and having involved in accidents have direct relation with employee retention. But 'salary' with very small p value, we can say it has the most influence.

c. **Does work related accidents cause employees to leave? Why and why not?**
Ans: Since 36% of employees with no accidents left and only 8% employees with accidents left, we can say accidents caused employees to stay. Reason can be, since they were not fired after the accident, they decided to stay.

d. **Do employees leave the company even after they got promotion? If yes, what is the cause?**
Ans: Following figure suggests that people who got promotion had lower mean of 'satisfaction level' 0.46 and compare to 0.61 and 'salary' mean of 0.263 as compare to 0.59 of whole data, which can be interpreted as the cause of them leaving the company.

```
satisfaction_level       0.460000
last_evaluation          0.588421
number_project           3.052632
average_montly_hours   177.736842
time_spend_company       3.263158
Work_accident            0.210526
promotion_last_5years    1.000000
department               2.842105
salary                   0.263158
left                     1.000000
dtype: float64
```

e. **Provide your recommendations to the management to retain the employees.**
Ans: We can say that high salary, promotion and less projects (less work load) have direct relation with employees staying in the company.

f. **Which department has highest rate of retention?**
Ans: From the figure below, we can say that department no 5 with left-stay ratio of 0.17, which is '**IT department**' has the highest retention rate.

```
      left    stayed
0  1014.0   3126.0
1   204.0    563.0
2   215.0    524.0
3   697.0   2023.0
4   555.0   1674.0
5    91.0    539.0
6   273.0    954.0
7   198.0    704.0
8   203.0    655.0
9   121.0    666.0
[[ 985.66171078 3154.33828922]
 [ 182.60930729  584.39069271]
 [ 175.9429962   563.0570038 ]
 [ 647.58450563 2072.41549437]
 [ 530.68597907 1698.31402093]
 [ 149.99199947  480.00800053]
 [ 292.12727515  934.87272485]
 [ 214.75045003  687.24954997]
 [ 204.27481832  653.72518168]
 [ 187.37095806  599.62904194]]
7.042130463822518e-15
```