LOGISTIC REGRESSION RESUT FOR ATTRITION DATASET:

STEPS:

- 1. Import pandas package.
- 2. Loading the dataset.
- 3. Cleaning the dataset.

```
In [38]: import os
              os.chdir(r"C:\Users\User\Downloads")
import pandas as pd
             dataset.head()
dataset-dataset.dropna()
             dataset.isna().sum()
Out[38]: Age
Attrition
BusinessTravel
             Department
DistanceFromHome
              Education
              {\sf EducationField}
              EmployeeCount
             EmployeeCount
EmployeeID
Gender
JobLevel
JobRole
MaritalStatus
              MonthlyIncome
              NumCompaniesWorked
Over18
PercentSalaryHike
              StandardHours
StockOptionLevel
              TotalWorkingYears
TrainingTimesLastYear
              YearsAtCompany
YearsSinceLastPromotion
              YearsWithCurrManager
```

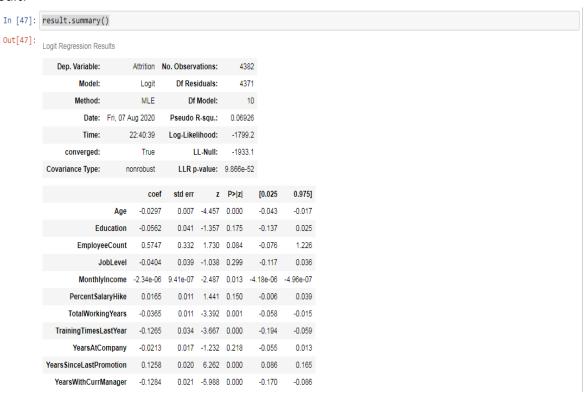
- 4. Assigning dependent variable as "attrition" and all other significant columns as independent variables.
- 5. Import statistical model and adding constants to dependent variable.

```
import statsmodels.api as sm
x1=sm.add_constant(x)
```

6. Applying logistic regression formula using "Logit" method and optimizing the result.

```
log=sm.Logit(y,x1)
result=log.fit()
```

7. Result:



As we can observe the p value for the following is less than 0.05 which means they are significant attributes.

- 1. Age
- 2. Monthly income
- 3. Total working years
- 4. Training times last year
- 5. Years since last promotion
- 6. Years with current manager