**Framing Hypothesis Statements**

import pandas as pd

dataset = pd.read\_excel("Attrition Analysis Data.xlsx", sheet\_name = 0)

dataset1=dataset[['MonthlyIncome','YearsAtCompany','Age','DistanceFromHome','PercentSalaryHike','JobLevel','NumCompaniesWorked','TotalWorkingYears','YearsSinceLastPromotion','TrainingTimesLastYear']].describe()

dataset1

Out[4]:

MonthlyIncome ... TrainingTimesLastYear

count 4410.000000 ... 4410.000000

mean 65029.312925 ... 2.799320

std 47068.888559 ... 1.288978

min 10090.000000 ... 0.000000

25% 29110.000000 ... 2.000000

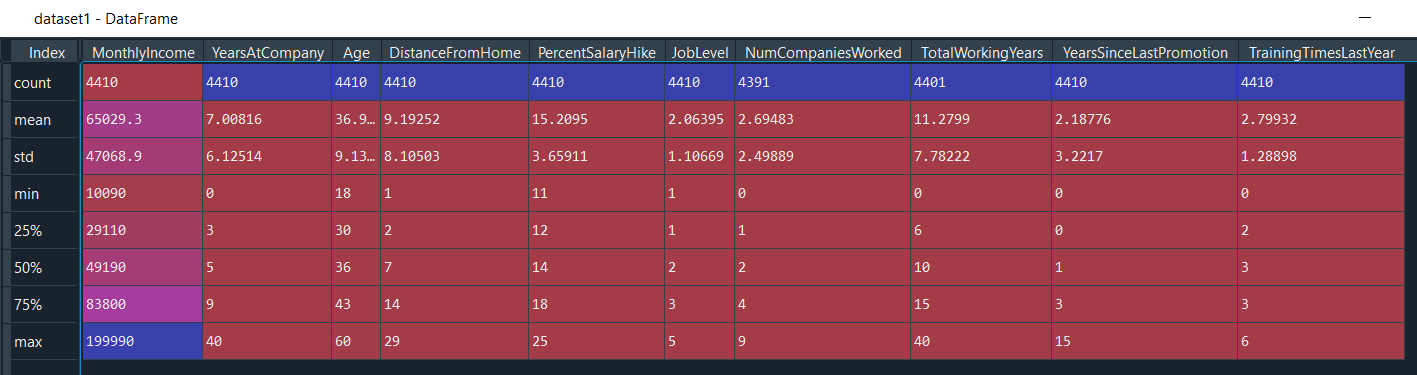
50% 49190.000000 ... 3.000000

75% 83800.000000 ... 3.000000

max 199990.000000 ... 6.000000

[8 rows x 10 columns]

Variable Explorer Image:



**Statement 1:**

If the monthly income of the employee is less than 65000, attrition is observed.

**Statement 2:**

If the employee has worked at the company for more than 7 years, attrition is observed.

**Statement 3:**

For employees whose age is less than 36 years, attrition is observed.

**Statement 4:**

For the employees whose distance from home is greater than 9km, attrition is observed.

**Statement 5:**

For the employees whose percent salary hike is lesser than 15%, attrition is observed.

**Statement 6:**

A high attrition rate is observed in the employees whose cadre/level is less than equal to 2.

**Statement 7:**

For employees having experience of 2+ years, a high attrition is observed.

**Statement 8:**

If total working years is more than 11 for an employee, then he is high probable for attrition.

**Statement 9:**

If the employee is working for more than 2 years since his last promotion, then is highly probable for attrition.

**Statement 10:**

For the employees whose training attempts are more than 2 in the past year, attrition rate is high.