*OPIM 5894-Seminar: Survival Analysis using SAS BASE*

*Project Report 1*

Submitted By: Team 6

Jagbir Kaur

  Janeet Arora

Namratha Kasineni

Raozhong Chen

Ritika Sinha

Shiqi Xu

Supreet Mangat

Swati Arora

Contents

[Project Overview: 3](#_Toc486335332)

[Business Case: 3](#_Toc486335333)

[Data Preprocessing: 4](#_Toc486335334)

[Data Exploration: 5](#_Toc486335335)

[Life Test Results for the Survival analysis: 11](#_Toc486335336)

[Identifying the significant covariates for the model 21](#_Toc486335337)

[Model Fit (Goodness of Fit) 22](#_Toc486335338)

[Complete list of coefficients with positive and negative effect: 23](#_Toc486335339)

[Major Findings in conclusion: 25](#_Toc486335340)

[Recommendations: 25](#_Toc486335341)

[References: 25](#_Toc486335342)

# Project Overview:

"Attrition" is one of the basic issue that is quite high in the business nowadays. Many well Trained and adapted workers leave the association to get a salary hike or get higher job satisfaction or better Job role. The pioneers and HR faculty attempt their best to reduce the attrition rate in their association so as to reap the benefits from all the investments and trainings that they bestowed at the employee.

FermaLogis, an extremely acclaimed pharmaceutics organization, is experiencing a similar stage now and needs to reveal the elements that prompt employee whittling down and investigate the imperative reasons that leads to attrition. To accomplish this objective, Larry Hansen, COO required to run Survival analysis on the information accumulated, as it is a standout method amongst the most utilized analysis procedures, particularly in Pharmaceutical industry.

# Business Case:

To Investigate and examine the main drivers of Attrition and present a report with the fascinating discoveries and recommendations to Fermalogis.

# Data Preprocessing:

Preprocessing, being the most critical action of any model building process, we began to investigate and break down the data. The dataset from FermaLogis had details of employees assembled over a period of 40 years. The most significant factors are listed beneath

|  |  |
| --- | --- |
| Attrition | Indicates that employee has left the organization or not |
| BusinessTravel | describes if the employee had opportunity to travel |
| EnvironmentSatisfaction | Rates the level of work environment satisfaction |
| JobLevel | States the level of Hierarchy in FermaLogis for an employee |
| JobRole | Role of the employee in the organization |
| OverTime | Employee works overtime or not |
| YearsAtCompany | Number of years at FermaLogis |

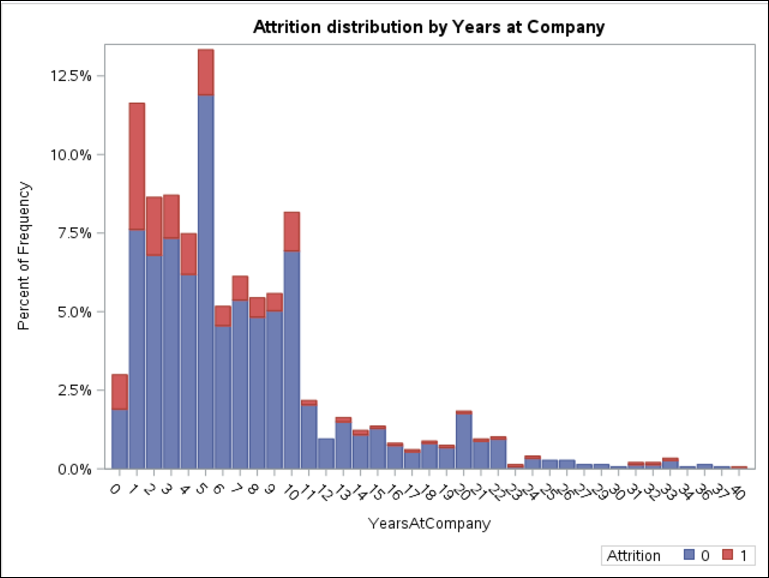
As we started to analyze the data, we noticed that most of them were categorical variables. Variables like Attrition, Gender and Marital status were recoded into numeric form for the better analysis.

|  |  |  |
| --- | --- | --- |
| **Variable** | **Original value** | **Recoded value** |
| Attrition | Yes | 1 |
|  | No | 0 |
| Gender | Male | 1 |
|  | Female | 2 |
| Marital Status | Single | 1 |
|  | Married | 2 |
|  | Divorced | 3 |
| Total Bonus | NA | Sum(bonus\_1-bonus\_40) |

We did not utilize variables like Employee number, Employee count, over 18, X, Standard hours as they were nit adding any value to the investigation and model building.

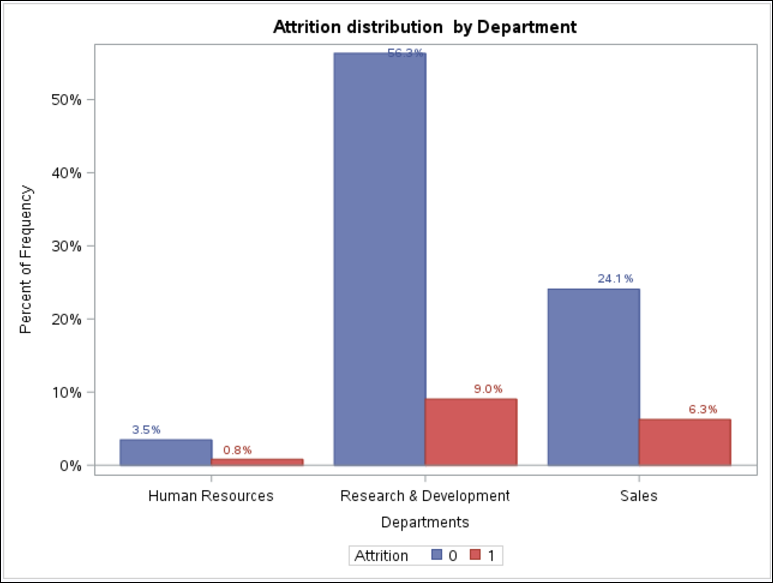
# Data Exploration:

Data exploration is a major step in data analysis and typically involves summarizing the main characteristics of a dataset. It is commonly conducted using visualization of the data. So, here are our results of data exploration.

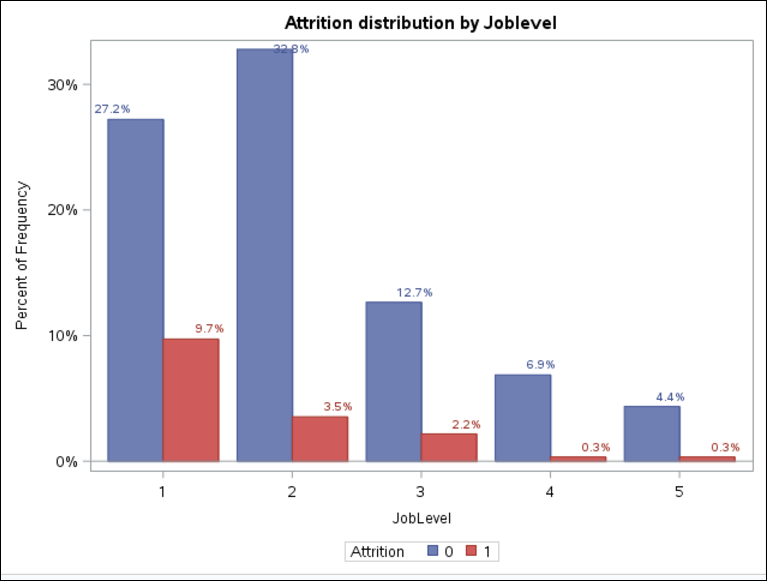


The employees in FermaLogis have a wide range when it comes to experience in the company itself. Most of the employees have been for 10 or less years in the company, whereas very few are with the FermaLogis for more than 10 years. It could be seen from the above graph, employees from 0 to 5 years duration have higher attrition rate and those with 1 year stay have the maximum attrition rate.

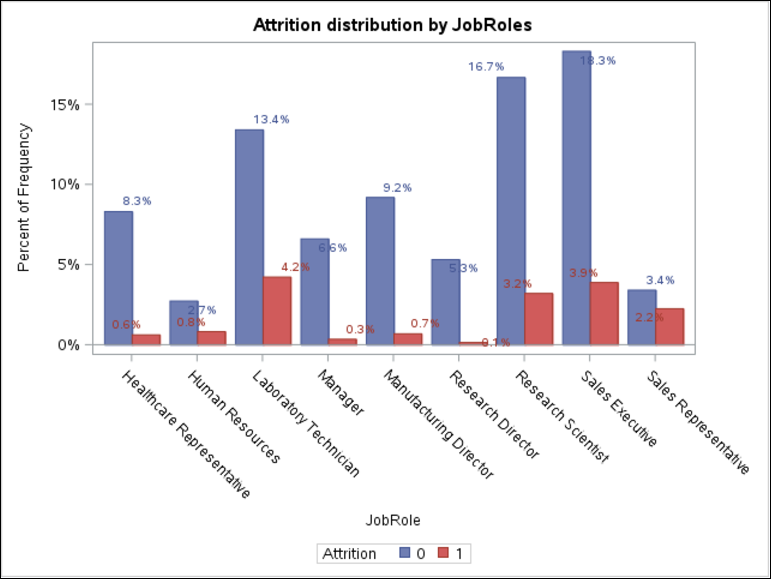
It was really exciting to explore the data provided by the company and find out the reasons behind the attrition of employees.

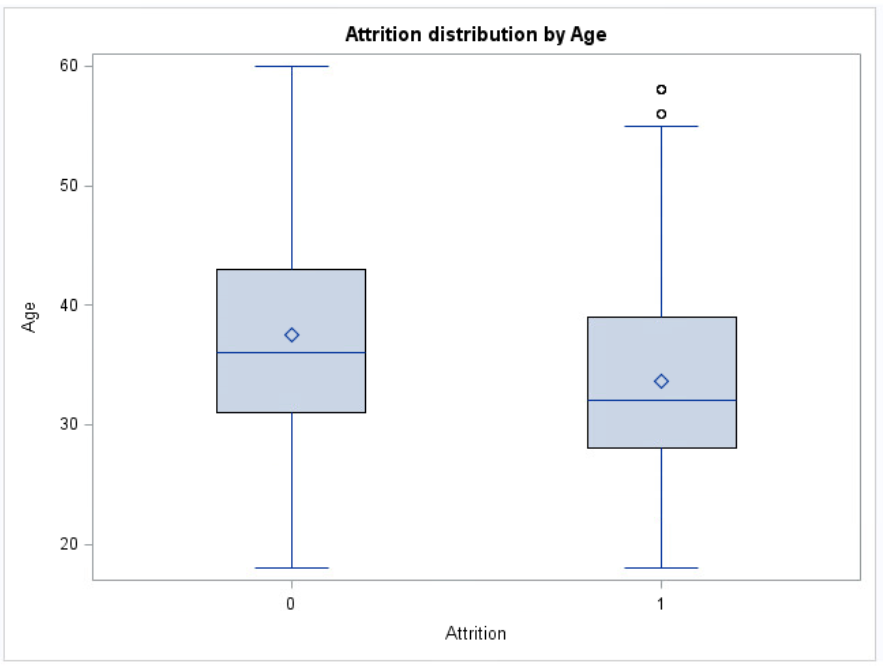
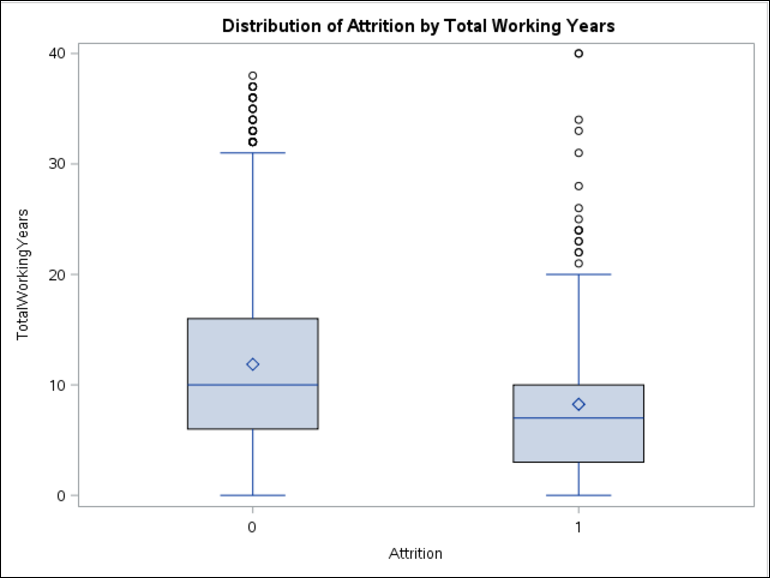


When comparing attrition rate across different Departments, we found out that around 20% of employees from each department have left the company. Which means that department is not a significant covariate in the attrition, it will be further explored in the model test for our dataset.



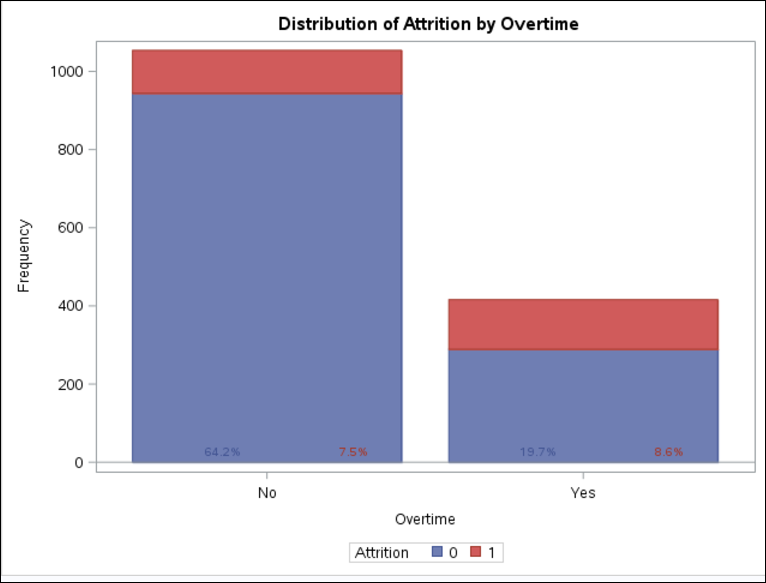
Here we tried to visualize the relation between company’s Attrition and the Job levels of the employees. We observed that lower job level employees are largely moving out of the company than the ones with higher job roles. To second this observation, the managers, research director, HR who generally at higher job level are less likely to leave as compared to others.



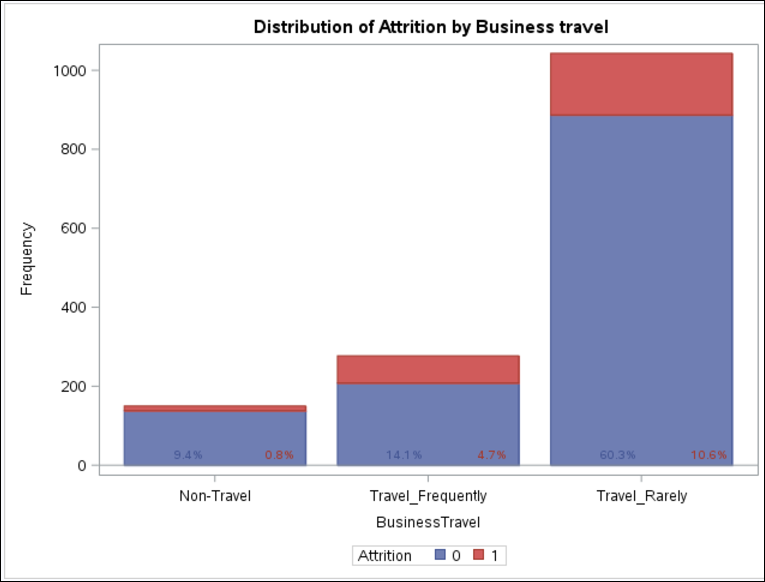
 

Overall distribution of employees specifies that young people are seeking opportunities outside the company. This again emphasizes on the point mentioned above, that employees with less work experience at the company are young employees and more prone to attrition.

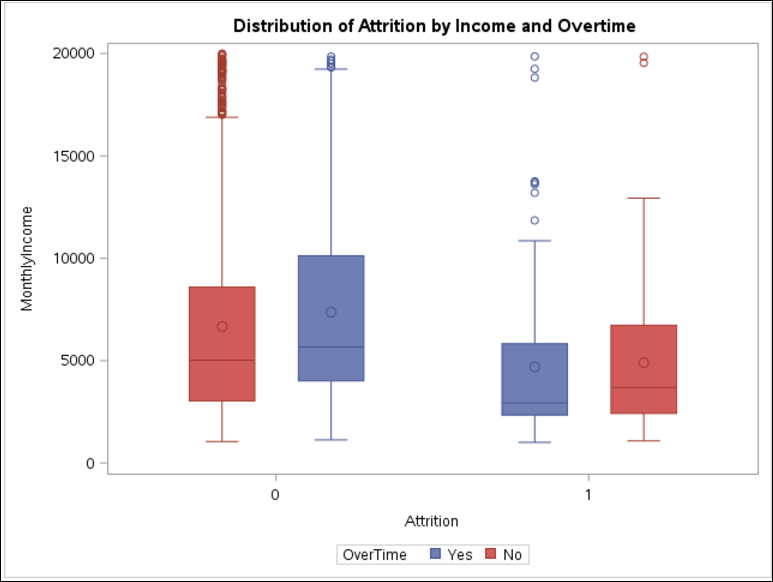
Let’s, explore some of the major reasons of employees leaving the company.



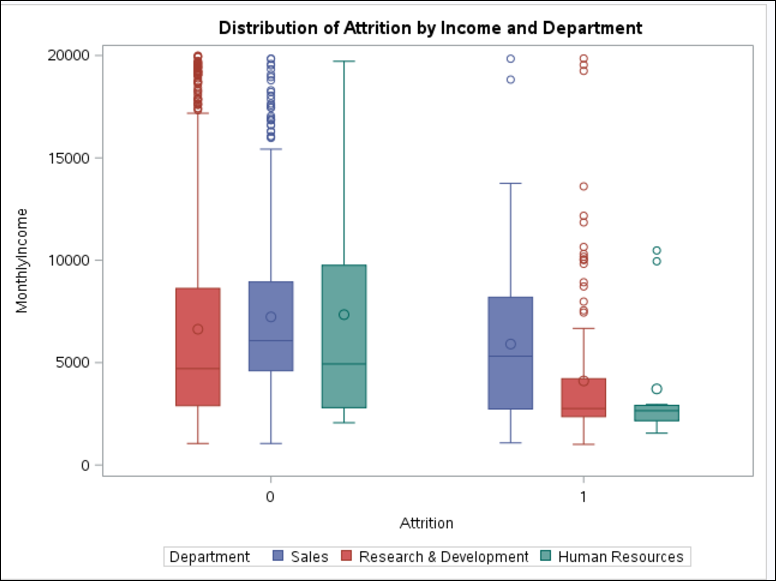
Employees with overtime has higher attrition rate. Around 45% of employees who worked overtime has left the company in contrast to only 11% for those who dint work overtime. This makes the point clear that people who work overtime has higher propensity to leave the job as compared to people with no overtime. This is a wonderful insight to suggest the management that people are working overtime have higher issues and should consider measures to reduce the overtime requirement and should also review their overtime compensation policies.



Here we tried to visualize the relation between Business travel of the employees and their Attrition rate. We observed that people who travel very frequently seem to be leaving the company. 25% of the total frequent travelers (4.7/ 18.8) left the company. On the contrary, only 14% who travel rarely and 7% who do not travel at all has left the company. This may be due to several reasons, one being disturbance of work life balance. Employees have to leave their family for business trips very often which encourage them to search for other options in the industry. Also, some employees are not comfortable travelling frequently. Management can rotate the employees for business trip, so that everyone gets the opportunity to explore and later ask employees to volunteer for travel.

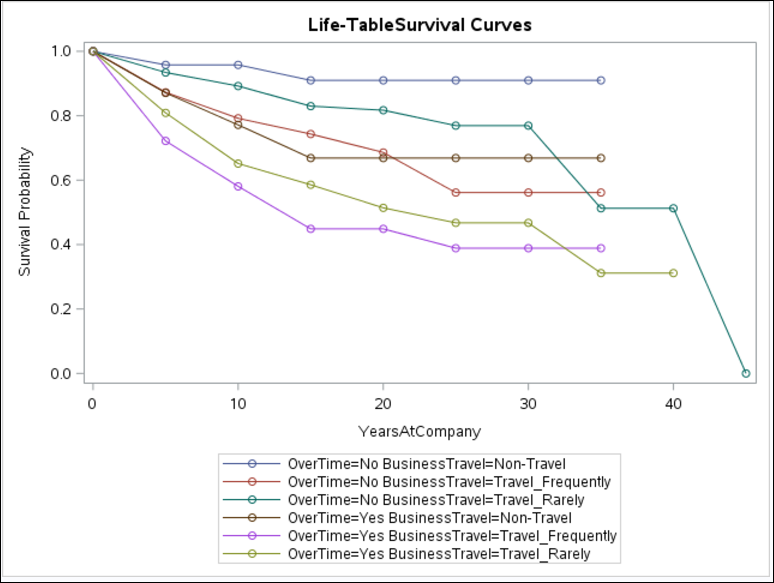


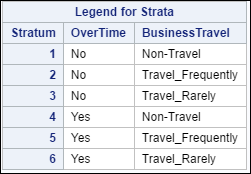
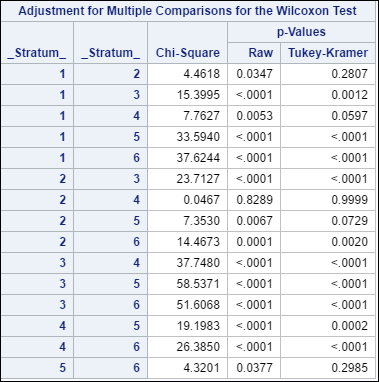
People with lower income have higher attrition irrespective of the overtime.



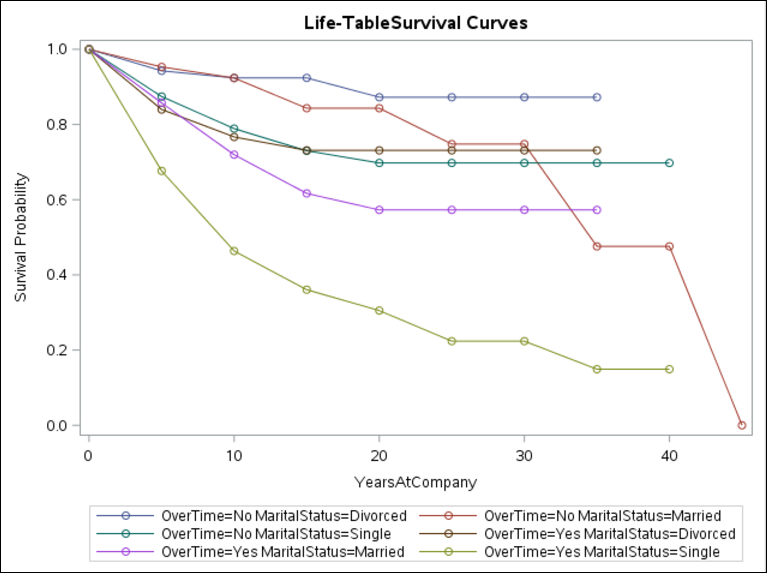
HR employees whose salaries are really low has greater propensity to leave, whereas those HR employees who are well paid are staying with the company. The same case is with Research and development employees. The Sales employees are leaving irrespective of their salary. So, management should consider the Sales department and take the survey from Sales employees about their expectations from job and what are their suggestion to make the sales job better.

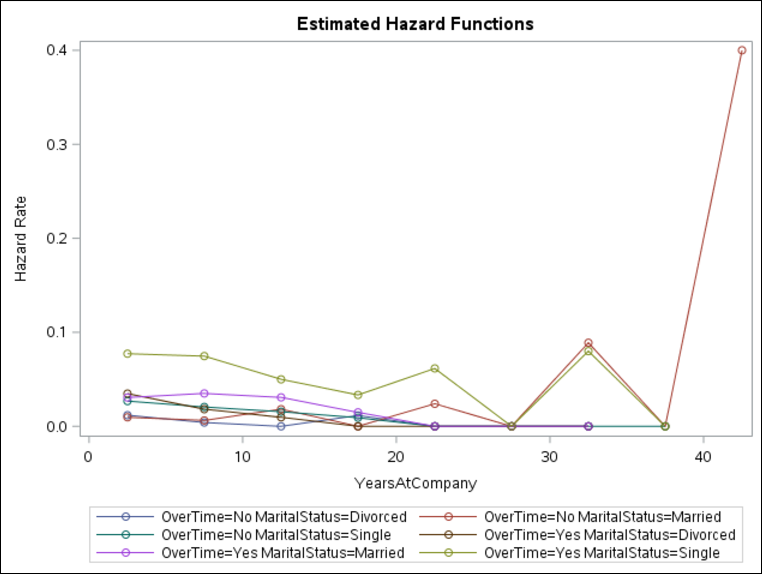
# Life Test Results for the Survival analysis:



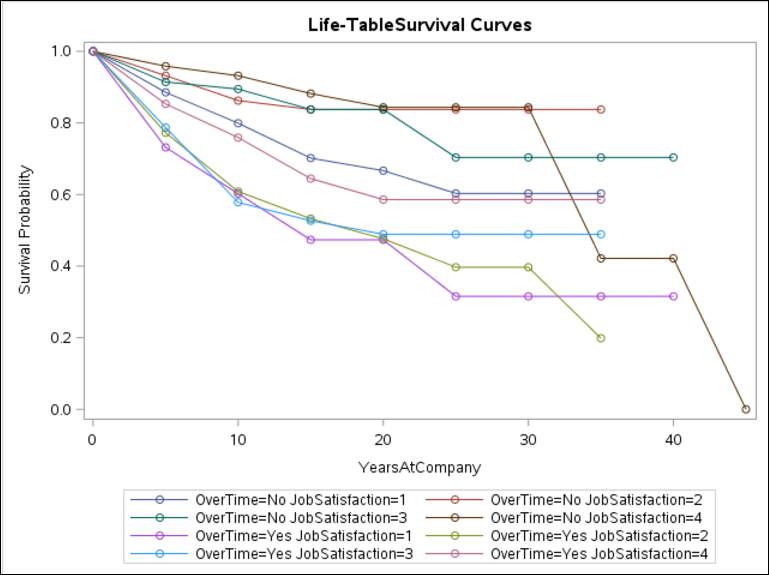
 

We applied the Lifetest on the overtime and business travel for Employees with respect to the Attrition and found out that employees who travel frequently and overtime has less survival chances in the company. Also, people who have no overtime and travel are staying in the company i.e. less prone to failure.



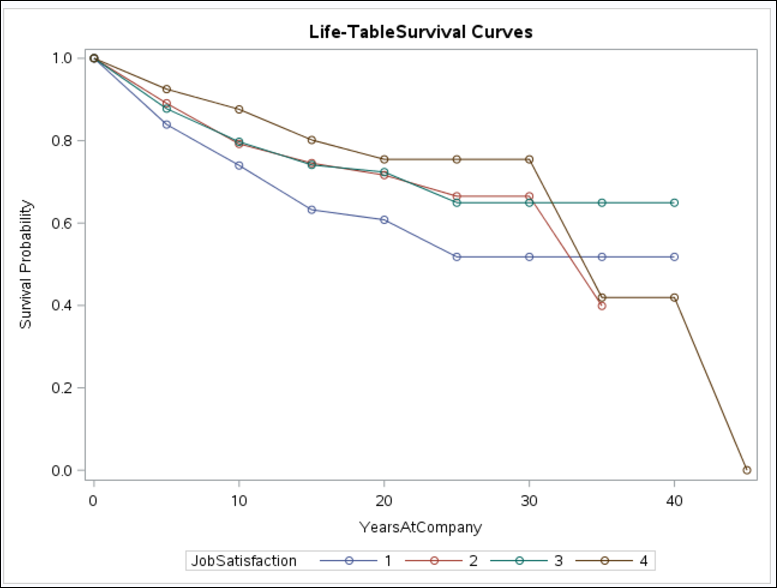


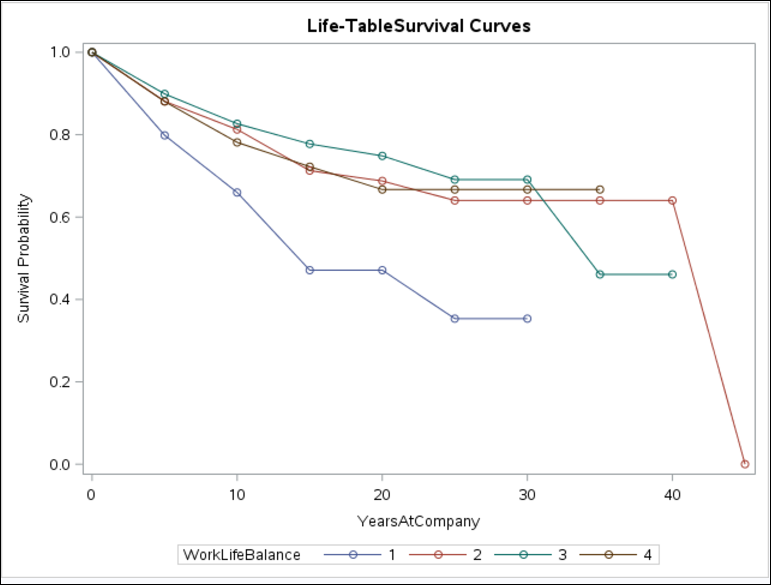
We plotted the survival and hazard plots for attrition with respect to overtime and marital status, hoping that might be married people with overtime are the ones with most hazard rate. But we found out that single who are working overtime are the ones with least survival rate/ most hazard rate.



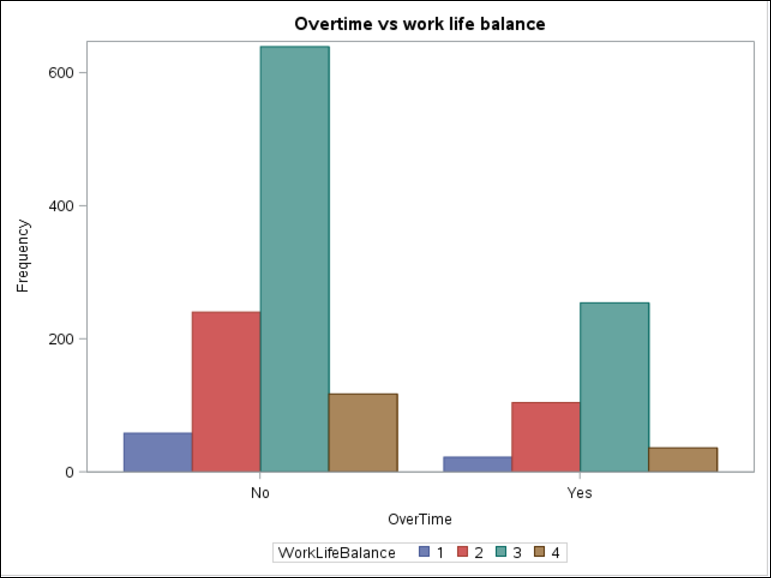
Employees with less job satisfaction and overtime has less probability of survival in the company whereas who are most satisfied with their job are staying with the company. Management should make a note to check with employees the reason they are not satisfied with their job and should adopt adequate measures to overcome those causes.

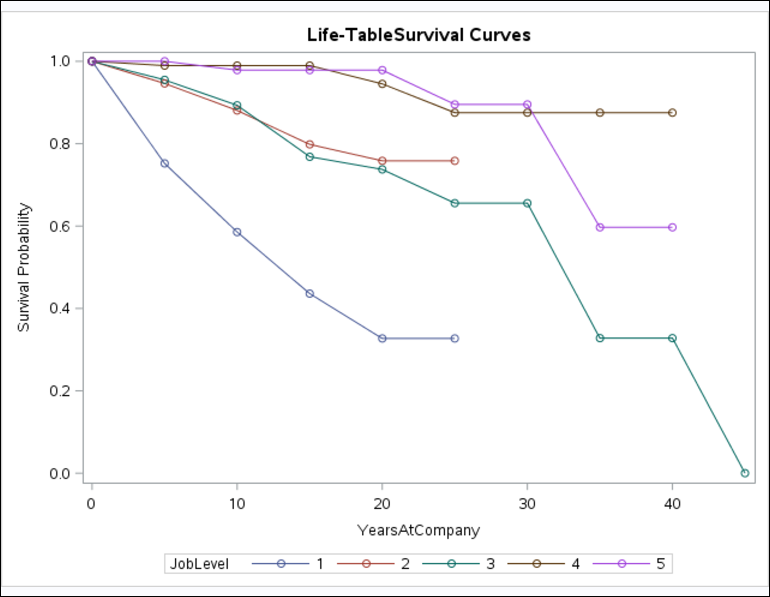
We could seen from the graph shown below that attrition is directly dependent on the job satisfaction of the employees.



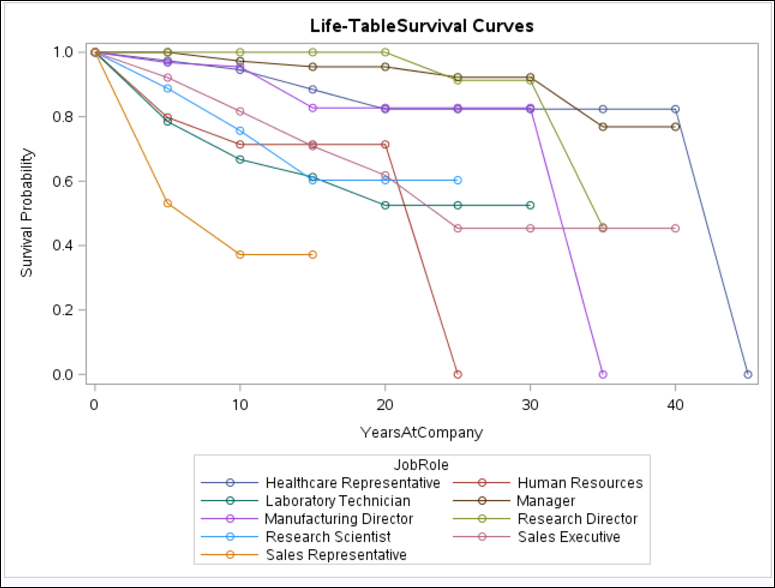


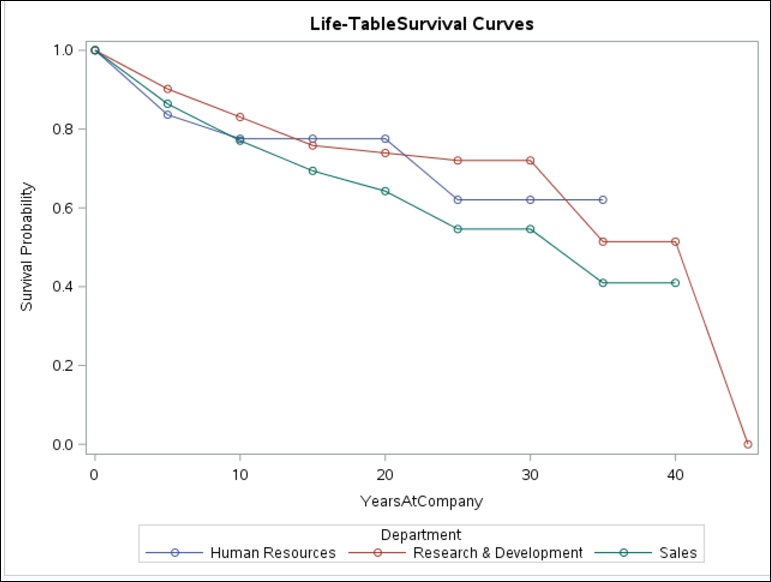
The survival plot for work life balance with respect to survival probability shows that the employees with least (or no) work life balance has higher hazard rate than others. We also tried to find the relationship between work life balance and overtime, and concluded that employees who work overtime does not seem to have best work life balance. So, management could think either to provide compensatory-offs to the overtime employees or other rewards to boost up their willingness to work with company. Also, reward the employees who are giving extra time and efforts in their work.



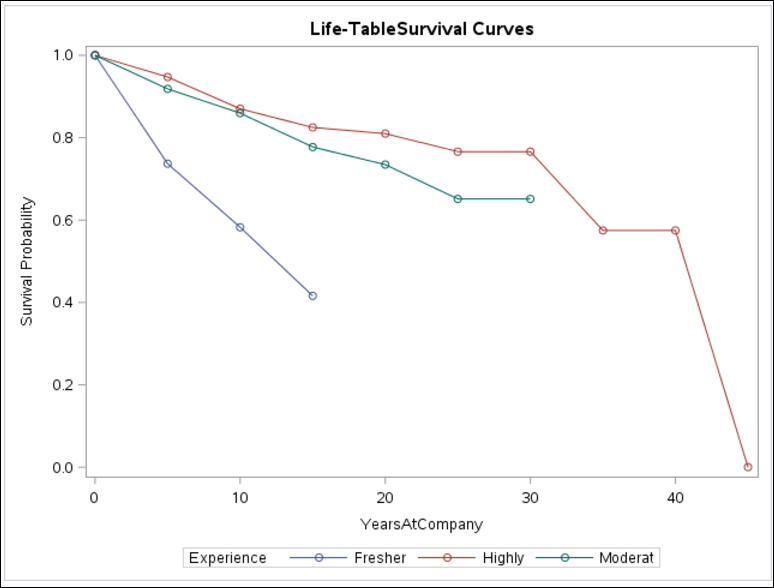


As stated before, the employees at lower job levels are less likely to survive in the company for longer duration. It is apparent from the survival plot, job level 1 employees have least survival probability. On considering the survival probability with the job role, we could observe that sales representatives are most prone to leave the company, followed by lab technicians, HR and research scientist. Also, directors and managers which must be at higher job level has the most survival probability.

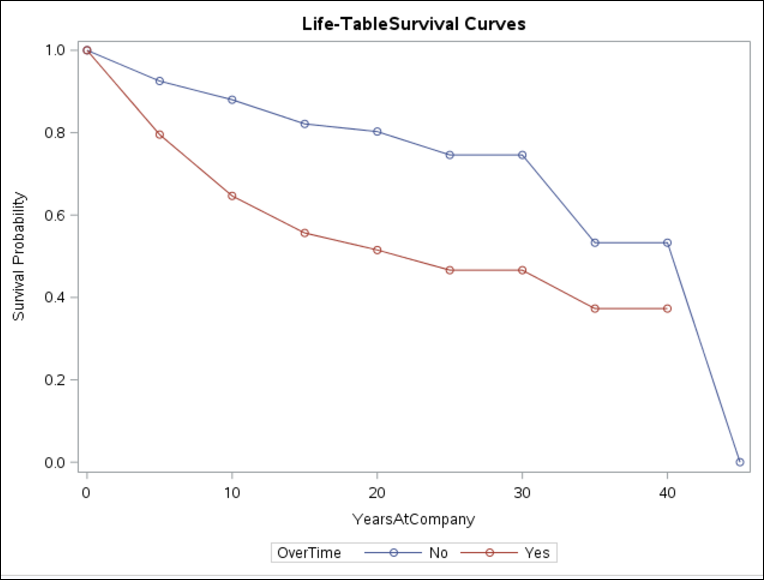




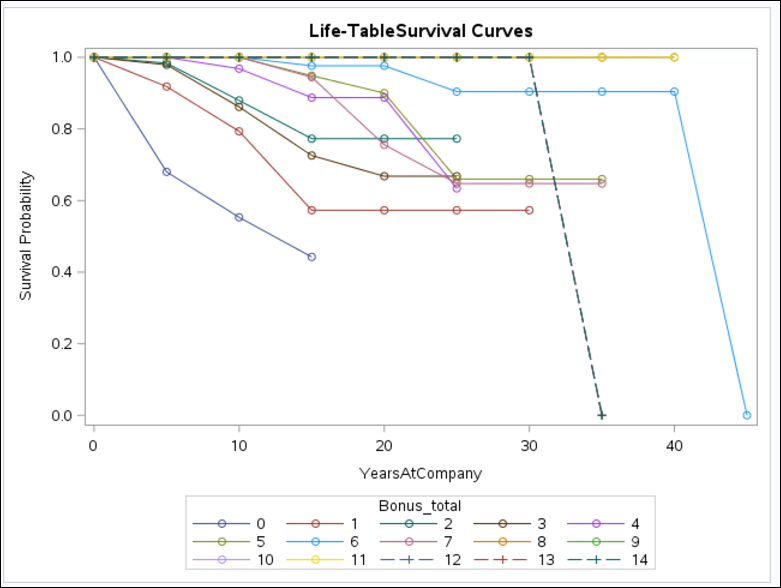
Department does not seem to play a major role in the attrition of the employees.



We tried to categorize the employees, as freshers (employees with age less than 30) moderate (between 30 to 45 years) and experienced (above 45 years). We figured out that freshers are leaving at much higher rate than the other two categories.

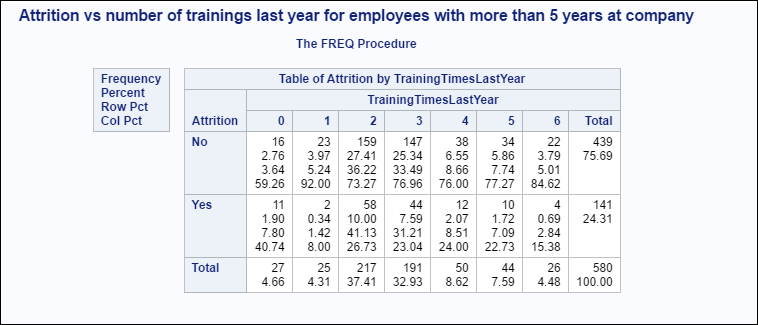


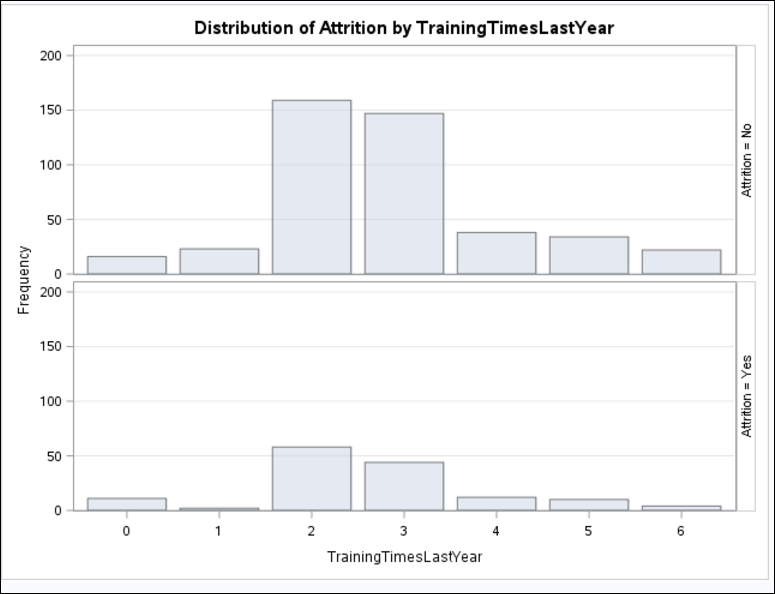
The above plot again corroborates the fact that employees who work overtime has very less survival probability over time as compared to ones who do not work overtime.



Our team was curious to see the impact of bonus on the attrition. As the company data has provided detailed list for 40 years of bonus for each employee. We decided to transpose the data into single column, with total number of bonuses an employee received in his/her tenure at company. For simplicity, we have clubbed data points having ‘NA’ value with 0.

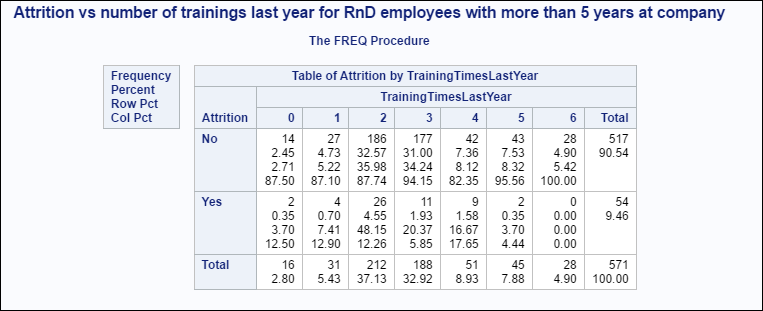
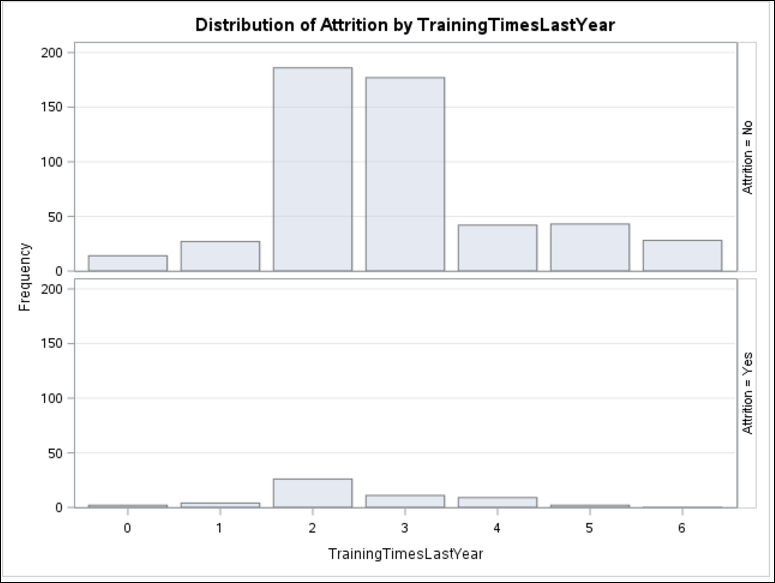
On plotting the survival graph, we observed that people who have not received bonus in 15 years of experience in the company has most hazard rate whereas people with more bonus in their career are less likely to leave.

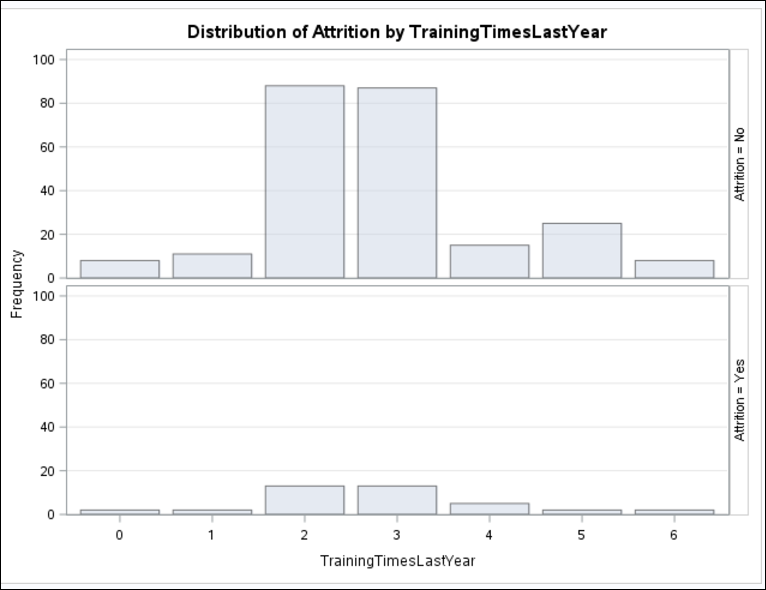
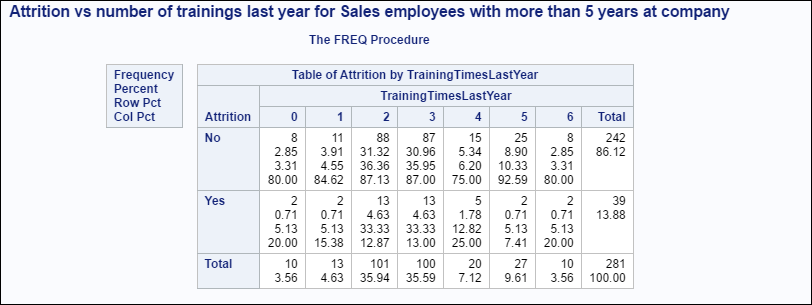


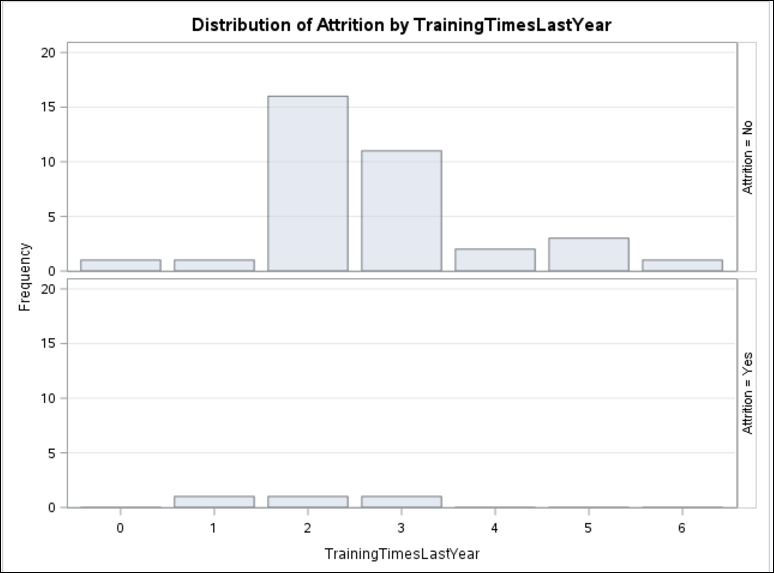
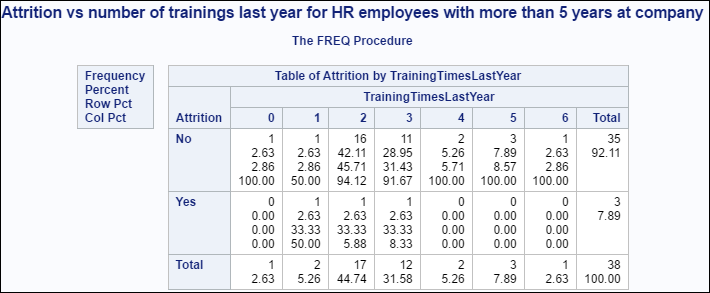


Since the COO of the Fermalogis expressed his concern for experienced employees who have been trained with Executive Training Program are leaving the company after becoming more skilled and competent. We decided to explore the data for employees with more than 5 years of experience in the company and number of trainings in last year. The employees with 0 training are more likely to leave. Also employees who attended 1-3 trainings are staying with the company. Hence, we could conclude that for experienced employees number of training is not the reason to leave the company.

To further analyze the experienced employees with number of trainings, we decided to visualize them department wise.

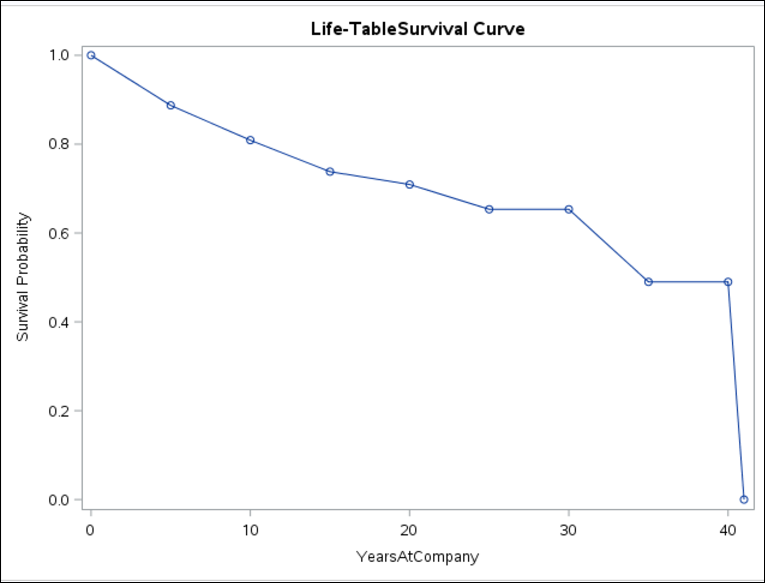
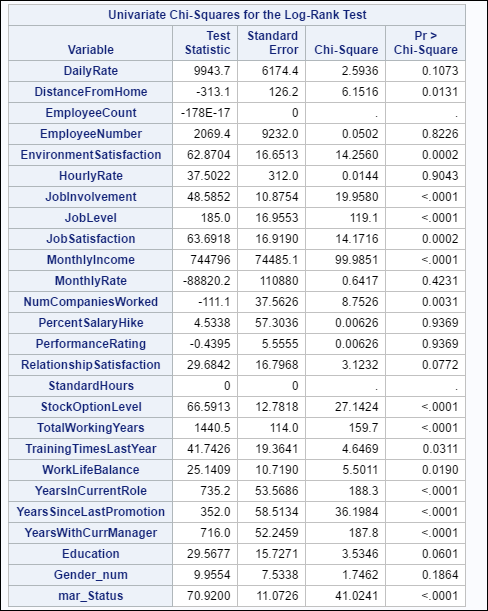
 





# Identifying the significant covariates for the model

The team has used LIFETEST to get an idea of the significant variables so that we can proceed with LIFEREG with different methods for model.



By looking at the chi square values we can identify the significant variables are: Distance from home, Environment Satisfaction, Job Involvement, Job Level, Job Satisfaction, Monthly Income, Number of companies worked, Stock Option, Total working years, Years since last promotion, years with current manager and marital status.

# Model Fit (Goodness of Fit)

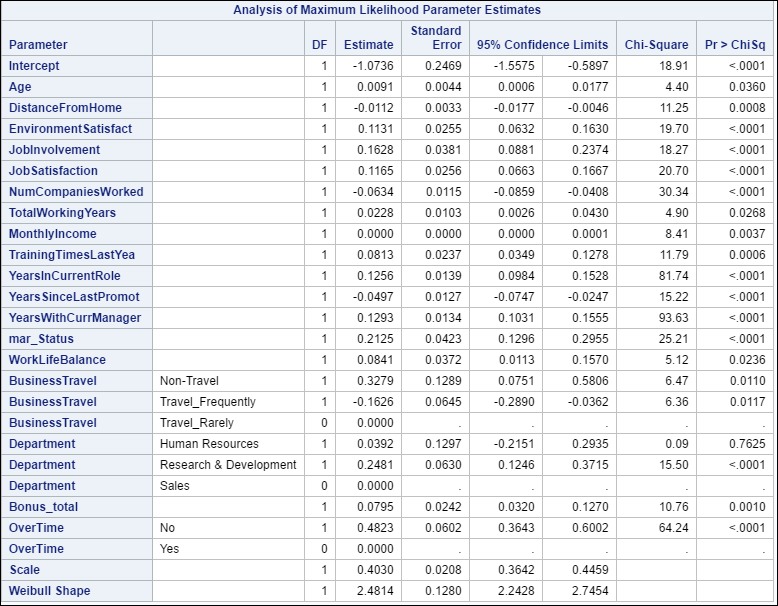
The team has used the AFT (Accelerated failure time) models to determine the important factors of employee’s attrition. To model the data, the team has considered all the possible distributions and selected the best hypothesis based on the log likelihood and fit statistics of the Hypothesis. We finalized on Weibull for better performance and simplicity.

|  |  |
| --- | --- |
| **Method** | **Log Likelihood** |
| Lognormal | -410.1804885 |
| Exponential | -512.9721128 |
| Weibull | -398.295252 |
| Gamma | -326.846816 |
| Normal | -769.4657793 |
| Logistic | -749.2256276 |

Removing insignificant covariates: DailyRate, Employeecount, EmployeeNumber, HourlyRate, MonthlyRate, PresentSalaryHike, PerformanceRating, RelationshipSatisfaction, Standard Hours, StockOptionLevel, WorkLifeBalance, Gender, JobRole, Education, EducationField, JobLevel



# Complete list of coefficients with positive and negative effect:

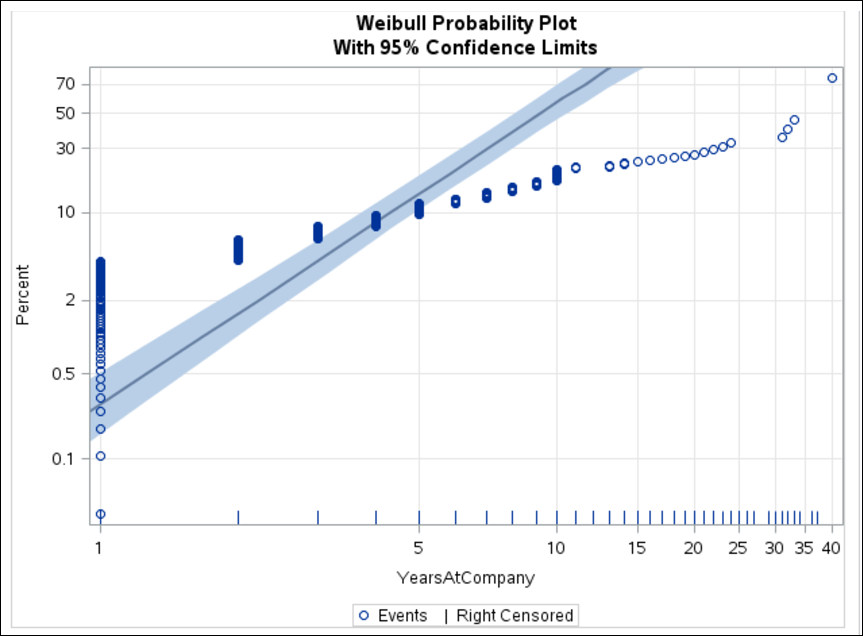


Positive Factors:

* Environment satisfaction is encouraging the employees to stay back
* Job involvement is required
* Non- travelling employees are happy
* Employees with no overtime

Negative Factors:

* Years since last promotion
* Frequent Travel
* Distance from home
* Number of companies worked



Even though the model is not perfect, it is performing considerable fair with narrow confidence interval on the data.

# Major Findings in conclusion:

1. It is observed that people who work overtime and travel occasionally are more inclined to leave the association.
2. Interestingly, people who work overtime and travel frequently are also leaving the organization.
3. People who receive lesser bonus and few promotions are prone to leave.
4. Surprisingly, it is observed that, the more the job involvement, lesser the job satisfaction.
5. Overtime accompanied with Low job satisfaction has less probability of survival in the company.

# Recommendations:

1. **SAY THANK YOU** - Prioritize employee recognition and you can ensure a positive, productive, innovative organizational climate.
2. **KNOW THE PULSE** -Frequent surveys can be conducted on the work environment satisfaction and one on one meetings with HR department addressing employee concerns.
3. "**Money isn't everything, but it certainly helps**," says Kingsley. " - the employee salaries must be reconsidered each year in view of their execution and endeavors..
4. **OFFER A SENSE OF OWNERSHIP** – This can be achieved my offering stocks and responsibilities to the individuals. This increases corporate morale and empowers the employees.
5. **WORK LIFE BALANCE** – Conduct sessions to bring awareness in the employees about work life balance and its importance
6. **PURSUE THE PASSION** – It is important to have career path discussions with the employees and make better utilization of their interests and efforts.

# References:

* <https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#lifereg_toc.htm>
* <https://www.google.com/>
* The Little SAS book – Lora D. Delwiche and Susan J. Slaughter