

Turnover Project

Priyajit Chowdhury

03/03/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Importing data set into R as excel file

Employee number and employee count are not considered as they are not related and unique number

```
library(readxl)
Turnover_issues<- read_excel("D:/Project/WA_Fn-UseC_-HR-Employee-Attrition.xls",
                             col_types = c("numeric", "text", "text",
                                             "numeric", "text", "numeric", "numeric",
                                             "text", "numeric",
                                             "text", "numeric", "numeric", "numeric",
                                             "text", "numeric", "text", "numeric",
                                             "numeric", "numeric", "text", "text",
                                             "numeric", "numeric",
                                             "numeric", "numeric", "numeric",
                                             "numeric", "numeric", "numeric",
                                             "numeric", "numeric", "numeric",
                                             "numeric"))
```

1. Finding missing values to avoid biased estimation and invalid conclusions

```
sum(is.na(Turnover_issues))
```

```
## [1] 0
```

2. Unclassing Dataset to remove information with equal name second argument

```
Turnover_issues_fac <- as.data.frame (unclass(Turnover_issues))
```

3. Dataset Structure

```
str(Turnover_issues_fac)
```

```
## 'data.frame':    1470 obs. of  33 variables:
## $ Age                : num  41 49 37 33 27 32 59 30 38 36 ...
## $ Attrition           : chr   "Yes" "No" "Yes" "No" ...
## $ BusinessTravel      : chr   "Travel_Rarely" "Travel_Frequently" "Travel_Rarely" "Travel_Frequently" ...
## $ DailyRate           : num  1102 279 1373 1392 591 ...
## $ Department          : chr   "Sales" "Research & Development" "Research & Development" "Research & Development" ...
## $ DistanceFromHome    : num    1 8 2 3 2 2 3 24 23 27 ...
## $ Education           : num    2 1 2 4 1 2 3 1 3 3 ...
## $ EducationField       : chr   "Life Sciences" "Life Sciences" "Other" "Life Sciences" ...
## $ EnvironmentSatisfaction : num    2 3 4 4 1 4 3 4 4 3 ...
## $ Gender              : chr   "Female" "Male" "Male" "Female" ...
## $ HourlyRate          : num   94 61 92 56 40 79 81 67 44 94 ...
## $ JobInvolvement       : num    3 2 2 3 3 3 4 3 2 3 ...
## $ JobLevel            : num    2 2 1 1 1 1 1 1 3 2 ...
## $ JobRole             : chr   "Sales Executive" "Research Scientist" "Laboratory Technician" "Research Scientist" ...
## $ JobSatisfaction      : num    4 2 3 3 2 4 1 3 3 3 ...
## $ MaritalStatus        : chr   "Single" "Married" "Single" "Married" ..
.
## $ MonthlyIncome       : num  5993 5130 2090 2909 3468 ...
## $ MonthlyRate         : num  19479 24907 2396 23159 16632 ...
## $ NumCompaniesWorked   : num    8 1 6 1 9 0 4 1 0 6 ...
## $ Over18              : chr   "Y" "Y" "Y" "Y" ...
## $ OverTime             : chr   "Yes" "No" "Yes" "Yes" ...
## $ PercentSalaryHike    : num   11 23 15 11 12 13 20 22 21 13 ...
## $ PerformanceRating    : num    3 4 3 3 3 3 4 4 4 3 ...
## $ RelationshipSatisfaction : num    1 4 2 3 4 3 1 2 2 2 ...
## $ StandardHours        : num   80 80 80 80 80 80 80 80 80 80 ...
## $ StockOptionLevel     : num    0 1 0 0 1 0 3 1 0 2 ...
## $ TotalWorkingYears    : num    8 10 7 8 6 8 12 1 10 17 ...
## $ TrainingTimesLastYear : num    0 3 3 3 3 2 3 2 2 3 ...
## $ WorkLifeBalance      : num    1 3 3 3 3 2 2 3 3 2 ...
## $ YearsAtCompany       : num    6 10 0 8 2 7 1 1 9 7 ...
## $ YearsInCurrentRole   : num    4 7 0 7 2 7 0 0 7 7 ...
## $ YearsSinceLastPromotion : num    0 1 0 3 2 3 0 0 1 7 ...
## $ YearsWithCurrManager : num    5 7 0 0 2 6 0 0 8 7 ...
```

Converting different attributes

4. Attrition where "Yes" as 1, otherwise 0

```
Attrition_Data<-ifelse(Turnover_issues_fac$Attrition=="Yes","1","0")
Turnover_issues_fac$Attrition<-Attrition_Data
Turnover_issues_fac$Attrition<-as.factor(Turnover_issues_fac$Attrition)
```

5. Business travel

```
BusinessTravel<-Turnover_issues$BusinessTravel
BusinessTravel_factor<-factor(BusinessTravel)
str(BusinessTravel_factor)
```

```
## Factor w/ 3 levels "Non-Travel","Travel_Frequently",...: 3 2 3 2 3 2 3 3 2 3 ...
```

```
Turnover_issues_fac$BusinessTravel<-BusinessTravel_factor
```

6. Department

```
Department<-Turnover_issues$Department  
Department_factor<-factor(Department)  
str(Department_factor)
```

```
## Factor w/ 3 levels "Human Resources",...: 3 2 2 2 2 2 2 2 2 2 ...
```

```
Turnover_issues_fac$Department<-Department_factor  
str(Turnover_issues_fac$Department)
```

```
## Factor w/ 3 levels "Human Resources",...: 3 2 2 2 2 2 2 2 2 2 ...
```

7. Education Field

```
EducationField<-Turnover_issues$EducationField  
EducationField_factor<-factor(EducationField)  
str(EducationField_factor)
```

```
## Factor w/ 6 levels "Human Resources",...: 2 2 5 2 4 2 4 2 2 4 ...
```

```
Turnover_issues_fac$EducationField<-EducationField_factor
```

8. Gender

```
Gender<-Turnover_issues$Gender  
Gender_factor<-factor(Gender)  
str(Gender_factor)
```

```
## Factor w/ 2 levels "Female","Male": 1 2 2 1 2 2 1 2 2 2 ...
```

```
Turnover_issues_fac$Gender<-Gender_factor
```

9. Job Role

```
JobRole<-Turnover_issues$JobRole  
JobRole_factor<-factor(JobRole)  
str(JobRole_factor)
```

```
## Factor w/ 9 levels "Healthcare Representative",...: 8 7 3 7 3 3 3 3 5 1 ..  
.
```

```
Turnover_issues_fac$JobRole<-JobRole_factor
```

10. MaritalStatus

```
MaritalStatus<-Turnover_issues$MaritalStatus  
MaritalStatus_factor<-factor(MaritalStatus)  
str(MaritalStatus_factor)
```

```
## Factor w/ 3 levels "Divorced","Married",...: 3 2 3 2 2 3 2 1 3 2 ...
```

```
Turnover_issues_fac$MaritalStatus<-MaritalStatus_factor
```

11. Over18

```
Over18<-Turnover_issues$Over18
Over18_factor<-factor(Over18)
str(Over18_factor)
```

```
## Factor w/ 1 level "Y": 1 1 1 1 1 1 1 1 1 1 ...
```

```
Turnover_issues_fac$Over18<-Over18_factor
```

12. OverTime where "Yes" as 1, otherwise 0

```
Overtime_Data<-ifelse(Turnover_issues_fac$OverTime=="Yes", "1", "0")
Turnover_issues_fac$OverTime<-Overtime_Data
Turnover_issues_fac$OverTime<-as.factor(Turnover_issues_fac$OverTime)
```

13. Changing numeric variables to factor variables

```
Turnover_issues_fac$StockOptionLevel<-as.factor(Turnover_issues_fac$StockOptionLevel)
Turnover_issues_fac$Education<-as.factor(Turnover_issues_fac$Education)
Turnover_issues_fac$EnvironmentSatisfaction<-as.factor(Turnover_issues_fac$EnvironmentSatisfaction)
Turnover_issues_fac$JobInvolvement<-as.factor(Turnover_issues_fac$JobInvolvement)
Turnover_issues_fac$JobLevel<-as.factor(Turnover_issues_fac$JobLevel)
Turnover_issues_fac$JobSatisfaction<-as.factor(Turnover_issues_fac$JobSatisfaction)
Turnover_issues_fac$TrainingTimesLastYear<-as.factor(Turnover_issues_fac$TrainingTimesLastYear)
Turnover_issues_fac$WorkLifeBalance<-as.factor(Turnover_issues_fac$WorkLifeBalance)
Turnover_issues_fac$NumCompaniesWorked<-as.factor(Turnover_issues_fac$NumCompaniesWorked)
Turnover_issues_fac$PerformanceRating<-as.factor(Turnover_issues_fac$PerformanceRating)
Turnover_issues_fac$RelationshipSatisfaction<-as.factor(Turnover_issues_fac$RelationshipSatisfaction)
```

14. Observing revised data structure

```
str(Turnover_issues_fac)
```

```
## 'data.frame': 1470 obs. of 33 variables:
## $ Age : num 41 49 37 33 27 32 59 30 38 36 ...
## $ Attrition : Factor w/ 2 levels "0","1": 2 1 2 1 1 1 1 1 1 1 ...
## $ BusinessTravel : Factor w/ 3 levels "Non-Travel","Travel_Frequently",...: 3 2 3 2 3 2 3 3 2 3 ...
## $ DailyRate : num 1102 279 1373 1392 591 ...
## $ Department : Factor w/ 3 levels "Human Resources",...: 3 2 2 2 2 2 2 2 2 ...
```

```

## $ DistanceFromHome      : num  1 8 2 3 2 2 3 24 23 27 ...
## $ Education              : Factor w/ 5 levels "1","2","3","4",...: 2 1 2
4 1 2 3 1 3 3 ...
## $ EducationField         : Factor w/ 6 levels "Human Resources",...: 2 2
5 2 4 2 4 2 2 4 ...
## $ EnvironmentSatisfaction : Factor w/ 4 levels "1","2","3","4": 2 3 4 4 1
4 3 4 4 3 ...
## $ Gender                 : Factor w/ 2 levels "Female","Male": 1 2 2 1 2
2 1 2 2 2 ...
## $ HourlyRate             : num  94 61 92 56 40 79 81 67 44 94 ...
## $ JobInvolvement         : Factor w/ 4 levels "1","2","3","4": 3 2 2 3 3
3 4 3 2 3 ...
## $ JobLevel               : Factor w/ 5 levels "1","2","3","4",...: 2 2 1
1 1 1 1 1 3 2 ...
## $ JobRole                : Factor w/ 9 levels "Healthcare Representative
",...: 8 7 3 7 3 3 3 3 5 1 ...
## $ JobSatisfaction        : Factor w/ 4 levels "1","2","3","4": 4 2 3 3 2
4 1 3 3 3 ...
## $ MaritalStatus          : Factor w/ 3 levels "Divorced","Married",...: 3
2 3 2 2 3 2 1 3 2 ...
## $ MonthlyIncome          : num  5993 5130 2090 2909 3468 ...
## $ MonthlyRate            : num  19479 24907 2396 23159 16632 ...
## $ NumCompaniesWorked     : Factor w/ 10 levels "0","1","2","3",...: 9 2 7
2 10 1 5 2 1 7 ...
## $ Over18                 : Factor w/ 1 level "Y": 1 1 1 1 1 1 1 1 1 1 ..
.
## $ OverTime               : Factor w/ 2 levels "0","1": 2 1 2 2 1 1 2 1 1
1 ...
## $ PercentSalaryHike      : num  11 23 15 11 12 13 20 22 21 13 ...
## $ PerformanceRating      : Factor w/ 2 levels "3","4": 1 2 1 1 1 1 2 2 2
1 ...
## $ RelationshipSatisfaction: Factor w/ 4 levels "1","2","3","4": 1 4 2 3 4
3 1 2 2 2 ...
## $ StandardHours          : num  80 80 80 80 80 80 80 80 80 80 ...
## $ StockOptionLevel       : Factor w/ 4 levels "0","1","2","3": 1 2 1 1 1 2
1 4 2 1 3 ...
## $ TotalWorkingYears      : num  8 10 7 8 6 8 12 1 10 17 ...
## $ TrainingTimesLastYear  : Factor w/ 7 levels "0","1","2","3",...: 1 4 4
4 4 3 4 3 3 4 ...
## $ WorkLifeBalance        : Factor w/ 4 levels "1","2","3","4": 1 3 3 3 3
2 2 3 3 2 ...
## $ YearsAtCompany         : num  6 10 0 8 2 7 1 1 9 7 ...
## $ YearsInCurrentRole     : num  4 7 0 7 2 7 0 0 7 7 ...
## $ YearsSinceLastPromotion : num  0 1 0 3 2 3 0 0 1 7 ...
## $ YearsWithCurrManager   : num  5 7 0 0 2 6 0 0 8 7 ...

```

Summarization of Dataset

```
summary(Turnover_issues_fac)
```

##	Age	Attrition	BusinessTravel	DailyRate	
##	Min. :18.00	0:1233	Non-Travel : 150	Min. : 102.0	
##	1st Qu.:30.00	1: 237	Travel_Frequently: 277	1st Qu.: 465.0	
##	Median :36.00		Travel_Rarely :1043	Median : 802.0	
##	Mean :36.92			Mean : 802.5	
##	3rd Qu.:43.00			3rd Qu.:1157.0	
##	Max. :60.00			Max. :1499.0	
##					
##		Department	DistanceFromHome	Education	
##	Human Resources	: 63	Min. : 1.000	1:170	
##	Research & Development	:961	1st Qu.: 2.000	2:282	
##	Sales	:446	Median : 7.000	3:572	
##			Mean : 9.193	4:398	
##			3rd Qu.:14.000	5: 48	
##			Max. :29.000		
##					
##		EducationField	EnvironmentSatisfaction	Gender	HourlyRate
##	Human Resources	: 27	1:284	Female:588	Min. : 30.00
##	Life Sciences	:606	2:287	Male :882	1st Qu.: 48.00
##	Marketing	:159	3:453		Median : 66.00
##	Medical	:464	4:446		Mean : 65.89
##	Other	: 82			3rd Qu.: 83.75
##	Technical Degree	:132			Max. :100.00
##					
##	JobInvolvement	JobLevel		JobRole	JobSatisfaction
##	1: 83	1:543	Sales Executive	:326	1:289
##	2:375	2:534	Research Scientist	:292	2:280
##	3:868	3:218	Laboratory Technician	:259	3:442
##	4:144	4:106	Manufacturing Director	:145	4:459
##		5: 69	Healthcare Representative	:131	
##			Manager	:102	
##			(Other)	:215	
##	MaritalStatus	MonthlyIncome	MonthlyRate	NumCompaniesWorked	Over18
##	Divorced:327	Min. : 1009	Min. : 2094	1 :521	Y:1470
##	Married :673	1st Qu.: 2911	1st Qu.: 8047	0 :197	
##	Single :470	Median : 4919	Median :14236	3 :159	
##		Mean : 6503	Mean :14313	2 :146	
##		3rd Qu.: 8379	3rd Qu.:20462	4 :139	
##		Max. :19999	Max. :26999	7 : 74	
##				(Other):234	
##	OverTime	PercentSalaryHike	PerformanceRating	RelationshipSatisfaction	
##	0:1054	Min. :11.00	3:1244	1:276	
##	1: 416	1st Qu.:12.00	4: 226	2:303	
##		Median :14.00		3:459	

```

##          Mean    :15.21                      4:432
##          3rd Qu.:18.00
##          Max.    :25.00
##
## StandardHours StockOptionLevel TotalWorkingYears TrainingTimesLastYear
## Min.    :80      0:631           Min.    : 0.00      0: 54
## 1st Qu.:80      1:596           1st Qu.: 6.00      1: 71
## Median :80      2:158           Median :10.00     2:547
## Mean    :80      3: 85           Mean    :11.28     3:491
## 3rd Qu.:80           3rd Qu.:15.00     4:123
## Max.    :80           Max.    :40.00     5:119
##                                           6: 65
## WorkLifeBalance YearsAtCompany   YearsInCurrentRole YearsSinceLastPromoti
on
## 1: 80           Min.    : 0.000   Min.    : 0.000   Min.    : 0.000
## 2:344           1st Qu.: 3.000   1st Qu.: 2.000   1st Qu.: 0.000
## 3:893           Median : 5.000   Median : 3.000   Median : 1.000
## 4:153           Mean    : 7.008   Mean    : 4.229   Mean    : 2.188
##                3rd Qu.: 9.000   3rd Qu.: 7.000   3rd Qu.: 3.000
##                Max.    :40.000   Max.    :18.000   Max.    :15.000
##
## YearsWithCurrManager
## Min.    : 0.000
## 1st Qu.: 2.000
## Median : 3.000
## Mean    : 4.123
## 3rd Qu.: 7.000
## Max.    :17.000
##

```

Observing top few outputs

```
head(Turnover_issues_fac)
```

```

##   Age Attrition   BusinessTravel DailyRate      Department
## 1  41         1      Travel_Rarely    1102             Sales
## 2  49         0 Travel_Frequently     279 Research & Development
## 3  37         1      Travel_Rarely    1373 Research & Development
## 4  33         0 Travel_Frequently    1392 Research & Development
## 5  27         0      Travel_Rarely     591 Research & Development
## 6  32         0 Travel_Frequently    1005 Research & Development
## DistanceFromHome Education EducationField EnvironmentSatisfaction Gender
## 1           1           2 Life Sciences             2 Female
## 2           8           1 Life Sciences             3 Male
## 3           2           2      Other                4 Male
## 4           3           4 Life Sciences             4 Female
## 5           2           1      Medical              1 Male
## 6           2           2 Life Sciences             4 Male
## HourlyRate JobInvolvement JobLevel      JobRole JobSatisfaction
## 1          94             3         2      Sales Executive          4

```

## 2	61	2	2	Research Scientist	2
## 3	92	2	1	Laboratory Technician	3
## 4	56	3	1	Research Scientist	3
## 5	40	3	1	Laboratory Technician	2
## 6	79	3	1	Laboratory Technician	4
##	MaritalStatus	MonthlyIncome	MonthlyRate	NumCompaniesWorked	Over18
me					OverTi
## 1	Single	5993	19479	8	Y
1					
## 2	Married	5130	24907	1	Y
0					
## 3	Single	2090	2396	6	Y
1					
## 4	Married	2909	23159	1	Y
1					
## 5	Married	3468	16632	9	Y
0					
## 6	Single	3068	11864	0	Y
0					
##	PercentSalaryHike	PerformanceRating	RelationshipSatisfaction	StandardHou	
rs					
## 1	11	3	1		
80					
## 2	23	4	4		
80					
## 3	15	3	2		
80					
## 4	11	3	3		
80					
## 5	12	3	4		
80					
## 6	13	3	3		
80					
##	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	
## 1	0	8	0	1	
## 2	1	10	3	3	
## 3	0	7	3	3	
## 4	0	8	3	3	
## 5	1	6	3	3	
## 6	0	8	2	2	
##	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion		
## 1	6	4	0		
## 2	10	7	1		
## 3	0	0	0		
## 4	8	7	3		
## 5	2	2	2		
## 6	7	7	3		
##	YearsWithCurrManager				
## 1	5				
## 2	7				


```
## 3          0
## 4          0
## 5          2
## 6          6
```

Descriptive Analysis

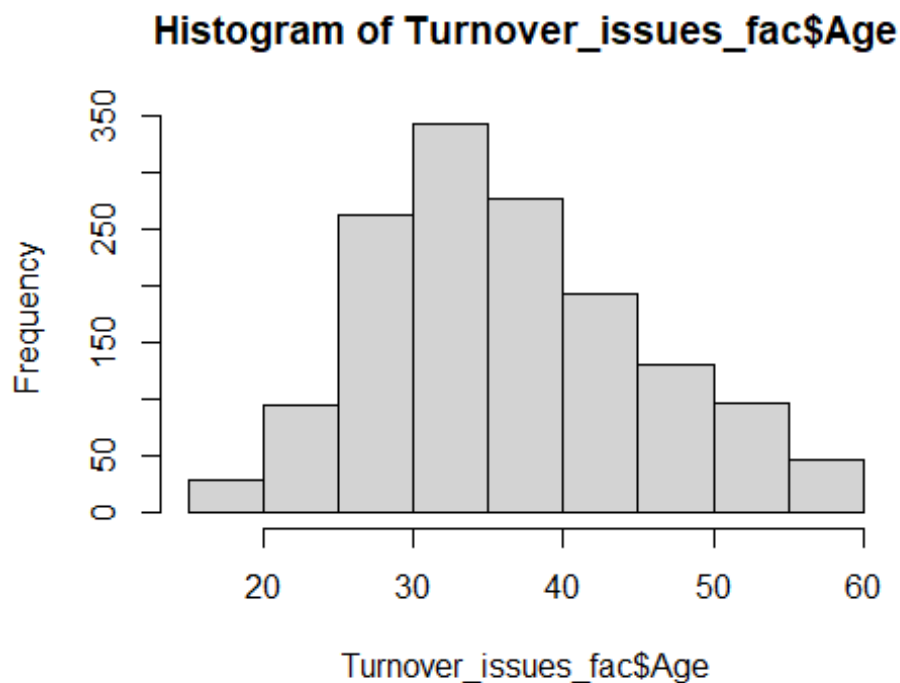
#A. Analyzing each variables

i. Age

```
summary(Turnover_issues_fac$Age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  18.00   30.00   36.00   36.92   43.00   60.00
```

```
hist(Turnover_issues_fac$Age)
```

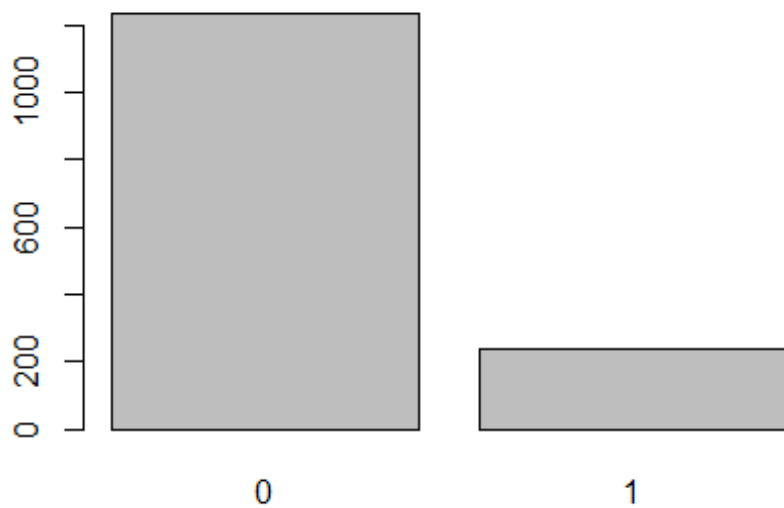


ii. Attrition

```
summary(Turnover_issues_fac$Attrition)
```

```
##      0      1
## 1233  237
```

```
plot(Turnover_issues_fac$Attrition)
```

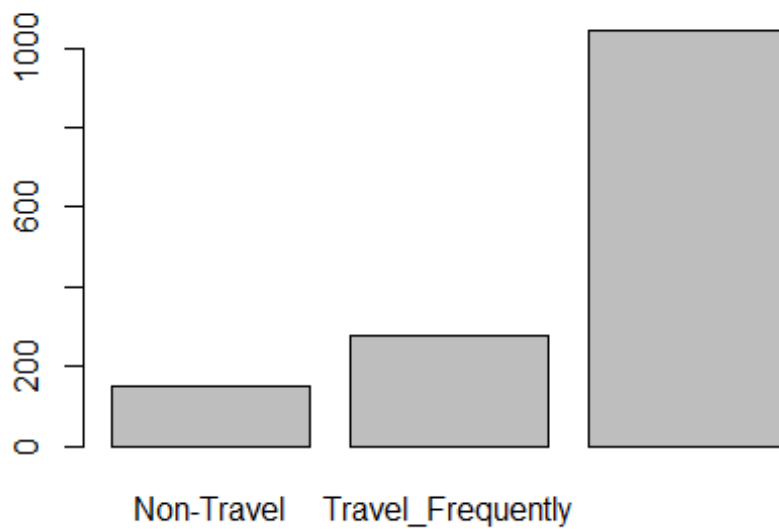


iii. Business Travel

```
summary(Turnover_issues_fac$BusinessTravel)
```

```
##           Non-Travel Travel_Frequently   Travel_Rarely  
##           150           277           1043
```

```
plot(Turnover_issues_fac$BusinessTravel)
```

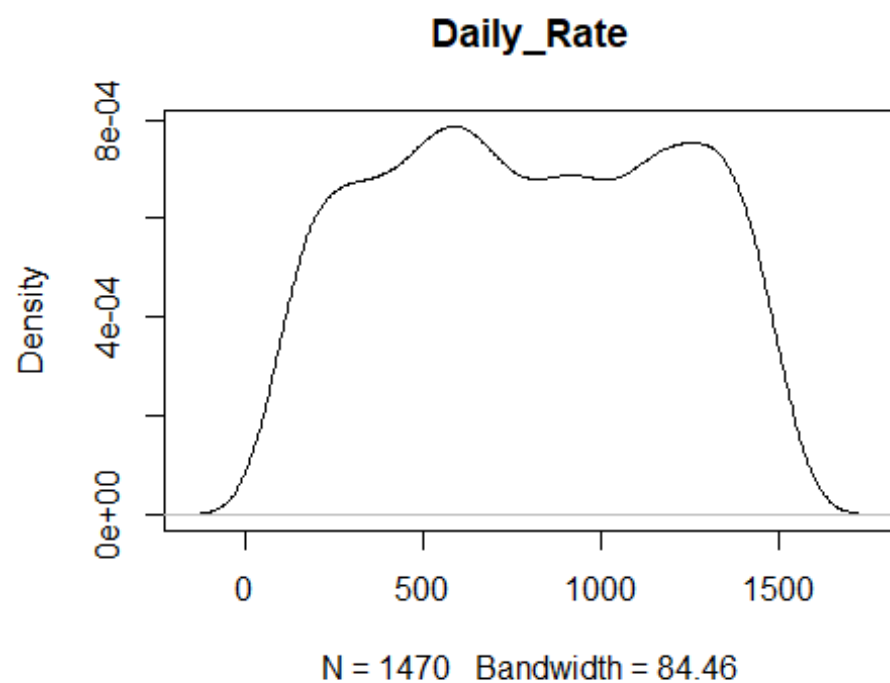


iv. Daily rate

```
summary(Turnover_issues_fac$DailyRate)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   102.0   465.0   802.0   802.5  1157.0  1499.0
```

```
plot(density(Turnover_issues_fac$DailyRate),main="Daily_Rate")
```

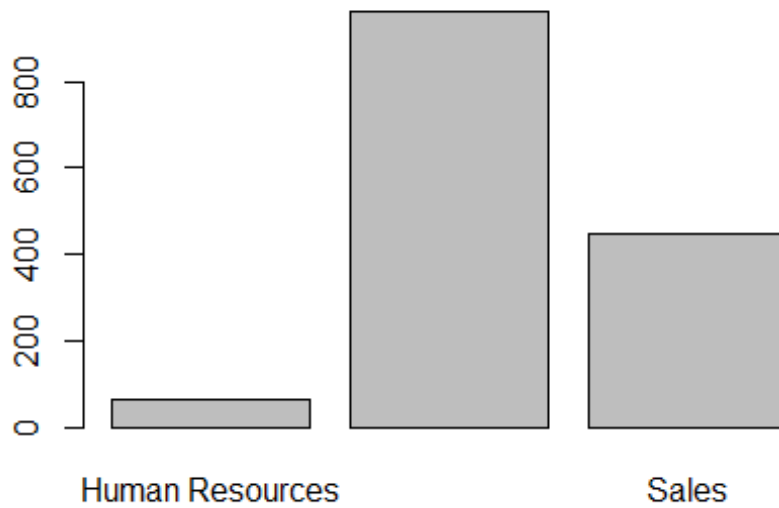


v. Department

```
summary(Turnover_issues_fac$Department)
```

##	Human Resources	Research & Development	Sales
##	63	961	446

```
plot(Turnover_issues_fac$Department)
```

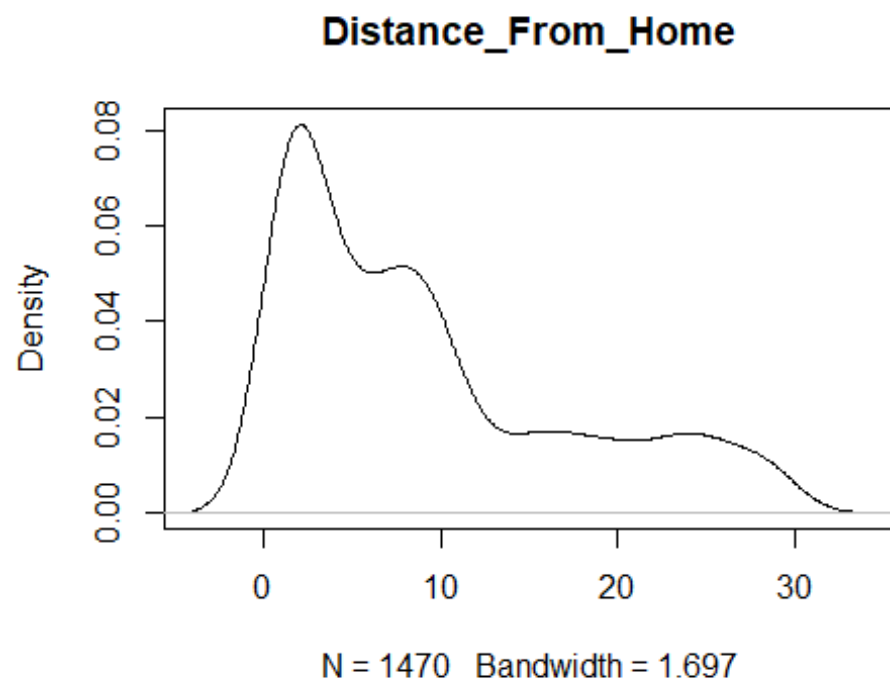


vi. Distance from home

```
summary(Turnover_issues_fac$DistanceFromHome)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   2.000   7.000   9.193  14.000  29.000
```

```
plot(density(Turnover_issues_fac$DistanceFromHome), main="Distance_From_Home"
)
```

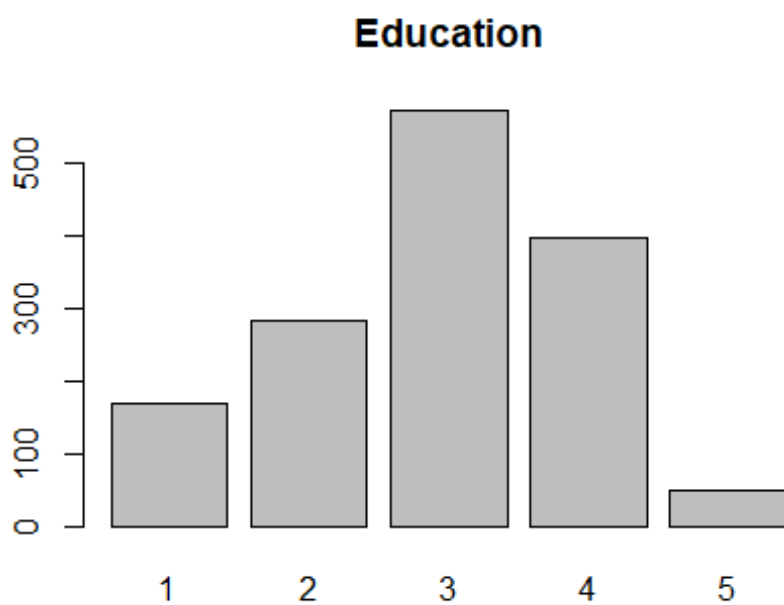


vii. Education

```
summary(Turnover_issues_fac$Education)
```

```
##    1    2    3    4    5  
## 170 282 572 398  48
```

```
plot(Turnover_issues_fac$Education, main = "Education")
```

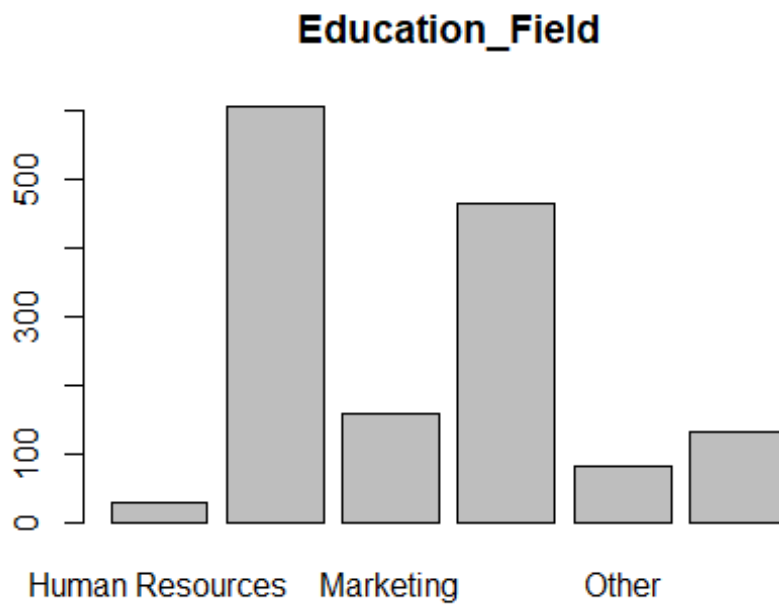


viii. Educationfield

```
summary(Turnover_issues_fac$EducationField)
```

```
## Human Resources    Life Sciences      Marketing      Medical
##           27           606           159           464
##      Other Technical Degree
##           82           132
```

```
plot(Turnover_issues_fac$EducationField, main = "Education_Field")
```

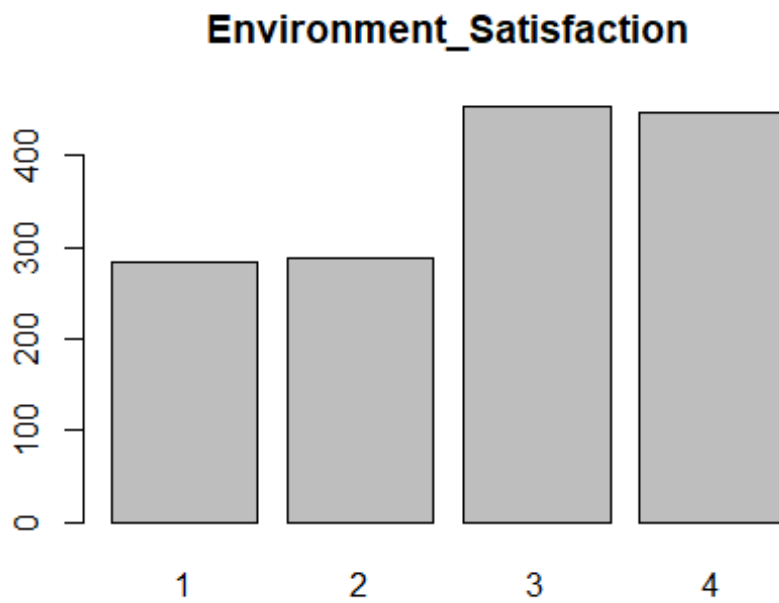


ix. Environment Satisfaction

```
summary(Turnover_issues_fac$EnvironmentSatisfaction)
```

```
##    1    2    3    4  
## 284 287 453 446
```

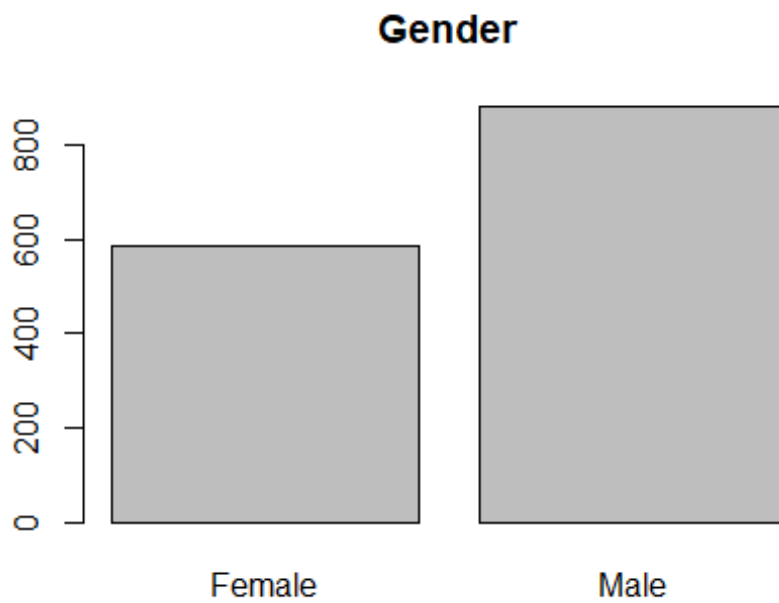
```
plot(Turnover_issues_fac$EnvironmentSatisfaction, main="Environment_Satisfaction")
```

```
x. Gender
summary(Turnover_issues_fac$Gender)

## Female    Male
##    588    882

plot(Turnover_issues_fac$Gender, main="Gender")
```

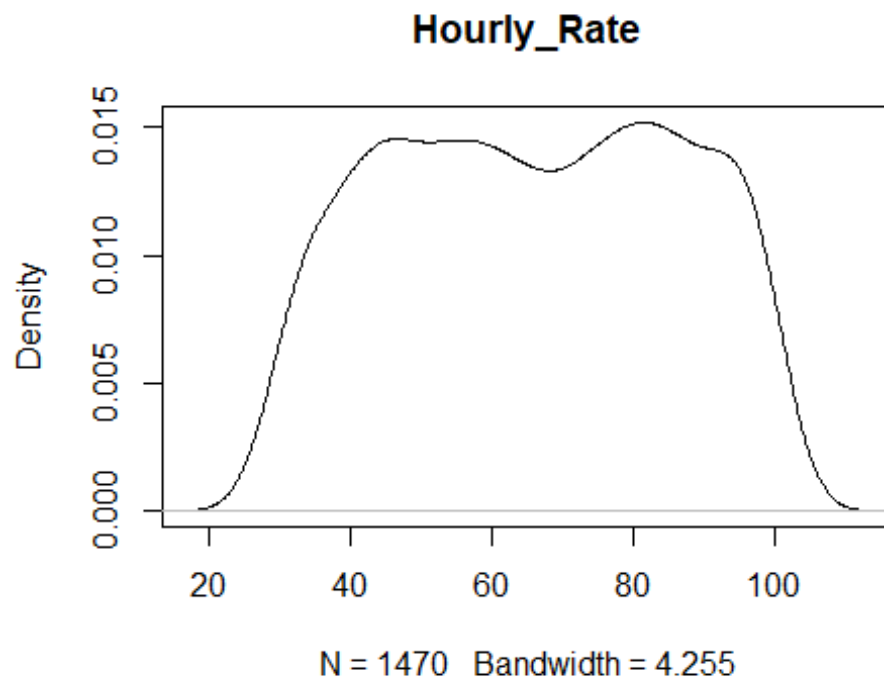


xi. Hourly Rate

```
summary(Turnover_issues_fac$HourlyRate)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  30.00   48.00   66.00   65.89   83.75   100.00
```

```
plot(density(Turnover_issues_fac$HourlyRate), main = "Hourly_Rate")
```

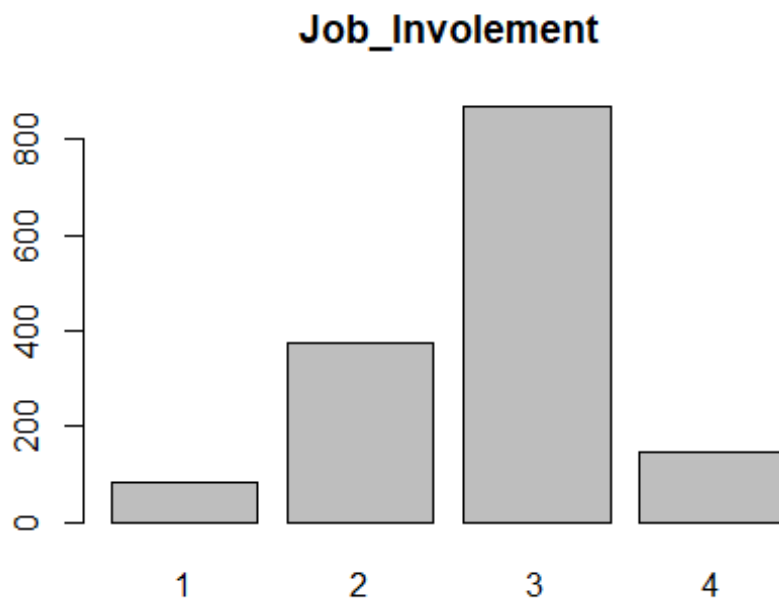


xii. Job involvement

```
summary(Turnover_issues_fac$JobInvolvement)
```

```
##    1    2    3    4  
##  83 375 868 144
```

```
plot(Turnover_issues_fac$JobInvolvement, main="Job_Involement")
```



xiii. Job Level

```
summary(Turnover_issues_fac$JobLevel)
```

```
##    1    2    3    4    5  
## 543 534 218 106  69
```

```
hist(Turnover_issues$JobLevel, main="Job_Level")
```

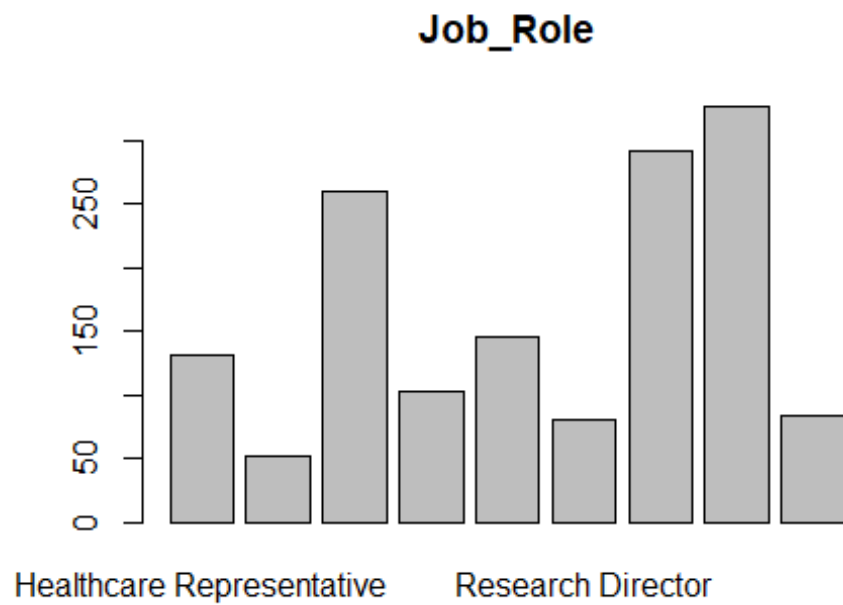


xiv. Job Role

```
summary(Turnover_issues_fac$JobRole)
```

## Healthcare Representative	Human Resources	Laboratory Technician
## 131	52	
259		
## Manager	Manufacturing Director	Research Director
## 102	145	
80		
## Research Scientist	Sales Executive	Sales Representative
## 292	326	
83		

```
plot(Turnover_issues_fac$JobRole, main="Job_Role")
```

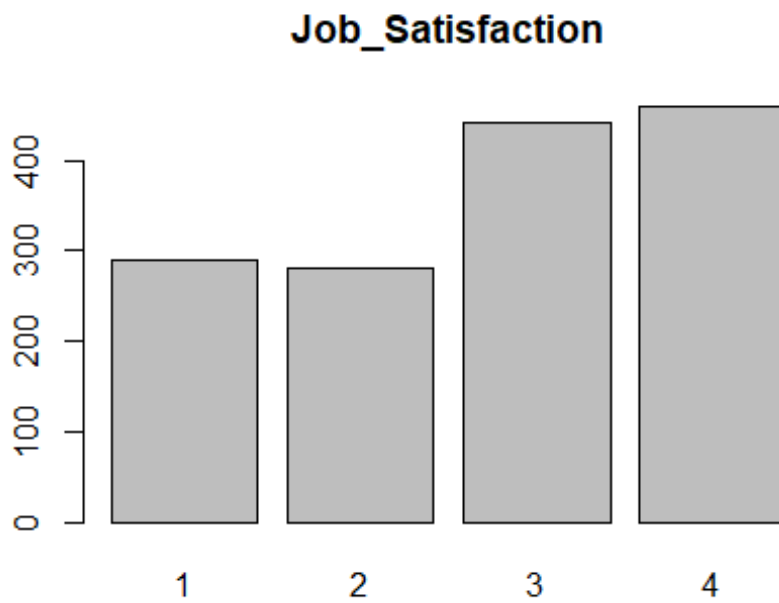


xv. Job Satisfaction

```
summary(Turnover_issues_fac$JobSatisfaction)
```

```
##    1    2    3    4  
## 289 280 442 459
```

```
plot(Turnover_issues_fac$JobSatisfaction, main= "Job_Satisfaction")
```

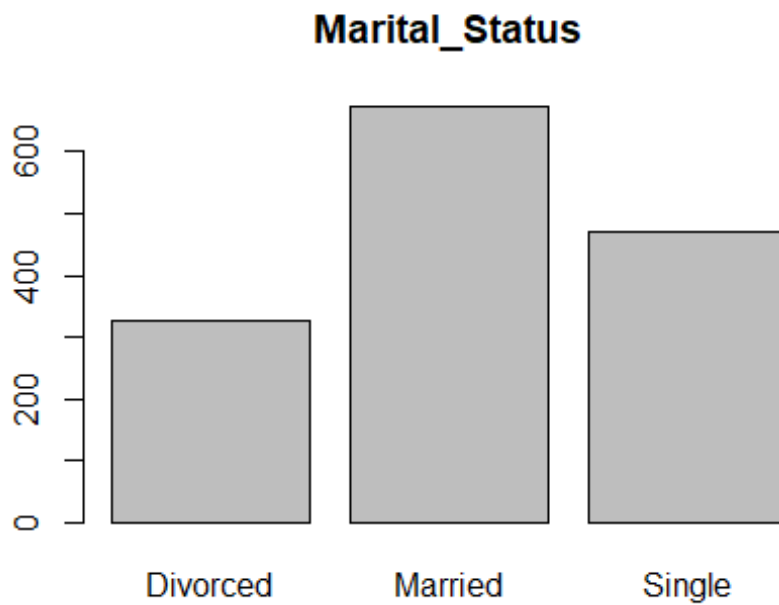


xvi. Marital Status

```
summary(Turnover_issues_fac$MaritalStatus)
```

```
## Divorced  Married   Single  
##      327      673      470
```

```
plot(Turnover_issues_fac$MaritalStatus, main="Marital_Status")
```

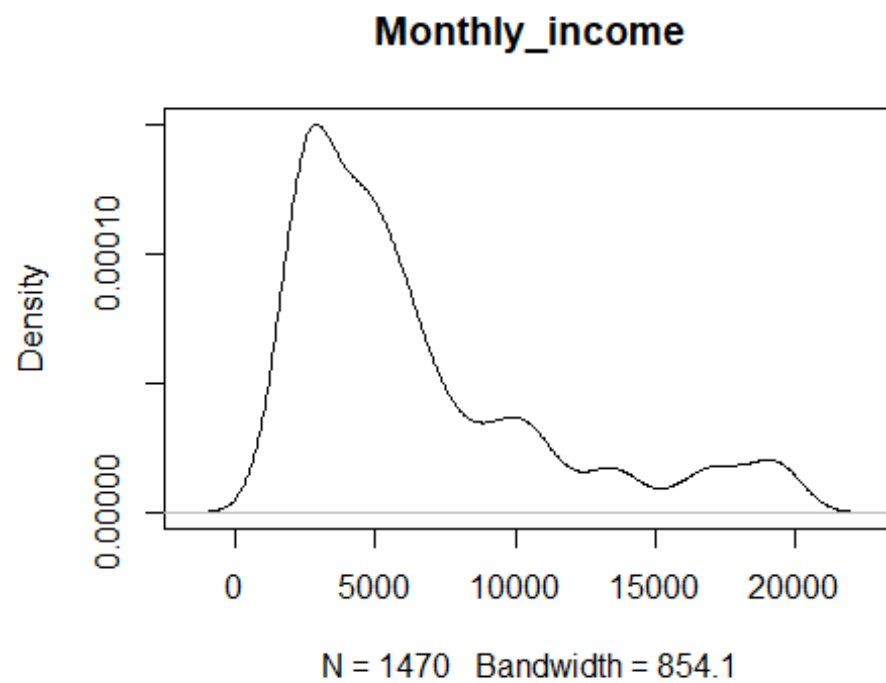


xvii. Monthly Income

```
summary(Turnover_issues_fac$MonthlyIncome)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1009   2911   4919   6503   8379   19999
```

```
plot(density(Turnover_issues_fac$MonthlyIncome), main = "Monthly_income")
```

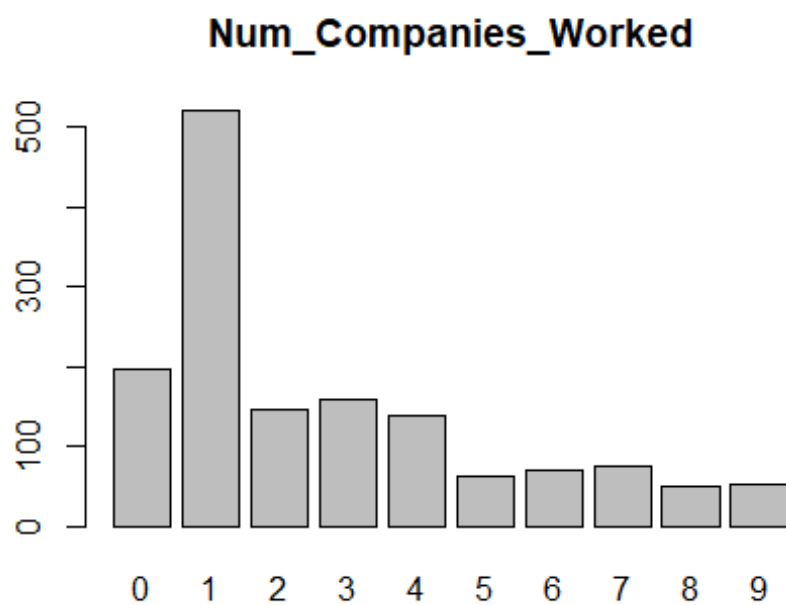



xviii. Num Companies Worked

```
summary(Turnover_issues_fac$NumCompaniesWorked)
```

```
##    0    1    2    3    4    5    6    7    8    9
## 197 521 146 159 139  63  70  74  49  52
```

```
plot(Turnover_issues_fac$NumCompaniesWorked, main="Num_Companies_Worked")
```

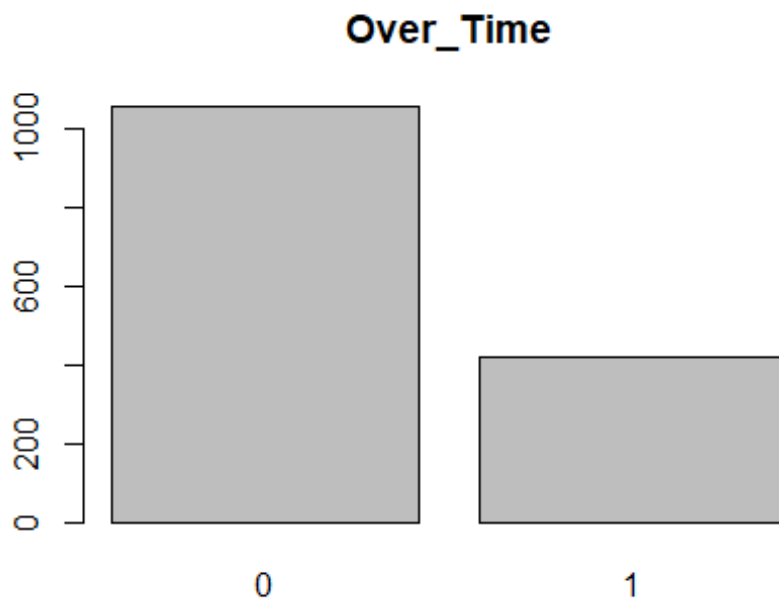


xix. Overtime

```
summary(Turnover_issues_fac$OverTime)
```

```
##      0      1  
## 1054  416
```

```
plot(Turnover_issues_fac$OverTime, main="Over_Time")
```

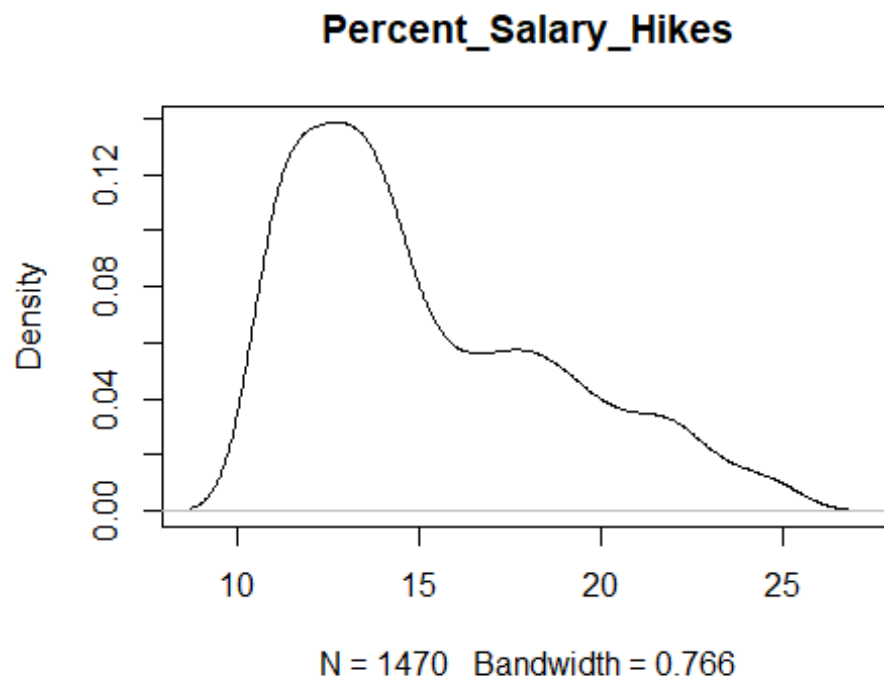


xx. Percent Salary Hike

```
summary(Turnover_issues_fac$PercentSalaryHike)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  11.00  12.00   14.00   15.21  18.00   25.00
```

```
plot(density(Turnover_issues_fac$PercentSalaryHike),main = "Percent_Salary_Hikes")
```

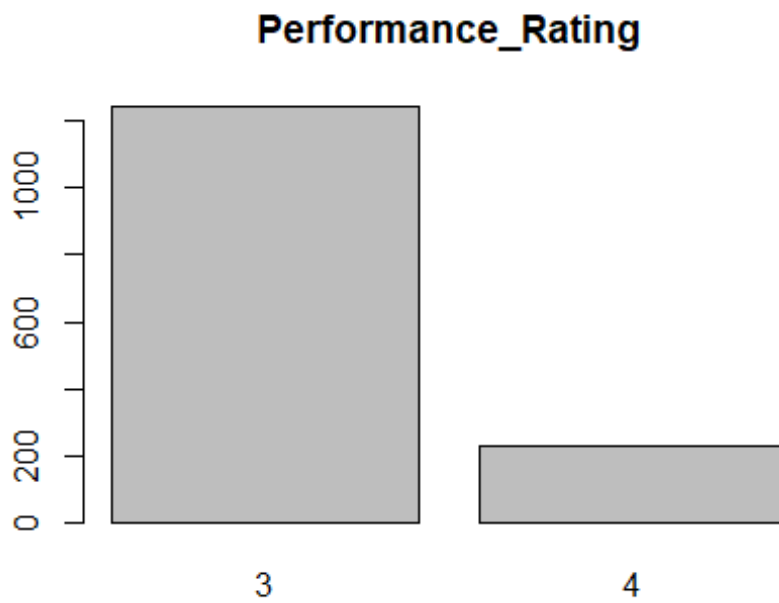


xxi. Performance Rating

```
summary(Turnover_issues_fac$PerformanceRating)
```

```
##      3      4  
## 1244  226
```

```
plot(Turnover_issues_fac$PerformanceRating, main="Performance_Rating")
```

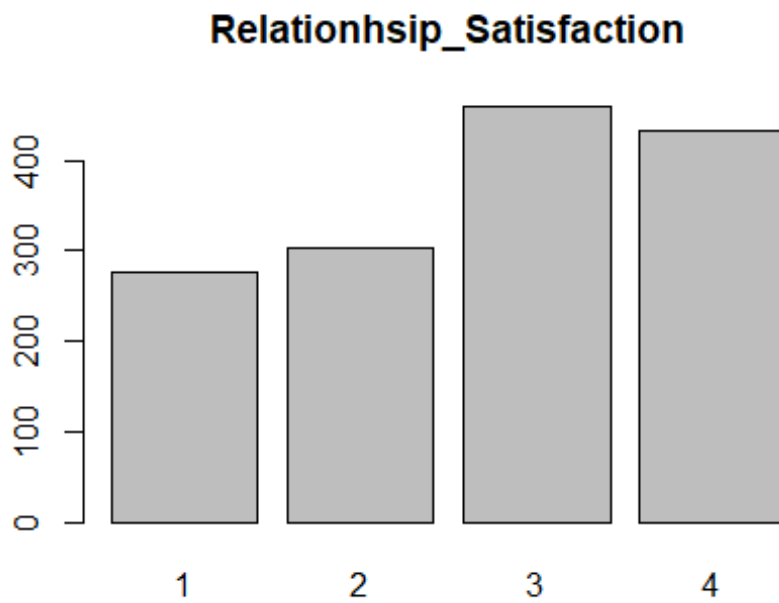


xxii.Relationship Satisfaction

```
summary(Turnover_issues_fac$RelationshipSatisfaction)
```

```
##    1    2    3    4  
## 276 303 459 432
```

```
plot(Turnover_issues_fac$RelationshipSatisfaction, main="Relationhsip_Satisfac  
tion")
```

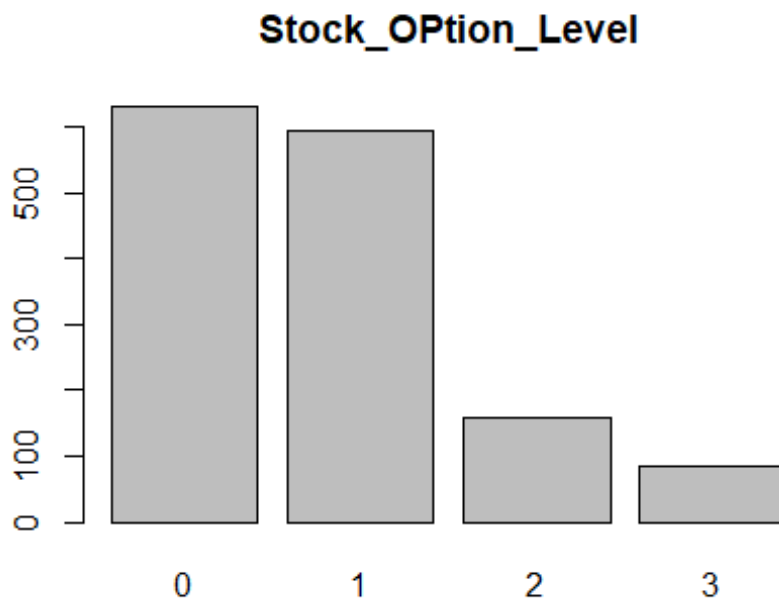


Standard hour input is 80 in all rows, so no analysis done

```
xxiii.      Stock Option Level
summary(Turnover_issues_fac$StockOptionLevel)

##    0    1    2    3
## 631 596 158   85

plot(Turnover_issues_fac$StockOptionLevel, main="Stock_OPTION_Level")
```



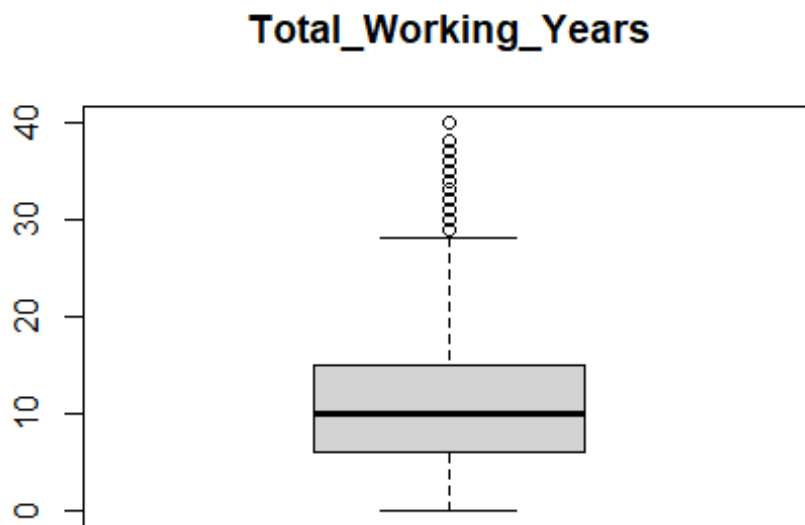
xiv. Total Working Years

```
summary(Turnover_issues_fac$TotalWorkingYears)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.00	6.00	10.00	11.28	15.00	40.00

Boxplot is used for better visualization

```
boxplot(Turnover_issues_fac$TotalWorkingYears, main="Total Working Years")
```



xxv. Training Times Last Year

```
summary(Turnover_issues_fac$TrainingTimesLastYear)
```

```
##    0    1    2    3    4    5    6  
##  54   71  547  491  123  119   65
```

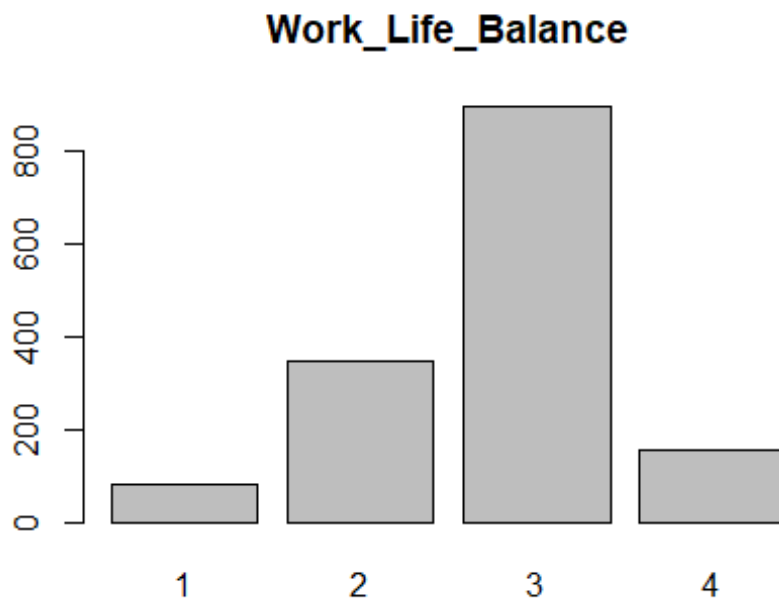
```
plot(Turnover_issues_fac$TrainingTimesLastYear, main = "Training_Times_Last_Year")
```




```
xxvi.Work Life Balance
summary(Turnover_issues_fac$WorkLifeBalance)

##    1    2    3    4
##  80 344 893 153

plot(Turnover_issues_fac$WorkLifeBalance, main = "Work_Life_Balance")
```

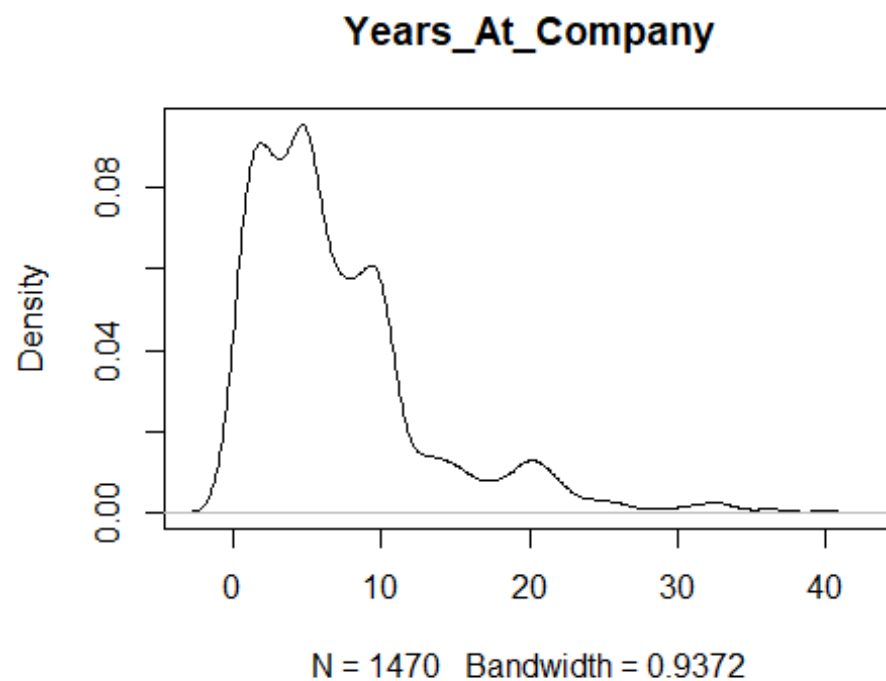


xxvii. Years at company

```
summary(Turnover_issues_fac$YearsAtCompany)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   3.000   5.000   7.008   9.000  40.000
```

```
plot(density(Turnover_issues_fac$YearsAtCompany), main = "Years_At_Company")
```

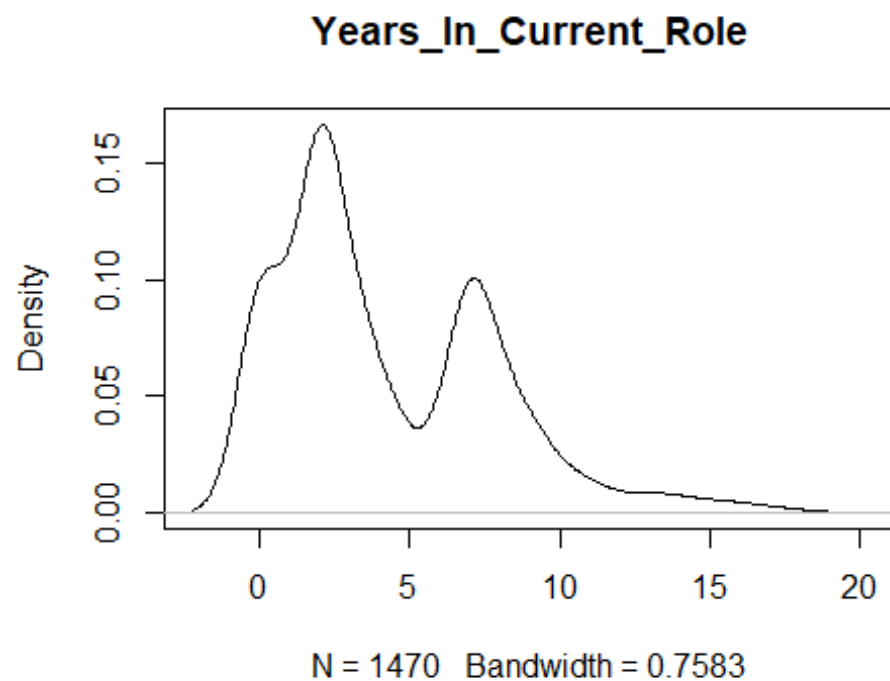


xxviii.Years in Current Role

```
summary(Turnover_issues_fac$YearsInCurrentRole)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   2.000   3.000   4.229   7.000  18.000
```

```
plot(density(Turnover_issues_fac$YearsInCurrentRole),main = "Years_In_Current_Role")
```

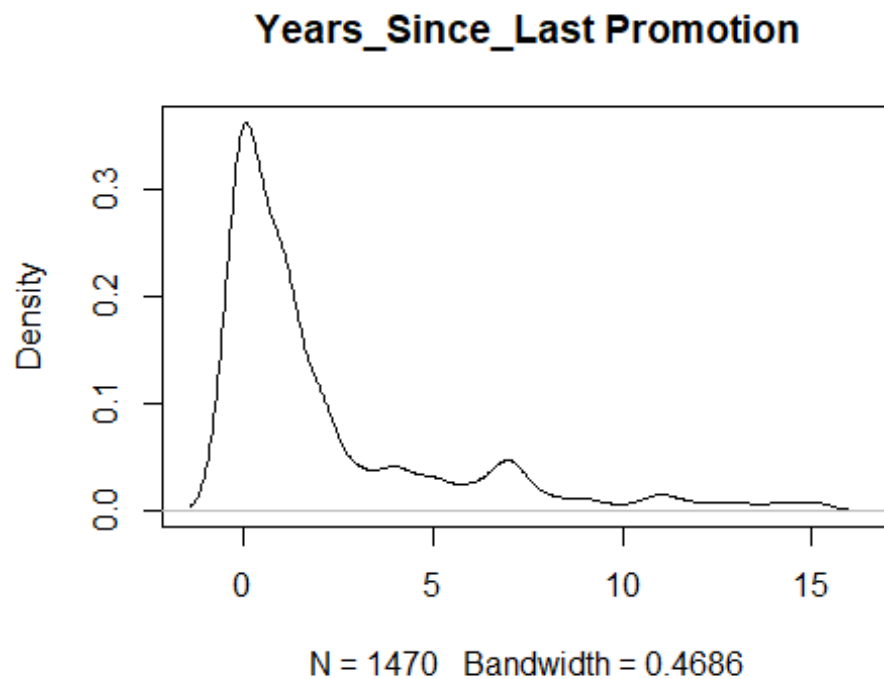


xix.Years Since Last Promotion

```
summary(Turnover_issues_fac$YearsSinceLastPromotion)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000  0.000   1.000   2.188  3.000  15.000
```

```
plot(density(Turnover_issues_fac$YearsSinceLastPromotion),main="Years_Since_L  
ast Promotion")
```

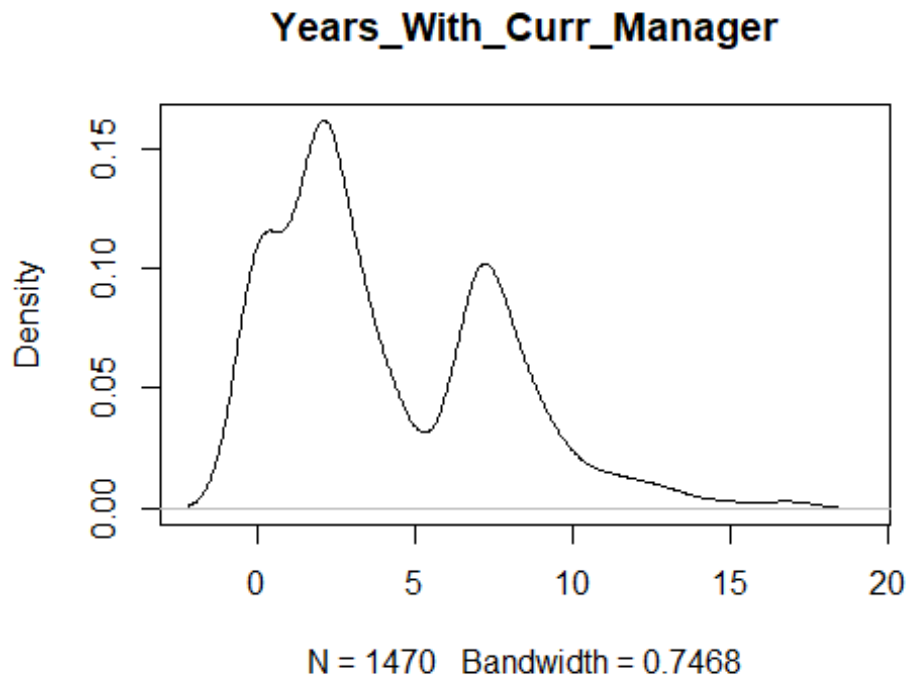


xxx. Years with Current Manager

```
summary(Turnover_issues_fac$YearsWithCurrManager)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   2.000   3.000   4.123   7.000  17.000
```

```
plot(density(Turnover_issues_fac$YearsWithCurrManager),main = "Years_With_Curr_Manager")
```



###B. Analyzing with two variables###

#Used cat for concatenating output, cut to convert continuous variable age into categorical variable and break for grouping which started from 18 to highest60#

```
Cat_Age<-cut(Turnover_issues_fac$Age, breaks = c(0,20,30,40,50,60), labels = c("20", "30", "40", "50", "60"))
```

#Used cat for concatenating output, cut to convert continuous variable Monthly Income into categorical variable and break for grouping which started from 1009 to highest 19999#

```
Cat_Monthly_Income<-cut(Turnover_issues_fac$MonthlyIncome, breaks = c(0,2000, 4000,6000,8000,10000,12000,14000,16000,18000,20000), labels = c("2000", "4000", "6000", "8000", "10000", "12000", "14000", "16000", "18000", "20000"))
```

#Used cat for concatenating output, cut to convert continuous variable Percent Salary Hike into categorical variable and break for grouping which started from 10 to highest 25#

```
Cat_Percent_Salary_Hike<-cut(Turnover_issues_fac$PercentSalaryHike, breaks = c(10,15,20,25), labels = c("10-15", "16-20", "21-25"))
```

The package "gmodels" need to be installed for utilizing CrossTable(...) function

```
# install.packages("gmodels")
library(gmodels)
```

1. Relationship of Age with Attrition

```
CrossTable(Cat_Age,Turnover_issues_fac$Attrition)
```

```
##
##
##   Cell Contents
## |-----|
## |                      N
## | Chi-square contribution
## |      N / Row Total
## |      N / Col Total
## |      N / Table Total
## |-----|
##
##
## Total Observations in Table:  1470
##
##
```

Cat_Age	Turnover_issues_fac\$Attrition		
	0	1	Row Total
20	12	16	28
	5.617	29.223	
	0.429	0.571	0.019
	0.010	0.068	
	0.008	0.011	
30	274	84	358
	2.300	11.967	
	0.765	0.235	0.244
	0.222	0.354	
	0.186	0.057	
40	534	85	619
	0.422	2.194	
	0.863	0.137	0.421
	0.433	0.359	
	0.363	0.058	
50	288	34	322
	1.188	6.182	
	0.894	0.106	0.219
	0.234	0.143	
	0.196	0.023	
60	125	18	143
	0.213	1.108	
	0.874	0.126	0.097
	0.101	0.076	
	0.085	0.012	

##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

2. Relationship of Business Travel with Attrition

CrossTable(Turnover_issues_fac\$BusinessTravel,Turnover_issues_fac\$Attrition)

##				
##				
##	Cell Contents			
##	-----			
##	N			
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table:	1470		
##				
##				
##		Turnover_issues_fac\$Attrition		
##	Turnover_issues_fac\$BusinessTravel	0	1	Row Total
##	-----	-----	-----	-----
##	Non-Travel	138	12	150
##		1.180	6.138	
##		0.920	0.080	0.102
##		0.112	0.051	
##		0.094	0.008	
##	-----	-----	-----	-----
##	Travel_Frequently	208	69	277
##		2.550	13.267	
##		0.751	0.249	0.188
##		0.169	0.291	
##		0.141	0.047	
##	-----	-----	-----	-----
##	Travel_Rarely	887	156	1043
##		0.169	0.879	
##		0.850	0.150	0.710
##		0.719	0.658	
##		0.603	0.106	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----


```
##
##
```

3. Relationship of Monthly Income with Attrition

```
CrossTable(Cat_Monthly_Income,Turnover_issues_fac$Attrition)
```

```
##
```

```
##
```

```
## Cell Contents
```

```
## |-----|
## |                      N |
## | Chi-square contribution |
## |      N / Row Total    |
## |      N / Col Total    |
## |      N / Table Total   |
## |-----|
```

```
##
```

```
##
```

```
## Total Observations in Table: 1470
```

```
##
```

```
##
```

```
##      Turnover_issues_fac$Attrition
## Cat_Monthly_Income      0      1 Row Total
## -----|-----|-----|
##           2000           15      18           33
##           5.808          30.218
##           0.455          0.545           0.022
##           0.012          0.076
##           0.010          0.012
## -----|-----|-----|
##           4000           391      119           510
##           3.162          16.448
##           0.767          0.233           0.347
##           0.317          0.502
##           0.266          0.081
## -----|-----|-----|
##           6000           329       42           371
##           1.020          5.306
##           0.887          0.113           0.252
##           0.267          0.177
##           0.224          0.029
## -----|-----|-----|
##           8000           157       18           175
##           0.711          3.698
##           0.897          0.103           0.119
##           0.127          0.076
##           0.107          0.012
## -----|-----|-----|
##           10000          85        15           100
##           0.015          0.078
```

##		0.850	0.150	0.068
##		0.069	0.063	
##		0.058	0.010	
##	-----	-----	-----	-----
##	12000	72	14	86
##		0.000	0.001	
##		0.837	0.163	0.059
##		0.058	0.059	
##		0.049	0.010	
##	-----	-----	-----	-----
##	14000	47	6	53
##		0.146	0.758	
##		0.887	0.113	0.036
##		0.038	0.025	
##		0.032	0.004	
##	-----	-----	-----	-----
##	16000	16	0	16
##		0.496	2.580	
##		1.000	0.000	0.011
##		0.013	0.000	
##		0.011	0.000	
##	-----	-----	-----	-----
##	18000	57	0	57
##		1.766	9.190	
##		1.000	0.000	0.039
##		0.046	0.000	
##		0.039	0.000	
##	-----	-----	-----	-----
##	20000	64	5	69
##		0.648	3.372	
##		0.928	0.072	0.047
##		0.052	0.021	
##		0.044	0.003	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

4. Relationship of department with Attrition

```
CrossTable(Turnover_issues_fac$Department,Turnover_issues_fac$Attrition)
```

##	
##	
##	Cell Contents
##	-----
##	N
##	Chi-square contribution
##	N / Row Total

```
## |          N / Col Total |
## |          N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$Department |          0 |          1 | Row Total |
## -----|-----|-----|
##              Human Resources |          51 |          12 |          63 |
##              |          0.064 |          0.334 |          |
##              |          0.810 |          0.190 |          0.043 |
##              |          0.041 |          0.051 |          |
##              |          0.035 |          0.008 |          |
## -----|-----|-----|
##              Research & Development |          828 |          133 |          961 |
##              |          0.597 |          3.106 |          |
##              |          0.862 |          0.138 |          0.654 |
##              |          0.672 |          0.561 |          |
##              |          0.563 |          0.090 |          |
## -----|-----|-----|
##              Sales |          354 |          92 |          446 |
##              |          1.079 |          5.615 |          |
##              |          0.794 |          0.206 |          0.303 |
##              |          0.287 |          0.388 |          |
##              |          0.241 |          0.063 |          |
## -----|-----|-----|
##              Column Total |          1233 |          237 |          1470 |
##              |          0.839 |          0.161 |          |
## -----|-----|-----|
##
##
```

5. Relationship of distance from home with Attrition

```
CrossTable(Turnover_issues_fac$DistanceFromHome,Turnover_issues_fac$Attrition
)
```

```
##
##
## Cell Contents
## |-----|
## |          N |
## | Chi-square contribution |
## |          N / Row Total |
## |          N / Col Total |
## |          N / Table Total |
## |-----|
##
```

```
##
## Total Observations in Table: 1470
##
```

		Turnover_issues_fac\$Attrition	
Turnover_issues_fac\$DistanceFromHome	0	1	Row Total

1	182	26	208
	0.325	1.693	
	0.875	0.125	0.141
	0.148	0.110	
	0.124	0.018	

2	183	28	211
	0.205	1.065	
	0.867	0.133	0.144
	0.148	0.118	
	0.124	0.019	

3	70	14	84
	0.003	0.015	
	0.833	0.167	0.057
	0.057	0.059	
	0.048	0.010	

4	55	9	64
	0.032	0.168	
	0.859	0.141	0.044
	0.045	0.038	
	0.037	0.006	

5	55	10	65
	0.004	0.022	
	0.846	0.154	0.044
	0.045	0.042	
	0.037	0.007	

6	52	7	59
	0.128	0.663	
	0.881	0.119	0.040
	0.042	0.030	
	0.035	0.005	

7	73	11	84
	0.092	0.477	
	0.869	0.131	0.057
	0.059	0.046	
	0.050	0.007	

8	70	10	80

##		0.125	0.651	
##		0.875	0.125	0.054
##		0.057	0.042	
##		0.048	0.007	
##				
##	9	67	18	85
##		0.259	1.347	
##		0.788	0.212	0.058
##		0.054	0.076	
##		0.046	0.012	
##				
##	10	75	11	86
##		0.114	0.592	
##		0.872	0.128	0.059
##		0.061	0.046	
##		0.051	0.007	
##				
##	11	25	4	29
##		0.019	0.098	
##		0.862	0.138	0.020
##		0.020	0.017	
##		0.017	0.003	
##				
##	12	14	6	20
##		0.459	2.389	
##		0.700	0.300	0.014
##		0.011	0.025	
##		0.010	0.004	
##				
##	13	13	6	19
##		0.541	2.815	
##		0.684	0.316	0.013
##		0.011	0.025	
##		0.009	0.004	
##				
##	14	17	4	21
##		0.021	0.111	
##		0.810	0.190	0.014
##		0.014	0.017	
##		0.012	0.003	
##				
##	15	21	5	26
##		0.030	0.156	
##		0.808	0.192	0.018
##		0.017	0.021	
##		0.014	0.003	
##				
##	16	25	7	32
##		0.126	0.657	
##		0.781	0.219	0.022

##		0.020	0.030	
##		0.017	0.005	
##				
##	17	15	5	20
##		0.188	0.978	
##		0.750	0.250	0.014
##		0.012	0.021	
##		0.010	0.003	
##				
##	18	22	4	26
##		0.002	0.009	
##		0.846	0.154	0.018
##		0.018	0.017	
##		0.015	0.003	
##				
##	19	19	3	22
##		0.016	0.084	
##		0.864	0.136	0.015
##		0.015	0.013	
##		0.013	0.002	
##				
##	20	21	4	25
##		0.000	0.000	
##		0.840	0.160	0.017
##		0.017	0.017	
##		0.014	0.003	
##				
##	21	15	3	18
##		0.001	0.003	
##		0.833	0.167	0.012
##		0.012	0.013	
##		0.010	0.002	
##				
##	22	13	6	19
##		0.541	2.815	
##		0.684	0.316	0.013
##		0.011	0.025	
##		0.009	0.004	
##				
##	23	22	5	27
##		0.018	0.096	
##		0.815	0.185	0.018
##		0.018	0.021	
##		0.015	0.003	
##				
##	24	16	12	28
##		2.386	12.413	
##		0.571	0.429	0.019
##		0.013	0.051	
##		0.011	0.008	

##	-----				
##		25	19	6	25
##			0.185	0.962	
##			0.760	0.240	0.017
##			0.015	0.025	
##			0.013	0.004	
##	-----				
##		26	22	3	25
##			0.051	0.264	
##			0.880	0.120	0.017
##			0.018	0.013	
##			0.015	0.002	
##	-----				
##		27	9	3	12
##			0.113	0.587	
##			0.750	0.250	0.008
##			0.007	0.013	
##			0.006	0.002	
##	-----				
##		28	21	2	23
##			0.151	0.787	
##			0.913	0.087	0.016
##			0.017	0.008	
##			0.014	0.001	
##	-----				
##		29	22	5	27
##			0.018	0.096	
##			0.815	0.185	0.018
##			0.018	0.021	
##			0.015	0.003	
##	-----				
##		Column Total	1233	237	1470
##			0.839	0.161	
##	-----				
##					
##					

6. Relationship of Education with attrition

CrossTable(Turnover_issues_fac\$Education,Turnover_issues_fac\$Attrition)

##	
##	
##	Cell Contents
##	-----
##	N
##	Chi-square contribution
##	N / Row Total
##	N / Col Total
##	N / Table Total
##	-----

```
##
##
## Total Observations in Table: 1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$Education
```

	0	1	Row Total
1	139	31	170
	0.090	0.471	
	0.818	0.182	0.116
	0.113	0.131	
	0.095	0.021	
2	238	44	282
	0.009	0.047	
	0.844	0.156	0.192
	0.193	0.186	
	0.162	0.030	
3	473	99	572
	0.096	0.498	
	0.827	0.173	0.389
	0.384	0.418	
	0.322	0.067	
4	340	58	398
	0.114	0.593	
	0.854	0.146	0.271
	0.276	0.245	
	0.231	0.039	
5	43	5	48
	0.186	0.969	
	0.896	0.104	0.033
	0.035	0.021	
	0.029	0.003	
Column Total	1233	237	1470
	0.839	0.161	

```
##
##
```

7. Relationship of Education Field with attrition

```
CrossTable(Turnover_issues_fac$EducationField,Turnover_issues_fac$Attrition)
```

```
##
##
## Cell Contents
```



```
## |-----|
## |               N |
## | Chi-square contribution |
## |       N / Row Total |
## |       N / Col Total |
## |       N / Table Total |
## |-----|
```

```
##
##
```

```
## Total Observations in Table:  1470
```

```
##
##
```

		Turnover_issues_fac\$Attrition	
Turnover_issues_fac\$EducationField	0	1	Row Total

Human Resources	20	7	27
	0.309	1.610	
	0.741	0.259	0.018
	0.016	0.030	
	0.014	0.005	

Life Sciences	517	89	606
	0.149	0.775	
	0.853	0.147	0.412
	0.419	0.376	
	0.352	0.061	

Marketing	124	35	159
	0.658	3.421	
	0.780	0.220	0.108
	0.101	0.148	
	0.084	0.024	

Medical	401	63	464
	0.358	1.864	
	0.864	0.136	0.316
	0.325	0.266	
	0.273	0.043	

Other	71	11	82
	0.072	0.373	
	0.866	0.134	0.056
	0.058	0.046	
	0.048	0.007	

Technical Degree	100	32	132
	1.038	5.398	
	0.758	0.242	0.090
	0.081	0.135	
	0.068	0.022	

```
## -----|-----|-----|
##               Column Total |      1233 |      237 |      1470 |
##               0.839 |      0.161 |
## -----|-----|-----|
##
##
```

8. Relationship of Environment Satisfaction with Attrition

```
CrossTable(Turnover_issues_fac$EnvironmentSatisfaction,Turnover_issues_fac$Attrition)
```

```
##
##
##   Cell Contents
## |-----|
## |               N
## | Chi-square contribution
## |       N / Row Total
## |       N / Col Total
## |       N / Table Total
## |-----|
##
##
## Total Observations in Table:  1470
##
##
##                                     | Turnover_issues_fac$Attritio
n
## Turnover_issues_fac$EnvironmentSatisfaction |      0 |      1 | Row
Total |
## -----|-----|-----|-----
-----|
##               1 |      212 |      72 |
284 |
##               |      2.884 |      15.006 |
|
##               |      0.746 |      0.254 |
0.193 |
##               |      0.172 |      0.304 |
|
##               |      0.144 |      0.049 |
|
## -----|-----|-----|-----
-----|
##               2 |      244 |      43 |
287 |
##               |      0.044 |      0.231 |
|
##               |      0.850 |      0.150 |
0.195 |
```

```

## | 0.198 | 0.181 |
## | 0.166 | 0.029 |
## -----|-----|-----|-----
-----|
## | 3 | 391 | 62 |
453 |
## | 0.320 | 1.667 |
## | 0.863 | 0.137 |
0.308 |
## | 0.317 | 0.262 |
## | 0.266 | 0.042 |
## -----|-----|-----|-----
-----|
## | 4 | 386 | 60 |
446 |
## | 0.379 | 1.971 |
## | 0.865 | 0.135 |
0.303 |
## | 0.313 | 0.253 |
## | 0.263 | 0.041 |
## -----|-----|-----|-----
-----|
## | Column Total | 1233 | 237 |
1470 |
## | 0.839 | 0.161 |
## -----|-----|-----|-----
-----|
##
##

```

9. Relationship of Gender with Attrition

```
CrossTable(Turnover_issues_fac$Gender,Turnover_issues_fac$Attrition)
```

```

##
##
## Cell Contents
## |-----|
## | N |
## | Chi-square contribution |
## | N / Row Total |
## | N / Col Total |

```

```
## |           N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$Gender |           0 |           1 | Row Total |
## -----|-----|-----|
##                Female |         501 |          87 |        588 |
##                |         0.123 |         0.642 |        |
##                |         0.852 |         0.148 |        0.400 |
##                |         0.406 |         0.367 |        |
##                |         0.341 |         0.059 |        |
## -----|-----|-----|
##                Male |         732 |         150 |        882 |
##                |         0.082 |         0.428 |        |
##                |         0.830 |         0.170 |        0.600 |
##                |         0.594 |         0.633 |        |
##                |         0.498 |         0.102 |        |
## -----|-----|-----|
##                Column Total |        1233 |         237 |        1470 |
##                |         0.839 |         0.161 |        |
## -----|-----|-----|
##
##
```

10. Relationship of Job Involvement with Attrition

```
CrossTable(Turnover_issues_fac$JobInvolvement,Turnover_issues_fac$Attrition)
```

```
##
##
## Cell Contents
## |-----|
## |           N |
## | Chi-square contribution |
## | N / Row Total |
## | N / Col Total |
## | N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$JobInvolvement |           0 |           1 | Row Total |
## -----|-----|-----|
##                1 |          55 |          28 |         83 |
```

##		3.070	15.969	
##		0.663	0.337	0.056
##		0.045	0.118	
##		0.037	0.019	
##	-----	-----	-----	-----
##	2	304	71	375
##		0.353	1.838	
##		0.811	0.189	0.255
##		0.247	0.300	
##		0.207	0.048	
##	-----	-----	-----	-----
##	3	743	125	868
##		0.307	1.596	
##		0.856	0.144	0.590
##		0.603	0.527	
##		0.505	0.085	
##	-----	-----	-----	-----
##	4	131	13	144
##		0.864	4.496	
##		0.910	0.090	0.098
##		0.106	0.055	
##		0.089	0.009	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

11. Relationship of Job level with Attrition

```
CrossTable(Turnover_issues_fac$JobLevel,Turnover_issues_fac$Attrition)
```

```
##
##
##   Cell Contents
## |-----|
## |              N |
## | Chi-square contribution |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$JobLevel | Turnover_issues_fac$Attrition
##                               0 | 1 | Row Total |
## -----|-----|-----|
```

##	1	400	143	543
##		6.752	35.128	
##		0.737	0.263	0.369
##		0.324	0.603	
##		0.272	0.097	
##	-----	-----	-----	-----
##	2	482	52	534
##		2.595	13.501	
##		0.903	0.097	0.363
##		0.391	0.219	
##		0.328	0.035	
##	-----	-----	-----	-----
##	3	186	32	218
##		0.054	0.282	
##		0.853	0.147	0.148
##		0.151	0.135	
##		0.127	0.022	
##	-----	-----	-----	-----
##	4	101	5	106
##		1.644	8.553	
##		0.953	0.047	0.072
##		0.082	0.021	
##		0.069	0.003	
##	-----	-----	-----	-----
##	5	64	5	69
##		0.648	3.372	
##		0.928	0.072	0.047
##		0.052	0.021	
##		0.044	0.003	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

12. Relationship of Percentage Job Role with Attrition

CrossTable(Turnover_issues_fac\$JobRole,Turnover_issues_fac\$Attrition)

##	
##	
##	Cell Contents
##	-----
##	N
##	Chi-square contribution
##	N / Row Total
##	N / Col Total
##	N / Table Total
##	-----
##	

##

Total Observations in Table: 1470

##

##

Turnover_issues_fac\$Attrition		Turnover_issues_fac\$JobRole	
0	1	0	1
Row Total			

Healthcare Representative	122	9	131
	1.337	6.956	
	0.931	0.069	0.089
	0.099	0.038	
	0.083	0.006	

Human Resources	40	12	52
	0.300	1.560	
	0.769	0.231	0.035
	0.032	0.051	
	0.027	0.008	

Laboratory Technician	197	62	259
	1.886	9.813	
	0.761	0.239	0.176
	0.160	0.262	
	0.134	0.042	

Manager	97	5	102
	1.531	7.965	
	0.951	0.049	0.069
	0.079	0.021	
	0.066	0.003	

Manufacturing Director	135	10	145
	1.471	7.655	
	0.931	0.069	0.099
	0.109	0.042	
	0.092	0.007	

Research Director	78	2	80
	1.770	9.208	
	0.975	0.025	0.054
	0.063	0.008	
	0.053	0.001	

Research Scientist	245	47	292
	0.000	0.000	
	0.839	0.161	0.199
	0.199	0.198	
	0.167	0.032	

Sales Executive	269	57	326

##		0.072	0.375	
##		0.825	0.175	0.222
##		0.218	0.241	
##		0.183	0.039	
##	-----	-----	-----	-----
##	Sales Representative	50	33	83
##		5.528	28.762	
##		0.602	0.398	0.056
##		0.041	0.139	
##		0.034	0.022	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

13. Relationship of Job Satisfaction with Attrition

CrossTable(Turnover_issues_fac\$JobSatisfaction,Turnover_issues_fac\$Attrition)

##				
##				
##	Cell Contents			
##	-----			
##		N		
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table:	1470		
##				
##				
##				
##	Turnover_issues_fac\$JobSatisfaction	Turnover_issues_fac\$Attrition		
##		0	1	Row Total
##	-----	-----	-----	-----
##	1	223	66	289
##		1.554	8.083	
##		0.772	0.228	0.197
##		0.181	0.278	
##		0.152	0.045	
##	-----	-----	-----	-----
##	2	234	46	280
##		0.003	0.016	
##		0.836	0.164	0.190
##		0.190	0.194	
##		0.159	0.031	
##	-----	-----	-----	-----

##	3	369	73	442
##		0.008	0.042	
##		0.835	0.165	0.301
##		0.299	0.308	
##		0.251	0.050	
##	-----			
##	4	407	52	459
##		1.257	6.542	
##		0.887	0.113	0.312
##		0.330	0.219	
##		0.277	0.035	
##	-----			
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----			
##				
##				

14. Relationship of Marital status with Attrition

```
CrossTable(Turnover_issues_fac$MaritalStatus,Turnover_issues_fac$Attrition)
```

##				
##				
##	Cell Contents			
##	-----			
##	N			
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table: 1470			
##				
##				
##	Turnover_issues_fac\$Attrition			
##	Turnover_issues_fac\$MaritalStatus	0	1	Row Total
##	-----			
##	Divorced	294	33	327
##		1.418	7.377	
##		0.899	0.101	0.222
##		0.238	0.139	
##		0.200	0.022	
##	-----			
##	Married	589	84	673
##		1.064	5.534	
##		0.875	0.125	0.458
##		0.478	0.354	
##		0.401	0.057	

##	-----	-----	-----	-----
##	Single	350	120	470
##		4.961	25.811	
##		0.745	0.255	0.320
##		0.284	0.506	
##		0.238	0.082	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

15. Relationship of number of companies worked with Attrition

CrossTable(Turnover_issues_fac\$NumCompaniesWorked, Turnover_issues_fac\$Attrition)

##				
##				
##	Cell Contents			
##	-----			
##	N			
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table:	1470		
##				
##				
##				
##	Turnover_issues_fac\$NumCompaniesWorked	Turnover_issues_fac\$Attrition		
##		0	1	Row Total
##	-----	-----	-----	-----
##	0	174	23	197
##		0.465	2.417	
##		0.883	0.117	0.134
##		0.141	0.097	
##		0.118	0.016	
##	-----	-----	-----	-----
##	1	423	98	521

##		0.449	2.334	
##		0.812	0.188	0.354
##		0.343	0.414	
##		0.288	0.067	
##	-----	-----	-----	-----
-				
##	2	130	16	146
##		0.464	2.414	
##		0.890	0.110	0.099
##		0.105	0.068	
##		0.088	0.011	
##	-----	-----	-----	-----
-				
##	3	143	16	159
##		0.696	3.621	
##		0.899	0.101	0.108
##		0.116	0.068	
##		0.097	0.011	
##	-----	-----	-----	-----
-				
##	4	122	17	139
##		0.251	1.306	
##		0.878	0.122	0.095
##		0.099	0.072	
##		0.083	0.012	
##	-----	-----	-----	-----
-				
##	5	47	16	63
##		0.646	3.361	

##		0.746	0.254	0.043
##		0.038	0.068	
##		0.032	0.011	
##	-----	-----	-----	-----
-				
##	6	54	16	70
##		0.379	1.969	
##		0.771	0.229	0.048
##		0.044	0.068	
##		0.037	0.011	
##	-----	-----	-----	-----
-				
##	7	57	17	74
##		0.414	2.154	
##		0.770	0.230	0.050
##		0.046	0.072	
##		0.039	0.012	
##	-----	-----	-----	-----
-				
##	8	43	6	49
##		0.088	0.457	
##		0.878	0.122	0.033
##		0.035	0.025	
##		0.029	0.004	
##	-----	-----	-----	-----
-				
##	9	40	12	52
##		0.300	1.560	
##		0.769	0.231	0.035

```

|
##          |      0.032 |      0.051 |
|
##          |      0.027 |      0.008 |
|
## -----|-----|-----|-----
-|
##          Column Total |      1233 |      237 |      1470
|
##          |      0.839 |      0.161 |
|
## -----|-----|-----|-----
-|
##
##

```

16. Relationship of overtime with Attrition

```
CrossTable(Turnover_issues_fac$OverTime,Turnover_issues_fac$Attrition)
```

```

##
##
##   Cell Contents
## |-----|
## |              N
## | Chi-square contribution
## |      N / Row Total
## |      N / Col Total
## |      N / Table Total
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$OverTime | Turnover_issues_fac$Attrition
##                               0      1 | Row Total |
## -----|-----|-----|-----
##                               0      1 |
##                               944    110 |      1054 |
##                               4.063    21.136 |
##                               0.896    0.104 |      0.717 |
##                               0.766    0.464 |
##                               0.642    0.075 |
## -----|-----|-----|-----
##                               1      1 |
##                               289    127 |      416 |
##                               10.293    53.552 |
##                               0.695    0.305 |      0.283 |
##                               0.234    0.536 |
##                               0.197    0.086 |
## -----|-----|-----|-----
##                               Column Total |      1233 |      237 |      1470 |

```

```
## | 0.839 | 0.161 |
## |-----|-----|-----|
##
##
```

17. Relationship of Percent Salary Hike with Attrition

```
CrossTable(Cat_Percent_Salary_Hike,Turnover_issues_fac$Attrition)
```

```
##
##
##   Cell Contents
## |-----|
## | N
## | Chi-square contribution
## | N / Row Total
## | N / Col Total
## | N / Table Total
## |-----|
##
##
## Total Observations in Table: 1470
##
##
##   Turnover_issues_fac$Attrition
## Cat_Percent_Salary_Hike      0      1  Row Total
## -----|-----|-----|
##           10-15      769      150      919
##           0.004      0.023
##           0.837      0.163      0.625
##           0.624      0.633
##           0.523      0.102
## -----|-----|-----|
##           16-20      323       57      380
##           0.057      0.297
##           0.850      0.150      0.259
##           0.262      0.241
##           0.220      0.039
## -----|-----|-----|
##           21-25      141       30      171
##           0.041      0.214
##           0.825      0.175      0.116
##           0.114      0.127
##           0.096      0.020
## -----|-----|-----|
##           Column Total      1233      237      1470
##           0.839      0.161
## -----|-----|-----|
##
##
```

18. Relationship of Performance rating with Attrition

```
CrossTable(Turnover_issues_fac$PerformanceRating,Turnover_issues_fac$Attrition)
```

```
##
##
##      Cell Contents
## |-----|
## |                      N
## | Chi-square contribution
## |      N / Row Total
## |      N / Col Total
## |      N / Table Total
## |-----|
##
##
## Total Observations in Table:  1470
##
##
##                               | Turnover_issues_fac$Attrition
## Turnover_issues_fac$PerformanceRating |      0 |      1 | Row Total
## |-----|-----|-----|-----|
## |
## |                      3 |    1044 |    200 |    1244
## |
## |                      |    0.000 |    0.002 |
## |
## |                      |    0.839 |    0.161 |    0.846
## |
## |                      |    0.847 |    0.844 |
## |
## |                      |    0.710 |    0.136 |
## |-----|-----|-----|-----|
## |
## |                      4 |     189 |     37 |     226
## |
## |                      |    0.002 |    0.009 |
## |
## |                      |    0.836 |    0.164 |    0.154
## |
## |                      |    0.153 |    0.156 |
## |
## |                      |    0.129 |    0.025 |
## |-----|-----|-----|-----|
## |
## |                      Column Total |    1233 |     237 |    1470
## |
```

```
## | 0.839 | 0.161 |
## -----|-----|-----|
##
##
```

19. Affect of Relationship Satisfactoin with Attrition

```
CrossTable(Turnover_issues_fac$RelationshipSatisfaction,Turnover_issues_fac$Attrition)
```

```
##
##
## Cell Contents
## |-----|
## | N
## | Chi-square contribution
## | N / Row Total
## | N / Col Total
## | N / Table Total
## |-----|
##
##
## Total Observations in Table: 1470
##
##
## | Turnover_issues_fac$Attriti
on
## Turnover_issues_fac$RelationshipSatisfaction | 0 | 1 | Row
Total |
## -----|-----|-----|
## |
## | 1 | 219 | 57 |
276 |
## | 0.675 | 3.513 |
|
## | 0.793 | 0.207 |
0.188 |
## | 0.178 | 0.241 |
|
## | 0.149 | 0.039 |
|
## -----|-----|-----|
## |
## | 2 | 258 | 45 |
303 |
## | 0.058 | 0.304 |
|
## | 0.851 | 0.149 |
0.206 |
```



```

##          |      0.209 |      0.190 |
|
##          |      0.176 |      0.031 |
|
## -----|-----|-----|-----
-----|
##          3 |      388 |      71 |
459 |
##          |      0.023 |      0.122 |
|
##          |      0.845 |      0.155 |
0.312 |
##          |      0.315 |      0.300 |
|
##          |      0.264 |      0.048 |
|
## -----|-----|-----|-----
-----|
##          4 |      368 |      64 |
432 |
##          |      0.088 |      0.458 |
|
##          |      0.852 |      0.148 |
0.294 |
##          |      0.298 |      0.270 |
|
##          |      0.250 |      0.044 |
|
## -----|-----|-----|-----
-----|
##          Column Total |      1233 |      237 |
1470 |
##          |      0.839 |      0.161 |
|
## -----|-----|-----|-----
-----|
##
##

```

20. Relationship of Stock Option Level with Attrition

```

CrossTable(Turnover_issues_fac$StockOptionLevel, Turnover_issues_fac$Attrition
)

```

```

##
##
##      Cell Contents
## |-----|
## |                      N |
## | Chi-square contribution |
## |          N / Row Total |

```

```
## |          N / Col Total |
## |          N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$StockOptionLevel |          0 |          1 | Row Total |
## -----|-----|-----|
##                               0 |          477 |          154 |          631 |
##                               5.162 |         26.853 |
##                               0.756 |          0.244 |          0.429 |
##                               0.387 |          0.650 |
##                               0.324 |          0.105 |
## -----|-----|-----|
##                               1 |          540 |           56 |          596 |
##                               3.215 |         16.726 |
##                               0.906 |          0.094 |          0.405 |
##                               0.438 |          0.236 |
##                               0.367 |          0.038 |
## -----|-----|-----|
##                               2 |          146 |           12 |          158 |
##                               1.370 |          7.126 |
##                               0.924 |          0.076 |          0.107 |
##                               0.118 |          0.051 |
##                               0.099 |          0.008 |
## -----|-----|-----|
##                               3 |           70 |           15 |           85 |
##                               0.024 |          0.123 |
##                               0.824 |          0.176 |          0.058 |
##                               0.057 |          0.063 |
##                               0.048 |          0.010 |
## -----|-----|-----|
##                               Column Total |          1233 |           237 |          1470 |
##                               0.839 |          0.161 |
## -----|-----|-----|
##
##
```

21. Relationship of Total Working Years with Attrition

```
CrossTable(Turnover_issues_fac$TotalWorkingYears,Turnover_issues_fac$Attritio
n)
```

```
##
##
## Cell Contents
## |-----|
## |          N |
```

```

## | Chi-square contribution |
## |           N / Row Total |
## |           N / Col Total |
## |           N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$Attrition
## Turnover_issues_fac$TotalWorkingYears |           0 |           1 | Row Total
|-----|-----|-----|-----|
## |
## |           0 |           6 |           5 |           11
## |
## |           |           1.128 |           5.870 |
## |
## |           |           0.545 |           0.455 |           0.007
## |
## |           |           0.005 |           0.021 |
## |
## |           |           0.004 |           0.003 |
## |-----|-----|-----|-----|
## |
## |           1 |           41 |           40 |           81
## |
## |           |           10.683 |           55.578 |
## |
## |           |           0.506 |           0.494 |           0.055
## |
## |           |           0.033 |           0.169 |
## |
## |           |           0.028 |           0.027 |
## |-----|-----|-----|-----|
## |
## |           2 |           22 |           9 |           31
## |
## |           |           0.616 |           3.205 |
## |
## |           |           0.710 |           0.290 |           0.021
## |
## |           |           0.018 |           0.038 |
## |
## |           |           0.015 |           0.006 |
## |-----|-----|-----|-----|

```

##	3	33	9	42
##		0.141	0.733	
##		0.786	0.214	0.029
##		0.027	0.038	
##		0.022	0.006	
##				
##	4	51	12	63
##		0.064	0.334	
##		0.810	0.190	0.043
##		0.041	0.051	
##		0.035	0.008	
##				
##	5	72	16	88
##		0.044	0.231	
##		0.818	0.182	0.060
##		0.058	0.068	
##		0.049	0.011	
##				
##	6	103	22	125
##		0.033	0.169	
##		0.824	0.176	0.085
##		0.084	0.093	
##		0.070	0.015	
##				
##	7	63	18	81

##		0.359	1.869	
##		0.778	0.222	0.055
##		0.051	0.076	
##		0.043	0.012	
##	-----	-----	-----	-----
##	8	87	16	103
##		0.004	0.022	
##		0.845	0.155	0.070
##		0.071	0.068	
##		0.059	0.011	
##	-----	-----	-----	-----
##	9	86	10	96
##		0.373	1.939	
##		0.896	0.104	0.065
##		0.070	0.042	
##		0.059	0.007	
##	-----	-----	-----	-----
##	10	177	25	202
##		0.338	1.758	
##		0.876	0.124	0.137
##		0.144	0.105	
##		0.120	0.017	
##	-----	-----	-----	-----
##	11	29	7	36
##		0.047	0.246	

##		0.806	0.194	0.024
##		0.024	0.030	
##		0.020	0.005	
##				
##	12	43	5	48
##		0.186	0.969	
##		0.896	0.104	0.033
##		0.035	0.021	
##		0.029	0.003	
##				
##	13	33	3	36
##		0.260	1.355	
##		0.917	0.083	0.024
##		0.027	0.013	
##		0.022	0.002	
##				
##	14	27	4	31
##		0.038	0.199	
##		0.871	0.129	0.021
##		0.022	0.017	
##		0.018	0.003	
##				
##	15	35	5	40
##		0.063	0.326	
##		0.875	0.125	0.027

##		0.028	0.021	
##		0.024	0.003	
##				
##	16	34	3	37
##		0.283	1.474	
##		0.919	0.081	0.025
##		0.028	0.013	
##		0.023	0.002	
##				
##	17	30	3	33
##		0.195	1.012	
##		0.909	0.091	0.022
##		0.024	0.013	
##		0.020	0.002	
##				
##	18	23	4	27
##		0.006	0.029	
##		0.852	0.148	0.018
##		0.019	0.017	
##		0.016	0.003	
##				
##	19	19	3	22
##		0.016	0.084	
##		0.864	0.136	0.015
##		0.015	0.013	

##		0.013	0.002	
##				
##	20	28	2	30
##		0.320	1.664	
##		0.933	0.067	0.020
##		0.023	0.008	
##		0.019	0.001	
##				
##	21	33	1	34
##		0.704	3.664	
##		0.971	0.029	0.023
##		0.027	0.004	
##		0.022	0.001	
##				
##	22	19	2	21
##		0.109	0.567	
##		0.905	0.095	0.014
##		0.015	0.008	
##		0.013	0.001	
##				
##	23	20	2	22
##		0.130	0.675	
##		0.909	0.091	0.015
##		0.016	0.008	
##		0.014	0.001	

##				
##	24	15	3	18
##		0.001	0.003	
##		0.833	0.167	0.012
##		0.012	0.013	
##		0.010	0.002	
##				
##	25	13	1	14
##		0.135	0.700	
##		0.929	0.071	0.010
##		0.011	0.004	
##		0.009	0.001	
##				
##	26	13	1	14
##		0.135	0.700	
##		0.929	0.071	0.010
##		0.011	0.004	
##		0.009	0.001	
##				
##	27	7	0	7
##		0.217	1.129	
##		1.000	0.000	0.005
##		0.006	0.000	
##		0.005	0.000	
##				

##	28	13	1	14
##		0.135	0.700	
##		0.929	0.071	0.010
##		0.011	0.004	
##		0.009	0.001	
##				
##	29	10	0	10
##		0.310	1.612	
##		1.000	0.000	0.007
##		0.008	0.000	
##		0.007	0.000	
##				
##	30	7	0	7
##		0.217	1.129	
##		1.000	0.000	0.005
##		0.006	0.000	
##		0.005	0.000	
##				
##	31	8	1	9
##		0.027	0.140	
##		0.889	0.111	0.006
##		0.006	0.004	
##		0.005	0.001	
##				
##	32	9	0	9

##		0.279	1.451	
##		1.000	0.000	0.006
##		0.007	0.000	
##		0.006	0.000	
##	-----	-----	-----	-----
##	33	6	1	7
##		0.003	0.015	
##		0.857	0.143	0.005
##		0.005	0.004	
##		0.004	0.001	
##	-----	-----	-----	-----
##	34	4	1	5
##		0.009	0.047	
##		0.800	0.200	0.003
##		0.003	0.004	
##		0.003	0.001	
##	-----	-----	-----	-----
##	35	3	0	3
##		0.093	0.484	
##		1.000	0.000	0.002
##		0.002	0.000	
##		0.002	0.000	
##	-----	-----	-----	-----
##	36	6	0	6
##		0.186	0.967	

##		1.000	0.000	0.004
##		0.005	0.000	
##		0.004	0.000	
##				
##	37	4	0	4
##		0.124	0.645	
##		1.000	0.000	0.003
##		0.003	0.000	
##		0.003	0.000	
##				
##	38	1	0	1
##		0.031	0.161	
##		1.000	0.000	0.001
##		0.001	0.000	
##		0.001	0.000	
##				
##	40	0	2	2
##		1.678	8.728	
##		0.000	1.000	0.001
##		0.000	0.008	
##		0.000	0.001	
##				
##	Column Total	1233	237	1470
##		0.839	0.161	
##				

```
|  
##  
##
```

22. Relationship of Training Times Last Year with Attrition

```
CrossTable(Turnover_issues_fac$TrainingTimesLastYear,Turnover_issues_fac$Attr  
ition)
```

```
##  
##  
##      Cell Contents  
## |-----|  
## |                      N |  
## | Chi-square contribution |  
## |      N / Row Total     |  
## |      N / Col Total     |  
## |      N / Table Total   |  
## |-----|  
##  
##  
## Total Observations in Table:  1470  
##  
##  
##                               | Turnover_issues_fac$Attrition  
## Turnover_issues_fac$TrainingTimesLastYear |      0 |      1 | Row To  
tal |  
## -----|-----|-----|-----  
----|  
##                               0 |      39 |      15 |  
54 |  
##                               | 0.875 | 4.550 |  
|  
##                               | 0.722 | 0.278 | 0.  
037 |  
##                               | 0.032 | 0.063 |  
|  
##                               | 0.027 | 0.010 |  
|  
## -----|-----|-----|-----  
----|  
##                               1 |      62 |      9 |  
71 |  
##                               | 0.101 | 0.523 |  
|  
##                               | 0.873 | 0.127 | 0.  
048 |  
##                               | 0.050 | 0.038 |  
|  
##                               | 0.042 | 0.006 |  
|
```

##	-----		-----		-----		-----

##		2		449		98	
547							
##				0.210		1.091	
##				0.821		0.179	0.
372							
##				0.364		0.414	
##				0.305		0.067	
##	-----		-----		-----		-----

##		3		422		69	
491							
##				0.251		1.304	
##				0.859		0.141	0.
334							
##				0.342		0.291	
##				0.287		0.047	
##	-----		-----		-----		-----

##		4		97		26	
123							
##				0.369		1.919	
##				0.789		0.211	0.
084							
##				0.079		0.110	
##				0.066		0.018	
##	-----		-----		-----		-----

##		5		105		14	
119							
##				0.269		1.402	
##				0.882		0.118	0.
081							
##				0.085		0.059	
##				0.071		0.010	
##	-----		-----		-----		-----

```
##          6 |          59 |          6 |
65 |
##          |          0.368 |          1.915 |
|
##          |          0.908 |          0.092 |          0.
044 |
##          |          0.048 |          0.025 |
|
##          |          0.040 |          0.004 |
|
## -----|-----|-----|-----
----|
##          Column Total |          1233 |          237 |          1
470 |
##          |          0.839 |          0.161 |
|
## -----|-----|-----|-----
----|
##
##
```

23. Relationship of Work Life Balance with Attrition

```
CrossTable(Turnover_issues_fac$WorkLifeBalance,Turnover_issues_fac$Attrition)
```

```
##
##
##   Cell Contents
## |-----|
## |          N |
## | Chi-square contribution |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##
##
## Turnover_issues_fac$WorkLifeBalance | Turnover_issues_fac$Attrition
##          0 |          1 | Row Total |
## -----|-----|-----|
##          1 |          55 |          25 |          80 |
##          2.183 |          11.355 |
##          0.688 |          0.312 |          0.054 |
##          0.045 |          0.105 |
##          0.037 |          0.017 |
## -----|-----|-----|
##          2 |          286 |          58 |          344 |
##          0.022 |          0.116 |
```

##		0.831	0.169	0.234
##		0.232	0.245	
##		0.195	0.039	
##	-----	-----	-----	-----
##	3	766	127	893
##		0.385	2.001	
##		0.858	0.142	0.607
##		0.621	0.536	
##		0.521	0.086	
##	-----	-----	-----	-----
##	4	126	27	153
##		0.042	0.221	
##		0.824	0.176	0.104
##		0.102	0.114	
##		0.086	0.018	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

24. Relationship of number of Years At Company with attrition

CrossTable(Turnover_issues_fac\$YearsAtCompany,Turnover_issues_fac\$Attrition)

##				
##				
##	Cell Contents			
##	-----			
##	N			
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table:	1470		
##				
##				
##	Turnover_issues_fac\$YearsAtCompany	Turnover_issues_fac\$Attrition		
##		0	1	Row Total
##	-----	-----	-----	-----
##	0	28	16	44
##		2.149	11.181	
##		0.636	0.364	0.030
##		0.023	0.068	
##		0.019	0.011	
##	-----	-----	-----	-----
##	1	112	59	171

##		6.888	35.833	
##		0.655	0.345	0.116
##		0.091	0.249	
##		0.076	0.040	
##				
##	2	100	27	127
##		0.400	2.079	
##		0.787	0.213	0.086
##		0.081	0.114	
##		0.068	0.018	
##				
##	3	108	20	128
##		0.004	0.020	
##		0.844	0.156	0.087
##		0.088	0.084	
##		0.073	0.014	
##				
##	4	91	19	110
##		0.017	0.090	
##		0.827	0.173	0.075
##		0.074	0.080	
##		0.062	0.013	
##				
##	5	175	21	196
##		0.683	3.556	
##		0.893	0.107	0.133
##		0.142	0.089	
##		0.119	0.014	
##				
##	6	67	9	76
##		0.166	0.864	
##		0.882	0.118	0.052
##		0.054	0.038	
##		0.046	0.006	
##				
##	7	79	11	90
##		0.163	0.849	
##		0.878	0.122	0.061
##		0.064	0.046	
##		0.054	0.007	
##				
##	8	71	9	80
##		0.226	1.178	
##		0.887	0.112	0.054
##		0.058	0.038	
##		0.048	0.006	
##				
##	9	74	8	82
##		0.396	2.061	
##		0.902	0.098	0.056

##		0.060	0.034	
##		0.050	0.005	
##				
##	10	102	18	120
##		0.018	0.094	
##		0.850	0.150	0.082
##		0.083	0.076	
##		0.069	0.012	
##				
##	11	30	2	32
##		0.372	1.935	
##		0.938	0.062	0.022
##		0.024	0.008	
##		0.020	0.001	
##				
##	12	14	0	14
##		0.434	2.257	
##		1.000	0.000	0.010
##		0.011	0.000	
##		0.010	0.000	
##				
##	13	22	2	24
##		0.174	0.903	
##		0.917	0.083	0.016
##		0.018	0.008	
##		0.015	0.001	
##				
##	14	16	2	18
##		0.054	0.280	
##		0.889	0.111	0.012
##		0.013	0.008	
##		0.011	0.001	
##				
##	15	19	1	20
##		0.295	1.535	
##		0.950	0.050	0.014
##		0.015	0.004	
##		0.013	0.001	
##				
##	16	11	1	12
##		0.087	0.452	
##		0.917	0.083	0.008
##		0.009	0.004	
##		0.007	0.001	
##				
##	17	8	1	9
##		0.027	0.140	
##		0.889	0.111	0.006
##		0.006	0.004	
##		0.005	0.001	

##				
##	18	12	1	13
##		0.110	0.573	
##		0.923	0.077	0.009
##		0.010	0.004	
##		0.008	0.001	
##	19	10	1	11
##		0.065	0.337	
##		0.909	0.091	0.007
##		0.008	0.004	
##		0.007	0.001	
##	20	26	1	27
##		0.496	2.583	
##		0.963	0.037	0.018
##		0.021	0.004	
##		0.018	0.001	
##	21	13	1	14
##		0.135	0.700	
##		0.929	0.071	0.010
##		0.011	0.004	
##		0.009	0.001	
##	22	14	1	15
##		0.160	0.832	
##		0.933	0.067	0.010
##		0.011	0.004	
##		0.010	0.001	
##	23	1	1	2
##		0.274	1.424	
##		0.500	0.500	0.001
##		0.001	0.004	
##		0.001	0.001	
##	24	5	1	6
##		0.000	0.001	
##		0.833	0.167	0.004
##		0.004	0.004	
##		0.003	0.001	
##	25	4	0	4
##		0.124	0.645	
##		1.000	0.000	0.003
##		0.003	0.000	
##		0.003	0.000	
##	26	4	0	4

##		0.124	0.645	
##		1.000	0.000	0.003
##		0.003	0.000	
##		0.003	0.000	
##				
##	27	2	0	2
##		0.062	0.322	
##		1.000	0.000	0.001
##		0.002	0.000	
##		0.001	0.000	
##				
##	29	2	0	2
##		0.062	0.322	
##		1.000	0.000	0.001
##		0.002	0.000	
##		0.001	0.000	
##				
##	30	1	0	1
##		0.031	0.161	
##		1.000	0.000	0.001
##		0.001	0.000	
##		0.001	0.000	
##				
##	31	2	1	3
##		0.106	0.551	
##		0.667	0.333	0.002
##		0.002	0.004	
##		0.001	0.001	
##				
##	32	2	1	3
##		0.106	0.551	
##		0.667	0.333	0.002
##		0.002	0.004	
##		0.001	0.001	
##				
##	33	4	1	5
##		0.009	0.047	
##		0.800	0.200	0.003
##		0.003	0.004	
##		0.003	0.001	
##				
##	34	1	0	1
##		0.031	0.161	
##		1.000	0.000	0.001
##		0.001	0.000	
##		0.001	0.000	
##				
##	36	2	0	2
##		0.062	0.322	
##		1.000	0.000	0.001

##		0.002	0.000	
##		0.001	0.000	
##	-----	-----	-----	-----
##	37	1	0	1
##		0.031	0.161	
##		1.000	0.000	0.001
##		0.001	0.000	
##		0.001	0.000	
##	-----	-----	-----	-----
##	40	0	1	1
##		0.839	4.364	
##		0.000	1.000	0.001
##		0.000	0.004	
##		0.000	0.001	
##	-----	-----	-----	-----
##	Column Total	1233	237	1470
##		0.839	0.161	
##	-----	-----	-----	-----
##				
##				

25. Relationship of number of Years In Current Role with attrition

CrossTable(Turnover_issues_fac\$YearsInCurrentRole,Turnover_issues_fac\$Attrition)

##				
##				
##	Cell Contents			
##	-----			
##	N			
##	Chi-square contribution			
##	N / Row Total			
##	N / Col Total			
##	N / Table Total			
##	-----			
##				
##				
##	Total Observations in Table:	1470		
##				
##				
##	Turnover_issues_fac\$YearsInCurrentRole	Turnover_issues_fac\$Attrition		
##		0	1	Row Total
##	-----	-----	-----	-----
##	-			
##	0	171	73	244
##		5.536	28.803	
##		0.701	0.299	0.166

##		0.139	0.308	
##		0.116	0.050	
##	-----	-----	-----	-----
-				
##	1	46	11	57
##		0.069	0.357	
##		0.807	0.193	0.039
##		0.037	0.046	
##		0.031	0.007	
##	-----	-----	-----	-----
-				
##	2	304	68	372
##		0.206	1.074	
##		0.817	0.183	0.253
##		0.247	0.287	
##		0.207	0.046	
##	-----	-----	-----	-----
-				
##	3	119	16	135
##		0.294	1.527	
##		0.881	0.119	0.092
##		0.097	0.068	
##		0.081	0.011	
##	-----	-----	-----	-----
-				
##	4	89	15	104
##		0.036	0.186	
##		0.856	0.144	0.071
##		0.072	0.063	

##		0.061	0.010	
##	-----	-----	-----	-----
-				
##	5	35	1	36
##		0.764	3.976	
##		0.972	0.028	0.024
##		0.028	0.004	
##		0.024	0.001	
##	-----	-----	-----	-----
-				
##	6	35	2	37
##		0.507	2.636	
##		0.946	0.054	0.025
##		0.028	0.008	
##		0.024	0.001	
##	-----	-----	-----	-----
-				
##	7	191	31	222
##		0.123	0.642	
##		0.860	0.140	0.151
##		0.155	0.131	
##		0.130	0.021	
##	-----	-----	-----	-----
-				
##	8	82	7	89
##		0.723	3.764	
##		0.921	0.079	0.061
##		0.067	0.030	
##		0.056	0.005	

##	-----	-----	-----	-----
-				
##	9	61	6	67
##		0.410	2.135	
##		0.910	0.090	0.046
##		0.049	0.025	
##		0.041	0.004	
##	-----	-----	-----	-----
-				
##	10	27	2	29
##		0.294	1.531	
##		0.931	0.069	0.020
##		0.022	0.008	
##		0.018	0.001	
##	-----	-----	-----	-----
-				
##	11	22	0	22
##		0.682	3.547	
##		1.000	0.000	0.015
##		0.018	0.000	
##		0.015	0.000	
##	-----	-----	-----	-----
-				
##	12	9	1	10
##		0.045	0.232	
##		0.900	0.100	0.007
##		0.007	0.004	
##		0.006	0.001	
##	-----	-----	-----	-----

-				
##	13	13	1	14
##		0.135	0.700	
##		0.929	0.071	0.010
##		0.011	0.004	
##		0.009	0.001	
##	-----	-----	-----	-----
-				
##	14	10	1	11
##		0.065	0.337	
##		0.909	0.091	0.007
##		0.008	0.004	
##		0.007	0.001	
##	-----	-----	-----	-----
-				
##	15	6	2	8
##		0.075	0.391	
##		0.750	0.250	0.005
##		0.005	0.008	
##		0.004	0.001	
##	-----	-----	-----	-----
-				
##	16	7	0	7
##		0.217	1.129	
##		1.000	0.000	0.005
##		0.006	0.000	
##		0.005	0.000	
##	-----	-----	-----	-----
-				
##	17	4	0	4

```

|
##          |      0.124 |      0.645 |
|
##          |      1.000 |      0.000 |      0.003
|
##          |      0.003 |      0.000 |
|
##          |      0.003 |      0.000 |
|
## -----|-----|-----|-----
-|
##          18 |          2 |          0 |          2
|
##          |      0.062 |      0.322 |
|
##          |      1.000 |      0.000 |      0.001
|
##          |      0.002 |      0.000 |
|
##          |      0.001 |      0.000 |
|
## -----|-----|-----|-----
-|
##          Column Total |      1233 |      237 |      1470
|
##          |      0.839 |      0.161 |
|
## -----|-----|-----|-----
-|
##
##

```

26. Relationship of Years Since Last Promotion with Attrition

```
CrossTable(Turnover_issues_fac$YearsSinceLastPromotion,Turnover_issues_fac$Attrition)
```

```

##
##
##      Cell Contents
## |-----|
## |              N |
## | Chi-square contribution |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table:  1470
##

```

Turnover_issues_fac\$YearsSinceLastPromotion	0	1	Row
581	471	110	
0.395	0.547	2.846	
	0.811	0.189	
	0.382	0.464	
	0.320	0.075	
357	308	49	
0.243	0.245	1.272	
	0.863	0.137	
	0.250	0.207	
	0.210	0.033	
159	132	27	
0.108	0.014	0.073	
	0.830	0.170	
	0.107	0.114	
	0.090	0.018	
52	43	9	
0.035	0.009	0.045	
	0.827	0.173	
	0.035	0.038	

##		0.029	0.006
##			
##	4	56	5
61		0.457	2.377
##		0.918	0.082
0.041		0.045	0.021
##		0.038	0.003
##			
##	5	43	2
45		0.732	3.806
##		0.956	0.044
0.031		0.035	0.008
##		0.029	0.001
##			
##	6	26	6
32		0.026	0.137
##		0.812	0.188
0.022		0.021	0.025
##		0.018	0.004
##			
##	7	60	16
76		0.220	1.146
##		0.789	0.211
0.052		0.049	0.068
##		0.041	0.011

##			
##	8	18	0
18		0.558	2.902
##		1.000	0.000
0.012		0.015	0.000
##		0.012	0.000
##			
##	9	13	4
17		0.111	0.578
##		0.765	0.235
0.012		0.011	0.017
##		0.009	0.003
##			
##	10	5	1
6		0.000	0.001
##		0.833	0.167
0.004		0.004	0.004
##		0.003	0.001
##			
##	11	22	2
24		0.174	0.903
##		0.917	0.083
0.016		0.018	0.008
##		0.015	0.001
##			

##	12	10	0
10			
##		0.310	1.612
##		1.000	0.000
0.007			
##		0.008	0.000
##		0.007	0.000
## -----	-----	-----	-----

##	13	8	2
10			
##		0.018	0.093
##		0.800	0.200
0.007			
##		0.006	0.008
##		0.005	0.001
## -----	-----	-----	-----

##	14	8	1
9			
##		0.027	0.140
##		0.889	0.111
0.006			
##		0.006	0.004
##		0.005	0.001
## -----	-----	-----	-----

##	15	10	3
13			
##		0.075	0.390
##		0.769	0.231
0.009			
##		0.008	0.013
##		0.007	0.002
## -----	-----	-----	-----

##	Column Total	1233	237

```

1470 |
## | 0.839 | 0.161 |
|
## -----|-----|-----|-----
-----|
##
##

```

Logistic Regressoin Model

#set.seed() function is used for generating randoms numbers#

#Looking for most significant variables with Attrition#

1. Dividing Data set into Training (70%) and Test (30%) set to predict Attrition. It ensures the data sets are marginally different, and thus, it may find out better solution.

```

set.seed(1234)
div <- sample(2, nrow(Turnover_issues_fac), replace = T, prob = c(0.7, 0.3))
train <- Turnover_issues_fac[div==1,]
test <- Turnover_issues_fac[div==2,]

```

2. Logistic Regression Model

#Logistic regression is used to find out probability of event success and event failure by categorizing data into discrete classes#

```

turnover_model <- glm(Attrition ~ Age+BusinessTravel+DailyRate+Department+Dis
tanceFromHome+Education+EducationField+EnvironmentSatisfaction+Gender+HourlyR
ate+JobInvolvement+JobLevel+JobRole+JobSatisfaction+MaritalStatus+MonthlyInco
me+NumCompaniesWorked+OverTime+PercentSalaryHike+PerformanceRating+Relationsh
ipSatisfaction+StockOptionLevel+TotalWorkingYears+TrainingTimesLastYear+WorkL
ifeBalance+YearsAtCompany+YearsInCurrentRole+YearsSinceLastPromotion+YearsWit
hCurrManager, data = train, family = 'binomial')
summary(turnover_model)

```

```

##
## Call:
## glm(formula = Attrition ~ Age + BusinessTravel + DailyRate +
##     Department + DistanceFromHome + Education + EducationField +
##     EnvironmentSatisfaction + Gender + HourlyRate + JobInvolvement +
##     JobLevel + JobRole + JobSatisfaction + MaritalStatus + MonthlyIncome +
##     NumCompaniesWorked + OverTime + PercentSalaryHike + PerformanceRating
+
##     RelationshipSatisfaction + StockOptionLevel + TotalWorkingYears +
##     TrainingTimesLastYear + WorkLifeBalance + YearsAtCompany +
##     YearsInCurrentRole + YearsSinceLastPromotion + YearsWithCurrManager,
##     family = "binomial", data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max

```

```

## -1.9532 -0.4283 -0.1854 -0.0455 3.5653
##
## Coefficients:
##
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.510e+00 4.874e+02 -0.017 0.986068
## Age -4.496e-02 1.819e-02 -2.472 0.013442 *
## BusinessTravelTravel_Frequently 1.929e+00 5.230e-01 3.689 0.000225 **
*
## BusinessTravelTravel_Rarely 9.431e-01 4.778e-01 1.974 0.048392 *
## DailyRate -4.426e-04 2.935e-04 -1.508 0.131567
## DepartmentResearch & Development 1.426e+01 4.874e+02 0.029 0.976665
## DepartmentSales 1.383e+01 4.874e+02 0.028 0.977356
## DistanceFromHome 4.688e-02 1.439e-02 3.259 0.001119 **
## Education2 -1.489e-01 4.113e-01 -0.362 0.717403
## Education3 5.125e-02 3.644e-01 0.141 0.888157
## Education4 1.067e-01 3.934e-01 0.271 0.786305
## Education5 -4.439e-01 9.727e-01 -0.456 0.648130
## EducationFieldLife Sciences -1.259e+00 1.073e+00 -1.173 0.240920
## EducationFieldMarketing -5.851e-01 1.139e+00 -0.514 0.607586
## EducationFieldMedical -1.049e+00 1.072e+00 -0.979 0.327710
## EducationFieldOther -9.540e-01 1.150e+00 -0.829 0.406925
## EducationFieldTechnical Degree -9.581e-02 1.071e+00 -0.089 0.928709
## EnvironmentSatisfaction2 -1.241e+00 3.558e-01 -3.488 0.000487 **
*
## EnvironmentSatisfaction3 -1.467e+00 3.434e-01 -4.273 1.93e-05 **
*
## EnvironmentSatisfaction4 -1.423e+00 3.368e-01 -4.226 2.38e-05 **
*
## GenderMale 4.569e-01 2.426e-01 1.883 0.059677 .
## HourlyRate 6.828e-04 5.813e-03 0.117 0.906498
## JobInvolvement2 -1.389e+00 4.498e-01 -3.088 0.002014 **
## JobInvolvement3 -1.549e+00 4.227e-01 -3.665 0.000247 **
*
## JobInvolvement4 -2.626e+00 6.338e-01 -4.144 3.42e-05 **
*
## JobLevel2 -1.195e+00 5.329e-01 -2.242 0.024947 *
## JobLevel3 -2.347e-03 8.754e-01 -0.003 0.997861
## JobLevel4 -8.183e-01 1.496e+00 -0.547 0.584482
## JobLevel5 2.687e+00 2.080e+00 1.292 0.196443
## JobRoleHuman Resources 1.482e+01 4.874e+02 0.030 0.975749
## JobRoleLaboratory Technician 1.044e+00 7.199e-01 1.450 0.147197
## JobRoleManager -2.187e+00 1.450e+00 -1.509 0.131395
## JobRoleManufacturing Director 2.902e-01 7.099e-01 0.409 0.682751
## JobRoleResearch Director -3.849e+00 1.538e+00 -2.502 0.012337 *
## JobRoleResearch Scientist -1.217e-01 7.270e-01 -0.167 0.867010
## JobRoleSales Executive 1.721e+00 1.523e+00 1.130 0.258636
## JobRoleSales Representative 2.100e+00 1.611e+00 1.304 0.192308
## JobSatisfaction2 -8.360e-01 3.583e-01 -2.333 0.019638 *
## JobSatisfaction3 -9.204e-01 3.135e-01 -2.935 0.003332 **
## JobSatisfaction4 -1.411e+00 3.306e-01 -4.270 1.96e-05 **

```



```

*
## MaritalStatusMarried      4.965e-01  3.500e-01  1.419 0.156039
## MaritalStatusSingle      5.186e-01  5.023e-01  1.032 0.301888
## MonthlyIncome            -1.619e-05  1.144e-04 -0.142 0.887462
## NumCompaniesWorked1      5.367e-01  3.962e-01  1.355 0.175496
## NumCompaniesWorked2      4.977e-01  5.565e-01  0.894 0.371166
## NumCompaniesWorked3      5.008e-01  5.841e-01  0.857 0.391265
## NumCompaniesWorked4      1.124e+00  5.926e-01  1.897 0.057858 .
## NumCompaniesWorked5      1.897e+00  6.286e-01  3.019 0.002540 **
## NumCompaniesWorked6      1.761e+00  5.959e-01  2.955 0.003128 **
## NumCompaniesWorked7      2.366e+00  6.190e-01  3.822 0.000132 **
*
## NumCompaniesWorked8      1.070e+00  7.732e-01  1.384 0.166210
## NumCompaniesWorked9      2.214e+00  6.903e-01  3.207 0.001339 **
## OverTime1                2.358e+00  2.615e-01  9.019 < 2e-16 **
*
## PercentSalaryHike        -4.344e-02  4.998e-02 -0.869 0.384810
## PerformanceRating4        3.390e-01  5.235e-01  0.647 0.517317
## RelationshipSatisfaction2  -6.249e-01  3.616e-01 -1.728 0.083914 .
## RelationshipSatisfaction3  -6.803e-01  3.269e-01 -2.081 0.037409 *
## RelationshipSatisfaction4  -9.421e-01  3.386e-01 -2.782 0.005405 **
## StockOptionLevel1        -1.239e+00  4.106e-01 -3.018 0.002543 **
## StockOptionLevel2        -1.123e+00  5.211e-01 -2.154 0.031210 *
## StockOptionLevel3        -7.314e-01  6.085e-01 -1.202 0.229405
## TotalWorkingYears        -4.059e-02  3.852e-02 -1.054 0.291998
## TrainingTimesLastYear1    -1.363e+00  7.161e-01 -1.903 0.057048 .
## TrainingTimesLastYear2    -1.692e+00  5.559e-01 -3.043 0.002339 **
## TrainingTimesLastYear3    -1.838e+00  5.664e-01 -3.245 0.001175 **
## TrainingTimesLastYear4    -1.250e+00  6.370e-01 -1.963 0.049702 *
## TrainingTimesLastYear5    -1.880e+00  6.899e-01 -2.725 0.006423 **
## TrainingTimesLastYear6    -2.297e+00  8.071e-01 -2.846 0.004427 **
## WorkLifeBalance2         -1.225e+00  5.137e-01 -2.385 0.017100 *
## WorkLifeBalance3         -1.686e+00  4.898e-01 -3.442 0.000577 **
*
## WorkLifeBalance4         -1.310e+00  5.868e-01 -2.233 0.025574 *
## YearsAtCompany            1.543e-01  4.888e-02  3.157 0.001595 **
## YearsInCurrentRole        -1.961e-01  6.409e-02 -3.061 0.002209 **
## YearsSinceLastPromotion    1.759e-01  5.580e-02  3.152 0.001619 **
## YearsWithCurrManager      -1.969e-01  6.090e-02 -3.234 0.001222 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 965.32  on 1041  degrees of freedom
## Residual deviance: 551.79  on  967  degrees of freedom
## AIC: 701.79
##
## Number of Fisher Scoring iterations: 14

```

#Picking significant variables with $p < 0.05$

#Age(-) #Business Travel(+) #Distance From Home(+) #Environment Satisfaction(-) #Job Involvement(-) #Job Level(-) #Job Role #Job Satisfaction(-) #NumCompaniesWorked(+) #Relationship Satisfaction(-) #Stock Option Level(-) #Work Life Balance(-) #Years At company(-) #Years In Current Role(-) #Years Since Last promotion(+) #Years With Curr Manager(-)

#Retaining statistically significant variables#

```
turnover_model <- glm(Attrition ~ Age+BusinessTravel+DistanceFromHome+EnvironmentSatisfaction+JobInvolvement+JobLevel+JobRole+JobSatisfaction+NumCompaniesWorked+OverTime+RelationshipSatisfaction+StockOptionLevel+WorkLifeBalance+YearsAtCompany+YearsInCurrentRole+YearsSinceLastPromotion+YearsWithCurrManager, data = train, family = 'binomial')
summary(turnover_model)
```

##

Call:

```
## glm(formula = Attrition ~ Age + BusinessTravel + DistanceFromHome +
##      EnvironmentSatisfaction + JobInvolvement + JobLevel + JobRole +
##      JobSatisfaction + NumCompaniesWorked + OverTime + RelationshipSatisfaction +
##      StockOptionLevel + WorkLifeBalance + YearsAtCompany + YearsInCurrentRole +
##      YearsSinceLastPromotion + YearsWithCurrManager, family = "binomial",
##      data = train)
```

##

Deviance Residuals:

```
##      Min       1Q   Median       3Q      Max
## -1.9634  -0.4605  -0.2251  -0.0667   3.4878
```

##

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
## (Intercept)	2.87199	1.13707	2.526	0.011544	*
## Age	-0.05516	0.01605	-3.437	0.000589	***
## BusinessTravelTravel_Frequently	1.87303	0.48791	3.839	0.000124	***
## BusinessTravelTravel_Rarely	0.93626	0.44817	2.089	0.036699	*
## DistanceFromHome	0.03935	0.01351	2.913	0.003578	**
## EnvironmentSatisfaction2	-1.03210	0.33038	-3.124	0.001784	**
## EnvironmentSatisfaction3	-1.18340	0.31206	-3.792	0.000149	***
## EnvironmentSatisfaction4	-1.17292	0.30854	-3.801	0.000144	***
## JobInvolvement2	-1.26975	0.41614	-3.051	0.002279	**
## JobInvolvement3	-1.46453	0.39158	-3.740	0.000184	***
## JobInvolvement4	-2.32429	0.58397	-3.980	6.89e-05	***
## JobLevel2	-1.41667	0.46964	-3.016	0.002557	**
## JobLevel3	-0.48716	0.59121	-0.824	0.409934	
## JobLevel4	-1.09975	0.97496	-1.128	0.259322	
## JobLevel5	2.07255	1.48456	1.396	0.162694	
## JobRoleHuman Resources	1.06807	0.77468	1.379	0.167983	

```

## JobRoleLaboratory Technician    0.85313    0.67929    1.256 0.209142
## JobRoleManager                  -2.42607    1.32285   -1.834 0.066658 .
## JobRoleManufacturing Director   0.40462    0.66768    0.606 0.544510
## JobRoleResearch Director        -3.86723    1.37053   -2.822 0.004777 **
## JobRoleResearch Scientist       -0.12954    0.68208   -0.190 0.849370
## JobRoleSales Executive           1.38834    0.54202    2.561 0.010424 *
## JobRoleSales Representative      1.53576    0.73450    2.091 0.036537 *
## JobSatisfaction2                -0.81683    0.34093   -2.396 0.016580 *
## JobSatisfaction3                -0.80298    0.29304   -2.740 0.006141 **
## JobSatisfaction4                -1.26441    0.30575   -4.135 3.54e-05 ***
## NumCompaniesWorked1             0.48123    0.37153    1.295 0.195224
## NumCompaniesWorked2             0.45447    0.51184    0.888 0.374589
## NumCompaniesWorked3             0.49750    0.52095    0.955 0.339587
## NumCompaniesWorked4             0.88082    0.54262    1.623 0.104535
## NumCompaniesWorked5             1.70565    0.57010    2.992 0.002773 **
## NumCompaniesWorked6             1.56968    0.53878    2.913 0.003575 **
## NumCompaniesWorked7             2.25332    0.57134    3.944 8.02e-05 ***
## NumCompaniesWorked8             1.10838    0.70989    1.561 0.118447
## NumCompaniesWorked9             2.02827    0.63774    3.180 0.001471 **
## OverTime1                       2.18701    0.23994    9.115 < 2e-16 ***
## RelationshipSatisfaction2        -0.57526    0.34051   -1.689 0.091140 .
## RelationshipSatisfaction3        -0.57523    0.30788   -1.868 0.061715 .
## RelationshipSatisfaction4        -0.84977    0.32147   -2.643 0.008207 **
## StockOptionLevel1              -1.31806    0.25107   -5.250 1.52e-07 ***
## StockOptionLevel2              -1.26911    0.39362   -3.224 0.001263 **
## StockOptionLevel3              -0.97094    0.48819   -1.989 0.046720 *
## WorkLifeBalance2               -1.23363    0.48571   -2.540 0.011090 *
## WorkLifeBalance3               -1.53361    0.45899   -3.341 0.000834 ***
## WorkLifeBalance4               -1.14309    0.55317   -2.066 0.038788 *
## YearsAtCompany                  0.10574    0.04242    2.492 0.012690 *
## YearsInCurrentRole              -0.14631    0.05897   -2.481 0.013100 *
## YearsSinceLastPromotion          0.16849    0.05169    3.259 0.001117 **
## YearsWithCurrManager            -0.18701    0.05749   -3.253 0.001143 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 965.32  on 1041  degrees of freedom
## Residual deviance: 590.60  on  993  degrees of freedom
## AIC: 688.6
##
## Number of Fisher Scoring iterations: 7

```

Prediction of associates leaving company

```

prediction <- predict (turnover_model, train, type = 'response')
head(prediction)

```

```

##           1           2           3           4           6           7
## 0.551320543 0.001127055 0.750483345 0.588867808 0.062987992 0.266209730

```

Observing few top output to predict attrition of associates

head(train)

```
##   Age Attrition   BusinessTravel DailyRate      Department
## 1  41         1   Travel_Rarely     1102           Sales
## 2  49         0 Travel_Frequently     279 Research & Development
## 3  37         1   Travel_Rarely     1373 Research & Development
## 4  33         0 Travel_Frequently     1392 Research & Development
## 6  32         0 Travel_Frequently     1005 Research & Development
## 7  59         0   Travel_Rarely     1324 Research & Development
##   DistanceFromHome Education EducationField EnvironmentSatisfaction Gender
## 1                 1         2 Life Sciences                2 Female
## 2                 8         1 Life Sciences                3  Male
## 3                 2         2      Other                  4  Male
## 4                 3         4 Life Sciences                4 Female
## 6                 2         2 Life Sciences                4  Male
## 7                 3         3      Medical                 3 Female
##   HourlyRate JobInvolvement JobLevel      JobRole JobSatisfaction
## 1         94             3         2   Sales Executive            4
## 2         61             2         2   Research Scientist        2
## 3         92             2         1 Laboratory Technician      3
## 4         56             3         1   Research Scientist        3
## 6         79             3         1 Laboratory Technician      4
## 7         81             4         1 Laboratory Technician      1
##   MaritalStatus MonthlyIncome MonthlyRate NumCompaniesWorked Over18 OverTi
me
## 1      Single      5993      19479            8      Y
1
## 2    Married      5130      24907            1      Y
0
## 3      Single      2090       2396            6      Y
1
## 4    Married      2909      23159            1      Y
1
## 6      Single      3068      11864            0      Y
0
## 7    Married      2670       9964            4      Y
1
##   PercentSalaryHike PerformanceRating RelationshipSatisfaction StandardHou
rs
## 1             11             3             1
80
## 2             23             4             4
80
## 3             15             3             2
80
## 4             11             3             3
80
## 6             13             3             3
80
```

```
## 7          20          4          1
80
## StockOptionLevel TotalWorkingYears TrainingTimesLastYear WorkLifeBalance
## 1          0          8          0          1
## 2          1         10          3          3
## 3          0          7          3          3
## 4          0          8          3          3
## 6          0          8          2          2
## 7          3         12          3          2
## YearsAtCompany YearsInCurrentRole YearsSinceLastPromotion
## 1          6          4          0
## 2         10          7          1
## 3          0          0          0
## 4          8          7          3
## 6          7          7          3
## 7          1          0          0
## YearsWithCurrManager
## 1          5
## 2          7
## 3          0
## 4          0
## 6          6
## 7          0
```

Confusion Matrix to find out misclassification error in training data

```
predic1 <- ifelse(prediction>0.5, 1, 0)
table1<-table(Predicted = predic1, Actual = train$Attrition)
table1

##           Actual
## Predicted    0    1
##           0 834  84
##           1  26  98
```

Number of associates who do not quit job are 834. Here it is predicted and truly do not happen. Again, 98 associates is predicted to quit and really it's happened. Moreover, 26 associates predicted to quit, but it's not happen. Furthermore, 84 associates predicted won't quit, but originally it's happened.

Misclassification Percentage in training data

```
1-sum(diag(table1))/sum(table1)

## [1] 0.1055662
```

So it will be 10.56%

Measuring misclassification error of Test Data

```
prediction1 <- predict(turnover_model, test, type = 'response')
predic2 <- ifelse(prediction1>0.5, 1, 0)
table2 <- table(Predicted = predic2, Actual = test$Attrition)
table2
```

```
##           Actual
## Predicted   0   1
##           0 353  31
##           1  20  24
```

Number of associates who do not quit job are 353. Here it is predicted and truly do not happen. Again, 24 associates is predicted to quit and really it's happened. Moreover, 20 associates predicted to quit, but it's not happen. Furthermore, 31 associates predicted won't quit, but originally it's happened.

Misclassification Percentage of test data

```
1-sum(diag(table2))/sum(table2)
```

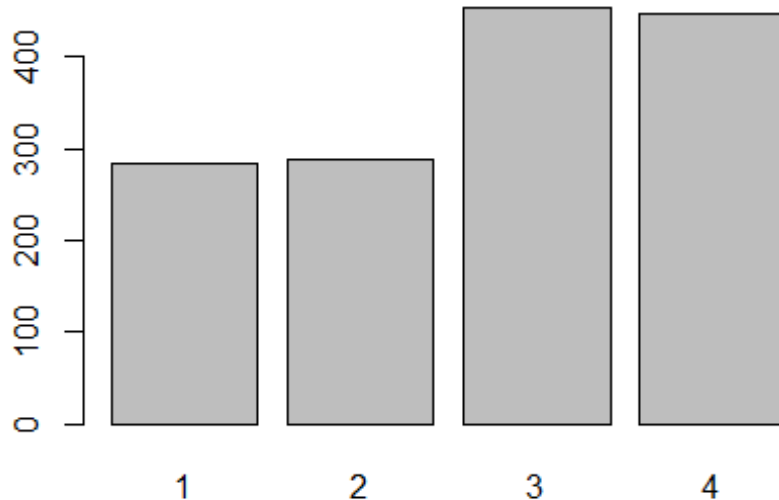
```
## [1] 0.1191589
```

So it will be 11.91%

Finding satisfaction of staff

1. Based on Environment Satisfaction

```
plot(Turnover_issues_fac$EnvironmentSatisfaction)
```



```
median(as.numeric(Turnover_issues_fac$EnvironmentSatisfaction))
```

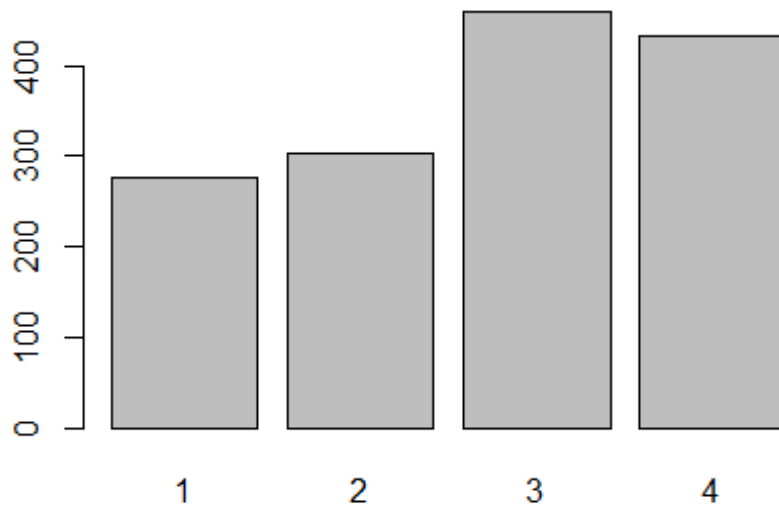
```
## [1] 3
```

```
mean(as.numeric(Turnover_issues_fac$EnvironmentSatisfaction))
```

```
## [1] 2.721769
```

2. Based on Relationship Satisfaction

```
plot(Turnover_issues_fac$RelationshipSatisfaction)
```



```
median(as.numeric(Turnover_issues_fac$RelationshipSatisfaction))
```

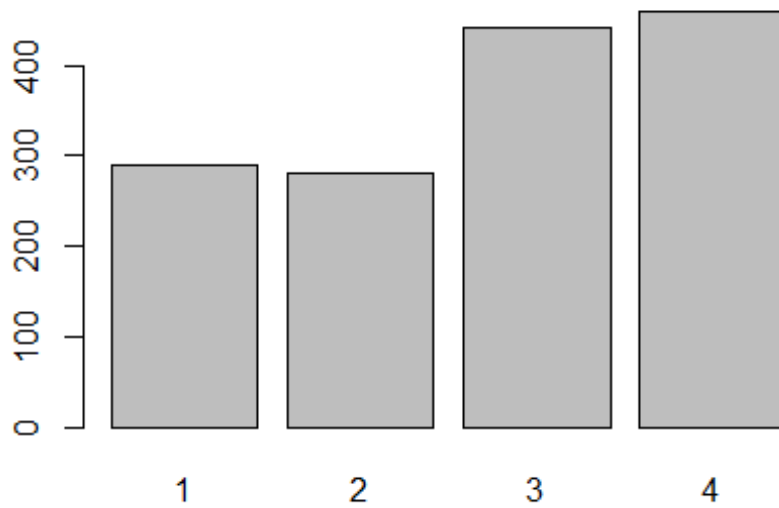
```
## [1] 3
```

```
mean(as.numeric(Turnover_issues_fac$RelationshipSatisfaction))
```

```
## [1] 2.712245
```

3. Based on Job Satisfaction

```
plot(Turnover_issues_fac$JobSatisfaction)
```



```
median(as.numeric(Turnover_issues_fac$JobSatisfaction))  
## [1] 3  
mean(as.numeric(Turnover_issues_fac$JobSatisfaction))  
## [1] 2.728571
```

Here, median is 3 out of 4 which is higher than mean 2.7 out of 4, so we can say that more than half of the associates have job satisfaction above average (2.7).