COMM1190 Individual Assignment T1 2022

Data, Insights and Decisions (University of New South Wales)

Table of Content

Introduction

The purpose of this report is to provide an analysis of factors contributing to job satisfaction at Globex Pharma ('Globex') and consider how to enhance this aspect within the workforce. In facilitating this analysis, driving factors will be considered in relation to attrition to analyse how employee morale is affected at Globex to better identify any solutions. In doing so, three key driving factors will be outlined, **supported by visual representation**, followed by any appropriate recommendations and solutions.

## Summary Statistics and Data Exploration

In determining the main factors that have bearing on job satisfaction at Globex, it becomes important to understand the dispersion of job roles within the company. Figure 1 presents a pie chart of the role composition at Globex.

Figure 1: Composition of Jobs at Globex Pharma
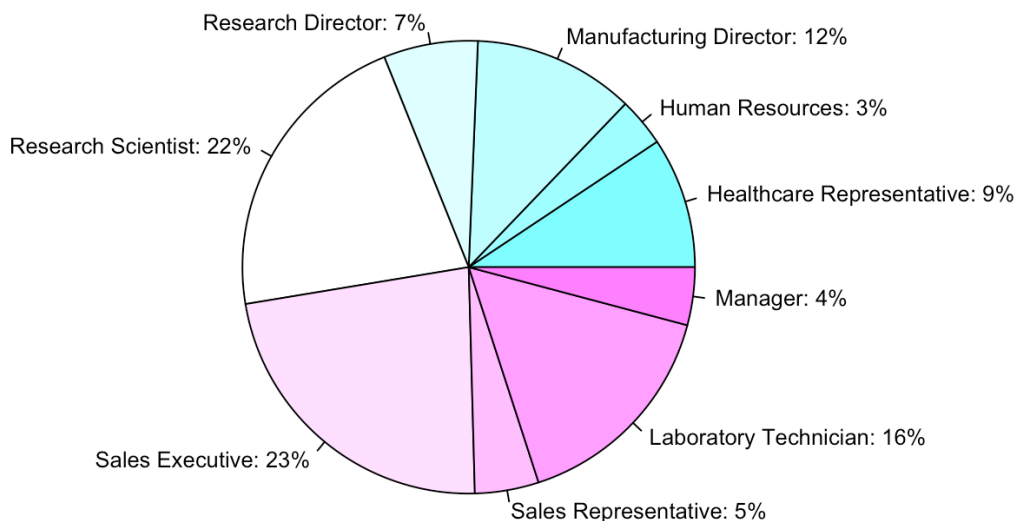
What becomes evident is the large dispersion of jobs at Globex, ranging across 9 job roles. In conducting this report, attrition will be identified as

one of the main variables in considering job satisfaction, which is presented in Figure 2.

Figure 2: Rate of Attrition
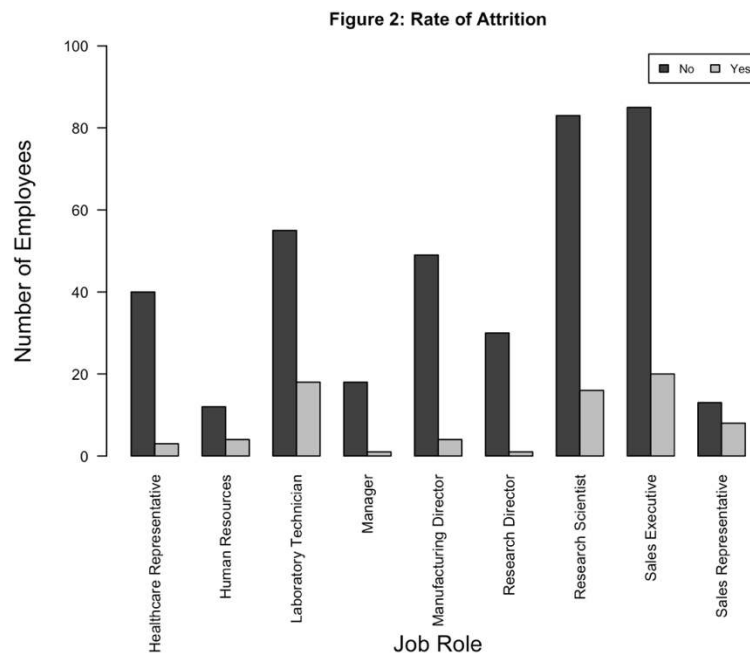
Factor Analysis

Relationship between Income and Attrition

Figure 2 indicates that the jobs with the highest number of attritions are Sales Executives, Laboratory Technicians, Research Scientists and Sales Representatives. One key driving factor which has been identified concerning this is the Monthly Income of employees, presented in Figure 3.
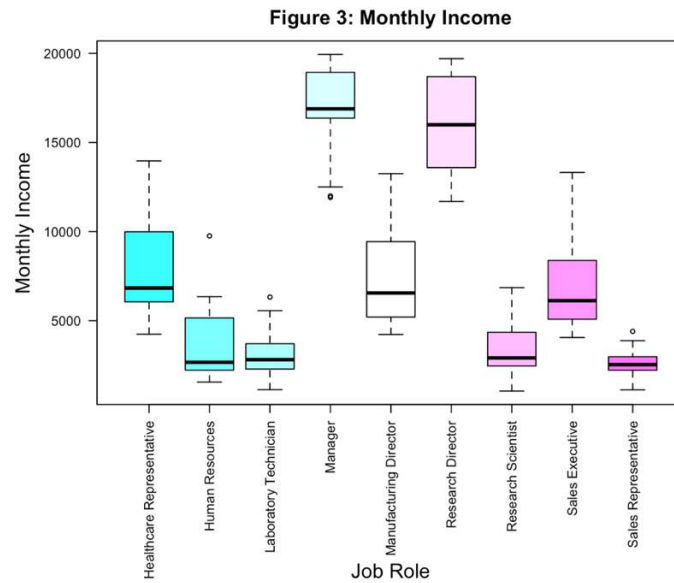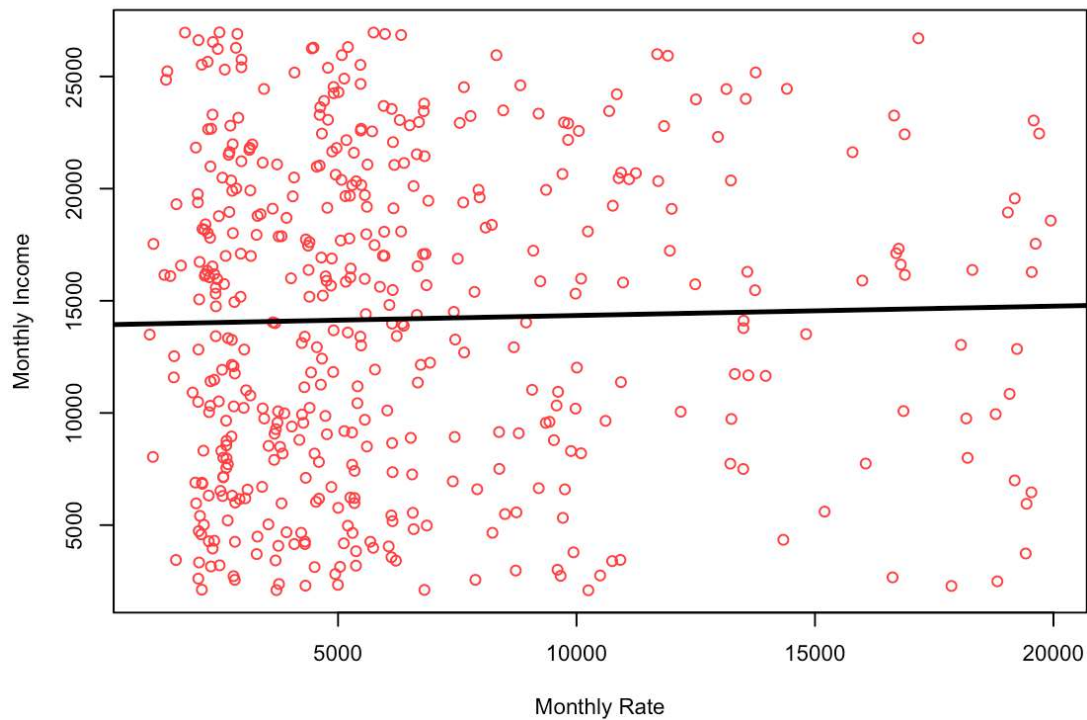
Figure 3: Monthly Income

Figure 3 shows that the highest paying job roles are Managers and Research Directors with a median Monthly Income of $16885 and $15992. The median monthly income of the previously mentioned jobs is $6120, $2811, $2909, and $2534. These results are suggested to be underwhelming when considering that the median Monthly Income of all employees is $5075. Subsequently, 46% of Job Roles have a median income of below $5075. This may suggest a large wage disparity amongst employees, resulting in high attrition.

4

Figure 4: Monthly Rate vs Monthly Income

Moreover, Figure 4 indicates a positive, yet weak correlation (r = 0.026) between the rate at which employees are valued compared to their Monthly Income. This further suggests a lack of consistency within their wage structure that allows for greater discrepancies, further contributing to low job satisfaction and high attrition.

Overtime

Another key driving factor that may contribute to high attrition can be identified as to whether employees have worked overtime. Figures 5 and 6 present the composition of Overtime amongst Job Roles and its effect on Monthly Income.

Figure 5: Overtime Worked vs Monthly Income
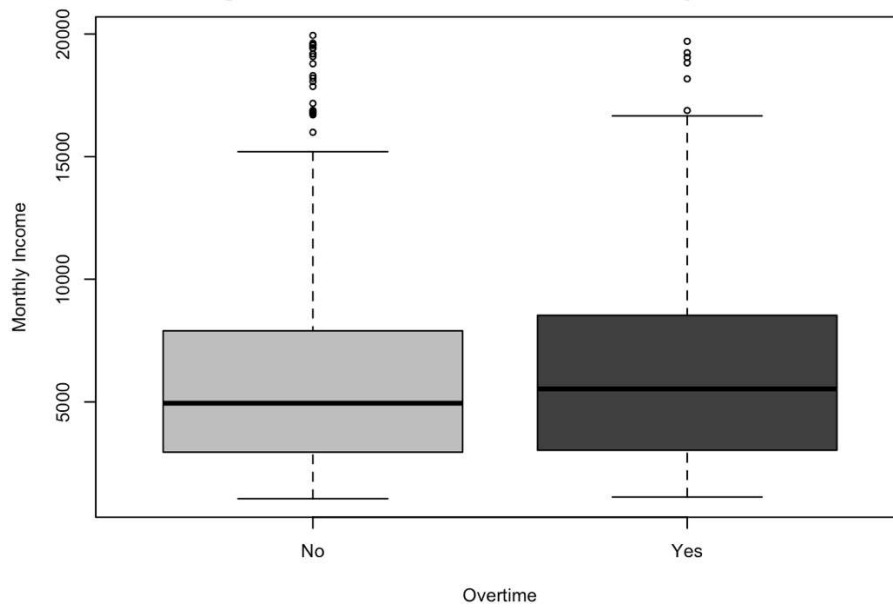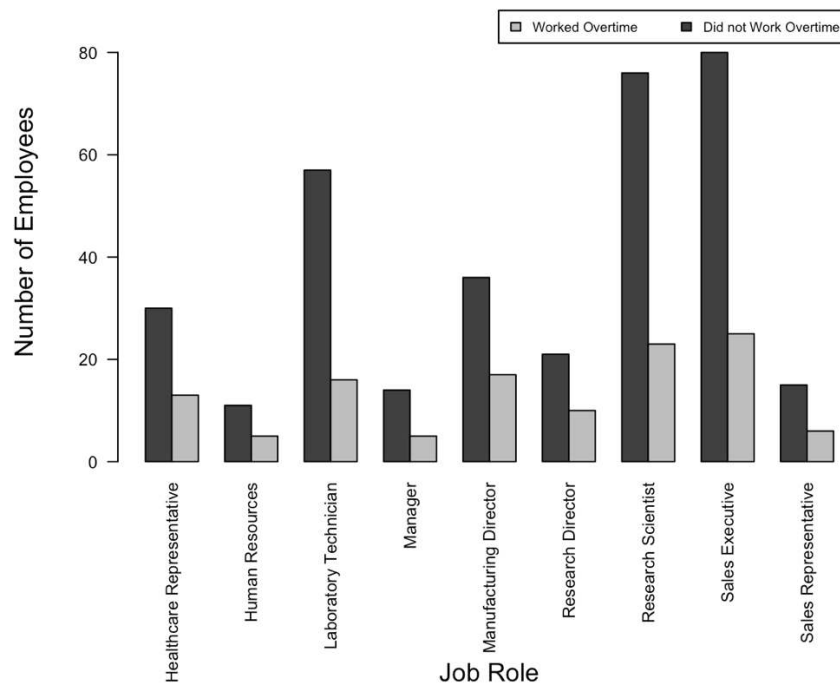

Figure 6: Worked Overtime vs Job Role

While the lowest paying jobs have proportionally fewer employees working overtime, Figure 5 indicates a larger IQR, suggesting greater variability amongst incomes and disproportionately affecting this group. Of those who did not work overtime, there are more top outliers which once again reflects the discrepancies of the wage structure.

Figure 7: Attrition Rate for Overtime Workers
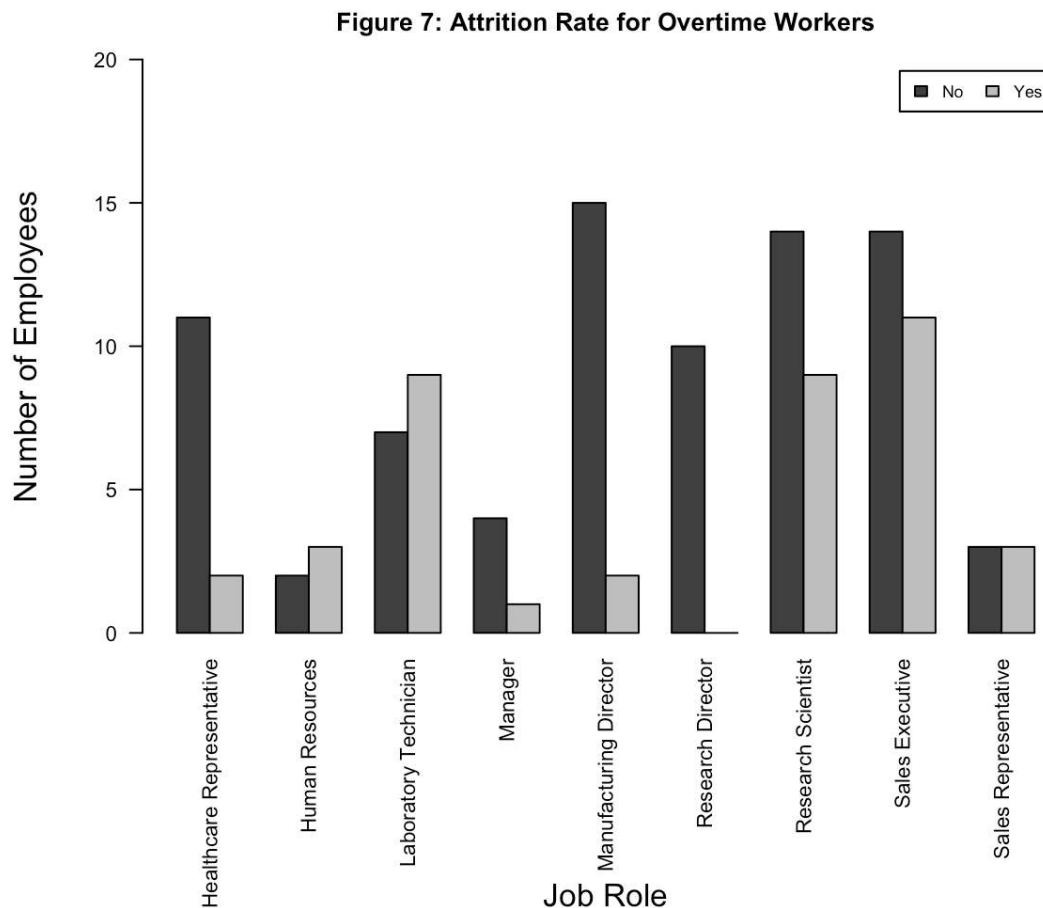
Figure 7 shows the rates of attrition amongst those employees who normally work overtime. It is indicated that the lowest-paid employees have the highest rates of attrition which may suggest both these factors contribute to job dissatisfaction.

Years at Company

Lastly, the years spent working for Globex is a key variable for job satisfaction which may reflect an imbalance of experience amongst employees.

Mean	6.770
Median	5
Kurtosis	7.830
Skewness	1.885

Figure 8: Histogram vs Normal PDF - Years At Company

Figure 9: QQ Plot for Years At Company against Normal Quantiles

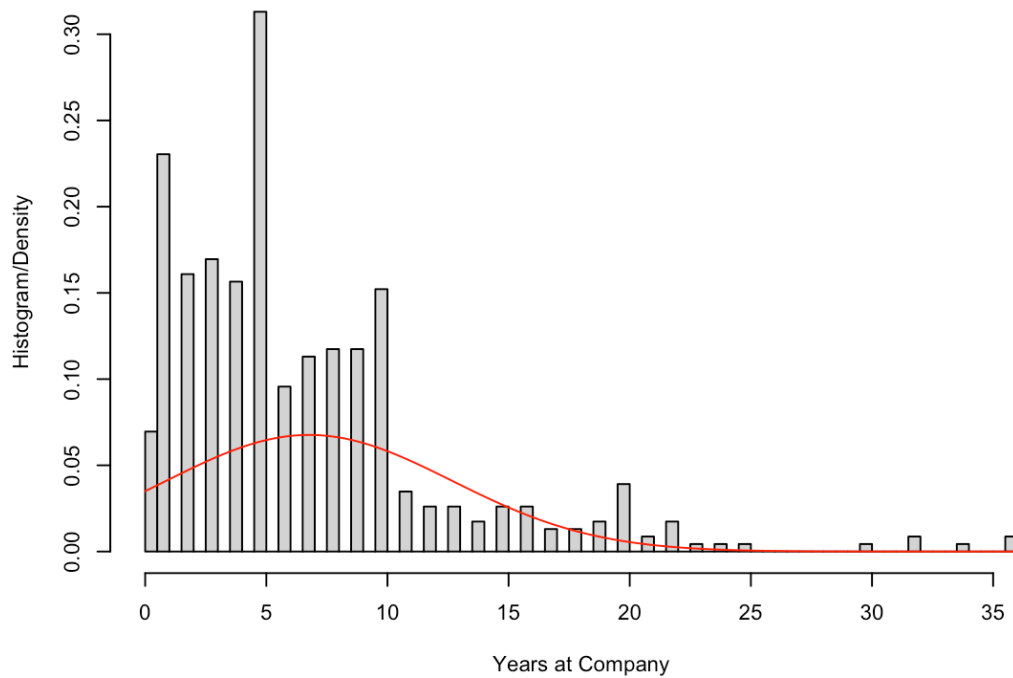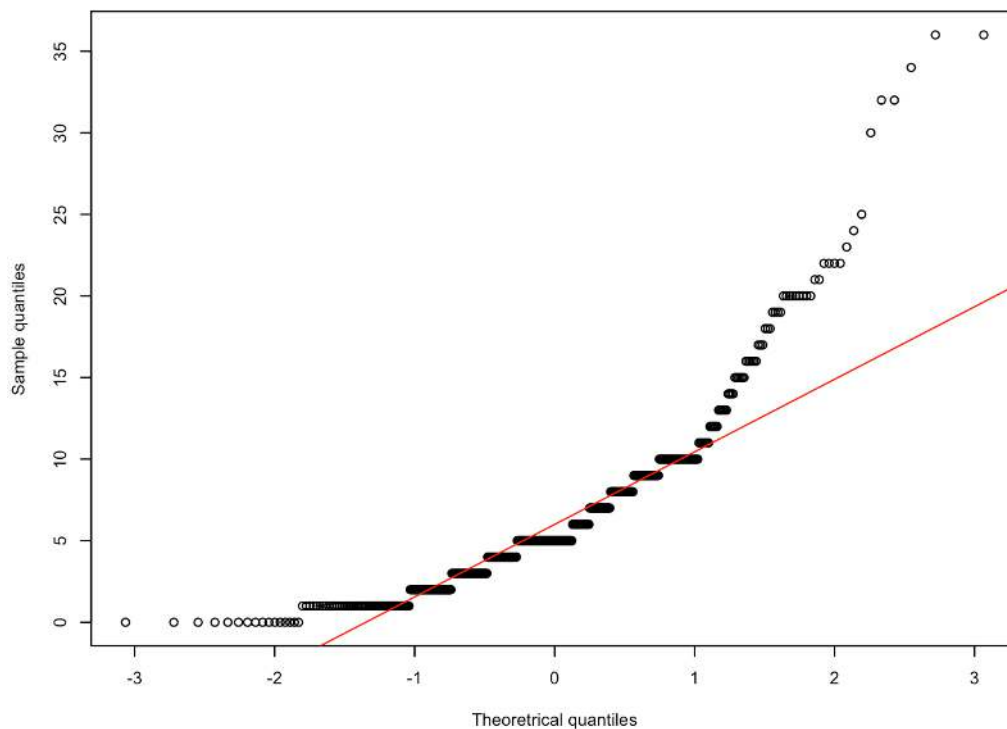Figure 8 reveals a positive skew in relation to years worked at Globex when compared to the histogram, revealing a large amount of employees not having worked at the company for very long. Moreover, the

8

distribution for Years at Company can be considered leptokurtic considering kurtosis>3, indicating heavy tails and a large probability of employees having worked at the company for more or less than the sample quintiles. This may allude to high staff turnover, and high dissatisfaction.

Recommendations

The findings would suggest job dissatisfaction across a range of areas at Globex resulting in attrition, including the wage structure, overtime hours, and a cultural imbalance at the company. Regarding the wage structure of the company, it is indicated that there is no linear or hierarchal fashion in which wages are calculated given the large variability amongst each job role. Furthermore, the data suggest a weak correlation between how employees are valued and their net monthly income, resulting in a lack of clarity and incentivising employees to leave. It is recommended that Globex adopts a hierarchal wage structure in which each role has a base salary and fixed yearly salary increases, to attract prospective employees and retain current ones.

The second factor relates to the level of attrition regarding overtime workers. It is revealed that of those who work overtime, there are higher proportions of attrition, particularly in the lowest-paying jobs where long hours are compounded on top of low salaries. It is recommended that a cap be placed on the number of hours worked by all employees to ensure their well-being.

Finally, the data suggests the median years working at Globex to be 5 years with kurtosis>3, indicating heavy tails and high staff turnover, resulting in a cultural imbalance. It is recommended that Globex work on retaining staff by creating a better work environment and culture, in

combination with a salary re-structure and cap on hours worked to increase job satisfaction.

Reference List

McNeese, B., 2016. *Are the Skewness and Kurtosis Useful Statistics?*. [online] SPC For Excel. Available at: <https://www.spcforexcel.com/knowledge/basic-statistics/are-skewness-and-kurtosis-useful-statistics> [Accessed 10 March 2022].

Villegas, A., 2022, *Week 3: Data Exploration & Visualisation II*, lecture notes, Data, Insights and Decisions COMM1190, University of New South Wales, delivered 1 March 2022.

Wilson, C., 2016. *Mean vs. median: What do they mean and when do you use them?*. [online] Michigan State University. Available at: <https://www.canr.msu.edu/news/mean_vs._median_what_do_they_mean_and_when_do_you_use_them> [Accessed 10 March 2022].

Appendix

```
1   #Extracting the data
2   download.file("https://raw.githubusercontent.com/dat-analytics/data_test/main/5359323.csv","5359323.csv")
3   Globex<-read.csv("5359323.csv")
4   Globex
5
6   #Extracting JobRole
7   JobRole<-Globex$JobRole
8
9   #Figure 1: Piechart Composition of Jobs
10  sum(Globex$JobRole=="Healthcare Representative")
11  sum(Globex$JobRole=="Human Resources")
12  sum(Globex$JobRole=="Laboratory Technician")
13  sum(Globex$JobRole=="Manager")
14  sum(Globex$JobRole=="Manufacturing Director")
15  sum(Globex$JobRole=="Research Director")
16  sum(Globex$JobRole=="Research Scientist")
17  sum(Globex$JobRole=="Sales Executive")
18  sum(Globex$JobRole=="Sales Representative")
19  slice<-c(43,16,53,31,99,105,21,73,19)
20  JobRoleLabel<-c("Healthcare Representative","Human Resources","Manufacturing Director","Research Director",
21                  "Research Scientist","Sales Executive","Sales Representative","Laboratory Technician","Manager")
22  percentage<-round(slice/sum(slice)*100)
23  cpercentage<-paste(JobRoleLabel,": ",percentage,"%",sep = "")
24  pie(slice,labels = cpercentage,col = cm.colors(length(cpercentage)))
25  title(main = "Figure 1: Composition of Jobs at Globex Pharma")
26
27  #Figure 2: Bar plot Rate of Attrition
28  JobAttrition<-table(Globex$Attrition,JobRole)
29  JobAttrition
30  par(mar=c(11,8,3,6),cex=0.7)
31  barplot(JobAttrition,beside = T,ylim = c(0,100),col=c("gray25","gray"),las=2)
32  title(xlab = "Job Role",line = 10,cex.lab=1.5)
33  title(ylab = "Number of Employees",line=4,cex.lab=1.5)
34  title(main = "Figure 2: Rate of Attrition",cex=1.5)
35  legend("topright",inset=.02,legend = c("No","Yes"),col = c("gray25","gray"),fill = c("gray25","gray"),
36         horiz = TRUE,cex = 0.8)
37
38  #Figure 3: Box plot Monthly Income
39  MonthlyIncome<-Globex$MonthlyIncome
40  par(mar=c(11,8,3,6),cex=0.70)
41  boxplot(MonthlyIncome~JobRole,col=cm.colors(length(JobRoleLabel)),xlab = "",ylab = "",
42         main="Figure 3: Monthly Income",cex.main=1.5,las=2,data = Globex)
43  title(xlab="Job Role",line=10,cex.lab=1.5)
44  title(ylab="Monthly Income",line = 4,cex.lab=1.5)
```

```r
45  |
46  #Median Incomes
47  median(MonthlyIncome[Globex$JobRole=="Research Scientist"])
48  median(MonthlyIncome[Globex$JobRole=="Sales Executive"])
49  median(MonthlyIncome[Globex$JobRole=="Manager"])
50  median(MonthlyIncome[Globex$JobRole=="Research Director"])
51  median(MonthlyIncome[Globex$JobRole=="Sales Representative"])
52  median(MonthlyIncome[Globex$JobRole=="Laboratory Technician"])
53  median(MonthlyIncome[Globex$JobRole=="Human Resources"])
54  median(MonthlyIncome[Globex$OverTime=="Yes"])
55  median(MonthlyIncome[Globex$OverTime=="No"])
56  median(MonthlyIncome)
57
58  #Figure 4: Scatterplot Monthly Rate vs Monthly Income
59  MonthlyRate<-Globex$MonthlyRate
60  MonthlyRate
61  cor(MonthlyIncome,MonthlyRate)
62  plot(MonthlyIncome,MonthlyRate,xlab = "Monthly Rate",ylab = "Monthly Income",
63      main = "Figure 4: Monthly Rate vs Monthly Income",cex.main=1.2,col="brown1")
64  abline(lm(Globex$MonthlyRate~Globex$MonthlyIncome),lwd=2.5)
65
66  #Figure 5: Box plot Overtime vs Monthly Income
67  Overtime<-Globex$OverTime
68  boxplot(MonthlyIncome~Overtime,col=c("gray","gray25"),xlab = "Overtime",ylab = "Monthly Income",
69         main="Figure 5: Overtime Worked vs Monthly Income",cex.main=1.5,data = Globex)
70
71  #Figure 6: Bar plot Overtime vs Job Role
72  JobOvertime<-table(Globex$OverTime,JobRole)
73  JobOvertime
74  par(mar=c(11,8,3,6),cex=0.7)
75  barplot(JobOvertime,beside = T,ylim = c(0,90),col = c("gray25","gray"),las=2)
76  title(xlab = "Job Role",line=10,cex.lab=1.5)
77  title(ylab = "Number of Employees",line = 4,cex.lab=1.5)
78  title(main="Figure 6: Worked Overtime vs Job Role",cex=1.5)
79  legend("topright",inset=.02,legend = c("Worked Overtime","Did not Work Overtime"),
80         col = c("gray","gray25"),fill=c("gray","gray25"),horiz = TRUE,cex=0.8)
81
82  #Figure 7: Bar plot Worked Overtime vs Attrition
83  WorkedOvertime<-subset(Globex,subset = Globex$OverTime=="Yes")
84  WorkedOvertime
85  OvertimeJobRoles<-table(WorkedOvertime$Attrition,WorkedOvertime$JobRole)
86  Overtime
87  par(mar=c(11,8,3,6),cex=0.7)
88  barplot(OvertimeJobRoles,beside = T,ylim = c(0,20),col = c("gray25","gray"),las=2)
89  title(xlab = "Job Role",line = 10,cex.lab=1.5)
```

```r
90  title(ylab = "Number of Employees",line = 4,cex.lab=1.5)
91  title(main = "Figure 7: Attrition Rate for Overtime Workers",cex=1.5)
92  legend("topright",inset = .02,legend = c("No","Yes"),col = c("gray25","gray"),
93         fill = c("gray25","gray"),horiz = TRUE,cex = 0.8)
94
95  #Descriptive Stats - Years At Company
96  YearsAtCompany<-Globex$YearsAtCompany
97  mean(YearsAtCompany)
98  median(YearsAtCompany)
99  kurtosis(YearsAtCompany)
100 skewness(YearsAtCompany)
101
102 #Figure 8: Histogram vs PDF - Years at Company
103 mu<-mean(YearsAtCompany)
104 sigma<-sd(YearsAtCompany)
105 hist<-hist(YearsAtCompany,breaks = 100,prob=TRUE,ann=FALSE)
106 title(xlab = "Years at Company",ylab = "Histogram/Density",
107       main = "Figure 8: Histogram vs Normal PDF - Years At Company")
108 x<-hist$mids
109 curve(dnorm(x,mean = mu,sd=sigma),add=TRUE,col="red")
110
111 #Figure 9: Q-Q Plot - Years at Company
112 options(repr.plot.width = 4,repr.plot.height = 3)
113 par(mar=c(5,5,4,5),cex=0.65)
114 qqnorm(YearsAtCompany,ann=FALSE)
115 qqline(YearsAtCompany,col="red")
116 title(xlab = "Theoretrical quantiles",ylab = "Sample quantiles",
117       main = "Figure 9: QQ Plot for Years At Company against Normal Quantiles")
```