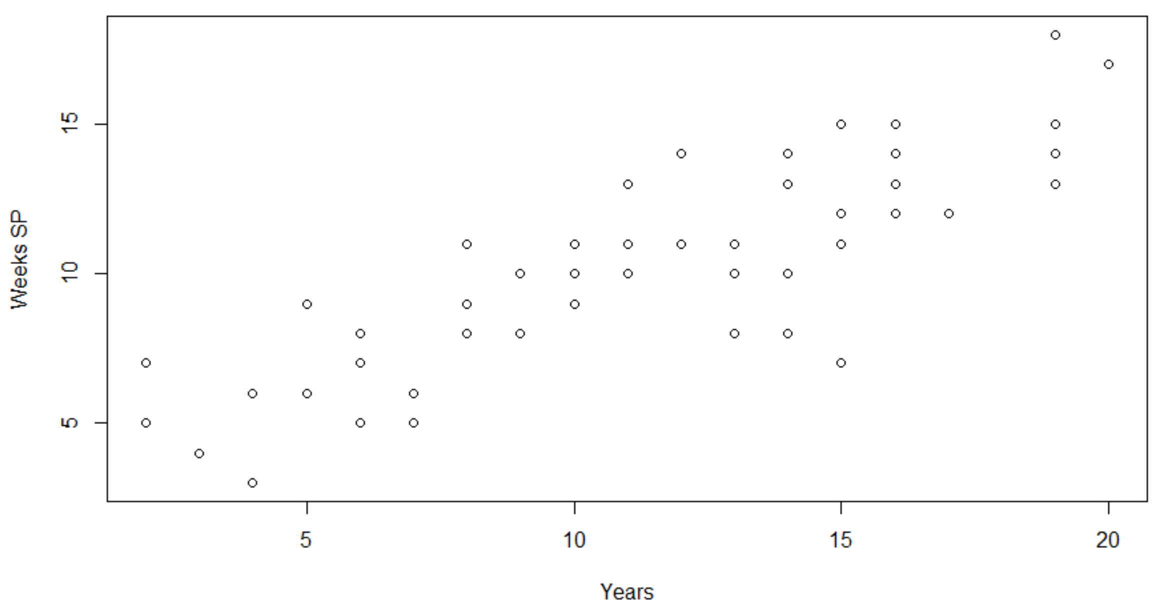
Mansi Borkar: CWID - 887480267

Case Study: Severance Pay

Severance pay is usually based on the length of employment for which an employee is eligible upon termination. Merging of two companies often leads to a reduction of the labor force from the two organizations. Because there is no hard and fast rule when it comes to severance packages, Laurier company, who recently acquired the Western Company, terminated 20 of the Western's employees and offered them severance packages equivalent to ex-Laurier employees. Bill Smith, an employee of Western Company, who was among the terminated, received five weeks of severance pay, is opposing his severance package as it is less when compared to the standard. To find out whether Severance Package received by Bill Smith is justified using data and statistical analysis, a statistician was brought in to verify his claim. The software used is R.

The statistician took a random sample of 50 ex-Laurier and considered different parameters:

The number of weeks of severance pay, Number of years with the company. The below Scatter Plot using data analysis shows a positive correlation between several weeks of severance pay and a number of years with the company.



a

Regression Model is a perfect fit for the current scenario that is to determine any underlying relations in the data and verify the claim made by Bill Smith as it involves one independent variable and one dependent variable. The regression analysis in R is used to predict the outcome of dependent variable y based on independent variable x. The goal is to build a mathematical model that defines y as a function of the variable x. The mathematical formula of the linear regression can be written as y = b0 + b1*x + e, where:

- b0 and b1 are known as the regression beta coefficients or parameters:
    - b0 is the intercept of the regression line; that is the predicted value when x = 0.

  - b1 is the slope of the regression line.
- e is the error term (also known as the residual errors), the part of y that can be explained by the regression

The estimated value of these parameters is then used to conclude.
Here the independent variable is a number of years with the company, and a dependent variable is a number of weeks of severance pay as severance pay depends on the number of years served. The relationship between the two is approximated by a straight line, as seen by the Scatter Plot above.

The mean and standard variation of the sample data is calculated. The critical value is calculated at a 5 percent significance level. The standard error is calculated for a number of weeks based on the count of entries in the sample data. The sum of the squares of the residual errors is called the Residual Sum of Squares. We infer from these values of R-squared that the model fits the data to an acceptable degree.

Using the result from the linear regression model and R calculations, we can conclude that Bill's claim is justified i.e. his offer of severance pay that includes 5 weeks of severance pay is less than what is offered to Laurier's employees when they were laid off, in contravention of the buyout agreement.

Statistical Calculations:

Mean Severance Pay	10.26 weeks
Standard Deviation of Severance Pay	3.41 weeks
Confidence Interval	8.79 to 9.94 weeks
Prediction Interval	5.47 to 13.26 weeks
Multiple R-squared	0.6903
Adjusted R-squared	0.6839

Appendix

#Reading csv file

```
dat1<- read.csv(file= "C:/Users/Mansi/Downloads/CaseStudyDataset.csv", header=T, sep=",")
dat1
head(dat1)
tail(dat1)
```

 # Scatter plot graph

```
plot(dat1$Years, dat1$Weeks_SP, xlab="Years", ylab="Weeks SP")
```

#Mean of Weeks and Years

```
mean_weeks <-mean(dat1$Weeks_SP)
mean_weeks
mean_years <-mean(dat1$Years)
mean_years
```

#Standard Deviation of Weeks and Years

```r
std_dev_weeks<- sd(dat1$Weeks_SP)
std_dev_weeks

std_dev_years<- sd(dat1$Years)
std_dev_years


fit<- lm(Weeks_SP ~ Years, data=dat1)
summary(fit)

# Calculating critical value and interval (Lower_Limit & Upper Limit)
n <- dim(dat1)[1]
CV <- qt(1-0.05/2, n-2)
L<- mean_weeks- CV*(std_dev_weeks/sqrt(n))
round(L,2)
U<- mean_weeks+ CV*(std_dev_weeks/sqrt(n))
round(U,2)

#Confidence Interval and Prediction Interval
newdata1<- data.frame(Years=10)

pred.clim1 <- predict(fit, newdata1, interval="confidence")
pred.clim1

pred.plim1 <- predict(fit, newdata1, interval="prediction")
pred.plim1
```