

## Simple Linear Regression in R

### Example 1 (Simple Linear Regression):

Fit a straight line to  $Y = (\text{number of}) \text{ Jobs}$  as a function of  $X = \text{GPA}$  to the following data:

GPA	Jobs
3.5	24
3.3	23
2.6	18
2.8	20
3	22
2.3	15
2.4	17
2.7	17
3.2	22
3.5	25
2.9	21
3.5	25
2.2	13
3.8	27
2.7	18

Open the data file JOB\_GPA.csv in R.

```
d1<-read.csv("g:/Stats24x7/R/JOB_GPA.csv",header=TRUE)
> attach(d1)
> names(d1)
[1] "GPA" "Jobs"
```

Read the datafile JOB\_GPA.csv  
in R.  
Attach the dataframe d2.  
  
Look at the variable names.

```
> l1 <- lm(Jobs~GPA, data = d1)
```

Fit Job = a + b GPA in R.

```
> l1
```

OUTPUT from R.  
R created an object l1.

Call:

```
lm(formula = Jobs ~ GPA, data = d1)
```

Coefficients:

(Intercept)	GPA
-3.452	8.081

## Simple Linear Regression in R

```
> summary(l1)
```

Call:

```
lm(formula = Jobs ~ GPA, data = d1)
```

Residuals:

```
   Min     1Q  Median     3Q      Max
-1.3657 -0.3859 -0.1335  0.6343  1.2101
```

Coefficients:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.4518     1.3796  -2.502  0.0265 *
GPA           8.0806     0.4602  17.558 1.95e-10 ***
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8456 on 13 degrees of freedom  
Multiple R-squared: 0.9595, Adjusted R-squared: 0.9564  
F-statistic: 308.3 on 1 and 13 DF, p-value: 1.951e-10

```
> names(l1)
```

```
[1] "coefficients" "residuals" "effects" "rank"
[5] "fitted.values" "assign" "qr" "df.residual"
[9] "xlevels" "call" "terms" "model"
```

```
> qqnorm(l1$residuals)
```

```
> qqline(l1$residuals)
```

OUTPUT from R.

Names of the object l1:  
Assess normality of residuals  
by drawing Q-Q plot in R.

