

Linear regression with r using income data

Data: csv file

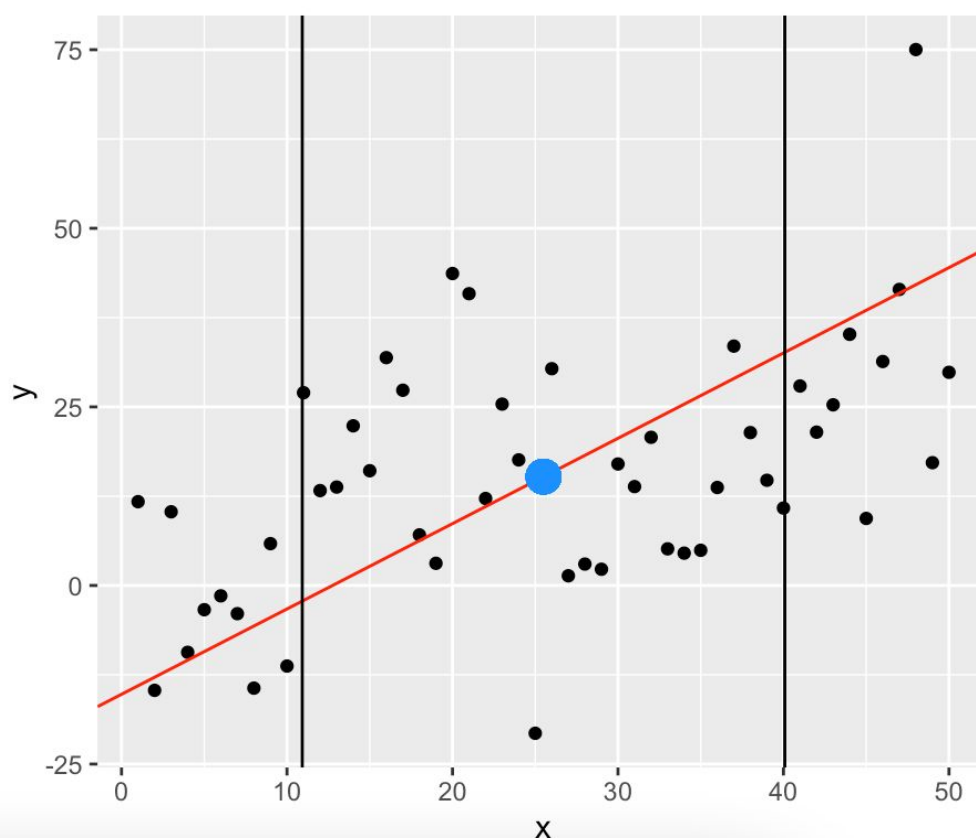
R studio console

```
Console Terminal x Jobs x
~/
> library(readr)
> income_data_csv_income_data_csv <- read_csv("Desktop/income.data.csv - income.data.csv.csv")

— Column specification —
cols(
  X1 = col_double(),
  income = col_double(),
  happiness = col_double()
)

Warning message:
Missing column names filled in: 'X1' [1]
> View(income_data_csv_income_data_csv)
> library(ggplot2)
> library(dplyr)
> library(magrittr)
> dat <- data.frame(x=x, y=y)
>
> p.sd <- ggplot(dat %>% mutate(slope.sd=sd(y)/sd(x),
+                               icept.sd= mean(y)- (sd(y)/sd(x))*mean(x)),
+               aes(x, y)) +
+   geom_point() +
+   geom_abline(aes(intercept=icept.sd, slope=slope.sd), col="red")+
+   geom_vline(xintercept = mean(x)+sd(x))+
+   geom_vline(xintercept = mean(x)-sd(x))+
+   geom_point(aes(x=mean(x), y=mean(y)), colour="dodgerblue", size=5)+
+   ggtitle("SD Line for simulated linear data with r = 0.5")
>
> p.sd
> |
```

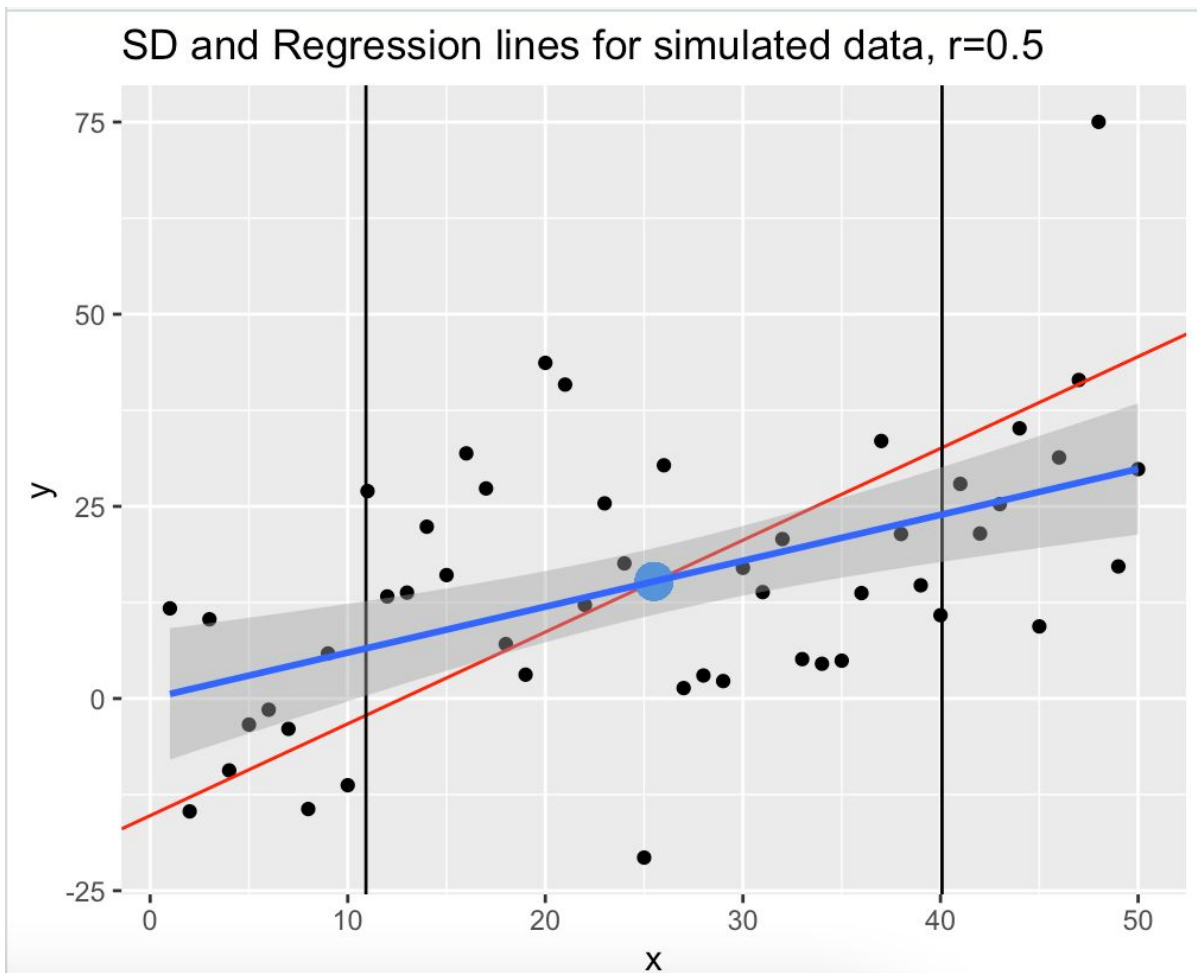
SD Line for simulated linear data with r = 0.5



```

> p.r <- p.sd + geom_smooth(method='lm', formula=y~x) + ggtitle("SD and Regression lines for
simulated data, r=0.5")
> p.r
>

```



```

> x <- 0:50
> r <- seq(0, 1, by=0.05)
>
> y.vals <- unlist(purrr::map(r, ~gen.y(x, .)))
> x.vals <- rep(x, length(r))
> r.vals <- rep(r, each=length(x))
>
>
> data <- data.frame(x.vals=x.vals, y.vals=y.vals, r.vals=r.vals)
> head(data)
  x.vals  y.vals r.vals
1      0 14.286094     0
2      1 11.762617     0
3      2 13.736524     0
4      3 11.510834     0
5      4  4.608008     0
6      5 -17.264810     0

```

```

> p.set <- ggplot( data %>% group_by(r.vals) %>% mutate(slope.sd=sd(y.vals)/sd(x.vals),
+                                                     icept.sd= mean(y.vals)- (sd(y.vals)/sd(x.vals))*mean(x.vals),
+                                                     slope.r=r.vals*sd(y.vals)/sd(x.vals),
+                                                     icept.r= mean(y.vals)- (r.vals*sd(y.vals)/sd(x.vals))*mean(x.vals))
+                                                     , aes(x.vals, y.vals)) +
+   geom_point() +
+   geom_abline(aes(intercept=icept.sd, slope=slope.sd), col="red") +
+   geom_abline(aes(intercept=icept.r, slope=slope.r), col="blue")
>
>
> p.set
>

```

