Lsn 15

Clark

Admin

Let's reconsider the Salary Discrimination dataset

```
salary.dat<-read.table("http://www.isi-stats.com/isi2/data/Wages.txt",header=T)
```

Instead of looking at College educated vs not college educated, we now consider the full dataset.

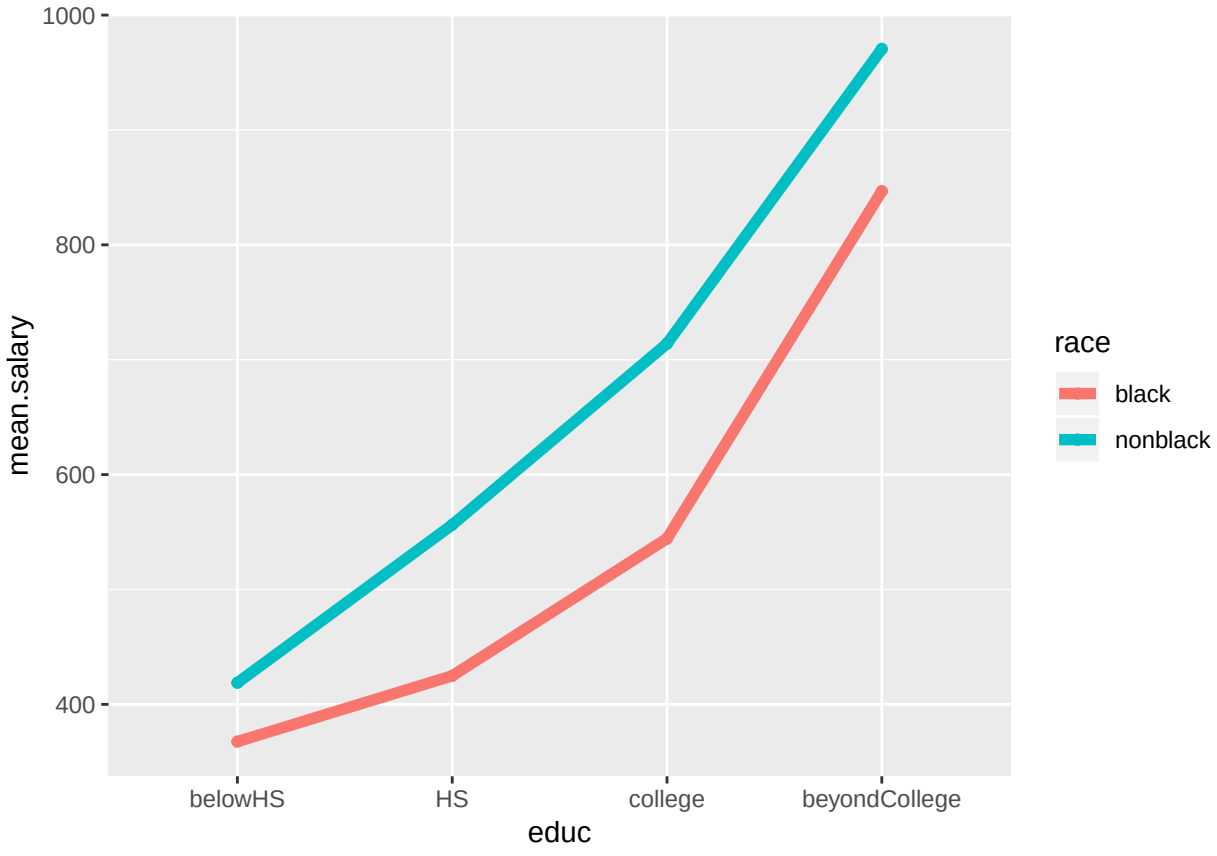
```
levels(salary.dat$educ)
```

```
## [1] "belowHS"       "beyondCollege" "college"       "HS"
```

```
gr.means=salary.dat%>%group_by(educ,race)%>%summarize(mean.salary=mean(wage))
```

What do we see?

```
gr.means$educ<-factor(gr.means$educ,levels=c("belowHS","HS","college","beyondCollege"))
gr.means %>% ggplot(aes(x=educ,y=mean.salary,color=race))+
  geom_line(aes(group=race),lwd=2)+geom_point()
```



A statistical model:

ANOVA table:

To fit the model we use:

```
contrasts(salary.dat$race)=contr.sum
contrasts(salary.dat$educ)=contr.sum
inter.lm<-lm(wage~race*educ,data=salary.dat)
coef(inter.lm)
```

```
## (Intercept)       race1       educ1       educ2       educ3 race1:educ1
##  605.387773  -59.514079 -212.189614  303.256089   23.755534   33.916246
## race1:educ2 race1:educ3
##   -2.381393  -25.316463
```

Getting the fits is a bit of a pain but we can do it:

To fit the ANOVA model we note that we are now interested in Type III Sums of squares.

```
library(car)
```

```
## Warning: package 'car' was built under R version 3.5.1
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode

## The following object is masked from 'package:purrr':
##
##     some
```

```
Anova(inter.lm,type=3)
```

```
## Anova Table (Type III tests)
##
## Response: wage
##                 Sum Sq    Df   F value    Pr(>F)
## (Intercept) 1213063109     1 6926.7761 < 2.2e-16 ***
## race          11723453     1   66.9427 2.924e-16 ***
## educ          97175833     3  184.9630 < 2.2e-16 ***
## race:educ      1641649     3    3.1247   0.02473 *
## Residuals   4487270185 25623
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

What does this mean?