

Math130Project

Oscar Diaz-Garcia

2022-09-23

INTRODUCTION

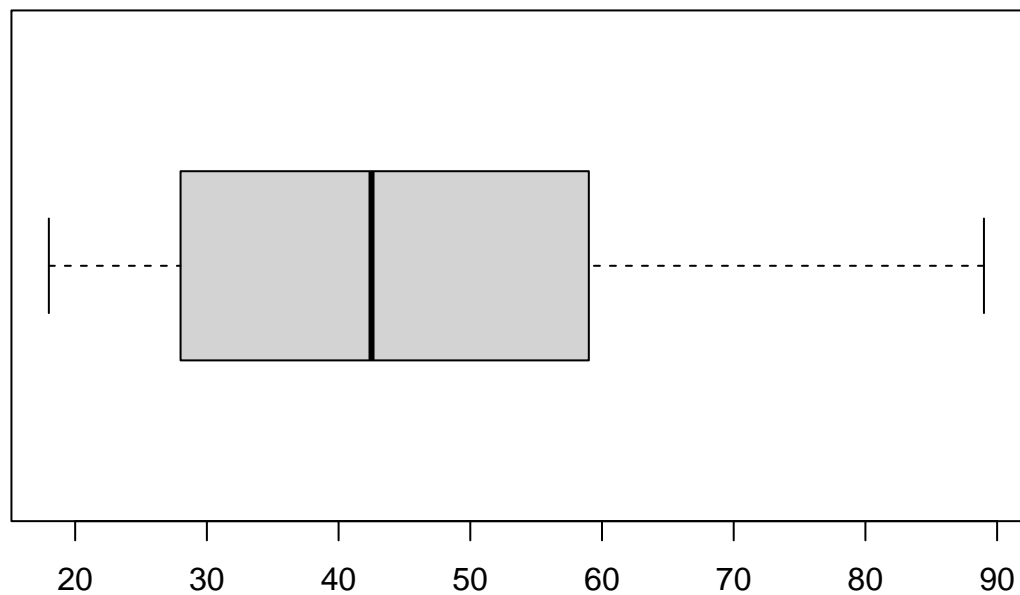
I will be analyzing the Depression data set. It consists of 294 observations and 37 variables. The variables I chose to explore are age and education level. Is there correlation between age and education level in the data set?

UNIVARIANT ANALYSIS

```
summary(depress$age)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	18.00	28.00	42.50	44.41	59.00	89.00

Age Range of Individuals

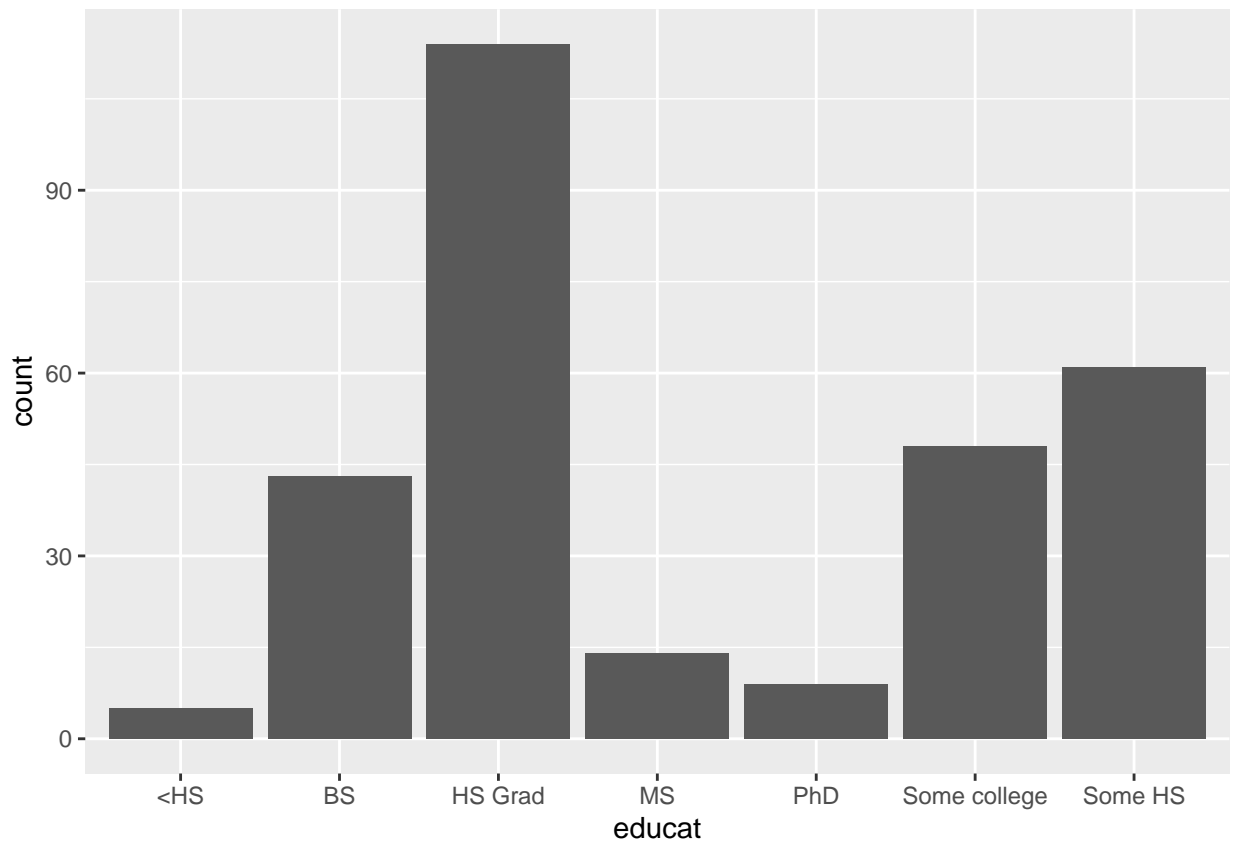


This boxplot displays that the youngest individual being 18 while the oldest being 89. The mean age is 44.

```
table(depress$educat)
```

```
##  
##      <HS      BS      HS Grad      MS      PhD Some college  
##       5      43      114      14      9      48  
##   Some HS  
##      61
```

```
ggplot(depress, aes(x=educat)) + geom_bar()
```

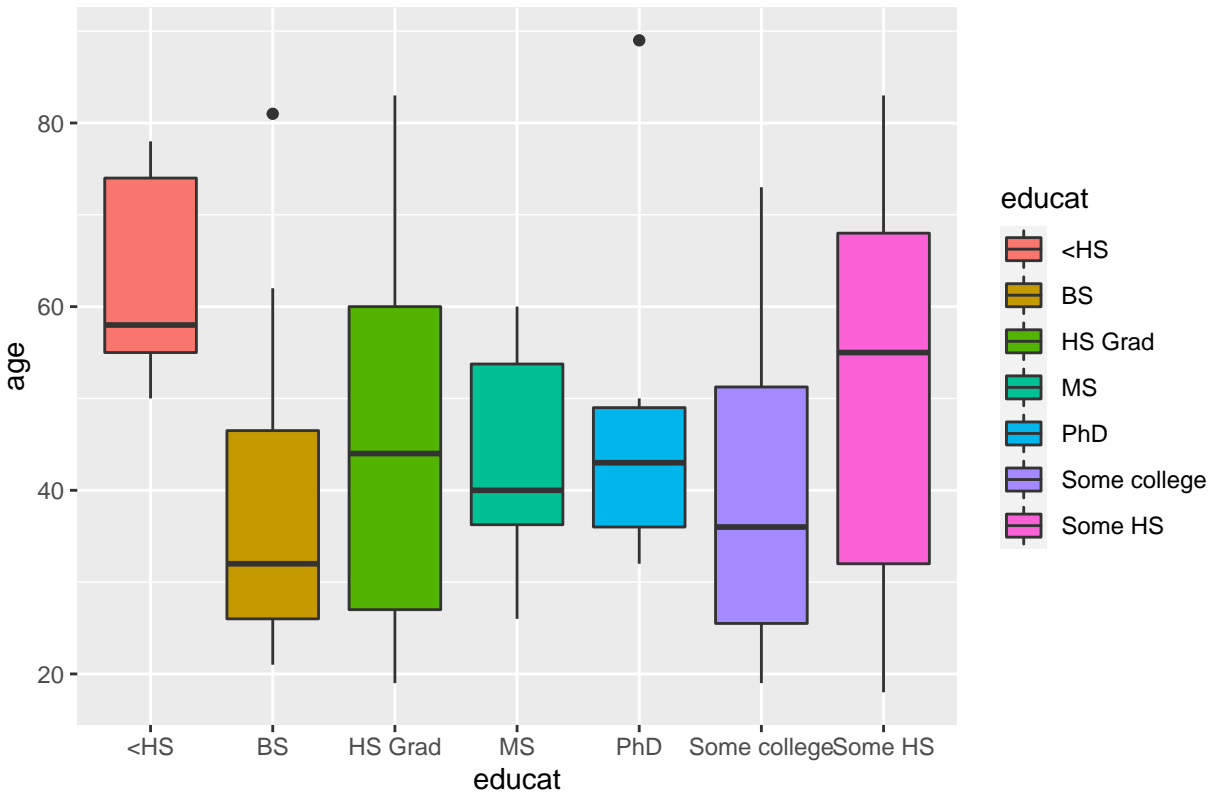


BIVARIANT ANALYSIS

Age vs Education

```
ggplot(depress, aes(x=educat, y=age, fill=educat)) + geom_boxplot() + ggtitle("Age vs. Education")
```

Age vs. Education



```
round(prop.table(table(depress$age, depress$educat), margin=1),3)
```

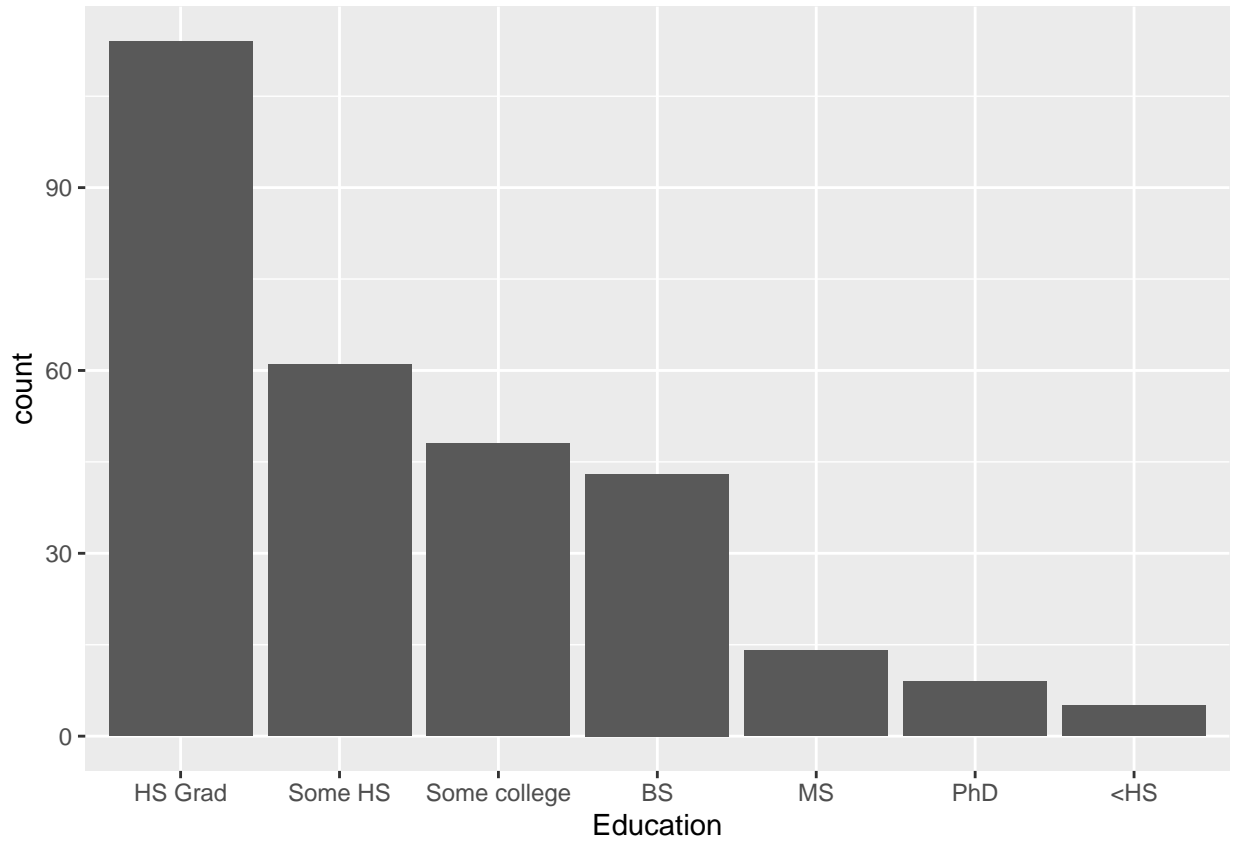
```
##
##      <HS  BS HS Grad  MS  PhD Some college Some HS
## 18 0.000 0.000  0.000 0.000 0.000      0.000  1.000
## 19 0.000 0.000  0.400 0.000 0.000      0.400  0.200
## 20 0.000 0.000  0.500 0.000 0.000      0.333  0.167
## 21 0.000 0.333  0.333 0.000 0.000      0.167  0.167
## 22 0.000 0.111  0.667 0.000 0.000      0.111  0.111
## 23 0.000 0.182  0.273 0.000 0.000      0.455  0.091
## 24 0.000 0.333  0.444 0.000 0.000      0.111  0.111
## 25 0.000 0.000  0.833 0.000 0.000      0.000  0.167
## 26 0.000 0.556  0.222 0.111 0.000      0.111  0.000
## 27 0.000 0.000  0.750 0.000 0.000      0.250  0.000
## 28 0.000 0.400  0.200 0.200 0.000      0.000  0.200
## 29 0.000 0.500  0.250 0.000 0.000      0.250  0.000
## 30 0.000 0.167  0.333 0.000 0.000      0.333  0.167
## 31 0.000 0.000  0.600 0.000 0.000      0.200  0.200
## 32 0.000 0.500  0.200 0.000 0.100      0.000  0.200
## 33 0.000 0.200  0.200 0.000 0.000      0.400  0.200
## 34 0.000 0.000  0.333 0.111 0.000      0.333  0.222
## 35 0.000 0.000  0.500 0.000 0.000      0.250  0.250
## 36 0.000 0.167  0.333 0.167 0.333      0.000  0.000
## 37 0.000 0.000  0.000 0.400 0.000      0.600  0.000
## 38 0.000 0.000  1.000 0.000 0.000      0.000  0.000
```

```

## 39 0.000 0.000 0.000 0.000 0.000 0.000 1.000
## 40 0.000 0.000 0.400 0.400 0.200 0.000 0.000
## 41 0.000 1.000 0.000 0.000 0.000 0.000 0.000
## 42 0.000 0.222 0.444 0.000 0.000 0.333 0.000
## 43 0.000 0.143 0.143 0.143 0.286 0.286 0.000
## 44 0.000 0.000 1.000 0.000 0.000 0.000 0.000
## 45 0.000 0.500 0.250 0.000 0.000 0.250 0.000
## 46 0.000 0.333 0.333 0.000 0.000 0.333 0.000
## 47 0.000 0.333 0.167 0.167 0.000 0.167 0.167
## 48 0.000 0.250 0.500 0.000 0.000 0.000 0.250
## 49 0.000 0.000 0.750 0.000 0.250 0.000 0.000
## 50 0.200 0.200 0.400 0.000 0.200 0.000 0.000
## 51 0.000 0.167 0.333 0.000 0.000 0.167 0.333
## 52 0.000 0.250 0.000 0.000 0.000 0.250 0.500
## 53 0.000 0.000 0.000 0.000 0.000 0.500 0.500
## 54 0.000 0.000 0.333 0.000 0.000 0.333 0.333
## 55 0.250 0.250 0.250 0.000 0.000 0.000 0.250
## 56 0.000 0.000 0.333 0.333 0.000 0.333 0.000
## 57 0.000 0.167 0.833 0.000 0.000 0.000 0.000
## 58 0.143 0.000 0.429 0.000 0.000 0.143 0.286
## 59 0.000 0.000 0.222 0.222 0.000 0.111 0.444
## 60 0.000 0.000 0.429 0.143 0.000 0.143 0.286
## 61 0.000 0.200 0.600 0.000 0.000 0.000 0.200
## 62 0.000 0.250 0.750 0.000 0.000 0.000 0.000
## 63 0.000 0.000 1.000 0.000 0.000 0.000 0.000
## 64 0.000 0.000 0.333 0.000 0.000 0.333 0.333
## 65 0.000 0.000 0.400 0.000 0.000 0.200 0.400
## 66 0.000 0.000 0.667 0.000 0.000 0.000 0.333
## 67 0.000 0.000 0.500 0.000 0.000 0.000 0.500
## 68 0.000 0.000 0.750 0.000 0.000 0.000 0.250
## 69 0.000 0.000 1.000 0.000 0.000 0.000 0.000
## 70 0.000 0.000 0.400 0.000 0.000 0.200 0.400
## 71 0.000 0.000 0.333 0.000 0.000 0.333 0.333
## 72 0.000 0.000 0.500 0.000 0.000 0.000 0.500
## 73 0.000 0.000 0.000 0.000 0.000 0.500 0.500
## 74 0.333 0.000 0.333 0.000 0.000 0.000 0.333
## 75 0.000 0.000 0.000 0.000 0.000 0.000 1.000
## 77 0.000 0.000 0.500 0.000 0.000 0.000 0.500
## 78 0.500 0.000 0.500 0.000 0.000 0.000 0.000
## 79 0.000 0.000 0.500 0.000 0.000 0.000 0.500
## 80 0.000 0.000 0.000 0.000 0.000 0.000 1.000
## 81 0.000 0.500 0.000 0.000 0.000 0.000 0.500
## 82 0.000 0.000 0.000 0.000 0.000 0.000 1.000
## 83 0.000 0.000 0.400 0.000 0.000 0.000 0.600
## 89 0.000 0.000 0.000 0.000 1.000 0.000 0.000

```

```
ggplot(depress, aes(x=forcats::fct_infreq(educat))) + geom_bar() + xlab("Education")
```



CONCLUSION

I discovered that there is no correlation between age and education level. There is no significance between the two variables.