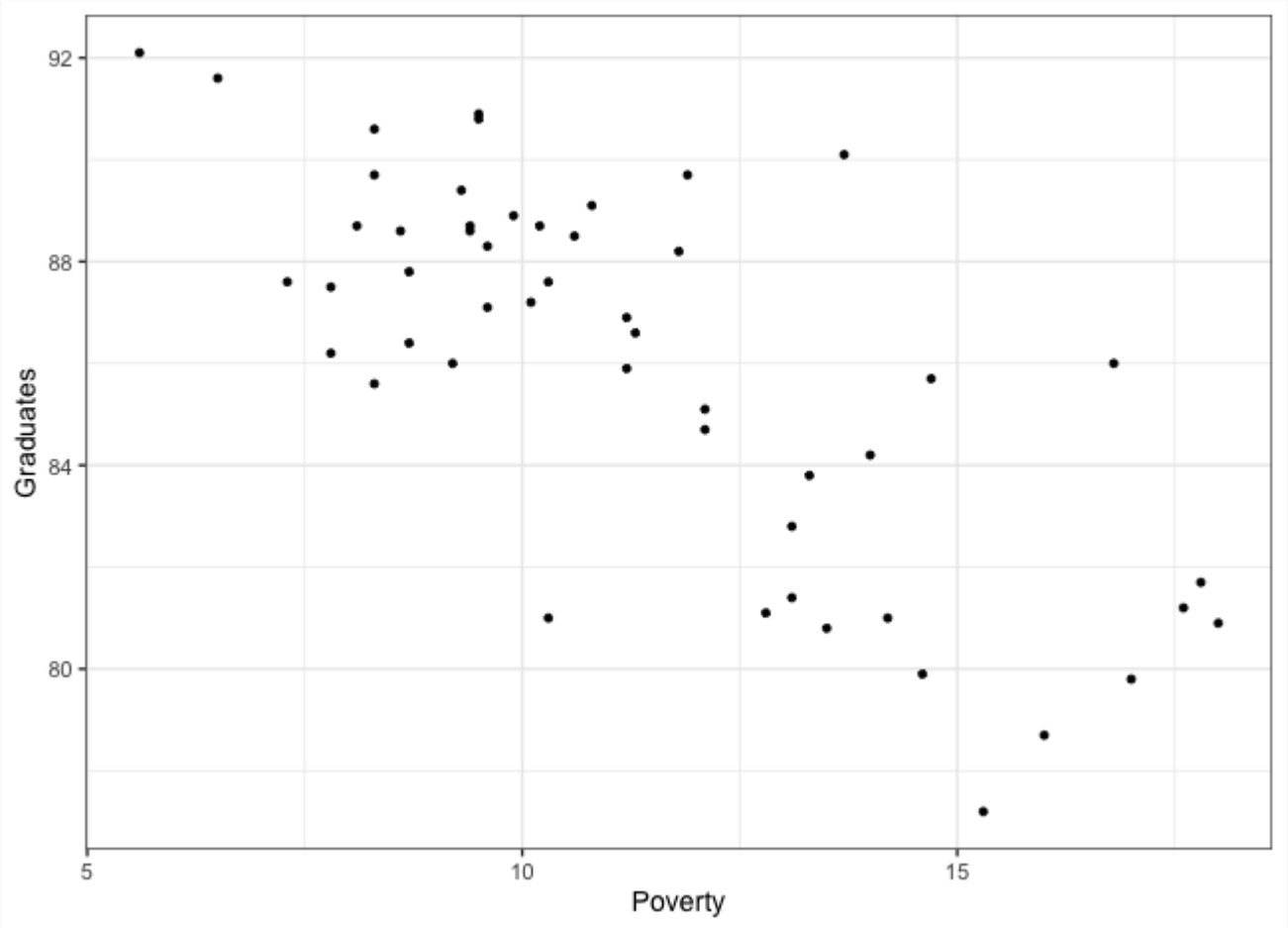
Simple Linear Regression

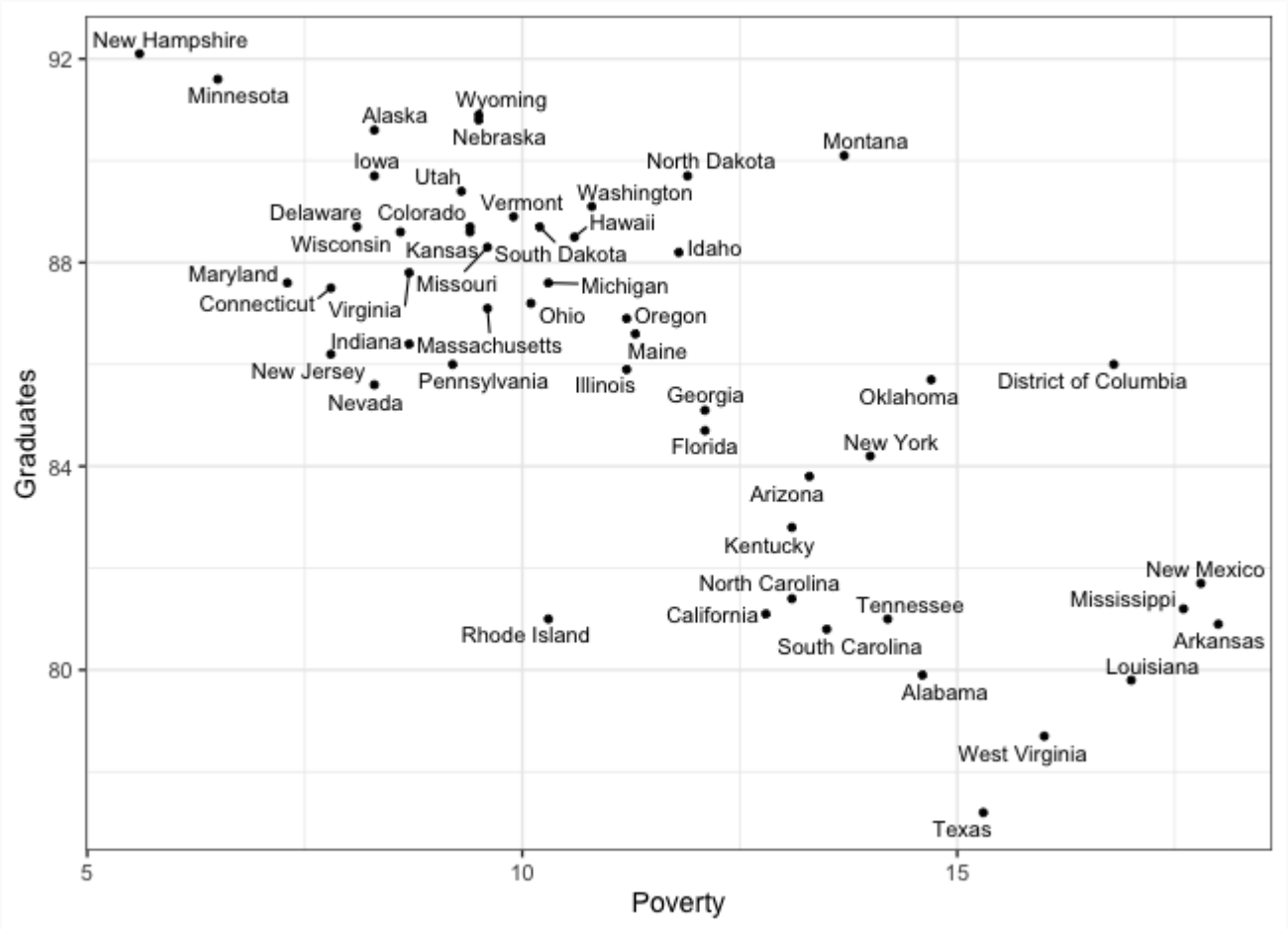
Poverty and Graduation

```
##          State Metropolitan.Residence White Graduates Poverty Percent
## 1      Alabama                    55.4  71.3      79.9    14.6    14.2
## 2       Alaska                    65.6  70.8      90.6     8.3    10.8
## 3      Arizona                    88.2  87.7      83.8    13.3    11.1
## 4     Arkansas                    52.5  81.0      80.9    18.0    12.1
## 5   California                    94.4  77.5      81.1    12.8    12.6
## 6     Colorado                    84.5  90.2      88.7     9.4     9.6
## 7  Connecticut                    87.7  85.4      87.5     7.8    12.1
## 8     Delaware                    80.1  76.3      88.7     8.1    13.1
```

Poverty and Graduation



Poverty and Graduation



Simple Linear Regression: $W_1 \sim W_2$

The Correlation Coefficient

The Correlation Coefficient in R

```
poverty %>%
  summarize(cor(Graduates, Poverty))
```

```
##   cor(Graduates, Poverty)
## 1                  -0.747
```

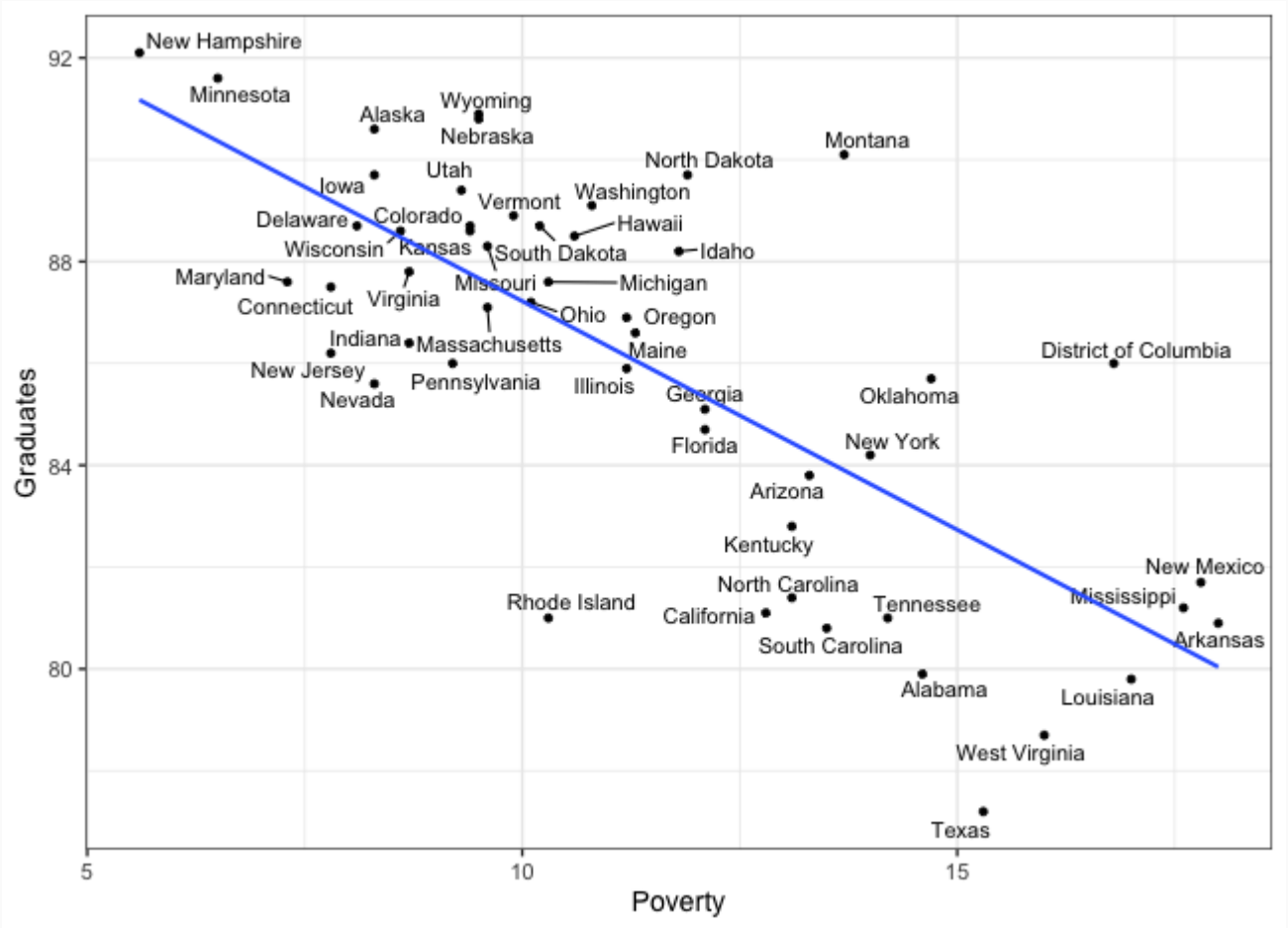
The Linear Model

The Linear Model in R

```
lm(Graduates ~ Poverty, data = poverty)
```

```
##
## Call:
## lm(formula = Graduates ~ Poverty, data = poverty)
##
## Coefficients:
## (Intercept)      Poverty
##      96.202       -0.898
```

Poverty and Graduation



Residuals

Residuals in R

```
m1 <- lm(Graduates ~ Poverty, data = poverty)
attributes(m1)
```

```
## $names
##  [1] "coefficients"  "residuals"     "effects"       "
##  [6] "assign"        "qr"            "df.residual"   "
## [11] "terms"         "model"
##
## $class
## [1] "lm"
```

```
m1$fitted
```

```
##    1    2    3    4    5    6    7    8    9   10   11
## 83.1 88.7 84.3 80.0 84.7 87.8 89.2 88.9 81.1 85.3 85.3
##   19   20   21   22   23   24   25   26   27   28   29
## 80.9 86.1 89.6 87.6 87.0 90.4 80.4 87.6 83.9 87.7 88.7
##   37   38   39   40   41   42   43   44   45   46   47
```

Residuals in R, cont.

```
m1$residuals
```

```
##       1       2       3       4       5       6
## -3.1927  1.8505 -0.4600  0.8602 -3.6089  0.9382 -1.698
##      12      13      14      15      16      17       1
##  1.8157  2.5932 -0.2456 -1.9903  0.9505  0.8382 -1.639
##      23      24      25      26      27      28       2
##  0.6463  1.2342  0.8011  0.7178  6.1992  3.1280 -3.149
##      34      35      36      37      38      39       4
## -3.0395  4.1830  0.0667  2.6971  0.7544 -1.9414 -5.953
##      45      46      47      48      49      50       5
##  1.5484  1.5871 -0.5903  2.5953 -3.1356  0.1199  3.228
```

Residual plot

```
m1 <- lm(Graduates ~ Poverty, data = poverty)
plot(m1, 1)
```