

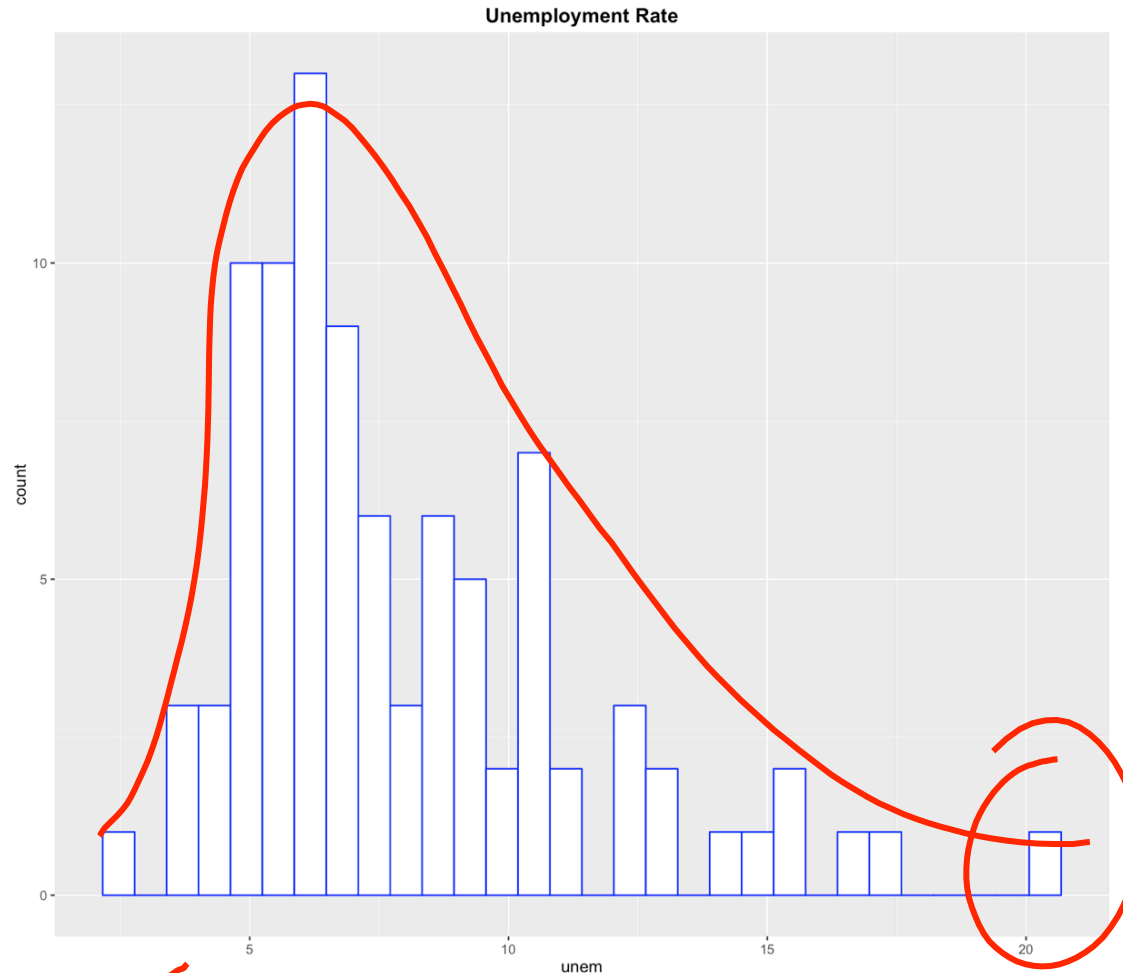
ANALYSIS OF PANEL DATA

An Introduction

datascience@berkeley

Using OLS Regression Model on Panel Data

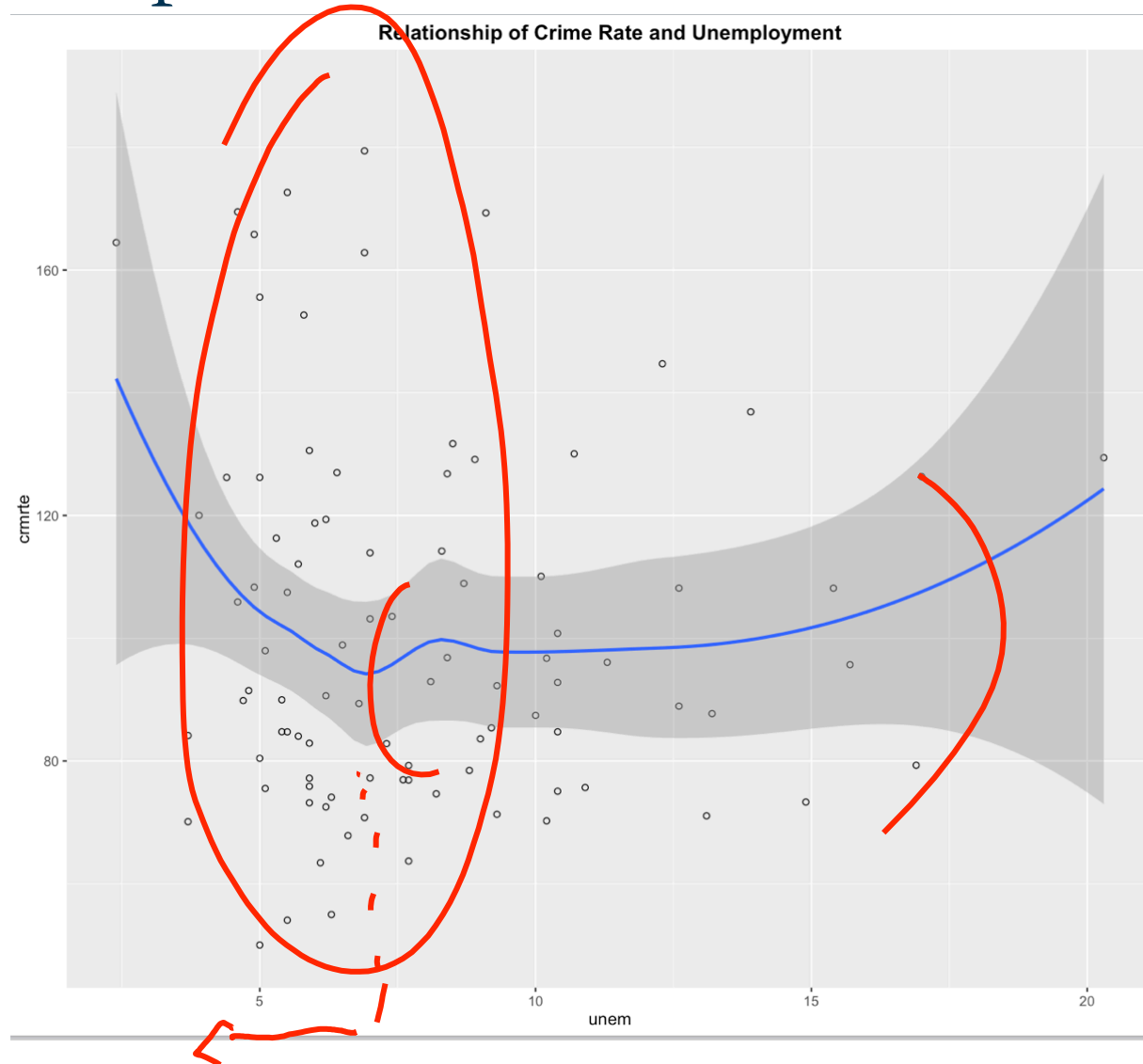
Histogram of the Variable of Interest: unem



```
> length(crime2$unem)
[1] 92
> summary(crime2$unem)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.400	5.500	6.950	7.972	9.475	20.300

Relationship Between **crmrate** and **unem**



OLS Regression

In fact, let's run a simple OLS regression of **crmrte** on **unem** using all of the observations in the dataset. Note that I didn't do much EDA before building the model, but I just want to use this model to illustrate a point later.

```
> ols.fit1 <- lm(crmrte ~ unem, data=crime2)
> summary(ols.fit1)
```

Call:
lm(formula = crmrte ~ unem, data = crime2)

Residuals:

Min	1Q	Median	3Q	Max
-51.686	-23.889	-7.961	17.522	78.297

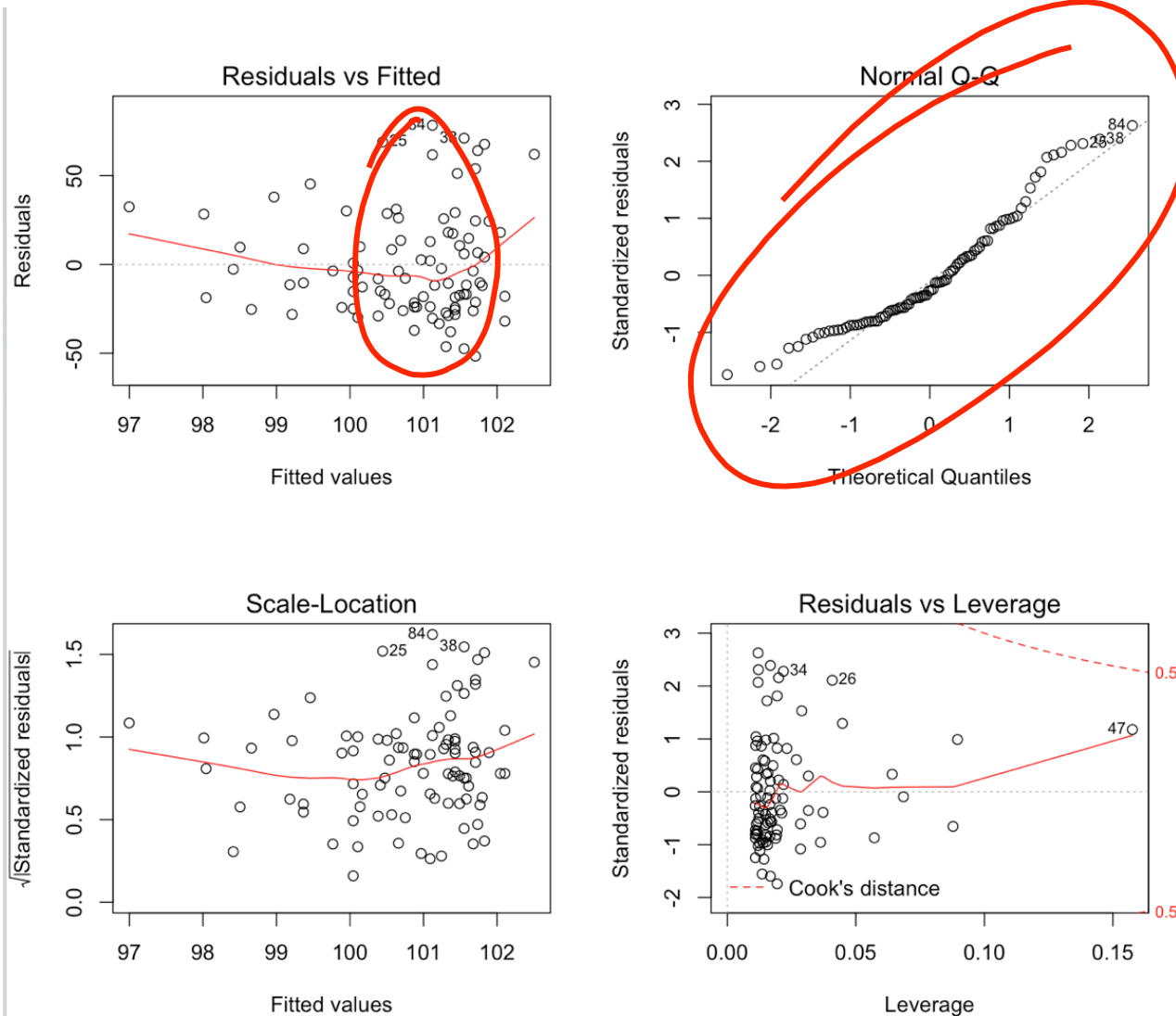
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	103.2434	8.0587	12.81	<2e-16 ***
unem	-0.3077	0.9317	-0.33	0.742

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

Residual standard error: 29.99 on 90 degrees of freedom
Multiple R-squared: 0.00121, Adjusted R-squared: -0.009888
F-statistic: 0.109 on 1 and 90 DF, p-value: 0.742

Residual Diagnostic Plots



Berkeley

SCHOOL OF
INFORMATION