

```

> # name: jblubau1_hw06_script.r
> # path: ~/Projects/learning/Statistics/STAT_604/Homework
> # created by: Joseph Blubaugh
> # created on: 24 Sept 2016
> # purpose: Homework 06
> # last ran:
> Sys.time()
[1] "2016-09-24 13:09:02 CDT"
>
> # 1) Clean the workspace
> ls(); rm(list = ls())
character(0)
>
> # 2) Load workspace from HW04
> load(file = "/home/jeston/Projects/learning/Statistics/STAT_604/Data/HW04.RData")
> ls()
[1] "Oklahoma"
>
> # 3) Output to pdf
> pdf(file = "/home/jeston/Projects/learning/Statistics/STAT_604/Homework/jblubau1_hw06_graph.pdf")
>
> # 4) Create histogram of Pupil/Teacher ratio
> ## a) Default histogram
> hist(Oklahoma$PTRatio, freq = FALSE, xlab = "Pupils/Teacher", main = "Pupil/Teacher Ratios in
Oklahoma Schools")
> ## b) Vector for breaks in increments of 5
> brks = seq(from = 0, to = 155, by = 5)
> ## c) Histogram with custom breaks
> hist(Oklahoma$PTRatio, breaks = brks, freq = FALSE,
+   xlab = "Pupils/Teacher", main = "Pupil/Teacher Ratios in Oklahoma Schools")
>
> # 5) Add maroon density line
> lines(density(na.omit(Oklahoma$PTRatio), bw = 5), col = "maroon")
>
> # 6) Add vertical line
> abline(v = mean(na.omit(Oklahoma$PTRatio)), col = 7)
>
> # 7) Scatterplot of Teacher to PTRatio
> plot(x = Oklahoma$Teachers, y = Oklahoma$PTRatio, pch = 3, col = "purple", xlim = c(0, 140),
+   xlab = "Teachers", ylab = "Pupil/Teacher Ratio")
>
> # 8) Add a fitted line
> abline(lm(PTRatio ~ Teachers, data = Oklahoma), col = "#FF9900")
> summary(lm(PTRatio ~ Teachers, data = Oklahoma))

```

Call:

```
lm(formula = PTRatio ~ Teachers, data = Oklahoma)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-14.627	-1.935	-0.015	1.666	140.506

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	14.383468	0.227431	63.243	< 2e-16 ***
Teachers	0.045986	0.007922	5.805	7.61e-09 ***

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

Residual standard error: 5.686 on 1778 degrees of freedom

(5 observations deleted due to missingness)

Multiple R-squared: 0.0186, Adjusted R-squared: 0.01805

F-statistic: 33.7 on 1 and 1778 DF, p-value: 7.612e-09

>

> # 9) Imbed the systemdate time

> text(x = 100, y = 150, labels = Sys.time())

>

> # 10) Make boxplots

> boxplot(x = Oklahoma[, 6:11], names = 7:12, col = "lightgreen", range = 0,

+ xlab = "Grades", ylab = "Students", main = "Tulsa County vs State")

>

> # 11) Add diamond points

> pts = colMeans(x = subset(Oklahoma, County == "TULSA COUNTY")[, 6:11], na.rm = TRUE)

> points(pts, pch = 23, col = "red", bg = "darkgreen", cex = 2)

>

> ## close graphics device

> dev.off()

RStudioGD