R and RStudio Cheatsheet

Preliminaries

Load package

```
library(hanoverbase)
```

Load Data

Data from package

```
data(counties)
View(counties)    # Only in console
help(counties)    # Only in console
```

Data Import

- Upload local file or use a URL
- File > Import > Dataset > From . . .
- Add import code to code chunk (exclude the View command)

Data Viewer

Order rows

- Click on column heading
- Click again to change direction

Filter

- Click Filter button on top left of viewer
- Use controls below column headings

Search

- Use the search box in top-right

RMarkdown Basics

Main section ## Heading here

Subsection ### Heading here

Italic *text here*

Bold **text here**

Numbered List 1. text here

Unnumbered List – text here

Blockquote > text here

Summaries

Numerical Variable

favstats

```
favstats(~pop2010, data=counties)
favstats(miles~direction, data=driving)  # miles by direction
favstats(~miles|direction, data=driving) # same thing
```

```
median(~female|state, data=counties) %>% sort()
iqr(~poverty|state, data=counties)
```

Categorical Variable

median frequency

```
tally(~state, data=counties)
tally(~state, data=counties) %>% sort()
```

```
tally(~genhealth, data=brfss, format="percent")
```

Two Variables

relative frequency Crosstabs

```
# Column-wise percents
tally(~genhealth|sex, data=brfss, format="percent", useNA="no")
# Total percents
tally(~genhealth+sex, data=brfss, format="percent", useNA="no")
```

```
cor(mort_rate~own_rate, data=guns)
```

Graphs

One Variable

Correlation Pie Chart

```
tally(~genhealth, data=brfss, useNA="no") %>% pie()
```

```
histogram(~pop2010, data=counties, breaks=20)
histogram(~pop2010, data=counties %>% filter(pop2010 <= 2e6))
```

Histogram Barchart

```
tally(~state, data=counties) %>% barchart()
tally(~state, data=counties) %>% sort() %>% barchart() # Pareto chart
```

```
sum(~own_rate|country, data=guns) %>% sort() %>% dotplot()
```

Two Variables

Labeled Dotplot

100% Stacked Barchart

```
healthVsExer <- tally(~genhealth|exerciseany, data=brfss,
    format="percent", useNA="no")
healthVsExer %>% t() %>% barchart(auto.key=list(space="right"))
```

```
xyplot(mort_rate~own_rate, data=guns)
ladd(panel.loess(x, y, col="magenta", lwd=2)) # add smooth fit line
```

Scatterplot

Boxplot

```
bwplot(state~female, data=counties)
```

Three Variables

Paneled Scatterplot

```
xyplot(mort_rate~own_rate|hdicat, data=guns)
```

Colors

list by name

```
colors() # in console
```

```
display.brewer.all()  # in console
```

palette list

pick a palette

```
brewer.pal(4, "Accent")
```

```
...plot...(..., col=name-or-palette, ...)
```

Graph Labeling

add in graph

main title

```
...plot...(..., main="Distribution of ...", ...)
```

```
...plot...(..., xlab="x label here", ...)
...plot...(..., ylab="y label here", ...)
```

Misc

Linear Modeling

axis labels

add to graph

```
ladd(panel.lmline(x, y, col="magenta", lwd=2))
```

```
fit <- lm(mort_rate~own_rate, data=guns)
coefficients(fit)
summary(fit)
```

get model diagnostics

```
xyplot(resid(fit)~fitted(fit))   # residual plot
ladd(panel.abline(h=0))
cor(mort_rate~own_rate, data=guns)^2   # r-squared
```