R Reference Card

Basics

`q()` quit R`<ESC>` cancel partially typed command

Getting Help

`apropos("frag")` list functions with `frag` as part of name
`?command`
`??`
`example(command)` run example code for `command`

Data

`data(dataset)` (re)load dataset from loaded package.
`nrow(dataset); ncol(dataset)` number of rows/cols
`dim(dataset)` number of rows and columns

Arithmetic and basic functions

`+ - * / ^` basic arithmetic operators
`%%` modulus operator
`log(x), exp(x)` natural logarithm and exponential
`sqrt(x)` square root
`sin(x), cos(x), tan(x)` trig functions
`choose(n,k)` $\binom{n}{k}$
`factorial(n)` $n!$
`abs(x)` $|x|$

Randomization/Simulation

`rflip(n)` simulate tossing n coins
`do(n) * command` repeat `command` n times
`sample(x,n)` sample n items from x without replacement
`resample(x,n)` sample n items from x with replacement
`shuffle(x,n)` sample n items from x with replacement
`shuffle(x)` same as `sample(x)`
`rbinom(n, size, prob )` n random draws from Binom(`size`, `prob`)
`rnorm(mean=`μ`, sd=`σ`)` n random draws from Norm(μ, σ) @

Formula Theme

The following syntax (often with some parts omitted) is used for graphical summaries, numerical summaries, and inference procedures.

```
fname( y ~ x | z, data=...,
                  groups=... )
```

For plots

- `y`: is y-axis variable
- `x`: is x-axis variable
- `z`: conditioning variable (separate panels)
- `groups`: conditioning variable (overlaid graphs)

For other things ?? can usually be read ?? is modeled by (or depends on) ??
separately for each ?? .
See the sampler for examples.

Distributions

```
pbinom(); pnorm();
pchisq(); pt()
qbinom(); qnorm();
qchisq(); qt()
plotDist();     # mosaic
```

Numerical Summaries

These functions have a formula interface to match plotting.

```
mean()      # mosaic augmented
median()    # mosaic augmented
sd()        # mosaic augmented
var()       # mosaic augmented
quantile()  # mosaic augmented
favstats()  # mosaic
tally()     # mosaic
```

Graphics (mostly lattice)

`lattice` is not the only option, but I find it works well because (a) it allows for easy multi-variable plots with good default settings, and (b) `lattice` uses the formula interface.

```
bwplot()
xyplot()
histogram()    # mosaic
qqmath()
densityplot()
plotFun()      # mosaic
```

```
ladd()         # mosaic
dotPlot()      # mosaic
bargraph()     # mosaic
mosaic()       # in vcd package
xhistogram()   # mosaic
xqqmath()      # mosaic
```

Inference

```
binom.test()  # mosaic augmented
prop.test()   # mosaic augmented
chisq.test()
t.test()
lm()          # linear models
anova( lm() )
summary( lm() )
```

```
makeFun( lm() )   # mosaic
resid( lm() )
plot( lm() )
TukeyHSD( lm () ) # mosaic aug
plot( TukeyHSD( lm() ) )
```

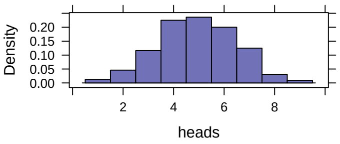
```
confint()      # mosaic augmented
pval()         # mosaic
fisher.test()
xchisq.test()  # mosaic
power.t.test()
power.prop.test()
wilcox.test()
```

Data

```
read.csv()
summary()
names()
head()
subset()
factor()
c()
cbind()
rbind()
```

```
merge()
```

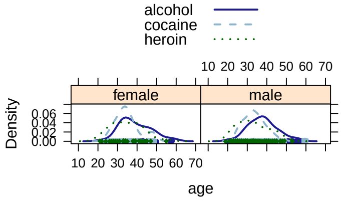
```
rflip(10)


Flipping 10 coins [ Prob(Heads) = 0

T H T H H T H H H T

Result: 6 heads.


do(2) * rflip(10)

   n heads tails
1 10     8     2
2 10     5     5

results <- do(1000) * rflip(10)
tally(~heads, data = results)


    1     2     3     4     5     6
   12    46   116   225   236   200
    7     8     9 Total
  125    31     9  1000

xhistogram(~heads, data = results,
    width = 1)
```



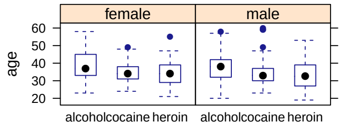
```
tally(~(heads > 8 | heads < 2),
    data = results)


 TRUE FALSE Total
   21   979  1000
```

```
tally(~substance + sex, data = HELPr

         sex
substance female male Total
  alcohol     36  141   177
  cocaine     41  111   152
  heroin      30   94   124
  Total      107  346   453

mean(age ~ substance, data = HELPrct

alcohol cocaine  heroin
  38.20   34.49   33.44

sd(age ~ substance, data = HELPrct)

alcohol cocaine  heroin
  7.652   6.693   7.986
```

```
densityplot(~age | sex, groups = substance,
    data = HELPrct, auto.key = TRUE)
```



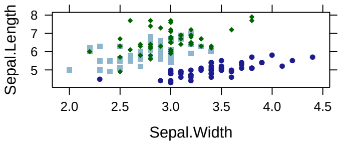
```
bwplot(age ~ substance | sex,
    data = HELPrct)
```



```
confint(t.test(~age, data = HELPrct))

Error:  'formula' missing or
incorrect
```

```
anova(lm(age ~ sex + substance,
    data = HELPrct))

Analysis of Variance Table

Response: age
           Df Sum Sq Mean Sq F value
sex         1     50      50    0.91
substance   2   1997     999   18.06
Residuals 449  24823      55
           Pr(>F)
sex          0.34
substance 2.8e-08
Residuals
```

```
xyplot(Sepal.Length ~ Sepal.Width,
    data = iris, groups = Species)
```


Sepal.Length
Sepal.Width