Simple Example

```
tab1 <- tableby(arm ~ sex + age, data=mockstudy)
summary(tab1, text=TRUE) # text version
as.data.frame(tab1) #Data frame version of table
```

Modifying Output

```
labels(mockstudy) <- c(age = 'Age, yrs', sex = "Gender") #Add labels to dataframe
labels(tab1) <- c(arm="Treatment Assignment", age="Baseline Age (yrs)") # or add labels to tableby object.
```

Change summary statistics globally

```
mycontrols <- tableby.control(test=FALSE, total=FALSE,
numeric.test="kwt", cat.test="chisq",
numeric.stats=c("N", "median", "q1q3"),
cat.stats=c("countpct"),
stats.labels=list(N='Count', median='Median', q1q3='Q1,Q3'))
tab2 <- tableby(arm ~ sex + age, data=mockstudy, control=mycontrols)
summary(tab2)
```

Change these settings directly in the tableby call

```
tab3 <- tableby(arm ~ sex + age, data=mockstudy, test=FALSE, total=FALSE,
numeric.stats=c("median","q1q3"), numeric.test="kwt")
summary(tab3)
```

Change summary statistics within the formula

```
tab.test <- tableby(arm ~ kwt(age) + anova(bmi) + notest(ast), data=mockstudy)
tests(tab.test)
summary(tab.test)
tab.test <- tableby(arm ~ kwt(ast, "Nmiss2","median") + anova(age, "N","mean") +
notest(bmi, "Nmiss","median"), data=mockstudy)
summary(tab.test)
```

Controlling Options for Categorical Tests

```
set.seed(100)
tab.catsim <- tableby(arm ~ sex + race, cat.test="fe", simulate.p.value=TRUE, B=500,
data=mockstudy) # simulations for these with 500 replicates for the Fisher's test (fe).
tests(tab.catsim)
cat.correct <- tableby(arm ~ sex + race, cat.test="chisq", subset = !grepl("^F", arm),
data=mockstudy) #use subset to ignore one of the three treatment arms.
tests(cat.correct)
```

Word version of the table

```
write2doc(tab1, "tab1.doc")
```

Summarize without a group/by variable

```
tab.noby <- tableby(~ bmi + sex + age, data=mockstudy)
summary(tab.noby)
```

Display footnotes which test was used

```
summary(tab.test, pfootnote=TRUE)
```

Summarize an ordered factor

```
mockstudy$age.ordnew <- ordered(c("a",NA,as.character(mockstudy$age.ord[-(1:2)])))
table(mockstudy$age.ord, mockstudy$sex)
```

Summarize a survival variable

```
survfit(Surv(fu.time, fu.stat)~sex, data=mockstudy) # The default is to show the median survival
survdiff(Surv(fu.time, fu.stat)~sex, data=mockstudy)
summary(tableby(sex ~ Surv(fu.time, fu.stat), data=mockstudy))
```

Subsetting

```
mytab <- tableby(arm ~ sex + alk.phos + age, data=mockstudy)
mytab2 <- mytab[c('age','sex','alk.phos')]
summary(mytab2)
summary(sort(mytab, decreasing = TRUE))
summary(mytab[mytab < 0.5])
```


Sponsored by **Readable.com**
Measure your website readability!
https://readable.com

Merge two tableby objects together

demographics

```
tab1 <- tableby(arm ~ sex + age, data=mockstudy,
control=tableby.control(numeric.stats=c("Nmiss","meansd"), total=-
FALSE))
```

lab data

```
tab2 <- tableby(arm ~ hgb + alk.phos, data=mockstudy,
control=tableby.control(numeric.stats=c("Nmiss","median","q1q3"),
numeric.test="kwt", total=FALSE))
tab12 <- merge(tab1, tab2)
class(tab12)
```

Modify how missing values are displayed

Show how many subjects have each variable (non-missing)

```
summary(tableby(sex ~ ast + age, data=mockstudy,
control=tableby.control(numeric.stats=c("N","median"), total=FAL-
SE)))
```

Always list the number of missing values

```
summary(tableby(sex ~ ast + age, data=mockstudy,
control=tableby.control(numeric.stats=c("Nmiss2","median"), total=-
FALSE)))
```

Only show the missing values if there are some (default)

```
summary(tableby(sex ~ ast + age, data=mockstudy,
control=tableby.control(numeric.stats=c("Nmiss","mean"),total=FA-
LSE)))
```

Don't show N at all

```
summary(tableby(sex ~ ast + age, data=mockstudy,
control=tableby.control(numeric.stats=c("mean"),total=FALSE)))
```

Modify the number of digits used

```
summary(tableby(arm ~ sex + age + fu.time, data=mockstudy),
digits=4, digits.p=2, digits.pct=1)
summary(tableby(arm ~ chisq(sex, digits.pct=1) + anova(age,
digits=4) +
anova(fu.time, digits = 1), data=mockstudy))
```

Tabulate data by a non-test group (strata)

```
summary(tableby(list(sex, ps) ~ age + bmi, strata = arm, data =
mockstudy))
```

Summary statistics

The default summary statistics, by varible type, are:

numeric.stats: Continuous variables will show by default Nmiss, meansd, range

cat.stats: Categorical and factor variables will show by default Nmiss, countpct

ordered.stats: Ordered factors will show by default Nmiss, countpct

surv.stats: Survival variables will show by default Nmiss, Nevents, medsurv

date.stats: Date variables will show by default Nmiss, median, range

There are a number of extra functions defined specifically for the tableby function.

N, Nmiss, Nmiss2, meansd, count, countN, countpct, countrowpct, q1q3, iqr

Testing options

anova, kwt, chisq, fe

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com