SumTableFunctions Vignette

Luke Fostvedt

November 7, 2019

Categorical and Continuous Summary Tables

This is a tutorial of some summary table functions created to simplify and standardize our reporting. While the functions have some edit checks, they wont catch everything. so it is important that the use confirm that the output is as expected. Two functions will be discussed here with the emphasis being mostly on the first one.

Warning: While the data are anonymous and further edits have been made to ensure they unidentifiable, they are based off real patient data and MUST NOT be distributed outside of pfizer.

Description of Each Function

A Flexible "One-Size-Fits-Most"" Table Function

The goal was to create a series of useful functions for making tables. They all eventually got wrapped up into this one function that hopefully make table creation simpler. There is some basic data manipulation that needs to occur before using the function.

- The continuous variables in your dataset much be either "numeric" or "integer".
- The categorical covariates should be defined as factors with the levels provided, otherwise the values in the dataset will print out (eg for Sex with values 1 and 2, the numbers will print rather than Male and Female).

These are really the only things necessary for the function to work. There are no package dependencies. You just need to source in the function in your R script.

```r
source("SummaryTableFunctions.R")
```

The function has several arguments, not all will be necessary depending on the table desired. The arguments are mostly defaulted to be `NULL` to allow for easier use. This way the user only needs to use the arguments that are necessary for the desired table.

```
pmx.sum.table.comb = function(
  data,
  cont.xvars=NULL, # variable(s) to be summarized
  cat.xvars=NULL,  # categorical variables to summarize
  sum.by=NULL, # what to summarize across
  more.variables = NULL, # if summarize by subgroups, not recommended
  digits=NULL, # number of significant digit(s);
                # need one defined for each sum.by variabel
  LaTeX=T, # if the output should be printed with LaTeX formatting
  Cont.Var.Name=NULL, #name of the
          # summary variable(s); need one for each sum.by variable
  Cat.Var.Name=NULL, #
  PERCENT=TRUE,
  Perc.Digits = 0,  # number of significant digits for percentages, default 0
  funs.summary=c("n","median","mean","sd","range","geomean","geosd","geocv","miss","se")
)
```

A description of each argument is provided in the following bulleted list:

- data: the dataset and must be a data.frame or tibble.
- cont.xvars: a vector with the names of the continuous variables to be summarize.
  - The vector is looking for column names in the dataset.
  - It will accept either integer or numeric variables.
- cat.xvars: This is a vector with the names of the categorical variables to be summarize.
  - The vector is looking for column names in the dataset.
  - It is recommended that the variables listed are already of class "factor".
- sum.by: This is the variables which will comprise the columns in the final table. It is best if this is a factor.
  - This can only be of length 1.
- more.variables: vector with variable names if the summaries of the continuous variables are desired to be broken down by subgroups.
  - It breaks down the subgroups for continuous variables.
  - the break-down is one variable at a time presented in new rows.
  - This does not affect categorical summaries.
  - If only categorical summaries are desired, this should be NULL. Otherwise it is ignored.
- digits: This is a vector with the number of significant digits for the summary.
  - It must be of the same length as cont.vars.
  - The value should correspond to the values of the raw data (eg Age would have 0 since it an integer).
  - The Mean, Standard Deviation, Geometric Mean, Geometric Standard Deviation, and Geometric CV will have one more significant digit than the input value.
- LaTeX: This is a TRUE/FALSE argument
  - TRUE will add LaTeX formatting as well as some LaTeX table indentation
  - FALSE will do just return the table
- Cont.Var.Name: This is the vector with the names that will print out for each continuous variable
  - It must be the same length as cont.xvars
- Cat.Var.Name: This is the vector with the names that will print out for each categorical variable

- It must be the same length as cat.xvars
- PERCENT: this is TRUE/FALSE.
    - TRUE will print the % sign [eg (50%)]
    - FALSE will just print the number in parentheses [eg (50)]
- Perc.Digits: This is the precision desired for the percentage calculations (default is 0 and results in integer output)
- funs.summary: This is the list of available summary functions available.
    - The output from each summary function will print in the order given by the user in the list
    - If both `mean` and `sd` are requested, a contracted "Mean (Std. Dev.)" will print.
    - If both `geomean` and `geosd` are requested, a contracted "GeoMean (GeoSD)" will print.

The Legacy tabcovsum() Function

The `tabcovsum()` function was created several years ago. It is included in this script of functions with the intent of providing a QC'd version that is backwards compatible. It is recommended to use the `pmx.sum.table.comb()` function instead.

```
tabcovsum = function(
 x1=NULL, # The variable for the rows
 x2=NULL, # the variable for the columns
 x1names=NULL, # The names of the levels for each row variable for tabulation
 x2names=NULL, # The name that will be output of the variable which will be tabulated
 Total=F  #determines if a column for total tabulation should be included.
          #default is false for backwards compatibility
)
```

A description of each argument is provided in the following bulleted list:

- x1: variable name (eg data$RowVAR1)
    - The unique values (or levels) will form the row names
- x2: variable name (eg data$ColVAR1)
    - The levels for this variable will form the column names
- x1.names: This is a vector with the names of the categorical variables to be summarize.
    - The vector will be the rownames that output.
- x2.names: character vector, length=1.
    - This will be the name of the variable that prints
    - This can only be of length 1.
- Total: TRUE/FALSE. If a column with the totals across all columns is desired.
    - default is FALSE

Data Preparation for Examples

```r
source("SummaryTableFunctions.R")
d <- read.csv("https://raw.githubusercontent.com/fostvedt/ACoP-ordinal-regression/master/AEs-ordinal-da
#write.csv(d.demog,"Example-dataset.csv",row.names = F,quote = F)
```

The functions work best when categorical covariates are factors with the levels as desired in the final report.
Below we will define as factors the categorical variables in our dataset.

```r
d$SMOK <- factor(d$SMOK,levels=c(1:3),c("Never Smoked","Former Smoker","Current Smoker"))
d$SEX <-  factor(d$SEX,levels=c(1,2),c("Male","Female"))
d$DOSE <- factor(d$DOSE,levels=c(0,10,30,50,100),
               labels = c("0 mg","10 gm","30 mg","50 mg","100 mg"))
#d$CMAXSS <- as.numeric(d$CMAXSS)
#d$BIO <- as.numeric(d$BIO)
#d$BIO2 <- as.numeric(d$BIO2)
```

pmx.sum.table.comb() Examples

Combined table with continuous and categorical covariates

The following table will be useful for baseline demographics summaries.

```r
b1 <- pmx.sum.table.comb(data = d,
    cont.xvars=c("AGE","WT","BIO","CMAXSS"),    # variable(s) to be summarized
    cat.xvars=c("SEX","SMOK"),     # categorical variables to be summarized
    sum.by=c("DOSE"),     # what to summarize across
#   more.variables = NULL, # if subgroups are to be summarized within each variable
    digits=c(0,1,1,1),     # number of significant digit(s);
                           #need one defined for each sum.by variabel
    LaTeX=T,               # if the output should be printed with LaTeX formatting
    Cont.Var.Name=c("Age (yrs)","Body Weight (kg)","BiomArker",
              "Steady-State Cmax"),
                      #name of the summary variable(s);
                      # need one for each sum.by variable
    Cat.Var.Name=c("Sex","Smoking Status"), #
    PERCENT=FALSE,
    Perc.Digits = 1,
    funs.summary=c("median","mean","sd","range","n","miss"))
#write.csv(b1,"Continuous-Categorical-example.csv",quote=F,row.names = F)
```

Table 1: Summary Table of demographics

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Age (yrs)						
Median	56	56	52	58	55	56
Mean (Std. Dev.)	54.9 (10.5)	54.2 (9.4)	50.7 (11.5)	54.0 (11.5)	54.3 (8.7)	53.6 (10.5)
Range (Min; Max)	(33; 70)	(35; 68)	(19; 70)	(25; 71)	(30; 73)	(19; 73)
N (%)	39 (19.5)	31 (15.5)	39 (19.5)	48 (24.0)	43 (21.5)	200 (100.0)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Body Weight (kg)						
Median	70.7	71.5	73.5	68.8	72.0	71.3
Mean (Std. Dev.)	70.91 (12.99)	73.03 (18.55)	74.80 (14.91)	71.89 (11.80)	75.70 (16.94)	73.26 (14.94)
Range (Min; Max)	(48.0; 110.0)	(41.5; 128.4)	(49.0; 107.8)	(43.0; 105.8)	(47.0; 121.0)	(41.5; 128.4)
N (%)	39 (19.5)	31 (15.5)	39 (19.5)	48 (24.0)	43 (21.5)	200 (100.0)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
BiomArker						
Median	18.1	13.7	16.3	16.8	18.4	16.2
Mean (Std. Dev.)	17.55 (7.43)	13.81 (6.11)	15.87 (5.77)	16.53 (5.92)	18.02 (6.50)	16.50 (6.45)
Range (Min; Max)	(1.5; 31.4)	(4.0; 29.6)	(5.0; 26.5)	(0.9; 30.7)	(4.0; 33.1)	(0.9; 33.1)
N (%)	39 (19.5)	31 (15.5)	39 (19.5)	48 (24.0)	43 (21.5)	200 (100.0)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Steady-State Cmax						
Median	0.0	7.7	15.6	16.3	28.5	15.0
Mean (Std. Dev.)	0.00 (0.00)	7.90 (1.96)	15.55 (2.32)	18.15 (5.67)	30.16 (7.26)	15.10 (11.15)
Range (Min; Max)	(0.0; 0.0)	(4.9; 14.5)	(9.6; 19.6)	(13.2; 42.1)	(21.6; 63.1)	(0.0; 63.1)
N (%)	39 (19.5)	31 (15.5)	39 (19.5)	48 (24.0)	43 (21.5)	200 (100.0)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Sex; N(%)						
Male	9 (23.1)	7 (22.6)	6 (15.4)	11 (22.9)	10 (23.3)	43 (21.5)
Female	30 (76.9)	24 (77.4)	33 (84.6)	37 (77.1)	33 (76.7)	157 (78.5)
Smoking Status; N(%)						
Never Smoked	28 (71.8)	23 (74.2)	30 (76.9)	35 (72.9)	32 (74.4)	148 (74.0)
Former Smoker	6 (15.4)	3 (9.7)	5 (12.8)	8 (16.7)	7 (16.3)	29 (14.5)
Current Smoker	5 (12.8)	5 (16.1)	4 (10.3)	5 (10.4)	4 (9.3)	23 (11.5)

Table With Different Summary Functions

```
cont.fun<-pmx.sum.table.comb(data = d,
      cont.xvars=c("AGE"), # variable(s) to be summarized
#       cat.xvars=NULL,  # categorical variables to summarize
      sum.by=c("DOSE"), # what to summarize across
#       more.variables = NULL, # if summarize by subgroups, not recommended
      digits=c(0), # number of significant digit(s);
                   # need one defined for each sum.by variabel
      LaTeX=T, # if the output should be printed with LaTeX formatting
      Cont.Var.Name=c("Age (yr)"), #name of the
             # summary variable(s); need one for each sum.by variable
#       Cat.Var.Name=NULL, #
      PERCENT=TRUE,
      Perc.Digits = 1,
      funs.summary=c("median","mean","sd","range","n","miss","geomean","geosd","geocv"))
#write.csv(cont.fun,"Continuous-functions-example.csv",quote=F,row.names = F)
```

Table 2: Summary Table of Age with All Summary Functions

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Age (yr)						
Median	56	56	52	58	55	56
Mean (Std. Dev.)	54.9 (10.5)	54.2 (9.4)	50.7 (11.5)	54.0 (11.5)	54.3 (8.7)	53.6 (10.5)
Range (Min; Max)	(33; 70)	(35; 68)	(19; 70)	(25; 71)	(30; 73)	(19; 73)
N (%)	39 (19.5%)	31 (15.5%)	39 (19.5%)	48 (24.0%)	43 (21.5%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
GeoMean (GeoSD)	53.8 (1.2)	53.4 (1.2)	49.2 (1.3)	52.5 (1.3)	53.6 (1.2)	52.5 (1.2)
GeoCV	0.2	0.2	0.3	0.3	0.2	0.2

Using kable: Table With Different Order of Functions

```
cont.fun1<-pmx.sum.table.comb(data = d,
      cont.xvars=c("AGE"), # variable(s) to be summarized
#      cat.xvars=NULL,  # categorical variables to summarize
      sum.by=c("DOSE"), # what to summarize across
#      more.variables = NULL, # if summarize by subgroups, not recommended
      digits=c(0), # number of significant digit(s);
                  # need one defined for each sum.by variabel
      LaTeX=F, # if the output should be printed with LaTeX formatting
      Cont.Var.Name=c("Age (yrs)"), #name of the
            # summary variable(s); need one for each sum.by variable
#      Cat.Var.Name=NULL, #
      PERCENT=TRUE,
      Perc.Digits = 1,
      funs.summary=c("n","geomean","geosd","miss","median","geocv","mean","sd","range"))
#write.csv(cont.fun1,"Continuous-functions-different-order-example.csv",quote=F,row.names = F)
```

Table 3: Summary Table of Age With Different Ordering of Summary Functions

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Age (yrs)						
N (%)	39 (19.5%)	31 (15.5%)	39 (19.5%)	48 (24.0%)	43 (21.5%)	200 (100.0%)
GeoMean (GeoSD)	53.8 (1.2)	53.4 (1.2)	49.2 (1.3)	52.5 (1.3)	53.6 (1.2)	52.5 (1.2)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Median	56	56	52	58	55	56
GeoCV	0.2	0.2	0.3	0.3	0.2	0.2
Mean (Std. Dev.)	54.9 (10.5)	54.2 (9.4)	50.7 (11.5)	54.0 (11.5)	54.3 (8.7)	53.6 (10.5)
Range (Min; Max)	(33; 70)	(35; 68)	(19; 70)	(25; 71)	(30; 73)	(19; 73)

Table With Several Categorical covariates

```
cat1<-pmx.sum.table.comb(data = d,
#      cont.xvars=NULL, # variable(s) to be summarized
       cat.xvars=c("SEX","SMOK"),  # categorical variables to summarize
       sum.by=c("DOSE"), # what to summarize across
#      more.variables = NULL, # if summarize by subgroups, not recommended
#      digits=NULL, # number of significant digit(s);
                    # need one defined for each sum.by variabel
       LaTeX=T, # if the output should be printed with LaTeX formatting
#      Cont.Var.Name=NULL, #name of the
              # summary variable(s); need one for each sum.by variable
       Cat.Var.Name=c("Sex","Smoking Status"), #
       PERCENT=TRUE,
       Perc.Digits = 1)
#write.csv(cat1,"Categorical-only-example.csv",quote=F,row.names = F)
```

Table 4: Summary Table of Categorical Demographics

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Sex; N(%)						
Male	9 (23.1%)	7 (22.6%)	6 (15.4%)	11 (22.9%)	10 (23.3%)	43 (21.5%)
Female	30 (76.9%)	24 (77.4%)	33 (84.6%)	37 (77.1%)	33 (76.7%)	157 (78.5%)
Smoking Status; N(%)						
Never Smoked	28 (71.8%)	23 (74.2%)	30 (76.9%)	35 (72.9%)	32 (74.4%)	148 (74.0%)
Former Smoker	6 (15.4%)	3 (9.7%)	5 (12.8%)	8 (16.7%)	7 (16.3%)	29 (14.5%)
Current Smoker	5 (12.8%)	5 (16.1%)	4 (10.3%)	5 (10.4%)	4 (9.3%)	23 (11.5%)

Table With Several Continuous covariates

In the event comparison by several categorical covariates are desired, a separate covariate should be constructed.

```
d$SMOK.SEX <- factor(interaction(d$SMOK,d$SEX))
levels(d$SMOK.SEX) <- c("No-Smoker Male","No-Smoker Female","Fo-SmokerMale","Fo-Smoker Female",
                        "Smoker  Male","Smoker Female")
```

```
a4 <- pmx.sum.table.comb(d,
                         cont.xvars=c("BIO","BIO2","CMAXSS","CAVGSS"),
#                         cat.xvars = NULL,
                         sum.by=c("SMOK.SEX"),
#                         more.variables = c(),
                         digits=c(2,2,1,1),
                         LaTeX=T,
                         PERCENT=T,
                         Perc.Digits = 1,
                         Cont.Var.Name=c("Biomarker 1","Biomarker 2","Steady-State Cmax","Steady-State Cavg"),
#                         Cat.Var.Name = NULL,
                         funs.summary=c("median","range","mean","sd","n","miss"))
#write.csv(a4,"Multiple-continuous-example.csv",quote=F,row.names = F)
```

Table 5: Summary Table With Combination of Variables

Group	No-Smoker Male	No-Smoker Female	Fo-SmokerMale	Fo-Smoker Female	Smoker Male	Smoker Female	Total
Biomarker 1							
Median	15.83	14.91	17.58	16.38	16.36	14.94	16.16
Range (Min; Max)	(5.33; 33.11)	(5.38; 23.67)	(4.02; 20.78)	(0.95; 31.65)	(6.41; 30.65)	(6.59; 23.88)	(0.95; 33.11)
Mean (Std. Dev.)	16.493 (6.814)	15.063 (5.921)	15.130 (5.687)	16.766 (6.657)	17.152 (6.662)	15.248 (4.895)	16.498 (6.452)
N (%)	21 (10.5%)	11 (5.5%)	11 (5.5%)	127 (63.5%)	18 (9.0%)	12 (6.0%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Biomarker 2							
Median	40.67	36.02	44.08	39.19	43.39	37.18	39.10
Range (Min; Max)	(19.31; 69.47)	(20.01; 56.70)	(18.47; 69.33)	(9.20; 76.25)	(18.78; 66.95)	(18.78; 53.21)	(9.20; 76.25)
Mean (Std. Dev.)	42.808 (13.897)	35.266 (13.411)	44.817 (15.835)	40.668 (13.825)	41.924 (12.309)	37.137 (8.262)	40.725 (13.518)
N (%)	21 (10.5%)	11 (5.5%)	11 (5.5%)	127 (63.5%)	18 (9.0%)	12 (6.0%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Steady-State Cmax							
Median	15.4	15.7	15.8	14.5	16.8	10.2	15.0
Range (Min; Max)	(0.0; 63.1)	(0.0; 26.9)	(0.0; 38.6)	(0.0; 45.1)	(0.0; 39.0)	(0.0; 42.1)	(0.0; 63.1)
Mean (Std. Dev.)	16.11 (15.24)	14.74 (9.55)	15.94 (11.85)	14.87 (10.34)	15.69 (10.94)	14.43 (14.24)	15.10 (11.15)
N (%)	21 (10.5%)	11 (5.5%)	11 (5.5%)	127 (63.5%)	18 (9.0%)	12 (6.0%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Steady-State Cavg							
Median	8.8	10.4	10.6	10.5	10.6	7.0	10.3
Range (Min; Max)	(0.0; 19.7)	(0.0; 18.9)	(0.0; 20.8)	(0.0; 21.4)	(0.0; 18.9)	(0.0; 19.3)	(0.0; 21.4)
Mean (Std. Dev.)	9.11 (6.46)	9.76 (6.65)	9.94 (6.62)	9.39 (5.90)	9.72 (6.43)	7.77 (6.65)	9.35 (6.07)
N (%)	21 (10.5%)	11 (5.5%)	11 (5.5%)	127 (63.5%)	18 (9.0%)	12 (6.0%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Table With One Continuous Variable Broken Down by Subgroup

```
a5 <- pmx.sum.table.comb(d,
                cont.xvars=c("BIO"),
                sum.by=c("DOSE"),
                more.variables = c("SEX"),
                digits=c(1),
                LaTeX=T,
                PERCENT=T,
                Perc.Digits = 1,
                Cont.Var.Name="Biomarker",
```

```
                    funs.summary=c("median","range","mean","sd","n","miss"))
#write.csv(a5,"Single-group-continuous-example.csv",quote=F,row.names = F)
```

Table 6: Summary Table of Biomarker with Gender Breakdown

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Biomarker						
Median	18.1	13.7	16.3	16.8	18.4	16.2
Range (Min; Max)	(1.5; 31.4)	(4.0; 29.6)	(5.0; 26.5)	(0.9; 30.7)	(4.0; 33.1)	(0.9; 33.1)
Mean (Std. Dev.)	17.55 (7.43)	13.81 (6.11)	15.87 (5.77)	16.53 (5.92)	18.02 (6.50)	16.50 (6.45)
N (%)	39 (19.5%)	31 (15.5%)	39 (19.5%)	48 (24.0%)	43 (21.5%)	200 (100.0%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Biomarker: Male						
Median	15.6	12.5	15.7	16.9	15.8	15.8
Range (Min; Max)	(5.3; 27.2)	(4.0; 17.6)	(5.4; 21.9)	(8.0; 23.7)	(8.5; 33.1)	(4.0; 33.1)
Mean (Std. Dev.)	16.10 (6.28)	12.12 (5.06)	14.22 (6.65)	16.73 (4.77)	17.94 (7.75)	15.78 (6.22)
N (%)	9 (20.9%)	7 (16.3%)	6 (14.0%)	11 (25.6%)	10 (23.3%)	43 (100%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Biomarker: Female						
Median	18.7	13.8	16.3	16.8	18.9	16.3
Range (Min; Max)	(1.5; 31.4)	(4.2; 29.6)	(5.0; 26.5)	(0.9; 30.7)	(4.0; 31.6)	(0.9; 31.6)
Mean (Std. Dev.)	17.98 (7.78)	14.31 (6.40)	16.16 (5.66)	16.47 (6.27)	18.05 (6.21)	16.69 (6.52)
N (%)	30 (19.1%)	24 (15.3%)	33 (21.0%)	37 (23.6%)	33 (21.0%)	157 (100%)
Missing (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Table With One Continuous Variable Broken Down by Multiple Subgroups

```
a6 <- pmx.sum.table.comb(d,
                        cont.xvars=c("BIO"),
                        sum.by=c("DOSE"),
                        more.variables = c("SEX","SMOK"),
                        digits=c(2),
                        LaTeX=T,
                        PERCENT=FALSE,
                        Perc.Digits = 1,
                        Cont.Var.Name="Biomarker",
                        funs.summary=c("median","range","mean","sd","n","miss"))
#write.csv(a6,"multiple-group-continuous-example.csv",quote=F,row.names = F)
```

Table 7: Summary Table of Baseline Disease Severity Across Studies by Gender and Dose

Group	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Biomarker						
Median	18.12	13.74	16.28	16.82	18.41	16.16
Range (Min; Max)	(1.50; 31.44)	(4.02; 29.58)	(4.95; 26.49)	(0.95; 30.65)	(4.05; 33.11)	(0.95; 33.11)
Mean (Std. Dev.)	17.546 (7.429)	13.814 (6.112)	15.866 (5.767)	16.527 (5.916)	18.020 (6.502)	16.498 (6.452)
N (%)	39 (19.5)	31 (15.5)	39 (19.5)	48 (24.0)	43 (21.5)	200 (100.0)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Biomarker: Male						
Median	15.56	12.48	15.68	16.87	15.84	15.83
Range (Min; Max)	(5.33; 27.24)	(4.02; 17.58)	(5.38; 21.88)	(7.96; 23.67)	(8.47; 33.11)	(4.02; 33.11)
Mean (Std. Dev.)	16.097 (6.279)	12.124 (5.056)	14.224 (6.646)	16.728 (4.774)	17.938 (7.751)	15.779 (6.218)
N (%)	9 (20.9)	7 (16.3)	6 (14.0)	11 (25.6)	10 (23.3)	43 (100)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Biomarker: Female						
Median	18.75	13.81	16.28	16.76	18.92	16.28
Range (Min; Max)	(1.50; 31.44)	(4.17; 29.58)	(4.95; 26.49)	(0.95; 30.65)	(4.05; 31.65)	(0.95; 31.65)
Mean (Std. Dev.)	17.981 (7.784)	14.307 (6.398)	16.165 (5.657)	16.468 (6.273)	18.045 (6.212)	16.694 (6.520)
N (%)	30 (19.1)	24 (15.3)	33 (21.0)	37 (23.6)	33 (21.0)	157 (100)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Biomarker: Never Smoked						
Median	18.22	13.58	16.46	16.76	19.03	16.16

Range (Min; Max)	(1.50; 31.44)	(4.17; 29.58)	(4.95; 26.49)	(0.95; 29.11)	(4.05; 33.11)	(0.95; 33.11)
Mean (Std. Dev.)	17.777 (7.728)	13.540 (6.460)	16.154 (5.897)	16.314 (5.878)	19.090 (6.578)	16.727 (6.657)
N (%)	28 (18.9)	23 (15.5)	30 (20.3)	35 (23.6)	32 (21.6)	148 (100)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Biomarker: Former Smoker						
Median	19.71	15.80	16.28	15.87	15.29	16.28
Range (Min; Max)	(8.10; 26.23)	(12.48; 16.93)	(5.38; 21.79)	(9.35; 30.65)	(6.41; 22.02)	(5.38; 30.65)
Mean (Std. Dev.)	18.840 (6.738)	15.069 (2.317)	15.775 (6.651)	17.290 (7.441)	14.142 (6.424)	16.360 (6.366)
N (%)	6 (20.7)	3 (10.3)	5 (17.2)	8 (27.6)	7 (24.1)	29 (100)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Biomarker: Current Smoker						
Median	14.73	17.58	14.94	18.77	15.20	15.26
Range (Min; Max)	(6.59; 23.88)	(4.02; 21.13)	(7.51; 17.89)	(10.21; 20.38)	(13.82; 20.78)	(4.02; 23.88)
Mean (Std. Dev.)	14.704 (7.173)	14.320 (6.785)	13.819 (4.436)	16.801 (4.357)	16.250 (3.274)	15.191 (5.166)
N (%)	5 (21.7)	5 (21.7)	4 (17.4)	5 (21.7)	4 (17.4)	23 (100)
Missing (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Example Using tabcovsum()

It is advised not to use `tabcovsum` anymore. `pmx.sum.table.comb()` is much more flexible.

Making a Summary Table of a Single Variable With tabcovsum()

```
demogtab<-tabcovsum(
    x1=d$SMOK,  # categorical variable to summarize
    x2=d$DOSE,
    x1names = levels(d$SMOK),
    x1vname = "Smoking Status",
    Total=F
  )
#write.csv(demogtab,"Categorical-Race-tabcovsum.csv",quote=F,row.names = F)
```

Table 8: Summary of Smoking Status by Dose

Variable	Category	0 mg	10 gm	30 mg	50 mg	100 mg
Smoking Status	Never Smoked	28 (72%)	23 (74%)	30 (77%)	35 (73%)	32 (74%)
	Former Smoker	6 (15%)	3 (10%)	5 (13%)	8 (17%)	7 (16%)
	Current Smoker	5 (13%)	5 (16%)	4 (10%)	5 (10%)	4 (9%)

Making a Full Summary Table With tabcovsum()

```
demvars1 <- rbind(
  tabcovsum(rep(1,nrow(d)),d$DOSE,c("n"),"Sample Size",Total=TRUE)
  ,tabcovsum(d$SEX,d$DOSE,levels(d$SEX),"Sex",Total=TRUE)
  ,tabcovsum(d$SMOK,d$DOSE,c("Never Smoker","Former Smoker","Smoker"),"Smoking Status",Total=TRUE)
)
#write.csv(demvars1,"Categorical-cov-summary-table.csv",quote=F,row.names = F)
```

Table 9: Summary of Categorical Covariates

Variable	Category	0 mg	10 gm	30 mg	50 mg	100 mg	Total
Sample Size	n	39 (100%)	31 (100%)	39 (100%)	48 (100%)	43 (100%)	200 (100%)
Sex	Male	9 (23%)	7 (23%)	6 (15%)	11 (23%)	10 (23%)	43 (22%)
	Female	30 (77%)	24 (77%)	33 (85%)	37 (77%)	33 (77%)	157 (78%)
Smoking Status	Never Smoker	28 (72%)	23 (74%)	30 (77%)	35 (73%)	32 (74%)	148 (74%)
	Former Smoker	6 (15%)	3 (10%)	5 (13%)	8 (17%)	7 (16%)	29 (14%)
	Smoker	5 (13%)	5 (16%)	4 (10%)	5 (10%)	4 (9%)	23 (12%)