```sas
title bold 'Proc Means';
proc odstext print;
p'Sample DataSets'/ styles={ color=black font_size=12pt};

/* Sample Dataset One*/

data one;
input family;
cards;
2
3
1
5
3
2
4
6
1
2
;
proc freq data=one;
tables family;
run;
title;



proc odstext print pagebreak=no;

p' Rather than printing out the entire empirical distribution function for a quantitative
variable in a data set, it is sometimes preferable to record just a few
numbers that summarize key features of the distribution. One possible choice is to record the mean and standard deviation of a
procedure proc means to do this in SAS.' / styles={ color=black font_size=12pt};


/*Sample Data set TWO */
proc odstext print;
p' Following is The NEWCOMB Dataset for Analysis...'/ styles={ color=black font_size=12pt};

data newcomb;
input light;
cards;
-44
-2
16
16
19
20
21
21
22
22
23
23
23
24
24
24
24
24
25
25
25
25
25
26
26
26
26
26
27
27
27
27
27
27
```

```
28
28
28
28
28
28
28
29
29
29
29
29
30
30
30
31
31
32
32
32
32
32
33
33
34
36
36
36
36
37
39
40
;
```

```
proc freq data=newcomb;
tables light;
run;


options nodate nonumber;
data newcomb;

input light;
cards;
proc means ;
title1 'Mean and standard deviation for Newcomb data';
var light;
footnote 'Figure: Output from running proc means on the Newcomb data.';
run;
title1;
footnote;

proc odstext print pagebreak;
p 'We describe here some of the more useful features of proc means.
proc means options keyword names;
var variables;
class variables;
freq variable;
weight variable;
output out = SASdataset keyword = names;
by variables;
Following are some of the options that can be used in the means statement.
data = SAS dataset
nway
noprint'/ styles={ color=black font_size=12pt};


proc odstext print;
p'n= number of observations on which the calculations are based.
nmiss= number of missing values.
mean= mean.
std= standard deviation.
min= smallest value.
max =largest value.
range= range = max – min.
sum= sum.
```

```
var= variance.
uss= uncorrected sum of squares = sum of squares of the data values.
css= corrected sum of squares = sum of squared deviations of the data values
from their mean.
stderr= standard error of the mean.
cv= coefficient of variation (percent) = standard error divided by the mean
and then multiplied by 100.
skewness measure of skewness.
kurtosis measure of kurtosis.
t= Student's t value for testing the hypothesis that the population mean is 0.
prt= probability of a greater absolute value of Student's t. '/ styles={ color=green fontweight=bold font_size=12pt};


proc odstext print;
p'In addition, observations that depend
on the class variables are added. The values of _type_ are more complicated
to explain, while _freq_ retains its earlier interpretation. To suppress the
addition of these observations, which is what we typically want to do, use
the nway option in the proc means statement. For example, if the SAS
data set three contains'/ styles={ color=black font_size=12pt};

data three nway;
input variable a b x y;
cards;
obs1 -1 1 6 0
obs2 1 2 3 -1
obs3 -1 1 2 2
obs4 1 2 0 3
obs5 -1 2 1 4
obs6 1 2 3 3
;
proc means data=three nway ;
class a b;
var x y;
output out=four mean=mnx mny ;
proc print data=four;
run;




proc odstext print;
p'The variables in the class statement are used to form subgroups; i.e.
these variables are used to classify observations. The class variables may
be either numeric or character, but normally each variable takes a small
number of values or levels.'/ styles={ color=black font_size=12pt};

data five;
input variable x y;
cards;
obs1 1 1.1
obs2 1 2.2
obs3 2 -3.0
obs4 2 2.0
;
proc means data=five mean std cv;
class x;
footnote1 'form two class groups; when x = 1 and x = 2, and compute the mean, standard
deviation and coefficient of variation for each group. The output from
this program is given in Figure';
run;
footnote1;

proc odstext print;
p'The output statement requests that proc means output statistics to a
new SAS data set that is specified as SASdataset in out = SASdataset. The
list of statistics specifies which statistics are to be included in the output
data set. The names are used to give names to the output variables. For
example, if data set one contains variables x, y, then'/ styles={ color=black font_size=12pt};

proc means data=five;
var x y;
output out=six mean=mn1 mn2 std=std1 std2;

proc print data=six;
footnote2 'Figure: Data set created by output statement consisting of values of
```

```sas
statistics computed on primary data.';
run;
footnote2;
proc odstext print;
p'creates a SAS data set two with one observation and variables mn1, mn2,
std1, and std2, where mn1 is the mean of x, mn2 is the mean of y, std1
is the standard deviation of x, and std2 is the standard deviation of y.
Therefore the first name in a list is associated with the value of the statistic
for the first variable in the var list, the second name is associated with the
value of the statistic for the second variable in the var list, and so on.'/ styles={ color=black font_size=12pt};
```