Transposing Values Within Groups:

by Season Basin Name;

Each unique combination of BY values creates one row in the output table.

The VAR statement lists the column or columns to be transposed

The output table will include a separate column for each value of the ID column.

There can be only one ID column. The ID column values must be unique in the column or BY group.

PREFIX= option provides a prefix for each value of the ID column in the output table

The NAME = option names the column that identifies the source column containing the transposed values.

```
proc means data = myclass;
    var age height wt;
run;
```

Global statements
Title
Options
Libname
} don't need a run

```
title; footnote;
ODS NOPROCTITLE;   ← turns off procedure title
proc means data = mycars.
```

```
Mean min max maxdec=1;
where Type = "Wagon";
var avgmpg;
LABEL col-name = "label-text";
class type;
run;
```

Adding Temporary labels to Columns.

By default, PROC MEANS generates simple summary statistics for each numeric column in the input data.

```
CLASS col-name(s);
WAYS n;
```

use to specify the numeric columns do analyze

PROC MEANS calculates default statistics — frequency count (N), mean, standard deviation, minimum & maximum — for each of the columns listed in the VAR statement.

PROC FREQ DATA = input-table;

TABLES col-name(s);

RUN;

use to specify the frequency tables to include in the results.

By default, PROC FREQ creates a frequency table for each column in the input table.

/* examine extreme values */ in univariate

OUTPUT

out = output-table

<statistic = col-name>;

output out = heart_stats

mean = AvgWeight;

Double quotation marks must be used around name variables

Formats are an easy way to group data in procedures.

Segmenting Reports

ORDER = FREQ. do omit results by descending frequency

PROC SQL ..... : PROC FedSQL

Create table work.myclass as

QUERY

QUIT

A query consists of clauses.

There must be at least a SELECT clause & a FROM clause in the query.

PROC SQL creates a report by default.

Adding CREATE TABLE at the beginning of the query turns a report into a table.

Combining Tables with SQL

DATA step

* provides more control of reading, writing & manipulating data.
* can create multiple tables in one step.

⇒ includes looping & array processing.

On upcase (s.basin) = b.basin

Applying Temporary Label to Columns.

```
proc print data= ____
                 label;
                 split = "*";

label.
=

run;
```

PROC UNIVARIATE

also generates summary statistics for each numeric column in the data by default but it includes more detailed statistics related to distribution & extreme values. The WHERE statement can be used in

PROC PRINT
PROC MEANS
FREQ
UNIVARIATE

A SAS macro variable stores text that is substituted in your code when it runs. It's like an automatic find and replace

```
ods rtf file = 'c:\shoes.rtf';
proc print data = sashelp.shoes;
run;
ods rtf close
```

In order to send the PROC PRINT results to an RTF file, you must use the ODS statement that specifies the destination & the FILE = option to specify the name of the file ( ) is

The PROC SORT step overwrites the original data set by default, sorts the data in ascending order by variables

```
proc sort data = sashelp.cars;
    by descending Make;
run;
```

```
libname exlib xlsx "c:\class.xlsx";

proc export data = mydata.storms
            outfile = "storms.csv"
            dbms = csv;
run;

proc format;
    value MULT 1,3,5 = 'odd'
               2,4,6 = 'evens';
run

proc sort data = SASUSER.PROSECTS out=PSORT;
    by Product Code descending Date Cost;
Run;
```

OUT = option specifies the name of the sorted data set

```
proc means data = sashelp.baseball;
     class position;
     var nRuns nRBI;
run;
```

Subsetting IF statements can only appear in DATA steps.

In SAS, WHERE statements can be used in both DATA & PROC steps

```
where position in ('1B' '2B' '3B');

where position = '1B' or
      position = '2B' or
      position = '3B';

format Date date9. Amount comma8.;
```

The TITLE statement is a global statement

The ORDER = FREQ option must be specified in the PROC FREQ statement itself in order to create a report by descending frequency.

```
proc freq data = sashelp.cars order=freq;
   tables type;
run;

proc freq data = sashelp.shoes;
   tables product / nopercent;
run;
```

NOPERCENT suppresses the percent & cumulative percent statistics.

The TABLES statement

specifies the variables that are analyzed in the FREQ procedure. By default, a PROC FREQ step without a Tables statement creates a frequency table for each character & numeric variable in the input data set.

N	Mean	Std Dev	Minimum	Maximum

MAXDEC =

By default, PROC PRINT displays all variables and observations from a dataset. It doesnot display labels

it does display row numbers

The NOOBS option

the display of the observation number in the PROC PRINT report.

DATA step permanently adds labels & formats do a data set

The DATA step creates data & any attributes assigned by the LABEL & FORMAT statements in the DATA step are permanently stored with the data set

144-164

PROC CONTENTS

creates a report of the descriptor portion of a SAS data set

The first action in the execution phase of the DATA step is initializing the PDV, then the observations are read into the PDV, calculations in the program are executed & finally the observations in the PDV is written to the output data set

The CONTAINS operator tests whether a string is part of the variable's values.

where T ext contains 'error'

> Character & numeric variables can be created.

> The value of a variable can be updated

SAS uses the results of calculation to missing missing value is used in an arithmetic expression

The CATX function requires the delimiter to be specified first, followed by the variables you wish to concatenate

Not a compile-time-only statement

The KEEP statement controls which columns are in the output table