SP101-103 Getting Started with SAS Programming

101 SAS Programming

102 Understanding SAS Data, Accessing data thru Library, Importing data into SAS

103 Exploring Data, Filtering Rows, Formatting Columns

```
*libname_PG1.sas;
libname PG1 "/home/u58304328/EPG1V2/data";

%let outpath=/home/u58304328/EPG1V2/output;



* p101d02.sas Understanding SAS Program Syntax *;

/* Program copied by William  */
data mycars;
        set sashelp.cars;
        AvgMPG=mean(mpg_city, mpg_highway);
run;

title "Cars with Average MPG over 35";

proc print data=mycars;
        var make model type avgmpg;
        where avgMPG > 35;
run;

title "Average MPG by Car Type";
```

```
proc means data=mycars mean min max maxdec=1;
        var avgmpg;
        class type;
run;
```

Cars with Average MPG over 35

Obs	Make	Model	Type	AvgMPG
150	Honda	Civic Hybrid 4dr manual (gas/electric)	Hybrid	48.5
151	Honda	Insight 2dr (gas/electric)	Hybrid	63.0
156	Honda	Civic HX 2dr	Sedan	40.0
374	Toyota	Prius 4dr (gas/electric)	Hybrid	55.0
380	Toyota	Corolla CE 4dr	Sedan	36.0
381	Toyota	Corolla S 4dr	Sedan	36.0
382	Toyota	Corolla LE 4dr	Sedan	36.0
383	Toyota	Echo 2dr manual	Sedan	39.0
384	Toyota	Echo 2dr auto	Sedan	36.0
385	Toyota	Echo 4dr	Sedan	39.0
405	Volkswagen	Jetta GLS TDI 4dr	Sedan	42.0

Average MPG by Car Type

The MEANS Procedure

Analysis Variable : AvgMPG				
Type	N Obs	Mean	Minimum	Maximum
Hybrid	3	55.5	48.5	63.0
SUV	60	18.3	11.0	24.5
Sedan	262	24.9	15.5	42.0
Sports	49	21.9	16.0	29.0
Truck	24	18.8	15.0	26.5
Wagon	30	24.5	17.0	33.0

```
* p102d02.sas Using a library to read excel files ;*

* Complete the option statement;
options validvarname=v7;

* Complete the libname statement;
libname xlstorm xlsx "/home/u58304328/EPG1V2/data/storm.xlsx";

* Complete the data= option to reference the STORM_SUMMARY worksheet;
proc contents data=xlstorm.storm_summary;
run;

* Clear the xlstorm library;
libname xlstorm clear;
```

The CONTENTS Procedure

Data Set Name	XLSTORM.storm_summary	Observations	.
Member Type	DATA	Variables	12
Engine	XLSX	Indexes	0
Created	.	Observation Length	0
Last Modified	.	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	Default		
Encoding	Default		

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
3	Basin	Char	2	$2.	$2.	Basin
8	EndDate	Num	8	DATE9.		EndDate
10	Hem_EW	Char	1	$1.	$1.	Hem EW
9	Hem_NS	Char	1	$1.	$1.	Hem NS
11	Lat	Num	8	BEST.		Lat
12	Lon	Num	8	BEST.		Lon
5	MaxWindMPH	Num	8	BEST.		MaxWindMPH
6	MinPressure	Num	8	BEST.		MinPressure
2	Name	Char	12	$12.	$12.	Name
1	Season	Num	8	BEST.		Season
7	StartDate	Num	8	DATE9.		StartDate
4	Type	Char	2	$2.	$2.	Type

```
************************************************************;
*  p102d03.sas Demo                                        *;
*  1) Complete the PROC IMPORT step to read                *;
*     STORM_DAMAGE.CSV and create a temporary SAS table    *;
*     named STORM_DAMAGE_IMPORT. Replace the table if it   *;
*     exists.                                              *;
*  2) Complete the PROC CONTENTS step to examine the       *;
*     properties of STORM_DAMAGE_IMPORT.                   *;
*  3) Highlight the demo program and submit the selected   *;
*     code.                                                *;
************************************************************;

*Complete the PROC IMPORT step;
proc import datafile="/home/u58304328/EPG1V2/data/storm_damage.csv" dbms=csv
                        out=storm_damage_import replace;
run;

*Complete the PROC CONTENTS step;
proc contents data=storm_damage_import;
run;
```

The CONTENTS Procedure

Data Set Name	WORK.STORM_DAMAGE_IMPORT	Observations	38
Member Type	DATA	Variables	4
Engine	V9	Indexes	0
Created	04/26/2021 22:49:48	Observation Length	808
Last Modified	04/26/2021 22:49:48	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	162
Obs in First Data Page	38
Number of Data Set Repairs	0
Filename	/saswork/SAS_workE72700018B2D_odaws04-usw2.oda.sas.com/SAS_workD01400018B2D_odaws04-usw2.oda.sas.com/storm_damage_import.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	1074796953
Access Permission	rw-r--r--
Owner Name	u58304328
File Size	256KB
File Size (bytes)	262144

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
4	Cost	Num	8	BEST12.	BEST32.
2	Date	Num	8	DATE9.	DATE9.
1	Event	Char	22	$22.	$22.
3	Summary	Char	764	$764.	$764.

```
************************************************************;
*  p102p01.sas LESSON 2, PRACTICE 1                        *;
*    a) Complete the PROC IMPORT step to read              *;
*      EU_SPORT_TRADE.XLSX. Create a SAS table named       *;
*      EU_SPORT_TRADE and replace the table                *;
*      if it exists.                                       *;
*    b) Modify the PROC CONTENTS code to display the       *;
*      descriptor portion of the EU_SPORT_TRADE table.     *;
*      Submit the program, and then view the output data   *;
*      and the results.                                    *;
************************************************************;

proc import datafile="/home/u58304328/EPG1V2/data/eu_sport_trade.xlsx" dbms=xlsx
                        out=eu_sport_trade_xlsx replace;
run;

proc contents data=eu_sport_trade_xlsx;
run;
```

The CONTENTS Procedure

Data Set Name	WORK.EU_SPORT_TRADE_XLSX	Observations	6816
Member Type	DATA	Variables	6
Engine	V9	Indexes	0
Created	04/26/2021 22:50:57	Observation Length	72
Last Modified	04/26/2021 22:50:57	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	4
First Data Page	1
Max Obs per Page	1816
Obs in First Data Page	1780
Number of Data Set Repairs	0
Filename	/saswork/SAS_workE72700018B2D_odaws04-usw2.oda.sas.com/SAS_workD01400018B2D_odaws04-usw2.oda.sas.com/eu_sport_trade_xlsx.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	1074796941
Access Permission	rw-r--r--
Owner Name	u58304328
File Size	640KB
File Size (bytes)	655360

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
6	Amt_Export	Num	8	COMMA15.		Amt_Export
5	Amt_Import	Num	8	COMMA15.		Amt_Import
3	Country	Char	37	$37.	$37.	Country
2	Geo_Code	Char	2	$2.	$2.	Geo_Code
1	Sport_Product	Char	7	$7.	$7.	Sport_Product
4	Year	Num	8	BEST.		Year

```
*************************************************************;
*  p103a02.sas Activity 3.02                                *;
*    1) Uncomment each WHERE statement one at a time and  *;
*       run the step to observe the rows that are        *;
*       included in the results.                         *;
*    2) Comment all previous WHERE statements. Add a new  *;
*       WHERE statement to print storms that begin with   *;
*       Z. How many storms are included in the results?   *;
*************************************************************;


proc print data=pg1.storm_summary(obs=50);
            *where MinPressure is missing; /*same as MinPressure = .*/
            *where Type is not missing; /*same as Type ne " "*/
            *where MaxWindMPH between 150 and 155;
            *where Basin like "_I";
            where Name like "Z%";
run;
```

Obs	Season	Name	Basin	Type	MaxWindMPH	MinPressure	StartDate	EndDate	Hem_NS	Hem_EW	Lat	Lon
665	1987	ZUMAN	SP	NR	69	975	22APR1987	26APR1987	S	W	-14.6	-168.70
922	1990	ZOLA	WP	ET	86	960	16AUG1990	23AUG1990	N	E	28.3	134.90
993	1991	ZEKE	WP	TS	75	970	09JUL1991	15JUL1991	N	E	17.0	112.60
994	1991	ZELDA	WP	ET	69	975	27NOV1991	06DEC1991	N	E	10.4	161.20
1089	1992	ZACK	WP	TS	46	992	08OCT1992	16OCT1992	N	E	25.0	153.20
1090	1992	ZEKE	EP	TS	52	1000	25OCT1992	30OCT1992	N	W	19.3	-114.90
1172	1993	ZOLA	WP	ET	58	985	05SEP1993	14SEP1993	N	E	29.6	132.40
1263	1994	ZEKE	WP	ET	58	985	17JUL1994	25JUL1994	N	E	29.5	154.80
1264	1994	ZELDA	WP	ET	121	910	26OCT1994	09NOV1994	N	E	18.6	138.90
1340	1995	ZACK	WP	TS	104	950	25OCT1995	02NOV1995	N	E	14.8	112.50
1416	1996	ZAKA	SP	NR	40	995	09MAR1996	11MAR1996	S	W	-21.5	168.70
1417	1996	ZANE	WP	ET	92	950	23SEP1996	05OCT1996	N	E	20.9	126.50
1506	1997	ZITA	WP	TS	63	980	20AUG1997	24AUG1997	N	E	21.0	110.00
1582	1998	ZEB	WP	ET	127	900	10OCT1998	20OCT1998	N	E	16.2	123.80
1583	1998	ZUMAN	SP	NR	92	955	29MAR1998	10APR1998	S	W	-15.3	166.60
1653	1999	ZELIA	SI	NR	46	990	07OCT1998	10OCT1998	S	E	-14.4	94.60
1654	1999	ZIA	WP	ET	52	985	11SEP1999	17SEP1999	N	E	31.3	131.20
1984	2003	ZOE	SP	NR	150	890	23DEC2002	04JAN2003	S	E	-12.5	169.50
2165	2005	ZETA	NA	DS	63	994	30DEC2005	07JAN2006	N	W	25.0	-38.60
2327	2007	ZITA	SP	NR	69	975	20JAN2007	25JAN2007	S	W	-17.6	-153.40
2665	2011	ZAKA	SP	NR	58	985	05FEB2011	08FEB2011	S	W	-25.7	-177.20
2666	2011	ZELIA	SP	NR	115	943	12JAN2011	18JAN2011	S	E	-18.5	158.05
2848	2013	ZANE	SP	NR	75	981	28APR2013	02MAY2013	S	E	-14.1	148.90
3118	2016	ZENA	SP	NR	81	975	05APR2016	07APR2016	S	W	-18.8	175.30

```
*************************************************************;
*  p103a05.sas Activity 3.05                                  *;
*    1) Highlight the PROC PRINT step and run the             *;
*       selected code. Notice how the values of Lat, Lon,     *;
*       StartDate, and EndDate are displayed in the           *;
*       report.                                               *;
*    2) Change the width of the DATE format to 7 and run      *;
*       the PROC PRINT step. How does the display of          *;
*       StartDate and EndDate change?                         *;
*    3) Change the width of the DATE format to 11 and run     *;
*       the PROC PRINT step. How does the display of          *;
*       StartDate and EndDate change?                         *;
*    4) Highlight the PROC FREQ step and run the selected     *;
*       code. Notice that the report includes the number      *;
*       of storms for each StartDate.                         *;
*    5) Add a FORMAT statement to apply the MONNAME.          *;
*       format to StartDate and run the PROC FREQ step.       *;
*       How many rows are in the report?                      *;
*************************************************************;

proc print data=pg1.storm_summary(obs=20);
        format Lat Lon 4. StartDate EndDate date7.;
run;

proc freq data=pg1.storm_summary order=freq;
        tables StartDate;
        *Add a FORMAT statement;
        format StartDate Monname.;
run;
```

Obs	Season	Name	Basin	Type	MaxWindMPH	MinPressure	StartDate	EndDate	Hem_NS	Hem_EW	Lat	Lon
1	1980		na	TS	35	.	17JUL80	18NOV80	N	W	26	-91
2	1980		SP	NR	.	998	27MAR80	30MAR80	S	E	19	137
3	1980	AGATHA	EP	TS	115	.	09JUN80	15JUN80	N	W	13	-119
4	1980	ALBINE	SI	ET	.	.	27NOV79	06DEC79	S	E	19	137
5	1980	ALEX	WP	TS	40	998	09OCT80	14OCT80	N	E	27	141
6	1980	ALLEN	NA	TS	190	899	31JUL80	11AUG80	N	W	22	-86
7	1980	AMY	SI	NR	132	915	04JAN80	12JAN80	S	E	-19	120
8	1980	BERENICE	SI	TS	.	.	15DEC79	21DEC79	S	E	-19	120
9	1980	BETTY	WP	ET	115	925	28OCT80	08NOV80	N	E	14	128
10	1980	BLAS	EP	TS	58	.	16JUN80	19JUN80	N	W	12	-111
11	1980	BONNIE	NA	ET	98	975	14AUG80	19AUG80	N	W	24	-39
12	1980	BRIAN	SI	NR	115	930	18JAN80	27JAN80	S	E	-18	103
13	1980	CARMEN	WP	TS	69	985	05APR80	07APR80	N	E	17	178
14	1980	CARY	WP	TS	52	996	28OCT80	02NOV80	N	E	16	115
15	1980	CELIA	EP	TS	75	.	25JUN80	29JUN80	N	W	15	-108
16	1980	CHARLEY	NA	TS	81	989	20AUG80	25AUG80	N	W	39	-67
17	1980	CLARA	SI	NR	69	980	21JAN80	29JAN80	S	E	-19	111
18	1980	DANIELLE	NA	TS	58	1004	04SEP80	07SEP80	N	W	29	-93
19	1980	DARBY	EP	TS	52	.	01JUL80	03JUL80	N	W	13	-120
20	1980	DEAN	SI	NR	127	930	27JAN80	04FEB80	S	E	-19	118

The FREQ Procedure

StartDate	Frequency	Percent	Cumulative Frequency	Cumulative Percent
August	483	15.49	483	15.49
September	482	15.46	965	30.95
July	351	11.26	1316	42.21
October	324	10.39	1640	52.60
January	250	8.02	1890	60.62
February	224	7.18	2114	67.80
June	198	6.35	2312	74.15
November	187	6.00	2499	80.15
December	183	5.87	2682	86.02
March	181	5.81	2863	91.82
May	131	4.20	2994	96.02

```
************************************************************;
*  p103a06.sas Activity 3.06                                *;
*    1) Modify the OUT= option in the PROC SORT statement *;
*       to create a temporary table named STORM_SORT.     *;
*    2) Complete the WHERE and BY statements to answer    *;
*       the following question: Which storm in the North  *;
*       Atlantic basin (NA or na) had the strongest       *;
*       MaxWindMPH?                                       *;
************************************************************;


proc sort data=pg1.storm_summary out=storm_sort;
          where Basin in ("NA" "na");
          by descending MaxWindMPH;
run;
```

Table: WORK.STORM_SORT | View: Column names | Filter: (none)

Columns

- Select all
- Season
- Name
- Basin
- Type
- MaxWindMPH
- MinPressure
- StartDate
- EndDate
- Hem_NS
- Hem_EW
- Lat
- Lon

Total rows: 488 Total columns: 12 Rows 1-100

	Season	Name	Basin	Type	MaxWindMPH	MinPressure	StartDate	EndDate	Hem_NS	Hem_
1	1980	ALLEN	NA	TS	190	899	31JUL1980	11AUG1980	N	W
2	1988	GILBERT	NA	ET	184	888	08SEP1988	20SEP1988	N	W
3	2005	WILMA	NA	ET	184	882	15OCT2005	26OCT2005	N	W
4	1998	MITCH	NA	ET	178	905	22OCT1998	09NOV1998	N	W
5	2005	RITA	NA	DS	178	895	18SEP2005	26SEP2005	N	W
6	1992	ANDREW	NA	TS	173	922	16AUG1992	28AUG1992	N	W
7	2005	KATRINA	NA	ET	173	902	23AUG2005	31AUG2005	N	W
8	2007	DEAN	NA	TS	173	905	13AUG2007	23AUG2007	N	W
9	2007	FELIX	NA	DS	173	929	31AUG2007	06SEP2007	N	W
10	2003	ISABEL	NA	ET	167	915	06SEP2003	20SEP2003	N	W
11	2004	IVAN	NA	TS	167	910	02SEP2004	24SEP2004	N	W
12	2016	MATTHEW	na	ET	167	934	28SEP2016	10OCT2016	N	W
13	1989	HUGO	NA	ET	161	918	10SEP1989	25SEP1989	N	W
14	2005	EMILY	NA	TS	161	929	11JUL2005	21JUL2005	N	W

```
*************************************************;
*    p103d01.sas Exploring Data with Procedures          *;
*************************************************;
* Syntax                          *;
*                                 *;
*   PROC PRINT DATA=input-table(OBS=n);       *;
*     VAR col-name(s);                *;
*   RUN;                            *;
*                                 *;
*   PROC MEANS DATA=input-table;            *;
*     VAR col-name(s);                *;
*   RUN;                            *;
*                                 *;
*   PROC UNIVARIATE DATA=input-table;         *;
*     VAR col-name(s);                *;
*   RUN;                            *;
*                                 *;
*   PROC FREQ DATA=input-table;             *;
*     TABLES col-name(s);               *;
*   RUN;                            *;
*************************************************;


proc print data=sashelp.cars(obs=10);
          var Make Model Type MSRP;
run;


proc means data=sashelp.cars;
          var EngineSize Horsepower MPG_City MPG_Highway;
run;
```

```
proc univariate data=sashelp.cars;
        var MPG_Highway;
run;

proc freq data=sashelp.cars;
        tables Origin Type DriveTrain;
run;

************************************************************;
*  Demo                                    *;
*    1) Complete the PROC PRINT statement to list the data in *;
*       PG1.STORM_SUMMARY. Print the first 10 observations.   *;
*       Highlight the step and run the selected code.        *;
*    2) Add a VAR statement to include only the following     *;
*       columns: Season, Name, Basin, MaxWindMPH, MinPressure,*;
*       StartDate, and EndDate. Add "list first 10 rows" as a *;
*       comment before the PROC PRINT statement. Run the step.*;
*    3) Copy the PROC PRINT step and paste it at the end of   *;
*       the program. Change PRINT to MEANS. Remove the OBS=   *;
*       data set option to analyze all observations. Modify   *;
*       the VAR statement to calculate summary statistics for *;
*       MaxWindMPH and MinPressure. Add "calculate summary    *;
*       statistics" as a comment before the PROC MEANS       *;
*       statement. Highlight the step and run the selected    *;
*       code.                                   *;
*    4) Copy the PROC MEANS step and paste it at the end of   *;
*       the program. Change MEANS to UNIVARIATE. Add "examine *;
*       extreme values" as a comment before the PROC         *;
```

```
*      UNIVARIATE statement. Highlight the step and run the  *;
*      selected code.                                        *;
*    5) Copy the PROC UNIVARIATE step and paste it at the end *:
*      of the program. Change UNIVARIATE to FREQ. Change the *;
*      VAR statement to a TABLES statement to produce        *;
*      frequency tables for Basin, Type, and Season. Add     *;
*      "list unique values and frequencies" as a comment     *;
*      before the PROC FREQ statement. Highlight the step    *;
*      and run the selected code.                            *;
************************************************************;

/* list first 10 rows */
proc print data=pg1.storm_summary (obs=10);
          var  Season Name Basin MaxWindMPH MinPressure StartDate EndDate;
run;

/* calculate summary statistics */
proc means data=pg1.storm_summary;
          var  MaxWindMPH MinPressure;
run;

/* examine extreme values */
proc univariate data=pg1.storm_summary;
          var  MaxWindMPH MinPressure;
run;

/* list unique values and frequencies */
proc freq data=pg1.storm_summary;
          table   Basin Type Season;
```

```
run;


/* list first 10 rows */
proc print data=pg1.np_summary (obs=10);
        var   Reg Type;
run;

/* calculate summary statistics */
proc means data=pg1.storm_summary;
        var  MaxWindMPH MinPressure;
run;

/* examine extreme values */
proc univariate data=pg1.np_summary;
        var  Acres;
run;

/* list unique values and frequencies */
proc freq data=pg1.np_summary;
        table   Reg Type;
run;


/* list unique values and frequencies */
proc freq data=pg1.np_species;
        table   Abundance Conservation_Status;
        WHERE  Species_ID like "YOSE%" and Category = "Mammal";
run;
```

```
/* list unique values and frequencies */
proc print data=pg1.np_species;
	var	Species_ID Category Scientific_Name Common_Names;
	WHERE  Species_ID like "YOSE%" and Category = "Mammal";
run;

%let ParkCode=YOSE;
%let SpeciesCat=Mammal;
proc freq data=pg1.np_species;
	table   Abundance Conservation_Status;
	WHERE  Species_ID like "&ParkCode%" and Category = "&SpeciesCat";
run;

/* list unique values and frequencies */
proc print data=pg1.np_species;
	var	Species_ID Category Scientific_Name Common_Names;
	WHERE  Species_ID like "&ParkCode%" and Category = "&SpeciesCat";
run;

%let ParkCode=ZION;
%let SpeciesCat=Bird;
proc freq data=pg1.np_species;
	table   Abundance Conservation_Status;
	WHERE  Species_ID like "&ParkCode%" and Category = "&SpeciesCat";
run;

/* list unique values and frequencies */
proc print data=pg1.np_species;
```

```
var     Species_ID Category Scientific_Name Common_Names;
WHERE  Species_ID like "&ParkCode%" and Category = "&SpeciesCat";
run;
```

Obs	Make	Model	Type	MSRP
1	Acura	MDX	SUV	$36,945
2	Acura	RSX Type S 2dr	Sedan	$23,820
3	Acura	TSX 4dr	Sedan	$26,990
4	Acura	TL 4dr	Sedan	$33,195
5	Acura	3.5 RL 4dr	Sedan	$43,755
6	Acura	3.5 RL w/Navigation 4dr	Sedan	$46,100
7	Acura	NSX coupe 2dr manual S	Sports	$89,765
8	Audi	A4 1.8T 4dr	Sedan	$25,940
9	Audi	A41.8T convertible 2dr	Sedan	$35,940
10	Audi	A4 3.0 4dr	Sedan	$31,840

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
EngineSize	Engine Size (L)	428	3.1967290	1.1085947	1.3000000	8.3000000
Horsepower		428	215.8855140	71.8360316	73.0000000	500.0000000
MPG_City	MPG (City)	428	20.0607477	5.2382176	10.0000000	60.0000000
MPG_Highway	MPG (Highway)	428	26.8434579	5.7412007	12.0000000	66.0000000

The UNIVARIATE Procedure
Variable: MPG_Highway (MPG (Highway))

Moments			
N	428	Sum Weights	428
Mean	26.8434579	Sum Observations	11489
Std Deviation	5.74120072	Variance	32.9613857
Skewness	1.25239527	Kurtosis	6.04561068
Uncorrected SS	322479	Corrected SS	14074.5117
Coeff Variation	21.3877092	Std Error Mean	0.27751141

Basic Statistical Measures			
Location		Variability	
Mean	26.84346	Std Deviation	5.74120
Median	26.00000	Variance	32.96139
Mode	26.00000	Range	54.00000
		Interquartile Range	5.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	96.7292	Pr > |t|	<.0001
Sign	M	214	Pr >= |M|	<.0001
Signed Rank	S	45903	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	66
99%	44
95%	36
90%	34
75% Q3	29
50% Median	26
25% Q1	24
10%	20
5%	18
1%	16
0% Min	12

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
12	167	44	156
13	119	46	405
14	252	51	150
16	217	51	374
16	216	66	151

The FREQ Procedure

Origin	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Asia	158	36.92	158	36.92
Europe	123	28.74	281	65.65
USA	147	34.35	428	100.00

Type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Hybrid	3	0.70	3	0.70
SUV	60	14.02	63	14.72
Sedan	262	61.21	325	75.93
Sports	49	11.45	374	87.38
Truck	24	5.61	398	92.99
Wagon	30	7.01	428	100.00

DriveTrain	Frequency	Percent	Cumulative Frequency	Cumulative Percent
All	92	21.50	92	21.50
Front	226	52.80	318	74.30
Rear	110	25.70	428	100.00

Obs	Season	Name	Basin	MaxWindMPH	MinPressure	StartDate	EndDate
1	1980		na	35	.	17JUL1980	18NOV1980
2	1980		SP	.	998	27MAR1980	30MAR1980
3	1980	AGATHA	EP	115	.	09JUN1980	15JUN1980
4	1980	ALBINE	SI	.	.	27NOV1979	06DEC1979
5	1980	ALEX	WP	40	998	09OCT1980	14OCT1980
6	1980	ALLEN	NA	190	899	31JUL1980	11AUG1980
7	1980	AMY	SI	132	915	04JAN1980	12JAN1980
8	1980	BERENICE	SI	.	.	15DEC1979	21DEC1979
9	1980	BETTY	WP	115	925	28OCT1980	08NOV1980
10	1980	BLAS	EP	58	.	16JUN1980	19JUN1980

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
MaxWindMPH	3095	79.3179321	31.6853937	6.0000000	213.0000000
MinPressure	2922	961.8545517	288.6582966	-9999.00	1012.00

The UNIVARIATE Procedure
Variable: MaxWindMPH

Moments			
N	3095	Sum Weights	3095
Mean	79.3179321	Sum Observations	245489
Std Deviation	31.6853937	Variance	1003.96417
Skewness	0.5963944	Kurtosis	-0.3710172
Uncorrected SS	22577945	Corrected SS	3106265.15
Coeff Variation	39.947327	Std Error Mean	0.56954597

Basic Statistical Measures			
Location		Variability	
Mean	79.31793	Std Deviation	31.68539
Median	75.00000	Variance	1004
Mode	52.00000	Range	207.00000
		Interquartile Range	52.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	139.2652	Pr > |t|	<.0001
Sign	M	1547.5	Pr >= |M|	<.0001
Signed Rank	S	2395530	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	213
99%	155
95%	138
90%	127
75% Q3	104
50% Median	75
25% Q1	52
10%	40
5%	40
1%	33
0% Min	6

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
6	2659	184	702
17	1960	184	1477
23	2757	184	2164
23	1366	190	6
23	1103	213	3017

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	23	0.74	100.00

The UNIVARIATE Procedure
Variable: MinPressure

Moments			
N	2922	Sum Weights	2922
Mean	961.854552	Sum Observations	2810539
Std Deviation	288.658297	Variance	83323.6122
Skewness	-37.521596	Kurtosis	1423.39373
Uncorrected SS	2946718001	Corrected SS	243388271
Coeff Variation	30.0105973	Std Error Mean	5.34003297

Basic Statistical Measures			
Location		Variability	
Mean	961.8546	Std Deviation	288.65830
Median	977.0000	Variance	83324
Mode	980.0000	Range	11011
		Interquartile Range	42.00000

Note: The mode displayed is the smallest of 2 modes with a count of 149.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	180.1215	Pr > |t|	<.0001
Sign	M	1459	Pr >= |M|	<.0001
Signed Rank	S	2129409	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1012
99%	1007
95%	1003
90%	1000
75% Q3	992
50% Median	977
25% Q1	950
10%	930
5%	916
1%	900
0% Min	-9999

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
-9999	3111	1009	2006
-9999	3109	1010	851
100	3117	1010	1103
872	3017	1010	3049
880	414	1012	2931

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	196	6.29	100.00

The FREQ Procedure

Basin	Frequency	Percent	Cumulative Frequency	Cumulative Percent
EP	671	21.52	671	21.52
NA	472	15.14	1143	36.66
NI	84	2.69	1227	39.35
SI	588	18.86	1815	58.21
SP	359	11.51	2174	69.72
WP	928	29.76	3102	99.49
na	16	0.51	3118	100.00

Type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
DS	293	9.40	293	9.40
ET	761	24.41	1054	33.80
NR	702	22.51	1756	56.32
SS	5	0.16	1761	56.48
TS	1357	43.52	3118	100.00

Season	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1980	81	2.60	81	2.60
1981	88	2.82	169	5.42
1982	81	2.60	250	8.02
1983	77	2.47	327	10.49
1984	92	2.95	419	13.44
1985	93	2.98	512	16.42
1986	82	2.63	594	19.05
1987	71	2.28	665	21.33
1988	74	2.37	739	23.70
1989	93	2.98	832	26.68
1990	90	2.89	922	29.57
1991	72	2.31	994	31.88
1992	96	3.08	1090	34.96
1993	82	2.63	1172	37.59
1994	92	2.95	1264	40.54
1995	76	2.44	1340	42.98

Obs	Reg	Type
1	A	NM
2	A	NP
3	A	NP
4	A	PRE
5	A	PRE
6	A	PRESERVE
7	IM	NM
8	IM	NM
9	IM	NM
10	IM	NM

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
MaxWindMPH	3095	79.3179321	31.6853937	6.0000000	213.0000000
MinPressure	2922	961.8545517	288.6582966	-9999.00	1012.00

The UNIVARIATE Procedure
Variable: Acres (Gross Acres)

Moments			
N	135	Sum Weights	135
Mean	285740.318	Sum Observations	38574942.9
Std Deviation	765563.33	Variance	5.86087E11
Skewness	5.38244467	Kurtosis	36.7127491
Uncorrected SS	8.95581E13	Corrected SS	7.85357E13
Coeff Variation	267.922754	Std Error Mean	65889.2006

Basic Statistical Measures			
Location		Variability	
Mean	285740.3	Std Deviation	765563
Median	33970.8	Variance	5.86087E11
Mode	.	Range	6587071
		Interquartile Range	209162

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	4.336679	Pr > |t|	<.0001
Sign	M	67.5	Pr >= |M|	<.0001
Signed Rank	S	4590	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	6587071.39
99%	3373063.14
95%	1542775.80
90%	761747.50
75% Q3	210281.92
50% Median	33970.84
25% Q1	1120.00
10%	281.78
5%	58.38
1%	1.00
0% Min	0.35

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.35	78	2219791	55
1.00	80	2523512	4
15.52	81	2697391	5
19.38	115	3373063	99
22.91	83	6587071	6

The FREQ Procedure

Region Code				
Reg	Frequency	Percent	Cumulative Frequency	Cumulative Percent
A	6	4.44	6	4.44
IM	52	38.52	58	42.96
MW	18	13.33	76	56.30
NC	1	0.74	77	57.04
NE	13	9.63	90	66.67
PW	23	17.04	113	83.70
SE	22	16.30	135	100.00

Type	Frequency	Percent	Cumulative Frequency	Cumulative Percent
NM	63	46.67	63	46.67
NP	51	37.78	114	84.44
NPRE	1	0.74	115	85.19
NS	10	7.41	125	92.59
PRE	3	2.22	128	94.81
PRESERVE	4	2.96	132	97.78
RIVERWAYS	1	0.74	133	98.52
RVR	2	1.48	135	100.00

The FREQ Procedure

Abundance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Abundant	1	7.14	1	7.14
Common	3	21.43	4	28.57
Rare	6	42.86	10	71.43
Uncommon	4	28.57	14	100.00
Frequency Missing = 2				

Conservation_Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Species of Concern	3	100.00	3	100.00
Frequency Missing = 13				

Obs	Species_ID	Category	Scientific_Name	Common_Names
17152	YOSE-1003	Mammal	Sus scrofa	Pig, Pig (Feral), Wild Boar, Wild Boar
17153	YOSE-1006	Mammal	Urocyon cinereoargenteus	Gray Fox
17154	YOSE-1007	Mammal	Vulpes vulpes necator	Sierra Nevada Red Fox
17155	YOSE-1014	Mammal	Martes americana	American Marten, Marten
17156	YOSE-1019	Mammal	Taxidea taxus	Badger
17157	YOSE-1023	Mammal	Ursus arctos	Brown Bear, Grizzly Bear
17158	YOSE-1033	Mammal	Myotis californicus	California Myotis
17159	YOSE-1038	Mammal	Myotis volans	Long-Legged Myotis, Long-Legged Myotis
17160	YOSE-1041	Mammal	Didelphis virginiana	Virginia Opossum
17161	YOSE-1056	Mammal	Peromyscus maniculatus	Deer Mouse, North American Deermouse
17162	YOSE-1058	Mammal	Phenacomys intermedius	Heather Vole, Western Heather Vole
17163	YOSE-1066	Mammal	Rattus rattus	Black Rat
17164	YOSE-1074	Mammal	Tamias merriami	Merriam's Chipmunk
17165	YOSE-1084	Mammal	Sorex palustris	Water Shrew
17166	YOSE-1085	Mammal	Sorex tenellus	Inyo Shrew
17167	YOSE-1087	Mammal	Scapanus latimanus	Broad-Footed Mole

The FREQ Procedure

Abundance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Abundant	1	7.14	1	7.14
Common	3	21.43	4	28.57
Rare	6	42.86	10	71.43
Uncommon	4	28.57	14	100.00
Frequency Missing = 2				

Conservation_Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Species of Concern	3	100.00	3	100.00
Frequency Missing = 13				

Obs	Species_ID	Category	Scientific_Name	Common_Names
17152	YOSE-1003	Mammal	Sus scrofa	Pig, Pig (Feral), Wild Boar, Wild Boar
17153	YOSE-1006	Mammal	Urocyon cinereoargenteus	Gray Fox
17154	YOSE-1007	Mammal	Vulpes vulpes necator	Sierra Nevada Red Fox
17155	YOSE-1014	Mammal	Martes americana	American Marten, Marten
17156	YOSE-1019	Mammal	Taxidea taxus	Badger
17157	YOSE-1023	Mammal	Ursus arctos	Brown Bear, Grizzly Bear
17158	YOSE-1033	Mammal	Myotis californicus	California Myotis
17159	YOSE-1038	Mammal	Myotis volans	Long-Legged Myotis, Long-Legged Myotis
17160	YOSE-1041	Mammal	Didelphis virginiana	Virginia Opossum
17161	YOSE-1056	Mammal	Peromyscus maniculatus	Deer Mouse, North American Deermouse
17162	YOSE-1058	Mammal	Phenacomys intermedius	Heather Vole, Western Heather Vole
17163	YOSE-1066	Mammal	Rattus rattus	Black Rat
17164	YOSE-1074	Mammal	Tamias merriami	Merriam's Chipmunk
17165	YOSE-1084	Mammal	Sorex palustris	Water Shrew
17166	YOSE-1085	Mammal	Sorex tenellus	Inyo Shrew
17167	YOSE-1087	Mammal	Scapanus latimanus	Broad-Footed Mole

The FREQ Procedure

Abundance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Abundant	1	3.45	1	3.45
Common	8	27.59	9	31.03
Occasional	8	27.59	17	58.62
Rare	4	13.79	21	72.41
Uncommon	8	27.59	29	100.00
Frequency Missing = 17				

Conservation_Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Endangered	1	14.29	1	14.29
In Recovery	1	14.29	2	28.57
Species of Concern	5	71.43	7	100.00
Frequency Missing = 39				

Obs	Species_ID	Category	Scientific_Name	Common_Names
17471	ZION-1094	Bird	Haliaeetus leucocephalus	Bald Eagle
17472	ZION-1109	Bird	Aythya collaris	Ring-Necked Duck
17473	ZION-1115	Bird	Clangula hyemalis	Long-Tailed Duck
17474	ZION-1117	Bird	Lophodytes cucullatus	Hooded Merganser
17475	ZION-1129	Bird	Calypte costae	Costa's Hummingbird
17476	ZION-1133	Bird	Selasphorus platycercus	Broad-Tailed Hummingbird
17477	ZION-1135	Bird	Antrostomus vociferus	Eastern Whip-Poor-Will
17478	ZION-1136	Bird	Chordeiles acutipennis	Lesser Nighthawk
17479	ZION-1137	Bird	Chordeiles minor	Common Nighthawk
17480	ZION-1144	Bird	Chlidonias niger	Black Tern
17481	ZION-1145	Bird	Chroicocephalus philadelphia	Bonaparte's Gull
17482	ZION-1146	Bird	Hydroprogne caspia	Caspian Tern
17483	ZION-1148	Bird	Larus californicus	California Gull
17484	ZION-1156	Bird	Calidris bairdii	Baird's Sandpiper
17485	ZION-1158	Bird	Calidris melanotos	Pectoral Sandpiper
17486	ZION-1173	Bird	Columbina inca	Inca Dove
17487	ZION-1181	Bird	Falco mexicanus	Prairie Falcon
17488	ZION-1201	Bird	Calcarius ornatus	Chestnut-Collared Longspur
17489	ZION-1211	Bird	Cinclus mexicanus	American Dipper
17490	ZION-1223	Bird	Artemisiospiza nevadensis	Sagebrush Sparrow
17491	ZION-1225	Bird	Chondestes grammacus	Lark Sparrow
17492	ZION-1226	Bird	Junco hyemalis caniceps	Dark-Eyed Junco (Gray-Headed)
17493	ZION-1228	Bird	Junco hyemalis mearnsi	Pink-Sided Junco
17494	ZION-1230	Bird	Melospiza georgiana	Swamp Sparrow
17495	ZION-1232	Bird	Melospiza melodia	Song Sparrow
17496	ZION-1233	Bird	Melozone aberti	Abert's Towhee
17497	ZION-1244	Bird	Zonotrichia atricapilla	Golden-Crowned Sparrow
17498	ZION-1247	Bird	Coccothraustes vespertinus	Evening Grosbeak
17499	ZION-1249	Bird	Leucosticte atrata	Black Rosy-Finch
17500	ZION-1263	Bird	Euphagus carolinus	Rusty Blackbird
17501	ZION-1266	Bird	Icterus cucullatus	Hooded Oriole
17502	ZION-1269	Bird	Quiscalus mexicanus	Great-Tailed Grackle
17503	ZION-1284	Bird	Cardellina pusilla	Wilson's Warbler
17504	ZION-1298	Bird	Setophaga graciae	Grace's Warbler
17505	ZION-1301	Bird	Setophaga ruticilla	American Redstart

```
************************************************************;
*  p103d02.sas Filtering Rows with Basic Operators          *;
************************************************************;
*  Syntax and Example                                       *;
*                                                           *;
*   WHERE expression;                                       *;
*                                                           *;
*   Basic Operators:                                        *;
*      = , EQ                                               *;
*      ^= , ~= , NE                                         *;
*      > , GT                                               *;
*      < , LT                                               *;
*      >= , GE                                              *;
*      <= , LE                                              *;
*   SAS Date Constant                                       *;
*      "ddmmmyyyy"d ("01JAN2015"d)                          *;
************************************************************;

proc print data=sashelp.cars;
          var Make Model Type MSRP MPG_City MPG_Highway;
          where Type="SUV" and MSRP <= 30000;
run;

************************************************************;
*  Demo                                                     *;
*   1) Write a PROC PRINT step to list the data in          *;
*     PG1.STORM_SUMMARY.                                    *;
*   2) Write a WHERE statement to include rows with         *;
*     MaxWindMPH values greater than or equal to 156        *;
```

```
*      (Category 5 storms). Highlight the PROC PRINT     *;
*      step and run the selected code.                   *;
*    3) Modify the WHERE statement for each of the       *;
*      conditions below. Highlight the PROC PRINT step   *;
*      and run the selected code after each condition.   *;
*      a) Basin equal to WP (West Pacific)               *;
*      b) Basin equal to SI or NI (South Indian or North *;
*        Indian)                                         *;
*      c) StartDate on or after January 1, 2010          *;
*      d) Type equal to TS (tropical storm) and Hem_EW   *;
*        equal to W (west)                               *;
*      e) MaxWindMPH greater than 156 or MinPressure     *;
*        less than 920                                   *;
*    4) In the final WHERE statement, are missing values  *;
*      included for MinPressure? How can you exclude     *;
*      missing values?                                   *;
************************************************************;

proc print data=pg1.storm_summary;
   *add WHERE statement;
          WHERE MaxWindMPH >= 156;
run;

proc print data=pg1.storm_summary;
   *add WHERE statement;
          WHERE Basin = "WP";
run;

proc print data=pg1.storm_summary;
```

```
   *add WHERE statement;
        WHERE Basin in ("SI" "NI");
run;

proc print data=pg1.storm_summary;
   *add WHERE statement;
        WHERE Startdate >= "01jan2010"d;
run;

proc print data=pg1.storm_summary;
   *add WHERE statement;
        WHERE Type = "TS" and HEM_EW = "W";
run;

proc print data=pg1.storm_summary;
   *add WHERE statement;
        WHERE MaxWindMPH > 156 or MinPressure < 920;
run;

proc print data=pg1.storm_summary;
   *add WHERE statement;
        WHERE MaxWindMPH > 156 or 0<MinPressure< 920;
run;
```

Obs	Make	Model	Type	MSRP	MPG_City	MPG_Highway
48	Buick	Rendezvous CX	SUV	$26,545	19	26
67	Chevrolet	Tracker	SUV	$20,255	19	22
121	Ford	Explorer XLT V6	SUV	$29,670	15	20
122	Ford	Escape XLS	SUV	$22,515	18	23
152	Honda	Pilot LX	SUV	$27,560	17	22
153	Honda	CR-V LX	SUV	$19,860	21	25
154	Honda	Element LX	SUV	$18,690	21	24
168	Hyundai	Santa Fe GLS	SUV	$21,589	20	26
189	Isuzu	Rodeo S	SUV	$20,449	17	21
202	Jeep	Grand Cherokee Laredo	SUV	$27,905	16	21
203	Jeep	Liberty Sport	SUV	$20,130	20	24
204	Jeep	Wrangler Sahara convertible 2dr	SUV	$25,520	16	19
205	Kia	Sorento LX	SUV	$19,635	16	19
218	Land Rover	Freelander SE	SUV	$25,995	18	21
241	Mazda	Tribute DX 2.0	SUV	$21,087	22	25
278	Mercury	Mountaineer	SUV	$29,995	16	21
289	Mitsubishi	Outlander LS	SUV	$18,892	21	27
301	Nissan	Pathfinder SE	SUV	$27,339	16	21
302	Nissan	Xterra XE V6	SUV	$20,939	17	20
320	Pontiac	Aztekt	SUV	$21,595	19	26
345	Saturn	VUE	SUV	$20,585	21	26
366	Suzuki	XL-7 EX	SUV	$23,699	18	22
367	Suzuki	Vitara LX	SUV	$17,163	19	22
376	Toyota	4Runner SR5 V6	SUV	$27,710	18	21
377	Toyota	Highlander V6	SUV	$27,930	18	24
379	Toyota	RAV4	SUV	$20,290	22	27

Obs	Season	Name	Basin	Type	MaxWindMPH	MinPressure	StartDate	EndDate	Hem_NS	Hem_EW	Lat	Lon
6	1980	ALLEN	NA	TS	190	899	31JUL1980	11AUG1980	N	W	21.8	-86.4
702	1988	GILBERT	NA	ET	184	888	08SEP1988	20SEP1988	N	W	19.7	-83.8
787	1989	HUGO	NA	ET	161	918	10SEP1989	25SEP1989	N	W	14.6	-54.6
1001	1992	ANDREW	NA	TS	173	922	16AUG1992	28AUG1992	N	W	25.4	-75.8
1196	1994	EMILIA	EP	TS	161	926	16JUL1994	25JUL1994	N	W	11.7	-149.6
1202	1994	GILMA	EP	TS	161	920	21JUL1994	31JUL1994	N	W	12.1	-143.3
1212	1994	JOHN	EP	ET	173	929	11AUG1994	10SEP1994	N	W	14.2	-155.1

```
************************************************************;
*  p103d03.sas Filtering Rows Using Macro Variables              *;
************************************************************;
*  Syntax and Example                          *;
*                                    *;
*   %LET macrovar=value;                        *;
*                                    *;
*   Usage:                                  *;
*   WHERE numvar=&macrovar;                     *;
*   WHERE charvar="&macrovar";                  *;
*   WHERE datevar="&macrovar"d;                  *;
************************************************************;

%let CarType=Wagon;

proc print data=sashelp.cars;
          where Type="&CarType";
          var Type Make Model MSRP;
run;

proc means data=sashelp.cars;
          where Type="&CarType";
          var MSRP MPG_Highway;
run;

proc freq data=sashelp.cars;
   where Type="&CarType";
          tables Origin Make;
run;
```

```
*************************************************************;
*  Demo                                  *;
*    1) Highlight the demo program and run the selected   *;
*       code.                            *;
*    2) Write three %LET statements to create macro       *;
*       variables named WindSpeed, BasinCode, and Date.   *;
*       Set the initial values of the variables to match  *;
*       the WHERE statement.             *;
*    3) Modify the WHERE statement to reference the macro *;
*       variables. Highlight the demo program and run the *;
*       selected code. Verify that the same results are   *;
*       produced.                        *;
*    4) Change the values of the macro variables to       *;
*       values that you select. Possible values for Basin *;
*       include NA, WP, SP, WP, NI, and SI. Highlight the *;
*       demo program and run the selected code.          *;
*************************************************************;

proc print data=pg1.storm_summary;
        where MaxWindMPH>=156 and Basin="NA" and StartDate>="01JAN2000"d;
        var Basin Name StartDate EndDate MaxWindMPH;
run;

proc means data=pg1.storm_summary;
        where MaxWindMPH>=156 and Basin="NA" and StartDate>="01JAN2000"d;
        var MaxWindMPH MinPressure;
run;
```

```
%let WindSpeed=156;
%let BasinCode=NA;
%let Date=01JAN2000;

proc print data=pg1.storm_summary;
        where MaxWindMPH>=&WindSpeed and Basin="&BasinCode" and StartDate>="&Date"d;
        var Basin Name StartDate EndDate MaxWindMPH;
run;

proc means data=pg1.storm_summary;
        where MaxWindMPH>=&WindSpeed and Basin="&BasinCode" and StartDate>="&Date"d;
        var MaxWindMPH MinPressure;
run;

%let WindSpeed=100;
%let BasinCode=SI;
%let Date=01JAN2010;

proc print data=pg1.storm_summary;
        where MaxWindMPH>=&WindSpeed and Basin="&BasinCode" and StartDate>="&Date"d;
        var Basin Name StartDate EndDate MaxWindMPH;
run;

proc means data=pg1.storm_summary;
        where MaxWindMPH>=&WindSpeed and Basin="&BasinCode" and StartDate>="&Date"d;
        var MaxWindMPH MinPressure;
run;
```

Obs	Type	Make	Model	MSRP
25	Wagon	Audi	A6 3.0 Avant Quattro	$40,840
26	Wagon	Audi	S4 Avant Quattro	$49,090
46	Wagon	BMW	325xi Sport	$32,845
90	Wagon	Chevrolet	Malibu Maxx LS	$22,225
105	Wagon	Chrysler	Pacifica	$31,230
140	Wagon	Ford	Focus ZTW	$17,475
141	Wagon	Ford	Taurus SE	$22,290
186	Wagon	Infiniti	FX35	$34,895
187	Wagon	Infiniti	FX45	$36,395
215	Wagon	Kia	Rio Cinco	$11,905
229	Wagon	Lexus	IS 300 SportCross	$32,455
275	Wagon	Mercedes-Benz	C240	$33,780
276	Wagon	Mercedes-Benz	E320	$50,670
277	Wagon	Mercedes-Benz	E500	$60,670
286	Wagon	Mercury	Sable GS	$22,595
299	Wagon	Mitsubishi	Lancer Sportback LS	$17,495
316	Wagon	Nissan	Murano SL	$28,739
330	Wagon	Pontiac	Vibe	$17,045
344	Wagon	Saab	9-5 Aero	$40,845
352	Wagon	Saturn	L300 2	$23,560
354	Wagon	Scion	xB	$14,165
364	Wagon	Subaru	Forester X	$21,445
365	Wagon	Subaru	Outback	$23,895
373	Wagon	Suzuki	Aerio SX	$16,497
401	Wagon	Toyota	Matrix XR	$16,695
414	Wagon	Volkswagen	Jetta GL	$19,005
415	Wagon	Volkswagen	Passat GLS 1.8T	$24,955
416	Wagon	Volkswagen	Passat W8	$40,235
427	Wagon	Volvo	V40	$26,135
428	Wagon	Volvo	XC70	$35,145

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MSRP		30	28840.53	11834.00	11905.00	60670.00
MPG_Highway	MPG (Highway)	30	27.9000000	4.4127558	19.0000000	36.0000000

The FREQ Procedure

Origin	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Asia	11	36.67	11	36.67
Europe	12	40.00	23	76.67
USA	7	23.33	30	100.00

Make	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Audi	2	6.67	2	6.67
BMW	1	3.33	3	10.00
Chevrolet	1	3.33	4	13.33
Chrysler	1	3.33	5	16.67
Ford	2	6.67	7	23.33
Infiniti	2	6.67	9	30.00
Kia	1	3.33	10	33.33
Lexus	1	3.33	11	36.67
Mercedes-Benz	3	10.00	14	46.67
Mercury	1	3.33	15	50.00
Mitsubishi	1	3.33	16	53.33
Nissan	1	3.33	17	56.67
Pontiac	1	3.33	18	60.00
Saab	1	3.33	19	63.33
Saturn	1	3.33	20	66.67
Scion	1	3.33	21	70.00
Subaru	2	6.67	23	76.67
Suzuki	1	3.33	24	80.00
Toyota	1	3.33	25	83.33
Volkswagen	3	10.00	28	93.33
Volvo	2	6.67	30	100.00

Obs	Basin	Name	StartDate	EndDate	MaxWindMPH
1946	NA	ISABEL	06SEP2003	20SEP2003	167
2024	NA	IVAN	02SEP2004	24SEP2004	167
2086	NA	EMILY	11JUL2005	21JUL2005	161
2113	NA	KATRINA	23AUG2005	31AUG2005	173
2144	NA	RITA	18SEP2005	26SEP2005	178
2164	NA	WILMA	15OCT2005	26OCT2005	184
2262	NA	DEAN	13AUG2007	23AUG2007	173
2269	NA	FELIX	31AUG2007	06SEP2007	173

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
MaxWindMPH	8	172.0000000	7.1113591	161.0000000	184.0000000
MinPressure	8	908.3750000	16.1416719	882.0000000	929.0000000

Obs	Basin	Name	StartDate	EndDate	MaxWindMPH
1946	NA	ISABEL	06SEP2003	20SEP2003	167
2024	NA	IVAN	02SEP2004	24SEP2004	167
2086	NA	EMILY	11JUL2005	21JUL2005	161
2113	NA	KATRINA	23AUG2005	31AUG2005	173
2144	NA	RITA	18SEP2005	26SEP2005	178
2164	NA	WILMA	15OCT2005	26OCT2005	184
2262	NA	DEAN	13AUG2007	23AUG2007	173
2269	NA	FELIX	31AUG2007	06SEP2007	173

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
MaxWindMPH	8	172.0000000	7.1113591	161.0000000	184.0000000
MinPressure	8	908.3750000	16.1416719	882.0000000	929.0000000

Obs	Basin	Name	StartDate	EndDate	MaxWindMPH
2525	SI	EDZANI	04JAN2010	15JAN2010	138
2534	SI	GELANE	15FEB2010	25FEB2010	127
2597	SI	BIANCA	21JAN2011	30JAN2011	109
2694	SI	FUNSO	17JAN2012	01FEB2012	127
2697	SI	GIOVANNA	07FEB2012	22FEB2012	121
2755	SI	ANAIS	10OCT2012	19OCT2012	115
2764	SI	CLAUDIA	02DEC2012	13DEC2012	109
2775	SI	FELLENG	25JAN2013	05FEB2013	104
2814	SI	NARELLE	05JAN2013	15JAN2013	121
2826	SI	RUSTY	21FEB2013	01MAR2013	104
2852	SI	AMARA	15DEC2013	27DEC2013	127
2855	SI	BEJISA	27DEC2013	07JAN2014	109
2858	SI	BRUCE	16DEC2013	24DEC2013	138
2859	SI	CHRISTINE	25DEC2013	01JAN2014	104
2860	SI	COLIN	09JAN2014	16JAN2014	127
2880	SI	GILLIAN	06MAR2014	26MAR2014	127
2887	SI	HELLEN	26MAR2014	05APR2014	144
2941	SI	BANSI	08JAN2015	19JAN2015	138
2963	SI	EUNICE	24JAN2015	02FEB2015	144
2979	SI	IKOLA	04APR2015	08APR2015	109
2985	SI	KATE	21DEC2014	04JAN2015	104
3018	SI	QUANG	27APR2015	01MAY2015	115
3050	SI	EMERAUDE	14MAR2016	23MAR2016	127
3052	SI	FANTALA	10APR2016	26APR2016	155
3113	SI	URIAH	09FEB2016	25FEB2016	127

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
MaxWindMPH	25	122.8000000	14.5143607	104.0000000	155.0000000
MinPressure	25	932.6400000	14.4478372	910.0000000	953.0000000

```
*************************************************************;
*  p103d04.sas Formatting Data Values in Results             *;
*************************************************************;
*  Syntax and Example                                        *;
*                                                            *;
*   FORMAT col-name(s) format;                               *;
*                                                            *;
*   <$>format-name<w>.<d>                                    *;
*                                                            *;
*   Common formats:                                          *;
*     dollar10.2 -> $12,345.67                               *;
*     dollar10.  -> $12,346                                  *;
*     comma8.1   -> 9,876.5                                  *;
*     date7.     -> 01JAN17                                  *;
*     date9.     -> 01JAN2017                                *;
*     mmddyy10.  -> 12/31/2017                               *;
*     ddmmyy8.   -> 31/12/17                                 *;
*************************************************************;

proc print data=pg1.class_birthdate;
          format Height Weight 3. Birthdate date9.;
run;

*************************************************************;
*  Demo                                                      *;
*   1) Write a PROC PRINT step to list the data in           *;
*     PG1.STORM_DAMAGE. Highlight the step and run the       *;
*     selected code.                                         *;
*   2) Add a FORMAT statement to apply the MMDDYY10.         *;
```

```
*      format to Date and DOLLAR16. to Cost. Highlight   *;
*      the step and run the selected code.              *;
*    3) Change the width of MMDDYY to 8 and DOLLAR to 14. *;
*      Highlight the step and run the selected code.     *;
*      Change MMDDYY to 6 and DOLLAR to 10. Highlight    *;
*      the step and run the selected cpgode again. What  *;
*      happens to the formatted values?                  *;
************************************************************;


*Write a PROC PRINT step and FORMAT statement;
proc print data=pg1.storm_damage;
            format date mmddyy10. cost dollar16.;
            *format date mmddyy8. cost dollar14.;
            *format date mmddyy6. cost dollar10.;
run;
```

Obs	Name	Sex	Age	Height	Weight	Birthdate
1	Alfred	M	14	69	113	26OCT2004
2	Alice	F	13	57	84	16NOV2005
3	Barbara	F	13	65	98	15JAN2005
4	Carol	F	14	63	103	04JUL2004
5	Henry	M	14	64	103	01DEC2004
6	James	M	12	57	83	15JUN2006
7	Jane	F	12	60	85	13MAR2006
8	Janet	F	15	63	113	02APR2003
9	Jeffrey	M	13	63	84	26APR2005
10	John	M	12	59	100	23AUG2006
11	Joyce	F	11	51	51	03JAN2007
12	Judy	F	14	64	90	05DEC2004
13	Louise	F	12	56	77	08AUG2006
14	Mary	F	15	67	112	26MAR2003
15	Philip	M	16	72	150	21NOV2002
16	Robert	M	12	65	128	06JUN2006
17	Ronald	M	15	67	133	14OCT2003
18	Thomas	M	11	58	85	18MAR2007
19	William	M	15	67	112	28DEC2003

Obs	Event	Date	Summary	Cost
1	Hurricane Katrina	08/25/2005	Category 3 hurricane initially impacts the U.S. as a Category 1 near Miami, FL, then as a strong Category 3 along the eastern LA-western MS coastlines, resulting in severe storm surge damage (maximum surge probably exceeded 30 feet) along the LA-MS-AL coasts, wind damage, and the failure of parts of the levee system in New Orleans. Inland effects included high winds and some flooding in the states of AL, MS, FL, TN, KY, IN, OH, and GA.	$161,300,000,000
2	Hurricane Harvey	08/25/2017	Category 4 hurricane made landfall near Rockport, Texas causing widespread damage. Harvey's devastation was most pronounced due to the large region of extreme rainfall producing historic flooding across Houston and surrounding areas. More than 30 inches of rainfall fell on 6.9 million people, while 1.25 million experienced over 45 inches and 11,000 had over 50 inches, based on 7-day rainfall totals ending August 31. This historic U.S. rainfall caused massive flooding that displaced over 30,000 people and damaged or destroyed over 200,000 homes and businesses.	$125,000,000,000
3	Hurricane Maria	09/19/2017	Category 4 hurricane made landfall in southeast Puerto Rico after striking the U.S. Virgin Island of St. Croix. Maria's high winds caused widespread devastation to Puerto Rico's transportation, agriculture, communication and energy infrastructure. Extreme rainfall up to 37 inches caused widespread flooding and mudslides across the island. The interruption to commerce and standard living conditions will be sustained for a long period, as much of Puerto Rico's infrastructure is rebuilt. Maria tied Hurricane Wilma (2005) for the most rapid intensification, strengthening from tropical depression to a category 5 storm in 54 hours. Maria's landfall at Category 4 strength gives the U.S. a record three Category 4+ landfalls this year (Maria, Harvey, and Irma).	$90,000,000,000
4	Hurricane Sandy	10/30/2012	Extensive damage across several northeastern states (MD, DE, NJ, NY, CT, MA, RI) due to high wind and coastal storm surge, particularly NY and NJ. Damage from wind, rain and heavy snow also extended more broadly to other states (NC, VA, WV, OH, PA, NH), as Sandy merged with a developing Nor'easter. Sandy's impact on major population centers caused widespread interruption to critical water / electrical services and also caused 159 deaths (72 direct, 87 indirect). Sandy also caused the New York Stock Exchange to close for two consecutive business days, which last happened in 1888 due to a major winter storm.	$70,900,000,000
5	Hurricane Irma	09/06/2017	Category 4 hurricane made landfall at Cudjoe Key, Florida after devastating the U.S. Virgin Islands - St John and St Thomas - as a category 5 storm. The Florida Keys were heavily impacted, as 25% of buildings were destroyed while 65% were significantly damaged. Severe wind and storm surge damage also occurred along the coasts of Florida and South Carolina. Jacksonville, FL and Charleston, SC received near-historic levels of storm surge causing significant coastal flooding. Irma maintained a maximum sustained wind of 185 mph for 37 hours, the longest in the satellite era. Irma also was a category 5 storm for longer than all other Atlantic hurricanes except Ivan in 2004.	$50,000,000,000
6	Hurricane Andrew	08/23/1992	Category 5 hurricane hits Florida and later impacts Louisiana as a category 3. High winds damage or destroy over 125,000 homes and leave at least 160,000 people homeless in Dade County, Florida alone. Initially rated as a category 4, Andrew was later upgraded to a category 5 upon further analysis. Andrew joins Hurricane Camille (1969) and the Labor Day Hurricane (1935), as the only land falling category 5 hurricanes on record to affect the U.S. mainland. Adjusted to present-day dollars, Andrew is the 3rd most costly hurricane to impact the U.S. since 1980, after Katrina (2005) and Sandy (2012).	$48,300,000,000

```
*************************************************************;
*  p103d05.sas Identifying and Removing Duplicate Values           *;
*************************************************************;
*  Syntax and Example                          *;
*                                              *;
*   Remove duplicate rows:                     *;
*   PROC SORT DATA=input-table <OUT=output-table>      *;
*      NODUPKEY <DUPOUT=output-table>;         *;
*      BY _ALL_;                               *;
*   RUN;                                       *;
*                                              *;
*   Remove duplicate key values:               *;
*   PROC SORT DATA=input-table <OUT=output-table>      *;
*      NODUPKEY <DUPOUT=output-table>;         *;
*      BY <DESCENDING> col-name (s);           *;
*   RUN;                                       *;
*************************************************************;

*************************************************************;
*  Demo                                        *;
*   1) Modify the first PROC SORT step to sort by all    *;
*      columns and remove any duplicate rows. Write the  *;
*      removed rows to a table named STORM_DUPS.         *;
*      Highlight the step and run the selected code.     *;
*      Confirm that there are 50,575 rows in STORM_CLEAN *;
*      and 7 rows in STORM_DUPS.                         *;
*   2) Run the second PROC SORT step and confirm that    *;
*      the first row for each storm represents           *;
*      the minimum value of Pressure.                    *;
```

```
*      Note: Because storm names can be reused in       *;
*            multiple years and basins, unique storms    *;
*            are grouped by sorting by Season, Basin,    *;
*            and Name.                                   *;
*    3) Modify the third PROC SORT step to sort the      *;
*      MIN_PRESSURE table and keep the first row for     *;
*      each storm. You do not need to keep the removed   *;
*      duplicates. Highlight the step and run the        *;
*      selected code.                                    *;
************************************************************;

*Step 1;
proc sort data=pg1.storm_detail out=storm_clean nodupkey dupout=storm_dups;
          by _all_;
run;

*Step 2;
proc sort data=pg1.storm_detail out=min_pressure;
          where Pressure is not missing and Name is not missing;
          by descending Season Basin Name Pressure;
run;

*Step 3;
proc sort data=min_pressure nodupkey;
          by descending Season Basin Name;
run;
```

Table: WORK.STORM_CLEAN | View: Column names | Filter: (none)

Columns

- Select all
- Season
- Basin
- Sub_basin
- Name
- ISO_time
- Type
- Latitude
- Longitude
- Wind
- Pressure
- Hem_NS
- Hem_EW
- Region

Total rows: 50757 Total columns: 13 Rows 1-100

	Season	Basin	Sub_basin	Name	ISO_time	Type	Latitude	Longitude	Wind
1	2000	EP	CP	DANIEL	29JUL2000:00:00:00.00	TS	18	-141.3	80
2	2000	EP	CP	DANIEL	29JUL2000:06:00:00.00	TS	18.4	-142.9	65
3	2000	EP	CP	DANIEL	29JUL2000:12:00:00.00	TS	18.7	-144.6	65
4	2000	EP	CP	DANIEL	29JUL2000:18:00:00.00	TS	18.9	-146.3	65
5	2000	EP	CP	DANIEL	30JUL2000:00:00:00.00	TS	19.3	-147.7	60
6	2000	EP	CP	DANIEL	30JUL2000:06:00:00.00	TS	19.8	-148.9	55
7	2000	EP	CP	DANIEL	30JUL2000:12:00:00.00	TS	20.1	-150	55
8	2000	EP	CP	DANIEL	30JUL2000:18:00:00.00	TS	20.5	-151	50
9	2000	EP	CP	DANIEL	31JUL2000:00:00:00.00	TS	20.7	-151.8	55
10	2000	EP	CP	DANIEL	31JUL2000:06:00:00.00	TS	20.8	-152.5	50
11	2000	EP	CP	DANIEL	31JUL2000:12:00:00.00	TS	21	-153.1	45
12	2000	EP	CP	DANIEL	31JUL2000:18:00:00.00	TS	21.3	-153.8	60
13	2000	EP	CP	DANIEL	01AUG2000:00:00:00.0(	TS	21.8	-154.5	60
14	2000	EP	CP	DANIEL	01AUG2000:06:00:00.0(	TS	22.3	-155.2	50