Zeru Zhou

Question 1

(a)

Code:

```
libname pg1 base "E:\SAS\PG1\data" ;
proc print data = pg1.np_summary (OBS = 7);
run;
```

Output:

			The SAS System							
Obs	Reg	Type	ParkName	DayVisits	OtherLodging	OtherCamping	TentCampers	RVCampers	BackcountryCampers	Acres
1	A	NM	Cape Krusenstern National Monument	15,000	0	0	0	0	6,375	649,096.15
2	A	NP	Kenai Fjords National Park	346,534	0	0	1,514	0	648	669,650.05
3	A	NP	Kobuk Valley National Park	15,500	0	0	0	0	7,050	1,750,716.16
4	A	PRE	Yukon-Charley Rivers National Preserve	1,146	0	0	0	0	3,063	2,523,512.44
5	A	PRE	Bering Land Bridge National Preserve	2,642	0	0	0	0	1,123	2,697,391.01
6	A	PRESERVE	Noatak National Preserve	17,000	0	0	0	0	5,500	6,587,071.39
7	IM	NM	Alibates Flint Quarries National Monument	8,153	0	0	0	0	0	1,370.97

(b)

Code:

```
libname pg1 base "E:\SAS\PG1\data" ;
proc print data = pg1.np_summary (OBS = 7);
var Reg ParkName Type;
run;
```

Output:

		The SAS System	
Obs	Reg	ParkName	Type
1	A	Cape Krusenstern National Monument	NM
2	A	Kenai Fjords National Park	NP
3	A	Kobuk Valley National Park	NP
4	A	Yukon-Charley Rivers National Preserve	PRE
5	A	Bering Land Bridge National Preserve	PRE
6	A	Noatak National Preserve	PRESERVE
7	IM	Alibates Flint Quarries National Monument	NM

There is inconsistency: in the Type column, observation #4,5 have type "PRE", but observation #6 has type "PRESERVE". The inconsistency is that the representations are distinct.

(c)

Code:

```
proc freq data = pg1.np_summary;
   tables Reg Type;
   run;
```

Output:

The SAS System

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

For Reg, "NC" only appears once.

For Type, "NPRE" and "RIVERWAYS" only appear once.

(d)

Code:

```
proc means data = pg1.np_summary;
   var DayVisits TentCampers RVCampers;
   run;
```

Output:

The SAS System

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
DayVisits	Recreational Day Visitors	135	966022.48	1568838.29	1146.00	11312786.00
TentCampers	Tent Campers	135	23870.81	60590.83	0	490431.00
RVCampers	RV Campers	135	14761.33	40977.10	0	376744.00

The minimum value for the number of recreational day visitors is 1146.

The minimum value for the number of tent campers is 0.

(e)

Codes:

```
proc univariate data = pg1.np_summary;
   var DayVisits;
   run;
```

Output:

The SAS System

The UNIVARIATE Procedure
Variable: DayVisits (Recreational Day Visitors)

Moments			
N	135	Sum Weights	135
Mean	966022.481	Sum Observations	130413035
Std Deviation	1568838.29	Variance	2.46125E12
Skewness	3.23070233	Kurtosis	14.5979115
Uncorrected SS	4.5579E14	Corrected SS	3.29808E14
Coeff Variation	162.40184	Std Error Mean	135024.101

Basic Statistical Measures			
Location		**Variability**	
Mean	966022.5	Std Deviation	1568838
Median	388290.0	Variance	2.46125E12
Mode	.	Range	11311640
		Interquartile Range	1026396

Tests for Location: Mu0=0				
Test		**Statistic**	**p Value**	
Student's t	t	7.154445	Pr > |t|	<.0001
Sign	M	67.5	Pr >= |M|	<.0001
Signed Rank	S	4590	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	11312786
99%	5969811
95%	4517585
90%	2946681
75% Q3	1102148
50% Median	388290
25% Q1	75752
10%	28646
5%	15555
1%	2642
0% Min	1146

Extreme Observations			
Lowest		**Highest**	
Value	Obs	Value	Obs
1146	4	4771309	134
2642	5	4812930	80
8153	7	5028868	111
11953	21	5969811	47
15000	1	11312786	126

Three lowest values: 1146, 2642, 8153.

Three highest values: 11312786, 5969811, 5028868.

(f)

Code:

```
proc print data = pg1.np_summary;
  where DayVisits = 11312786;
  run;
```

Output:

The SAS System

Obs	Reg	Type	ParkName	DayVisits	OtherLodging	OtherCamping	TentCampers	RVCampers	BackcountryCampers	Acres
126	SE	NP	Great Smoky Mountains National Park	11,312,786	11,493	0	190,574	111,680	109,349	522,426.88

Question 2

(a)

Code:

```
proc print data=pg1.np_summary;
    var Type ParkName;
    where ParkName like "% Preserve %";
  run;
```

Output:

The SAS System

Obs	Type	ParkName
4	PRE	Yukon-Charley Rivers National Preserve
5	PRE	Bering Land Bridge National Preserve
6	PRESERVE	Noatak National Preserve
58	PRESERVE	Big Thicket National Preserve
74	PRE	Tallgrass Prairie National Preserve
113	PRESERVE	Mojave National Preserve
127	NPRE	Little River Canyon National Preserve
135	PRESERVE	Big Cypress National Preserve

PRE, PRESERVE, NPRE are currently being used to denote Preserves.

(b)

Code:

```
proc print data=pg1.np_summary;
    var Type ParkName DayVisits;
    where ParkName like "% Preserve %";
    where (DayVisits >= 5000) and (DayVisits <= 500000);
run;
```

Log note:

```
55   proc print data=pg1.np_summary;
56       var Type ParkName DayVisits;
57       where ParkName like "% Preserve %";
58       where (DayVisits >= 5000) and (DayVisits <= 500000);
NOTE: WHERE clause has been replaced.
59   run;

NOTE: There were 74 observations read from the data set PG1.NP_SUMMARY.
      WHERE (DayVisits>=5000 and DayVisits<=500000);
NOTE: PROCEDURE PRINT used (Total process time):
      real time               0.02 seconds
      cpu time                0.01 seconds
```

(c)

Code:

```
proc print data=pg1.np_summary;
    var Type ParkName DayVisits;
    where ParkName like "% Preserve %" and ((DayVisits >= 5000) and (DayVisits <= 500000));
run;
```

Output:

The SAS System

Obs	Type	ParkName	DayVisits
6	PRESERVE	Noatak National Preserve	17,000
58	PRESERVE	Big Thicket National Preserve	192,809
74	PRE	Tallgrass Prairie National Preserve	29,378
127	NPRE	Little River Canyon National Preserve	462,700

Question 3

(a)

Code:

```
%let regcode = "IM";
```

(b)

Code:

```
proc means data = pg1.np_summary;
  where REG = &regcode;
  var ACRES ;
  run;
```

Log note:

```
68    %let regcode = "IM";
69    proc means data = pg1.np_summary;
70    where REG = &regcode;
71    var ACRES ;
72    run;

NOTE: There were 52 observations read from the data set PG1.NP_SUMMARY.
      WHERE REG='IM';
NOTE: PROCEDURE MEANS used (Total process time):
      real time               0.01 seconds
      cpu time                0.01 seconds
```

Output:

The SAS System

The MEANS Procedure

Analysis Variable : Acres Gross Acres				
N	Mean	Std Dev	Minimum	Maximum
52	163119.69	378927.78	160.0000000	2219790.71

(c)

Code:

```
%let regcode = "NE";

proc means data = pg1.np_summary;
   where REG = &regcode;
   var ACRES ;
   run;
```

Log note:

```
73     %let regcode = "NE";
74
75     proc means data = pg1.np_summary;
76     where REG = &regcode;
77     var ACRES ;
78     run;

NOTE: There were 13 observations read from the data set PG1.NP_SUMMARY.
      WHERE REG='NE';
NOTE: PROCEDURE MEANS used (Total process time):
      real time             0.01 seconds
      cpu time              0.00 seconds
```

Output:

The SAS System

The MEANS Procedure

Analysis Variable : Acres Gross Acres				
N	Mean	Std Dev	Minimum	Maximum
13	32766.47	55777.78	0.3500000	199195.27

(d)

Code:

```
proc means data = pg1.np_summary;
    by REG;
    var ACRES ;
    run;
```

Log note:

```
79    %let regcode = "NE";
80
81    proc means data = pg1.np_summary;
82    by REG;
83    var ACRES ;
84    run;

NOTE: There were 135 observations read from the data set PG1.NP_SUMMARY.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds
```

Question 4

(a)

Code:

```
libname pg1 base "E:\SAS\PG1\data";
proc sort data = pg1.np_summary out = np_sorted;
    where Type in ("NP" "NM");
    by Reg descending DayVisits ;
    run;
```

Log note:

```
1      libname pg1 base "E:\SAS\PG1\data";
NOTE: Libref PG1 was successfully assigned as follows:
      Engine:         BASE
      Physical Name: E:\SAS\PG1\data
2      proc sort data = pg1.np_summary out = np_sorted;
3      where Type in ("NP" "NM");
4      by Reg descending DayVisits ;
5      run;

NOTE: There were 114 observations read from the data set PG1.NP_SUMMARY.
      WHERE Type in ('NM', 'NP');
NOTE: The data set WORK.NP_SORTED has 114 observations and 10 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.03 seconds
      cpu time            0.00 seconds
```

(b)

Code:

```
proc print data = work.np_sorted (OBS = 15);
    var Reg Type DayVisits ParkName;
run;
```

Output:

The SAS System

Obs	Reg	Type	DayVisits	ParkName
1	A	NP	346,534	Kenai Fjords National Park
2	A	NP	15,500	Kobuk Valley National Park
3	A	NM	15,000	Cape Krusenstern National Monument
4	IM	NP	5,969,811	Grand Canyon National Park
5	IM	NP	4,517,585	Rocky Mountain National Park
6	IM	NP	4,295,127	Zion National Park
7	IM	NP	4,257,177	Yellowstone National Park
8	IM	NP	3,270,076	Grand Teton National Park
9	IM	NP	2,946,681	Glacier National Park
10	IM	NP	2,365,110	Bryce Canyon National Park
11	IM	NP	1,585,718	Arches National Park
12	IM	NP	1,064,904	Capitol Reef National Park
13	IM	NM	899,676	Cedar Breaks National Monument
14	IM	NM	821,406	Canyon De Chelly National Monument
15	IM	NP	820,426	Saguaro National Park

Question 5

(a)

Code:

```
proc contents data = pg1.np_westweather ;
  run;
```

The format of the **DATE** variable is **YYMMDD10.**

(b)

Code:

```
proc print data = pg1.np_westweather (OBS = 5);
  format DATE DATE9. SNOW SNOWDEPTH 6.2 ;
  run;
```

Output:

Obs	STATION	NAME	UNITCODE	Year	Month	DATE	EVAP	EVAPMIN	EVAPMAX	PRECIP	SNOW	SNOWDEPTH	TEMPMAX	TEMPMIN	FOG	THUNDER	ICE	HAIL	RIME
1	USC00429717	ZION NATIONAL PARK, UT US	ZION	2015	1	01JAN2015	.	.	.	0.28	4.00	2.00	35	13	.	.	.	.	.
2	USC00429717	ZION NATIONAL PARK, UT US	ZION	2015	1	02JAN2015	.	.	.	0	0.00	0.00	40	7	.	.	.	.	.
3	USC00429717	ZION NATIONAL PARK, UT US	ZION	2015	1	03JAN2015	.	.	.	0	0.00	0.00	45	13	.	.	.	.	.
4	USC00429717	ZION NATIONAL PARK, UT US	ZION	2015	1	04JAN2015	.	.	.	0	0.00	0.00	50	17	.	.	.	.	.
5	USC00429717	ZION NATIONAL PARK, UT US	ZION	2015	1	05JAN2015	.	.	.	0	0.00	0.00	56	26	.	.	.	.	.

The SAS System

Question 6

Code:

```
proc sort data = pg1.np_westweather out = newyearsdays nodupkey dupout = others;
  by Name Year;
  run;
```

Log output:

```
14    proc sort data = pg1.np_westweather out = newyearsdays nodupkey dupout = others;
15    by Name Year;
16    run;

NOTE: There were 4355 observations read from the data set PG1.NP_WESTWEATHER.
NOTE: 4343 observations with duplicate key values were deleted.
NOTE: The data set WORK.NEWYEARSDAYS has 12 observations and 19 variables.
NOTE: The data set WORK.OTHERS has 4343 observations and 19 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds
```

There are 12 observations in table newyearsdays!