**STA502 Homework 7**

**Lina Lee**

**Question 1**

/\*header

Purpose: report statistics by year or quarter of gasoline price.

in 2009.

Input: gas.sas7bdat\*/

%let data\_folder=C:\Users\linal\Desktop\2018\STA502WHW7;

ods rtf file="C:\Users\linal\Desktop\2018\STA502\HW7\problem1" bodytitle style=journal;

/\*a, b\*/

proc means data=gas min max mean maxdec=2;

var GasPrice /\* Label: U.S. average price of unleaded regular gasoline (per gallon) , variable type: numeric\*/;

by year;

output out=gas2 min=gasmin max=gasmax mean=gasmean;

run;

proc report data=gas2;

title "the gasolin price statistics";

column year gasmin gasmax gasmean;

define gasmin/display "Minimum" format=4.2;

define gasmax/display "Maximum" format=4.2;

define gasmean/display "Mean" format=4.2;

run;

/\*c\*/

data gas3;

set gas;

do i=1to 3;

if month=i then quarter=1;

end;

do i=4 to 6;

if month=i then quarter=2;

end;

do i=7 to 9;

if month=i then quarter=3;

end;

do i=10 to 12;

if month=i then quarter=4;

end;

run;

proc means data=gas3;

var GasPrice;

by year quarter ;

output out=gas4 min=gasmin max=gasmax mean=gasmean;

run;

proc report data=gas4;

title "the gasolin price statistics";

column year quarter gasmin gasmax gasmean;

define year /display order;

define quarter /display order;

define gasmin/display "Minimum" format=10.3;

define gasmax/display "Maximum" format=10.3;

define gasmean/display "Mean" format=dollar10.3;

run;

ods rtf close;

***1\_Result:***

1. comment on above sas code

***(b) the gasoline price statistics***

| *Year* | *Minimum* | *Maximum* | *Mean* |
| --- | --- | --- | --- |
| 2003 | 1.47 | 1.75 | 1.59 |
| 2004 | 1.59 | 2.04 | 1.88 |
| 2005 | 1.82 | 2.93 | 2.30 |
| 2006 | 2.24 | 3.00 | 2.59 |
| 2007 | 2.27 | 3.13 | 2.80 |
| 2008 | 1.69 | 4.09 | 3.27 |
| 2009 | 1.79 | 2.66 | 2.35 |
| 2010 | 2.66 | 2.99 | 2.79 |
| 2011 | 3.09 | 3.93 | 3.53 |
| 2012 | 3.33 | 3.93 | 3.64 |
| 2013 | 3.25 | 3.74 | 3.53 |
| 2014 | 2.56 | 3.70 | 3.37 |

***(c)the gasolin price statistics***

| *Year* | *quarter* | *Minimum* | *Maximum* | *Mean* |
| --- | --- | --- | --- | --- |
| 2003 | 1 | 1.473 | 1.748 | $1.621 |
|  | 2 | 1.514 | 1.659 | $1.572 |
|  | 3 | 1.524 | 1.728 | $1.627 |
|  | 4 | 1.494 | 1.603 | $1.544 |
| 2004 | 1 | 1.592 | 1.766 | $1.677 |
|  | 2 | 1.833 | 2.041 | $1.961 |
|  | 3 | 1.891 | 1.939 | $1.909 |
|  | 4 | 1.882 | 2.029 | $1.974 |
| 2005 | 1 | 1.823 | 2.065 | $1.935 |
|  | 2 | 2.176 | 2.283 | $2.225 |
|  | 3 | 2.316 | 2.927 | $2.583 |
|  | 4 | 2.186 | 2.785 | $2.438 |
| 2006 | 1 | 2.310 | 2.401 | $2.342 |
|  | 2 | 2.757 | 2.947 | $2.874 |
|  | 3 | 2.589 | 2.999 | $2.858 |
|  | 4 | 2.241 | 2.334 | $2.282 |
| 2007 | 1 | 2.274 | 2.592 | $2.384 |
|  | 2 | 2.860 | 3.130 | $3.014 |
|  | 3 | 2.782 | 2.961 | $2.844 |
|  | 4 | 2.793 | 3.069 | $2.961 |
| 2008 | 1 | 3.033 | 3.258 | $3.113 |
|  | 2 | 3.441 | 4.065 | $3.757 |
|  | 3 | 3.698 | 4.090 | $3.858 |
|  | 4 | 1.689 | 3.173 | $2.338 |
| 2009 | 1 | 1.787 | 1.949 | $1.888 |
|  | 2 | 2.056 | 2.631 | $2.317 |
|  | 3 | 2.543 | 2.627 | $2.581 |
|  | 4 | 2.561 | 2.660 | $2.614 |
| 2010 | 1 | 2.659 | 2.780 | $2.723 |
|  | 2 | 2.736 | 2.869 | $2.821 |
|  | 3 | 2.704 | 2.745 | $2.728 |
|  | 4 | 2.795 | 2.985 | $2.877 |
| 2011 | 1 | 3.091 | 3.546 | $3.268 |
|  | 2 | 3.702 | 3.933 | $3.817 |
|  | 3 | 3.612 | 3.654 | $3.632 |
|  | 4 | 3.278 | 3.468 | $3.390 |
| 2012 | 1 | 3.399 | 3.868 | $3.613 |
|  | 2 | 3.552 | 3.927 | $3.757 |
|  | 3 | 3.451 | 3.856 | $3.671 |
|  | 4 | 3.331 | 3.786 | $3.535 |
| 2013 | 1 | 3.351 | 3.735 | $3.593 |
|  | 2 | 3.590 | 3.633 | $3.615 |
|  | 3 | 3.556 | 3.628 | $3.595 |
|  | 4 | 3.251 | 3.375 | $3.301 |
| 2014 | 1 | 3.320 | 3.532 | $3.405 |
|  | 2 | 3.659 | 3.695 | $3.682 |
|  | 3 | 3.403 | 3.633 | $3.506 |
|  | 4 | 2.560 | 3.182 | $2.876 |

**Question 2**

**SAS CODE**

%let data\_folder=C:\Users\linal\Desktop\2018\STA502WHW7;

ods rtf file="C:\Users\linal\Desktop\2018\STA502\HW7\problem2" bodytitle style=journal;

/\*header

Purpose: practice using RETAIN with data in countries across the world regarding the outbreak of swine u cases and deaths

in 2009.

Input: sff.sas7bdat\*/

/\* character variable- Name:Country, length:30 Name:Continent, length:13\*/

/\*(b)\*/

data sff\_1 ;

set sff;

retain cum\_sum;

cum\_sum=sum (cum\_sum,Oct\_d);

run;

data sff2;

set sff\_1;

keep cum\_sum Oct\_d;

run;

proc print data=sff\_1(keep=Oct\_d Continent cum\_sum FIRSTOBS=75 obs=80) ;

title " the cumulative sum of the number of cumulative deaths reported on the first day of the month

for Oc-tober across all countries";

run;

/\*c\*/

proc sort data=sff;

by continent;

run;

data sff\_count;

set sff;

by Continent;

retain ctycount;

if first.Continent then do;

ctycount=0;

end;

ctycount=ctycount+1;

if last.continent then output;

run;

proc print data=sff\_count(keep=Continent ctycount) ;

title "the number of countries within each continent";

run;

/\*d\*/

data sff\_report;

set sff;

if FirstDeath ne . and FirstCase =. then report=1;

run;

data sff\_report2 (keep=Continent country FirstCase Latest FirstDeath ) ;

set sff\_report;

if report ne 1 than delete;

run;

/\*(e)\*/

data sff\_nocase;

set sff;

if aug ne . then delete;

run;

proc sort data=sff\_nocase;

by continent;

run;

data sff\_count\_nocase(keep=continent ctycount\_aug\_no);

set sff\_nocase;

by Continent;

retain ctycount\_aug\_no;

if first.Continent then do;

ctycount\_aug\_no=0;

end;

ctycount\_aug\_no=ctycount\_aug\_no+1;

if last.continent then output;

label ctycount\_aug\_no="Number of countries reported no cases";

run;

data sff\_acase;

set sff;

if aug=. then delete;

run;

proc sort data=sff\_acase;

by continent;

run;

data sff\_count\_acase(keep=continent ctycount\_aug);

set sff\_acase;

by Continent;

retain ctycount\_aug;

if first.Continent then do;

ctycount\_aug=0;

end;

ctycount\_aug=ctycount\_aug+1;

if last.continent then output;

label ctycount\_aug="Number of countries reported cases";

run;

proc sort data=sff;

by continent;

run;

data sff\_cum(keep=continent cum\_sum);

set sff;

by continent;

retain cum\_sum;

if first.Continent then do;

cum\_sum=0;

end;

cum\_sum=sum (cum\_sum,aug);

if last.continent then output;

label cum\_sum="the summation of cumulative cases per continent";

run;

data merge1;

merge sff\_count\_acase sff\_count\_nocase sff\_cum;

by continent;

run;

proc print data=merge1 label;

run;

ods rtf close;

***2\_Result:***

***(a) comment on the above SAS code.***

1. ***the cumulative sum of the number of cumulative deaths reported on the first day of the monthfor Oc-tober across all countries***

| *Obs* | *Oct\_d* | *Continent* | *cum\_sum* |
| --- | --- | --- | --- |
| *75* | 18 | Australia | 1121 |
| *76* | . | Australia | 1121 |
| *77* | 2 | Australia | 1123 |
| *78* | 1 | Australia | 1124 |
| *79* | 1 | Australia | 1125 |
| *80* | . | Australia | 1125 |

***(c) the number of countries within each continent***

| *Obs* | *Continent* | *ctycount* |
| --- | --- | --- |
| *1* | Africa | 24 |
| *2* | Asia | 40 |
| *3* | Australia | 16 |
| *4* | Europe | 50 |
| *5* | North America | 35 |
| *6* | South America | 14 |

***(d) new dataset containing countries that reported a first death data, but reported no first case date***

| *Obs* | *Country* | *FirstCase* | *Latest* | *FirstDeath* | *Continent* |
| --- | --- | --- | --- | --- | --- |
| *1* | Madagascar | . | . | 18151 | Africa |
| *2* | Mozambique | . | . | 18156 | Africa |
| *3* | São Tomé and Príncipe | . | . | 18196 | Africa |
| *4* | Mongolia | . | . | 18196 | Asia |
| *5* | Belarus | . | . | 18207 | Europe |

***(e) the number of countries reported case or reported no case***

| *Obs* | *Continent* | *Number of countries reported cases* | *Number of countries reported no cases* | *the summation of cumulative cases per continent* |
| --- | --- | --- | --- | --- |
| *1* | Africa | 20 | 4 | 972 |
| *2* | Asia | 39 | 1 | 35932 |
| *3* | Australia | 13 | 3 | 25244 |
| *4* | Europe | 48 | 2 | 29725 |
| *5* | North America | 34 | 1 | 74408 |
| *6* | South America | 13 | 1 | 24145 |