**SAS code**

**/\*Problem Number 1\*/**

/\*header: print observations and provide statistics.

Author: Lina Lee

Purpose: read text file sllep.txt, print the first three observations from this dataset. And provide a table consisting of mean,standard deviation and the 5 number summary statistics.

input data:sleep.txt

input variable: Species

BodyWt

BrainWt

NonDreaming

Dreaming

TotalSleep

\*/

**/\*1(a)read data\*/**

data sleepinfile;

infile "C:\Users\leel3\HW1\sleep.txt" firstobs=2 ;

length Species $26;

input Species & BodyWt BrainWt NonDreaming Dreaming TotalSleep LifeSpan Gestation Predation Exposure Danger;

**/\*1(a)reassign missing code from -999 to.\*/**

if Species=-999 then Species=.;

if BodyWt=-999 then BodyWt=.;

if BrainWt=-999 then BrainWt=.;

if NonDreaming=-999 then NonDreaming=.;

if Dreaming=-999 then Dreaming=.;

if TotalSleep=-999 then TotalSleep=.;

if LifeSpan=-999 then LifeSpan=.;

if Gestation=-999 then Gestation=.;

if Predation=-999 then Predation=.;

if Exposure=-999 then Exposure=.;

if Danger=-999 then Danger=.;

run;

ods rtf file="C:\Users\leel3\HW1\problem1" bodytitle style=journal;

title "EX1 sleep of mammals";

**/\*1(a) print the firsrt three obs\*/**

proc print data=sleepinfile(obs=3);

**/\*1(b)provide table consisting of mean,std and 5 number summary for the gestation time\*/**

proc means data=sleepinfile mean std min Q1 median Q3 max maxdec=2;

var gestation;

run;

ods rtf close;

**/\*Problem Number 2\*/**

/\*header: provide table with statistics and finding the the estimate of parameter for simple linear regression model.

Author: Lina Lee

Purpose: provide a table consisting of mean, median, mode, standard deviation, variance, range and interquartile range. Fit a simple regression model using oxygen intake rate as responsible variable and time to run 1.5 miles as predictor variable and finding the estimated intercept and slope of the regression function

input data: written in datalines

input variable: Age

Weight

Oxygen

RunTime

RestPulse

RunPulse

MaxPulse

\*/

**/\*2(a)read data into SAS\*/**

data fitness;

input Age Weight Oxygen RunTime RestPulse RunPulse MaxPulse @@;

datalines;

44 89.47 44.609 11.37 62 178 182 40 75.07 45.313 10.07 62 185 185

44 85.84 54.297 8.65 45 156 168 42 68.15 59.571 8.17 40 166 172

38 89.02 49.874 9.22 55 178 180 47 77.45 44.811 11.63 58 176 176

40 75.98 45.681 11.95 70 176 180 43 81.19 49.091 10.85 64 162 170

44 81.42 39.442 13.08 63 174 176 38 81.87 60.055 8.63 48 170 186

44 73.03 50.541 10.13 45 168 168 45 87.66 37.388 14.03 56 186 192

45 66.45 44.754 11.12 51 176 176 47 79.15 47.273 10.60 47 162 164

54 83.12 51.855 10.33 50 166 170 49 81.42 49.156 8.95 44 180 185

51 69.63 40.836 10.95 57 168 172 51 77.91 46.672 10.00 48 162 168

48 91.63 46.774 10.25 48 162 164 49 73.37 50.388 10.08 67 168 168

57 73.37 39.407 12.63 58 174 176 54 79.38 46.080 11.17 62 156 165

52 76.32 45.441 9.63 48 164 166 50 70.87 54.625 8.92 48 146 155

51 67.25 45.118 11.08 48 172 172 54 91.63 39.203 12.88 44 168 172

51 73.71 45.790 10.47 59 186 188 57 59.08 50.545 9.93 49 148 155

49 76.32 48.673 9.40 56 186 188 48 61.24 47.920 11.50 52 170 176

52 82.78 47.467 10.50 53 170 172

;

run;

ods rtf file="C:\Users\leel3\HW1\problem2" bodytitle style=journal;

title "EX2\_Aerobic Fitness prediction";

**/\*2(b)summary statistics for heart rate while resting\*/**

proc univariate data=fitness ;

var RestPulse 4.;

run;

**/\*2(c) the estimated intercept and slope of the regression function\*/**

proc reg data=fitness;

model oxygen=RunTime;

run;

ods rtf close;

**/\*Problem Number 3\*/**

/\*header: print observations and count the number of people who will attend Thursday lunch. Count the number for each group(“regular”,”Early:,On-site).

Author: Lina Lee

Purpose: print 2,3, and 4th observations and count the number of people who will attend Thursday lunch. After generate three groups(“regular”,”Early:,On-site), count the number for each group.

input data: conference.txt

input variable: FirstName LastName

AttendID : Attendee ID

BusiPhone: Business Phone Number

HPhone: Home Phone Number

Mphone:Mobile Phone Number

ContactBusi: OK to contact attendee at business

ContactHome: OK to contact at home

ContactM: OK to contact at Mobile

RegiRate: registration rate

WedMix: will attend Wednesday Mixer

ThursLunch: will attend Thursday Lunch

Volunteer: whether willing to volunteer at the conference

**/\*3(a)read data\*/**

data conference;

infile "C:\Users\leel3\HW1\conference.txt";

input FirstName :$9. LastName :$8. AttendID $ BusiPhone :$13. HPhone :$13. Mphone :$13. ContactBusi $ ContactHome $ ContactM $ RegiRate WedMix $ ThursLunch $ Volunteer $;

**/\*make categories\*/**

if regirate=350 then groups1="regular";

else if regirate=200 then groups1="regular";

else if regirate=450 then groups1="regular";

else if regirate=295 then groups1="early";

else if regirate=150 then groups1="early";

else if regirate=395 then groups1="early";

else if regirate=550 then groups1="on-site";

run;

ods rtf file="C:\Users\leel3\HW1\problem3-groups1" bodytitle style=journal;

title "EX3\_conference";

**/\*3(a)print the 2,3 and 4th obs\*/**

proc print data=conference (firstobs=2 obs=4);

run;

**/\*3(b)count the number of people for Thursday Lunch\*/**

Proc Freq data=conference;

tables ThursLunch;

run;

**/\*3(c) count the number of people for each group\*/**

proc Freq data=conference;

tables groups1;

run;

ods rtf close;

***EX 1 (a) sleep of mammals***

| *Obs* | *Species* | *BodyWt* | *BrainWt* | *Non*  *Dreaming* | *Dreaming* | *TotalSleep* | *LifeSpan* | *Gestation* | *Predation* | *Exposure* | *Danger* |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *1* | African elephant | 6654.00 | 5712.0 | . | . | 3.3 | 38.6 | 645 | 3 | 5 | 3 |
| *2* | African giant pouched rat | 1.00 | 6.6 | 6.3 | 2 | 8.3 | 4.5 | 42 | 3 | 1 | 3 |
| *3* | Arctic Fox | 3.39 | 44.5 | . | . | 12.5 | 14.0 | 60 | 1 | 1 | 1 |

**Answersr for EX1(a):** the first threes observation includes African elephant, African giant pouched rat, Arctic Fox. Their body weights are 6654, 1.00 and 3.39 each. Their brain weight are 5712.0, 6.6 and 44.5 each. Totalsleep of them are 3.3,8.3 and 12.5 each.

***EX 1 (b) sleep of mammals***

|  |
| --- |
| ***The MEANS Procedure*** |

| *Analysis Variable : Gestation* | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| *Mean* | *Std Dev* | *Minimum* | *Lower Quartile* | *Median* | *Upper Quartile* | *Maximum* |
| 142.35 | 146.81 | 12.00 | 35.00 | 79.00 | 210.00 | 645.00 |

**Answer for EX 1(b):** Mean for gestations is 142.35, and Standard Deviation is 146.81. Minimum is 12.00, Lower Quartile is 35.00, Median is 79.00, Upper Quartile is 210.00 and Maximum is 645.00

***EX 2 (b)\_Aerobic Fitness prediction***

|  |
| --- |
| ***The UNIVARIATE Procedure*** |
| ***Variable: RestPulse*** |

| *Basic Statistical Measures* | | | |
| --- | --- | --- | --- |
| *Location* | | *Variability* | |
| *Mean* | 53.45161 | *Std Deviation* | 7.61944 |
| *Median* | 52.00000 | *Variance* | 58.05591 |
| *Mode* | 48.00000 | *Range* | 30.00000 |
|  |  | *Interquartile Range* | 11.00000 |

**Answer for EX2(b):** Mean for heart rate while resting is 53.4516, median for heart rate while resting is 52.0000. Mode is 48, standard deviation is 7.6194,range is 30.0000 and interquartile range is 11.0000.

***EX 2 (c)\_Aerobic Fitness prediction***

|  |
| --- |
| ***The REG Procedure*** |
| ***Model: MODEL1*** |
| ***Dependent Variable: Oxygen*** |

|  |  |
| --- | --- |
| *Number of Observations Read* | 31 |
| *Number of Observations Used* | 31 |

| *Analysis of Variance* | | | | | |
| --- | --- | --- | --- | --- | --- |
| *Source* | *DF* | *Sum of Squares* | *Mean Square* | *F Value* | *Pr > F* |
| *Model* | 1 | 632.90010 | 632.90010 | 84.01 | <.0001 |
| *Error* | 29 | 218.48144 | 7.53384 |  |  |
| *Corrected Total* | 30 | 851.38154 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| *Root MSE* | 2.74478 | *R-Square* | 0.7434 |
| *Dependent Mean* | 47.37581 | *Adj R-Sq* | 0.7345 |
| *Coeff Var* | 5.79364 |  |  |

| *Parameter Estimates* | | | | | |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *DF* | *Parameter Estimate* | *Standard Error* | *t Value* | *Pr > |t|* |
| *Intercept* | 1 | 82.42177 | 3.85530 | 21.38 | <.0001 |
| *RunTime* | 1 | -3.31056 | 0.36119 | -9.17 | <.0001 |

***EX 2 (c)\_Aerobic Fitness prediction***

|  |
| --- |
| ***The REG Procedure*** |
| ***Model: MODEL1*** |
| ***Dependent Variable: Oxygen*** |



**Answer for EX2 (c):**

The estimated intercept for the simple linear regression model is 82.42177. The estimated slope of the regression model is -3.3105

***EX 3 (a)\_conference***

| *Obs* | *FirstName* | *LastName* | *AttendID* | *BusiPhone* | *HPhone* | *Mphone* | *ContactBusi* |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *2* | CHRISTINE | GORDON | 1002 | (626)715-7752 | (626)739-2921 | (626)242-8025 | Yes |
| *3* | JOHN | MILLER | 1003 | (412)669-2745 | (412)606-1283 | (412)208-4521 | Yes |
| *4* | RONALD | MITCHELL | 1004 | (510)254-4432 | (510)359-9901 | (510)122-7204 | Yes |

| *Obs* | *ContactHome* | *ContactM* | *RegiRate* | *WedMix* | *ThursLunch* | *Volunteer* | *groups1* |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *2* | Yes | No | 200 | Yes | No | Yes | regular |
| *3* | No | Yes | 395 | Yes | Yes | No | early |
| *4* | Yes | Yes | 395 | No | Yes | Yes | early |

**Answer for EX 3(a):** The 2,3, and 4th observations includes Christine Gordon, John Miller, and Ronald Mitchell. Christine and John will attend Wednesday mixer, but Ronald will not attend it. Christine will not attend Thursday Lunch, but Ronald will attend it. Christine and Ronald are willing to volunteer at the conference, but John are not.

***EX 3 (b)\_conference***

|  |
| --- |
| ***The FREQ Procedure*** |

| *ThursLunch* | *Frequency* | *Percent* | *Cumulative Frequency* | *Cumulative Percent* |
| --- | --- | --- | --- | --- |
| *No* | 21 | 18.58 | 21 | 18.58 |
| *Yes* | 92 | 81.42 | 113 | 100.00 |

**Answer for EX3(b):** The number of people who will attend the Thursday lunch is 92. The number of people who will not attend the Thursday lunch 21.

***EX 3 (c)\_conference***

|  |
| --- |
| ***The FREQ Procedure*** |

| *groups1* | *Frequency* | *Percent* | *Cumulative Frequency* | *Cumulative Percent* |
| --- | --- | --- | --- | --- |
| *early* | 76 | 67.26 | 76 | 67.26 |
| *on-site* | 1 | 0.88 | 77 | 68.14 |
| *regular* | 36 | 31.86 | 113 | 100.00 |

**Answer for Exercise 3(c):** 76 people are in “early” group, 1 person is in “on-site group” and 36 people are in regular group.