STATISTICS 147 EXAM I, Part 1: SAS
Summer 2020; 60 pts; Show all work!

NAME: (Please Print) Wesley Chang

ID: (last 4 digits only) 0996

General Information:

- Complete the following using SAS.
- Save your SAS program in a file named *e1su20_XX.sas*, where XX = your initials.
- (2 pts) Include the following titles:

```
Statistics 147 Exam I, Part 1: SAS
Summer 2020
Your name
Question X      (where X = question number)
Part T          (where T = subpart number)
```

- Be sure to include code after the `goptions` to **change the SAS working directory** to the folder that contains your data files, `dogdiet.dat` and `dogjudge.dat`. This should allow you to read in files using only the file name in all of your `infile` statements.
- You should place comments in your SAS program file. (3 pts)
- Be sure to use an infile statement to read in your data. DO NOT COPY THE DATA INTO YOUR SAS PROGRAM!
- For any plots, place them in an **MS Word/.docx** document. Save it as **ex1SASsu20_XX.docx**, where XX = your initials. Be sure to make the question number very clear.
- When you have finished your exam, email the following files to lklei001@ucr.edu:
  - (i) Your completed exam PDF.
  - (ii) Your **SAS script/.sas** document.
  - (iii) Your **MS Word/.docx** document containing your plots.
  - (iv) DO NOT email/submit your SAS output! Luke will run your code for himself.
- Use the following options/goptions:

```
/* Set up format of the output */
options nocenter ps = 55 nocenter ls = 78 nodate nonumber formdlim='*';

/* Add line of code to clear all windows except the editor window */
DM log "odsresults; clear; out; clear; log; clear;";
ods graphics off;

/* goptions   formats the plot  */
goptions reset = all colors= (blue,red,green,purple)
       ftitle = swissb ftext=swissb htitle=1.5 htext = 1.0;

/* Place code to change the working directory below here  */
```

1. Carly and Shannon work in Research & Development for a premium dog food company that was interested in marketing a new food for adult dogs that are severely overweight. Carly and Shannon are considering 3 recipes for the new "diet" food and have decided to test them using severely overweight Labrador Retriever adult dogs (of approximately the same age and health condition). Twenty-four severely overweight Labs were selected at random and then each randomly assigned to one of the three recipes of dog food. After an appropriate amount of time, the weight loss, measured in pounds, was recorded.

   The following data has been saved in a file named **dogdiet.dat** which you should have downloaded from Blackboard (iLearn). (Each column represents a diet formula.)

```
file name: dogdiet.dat
F1    F2    F3
4.95  7.70  6.47
4.04  5.81  5.23
3.72  6.61  6.75
2.21  6.07  7.75
2.48  8.04  4.70
3.31  5.96  6.92
3.50  7.30  6.01
2.90  7.46  5.67
```

   NOTE: The 2 lines of headings are included in the data file.

   NOTE:

   - ♠ The data is located in a datafile named **dogdiet.dat**.
   - ♠ The headings are included in the data file. The actual data begins on line 3.

   (i) Write the SAS code to read in and print out the data. Use **diet_all** as your temporary SAS dataset name. Use nested do loops! **(NOTE: DO NOT COPY AND PASTE THE DATA INTO YOUR SAS PROGRAM. READ THE DATA IN FROM THE EXTERNAL DATA FILE!)** (5 pts)

   (ii) Using if-then-else structures, name the formulas as follows: (3 pts)

formula	1	2	3
form_name	**Formula1**	**Formula2**	**Formula3**

   (iii) Add the appropriate lines of code to sort the data by form_name. (2 pts)

   (iv) Jose and Dung decided they should generate some descriptive statistics for each of the formulas. Add the appropriate lines of code to your program to generate the mean, standard deviation and median for each of the formulas and complete the following table: (5 pts)

form_name	Mean	Standard Deviation	Median
Formula1	3.3887500	0.8818882	3.4050000
Formula2	6.8687500	0.8660326	6.9550000
Formula3	6.1875000	0.9858245	6.2400000

(v) Create a 3D vertical bar chart (high resolution) for the total weight loss of each of the dog formulas. (6 pts)

(a) Include the following title (as title6): **Weight Loss of the Dog Diets**

★ Rotate the title by 25%.

★ Use **swissb** as the font.

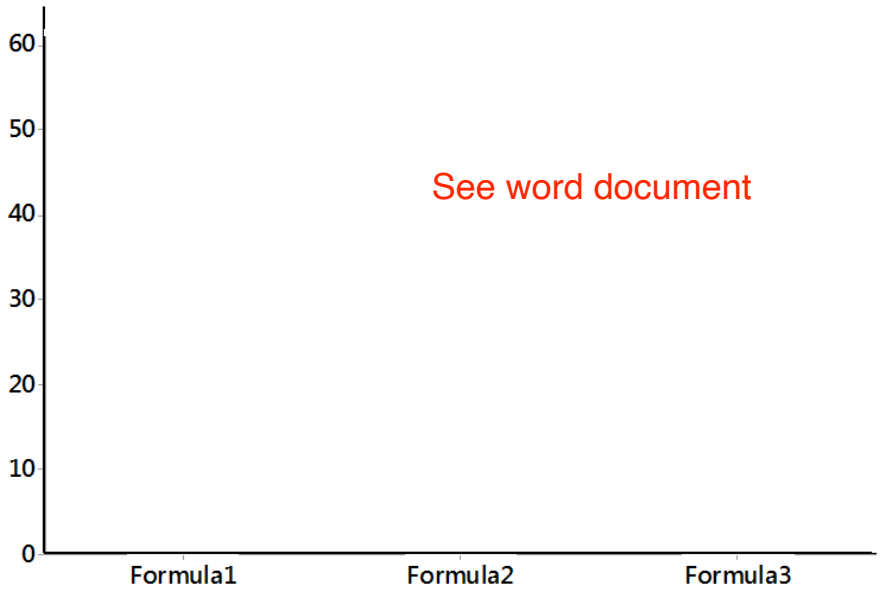
★ Use a height of 2..

(b) Make the colors of the bars: red, blue, and green

(c) Make the shape of the bars: cylinder

(d) Print the sum of the weights above each bar.

Sketch your chart in the space below. (6 pts)



2. Refer to Question 1. Create a new temporary SAS data set, named **onlyFormula1**, in which the data is restricted to the **Formula1** labs only. (Be sure to print the data, with a descriptive title, as a check.) (3 pts)

3. Refer to Question 1. Create a new temporary SAS data set, named **bothF1F3** in which the data is restricted to the weight loss observation for the **Formula1** and **Formula3** labradors. (Be sure to print the data, with a descriptive title, as a check.) (3 pts)

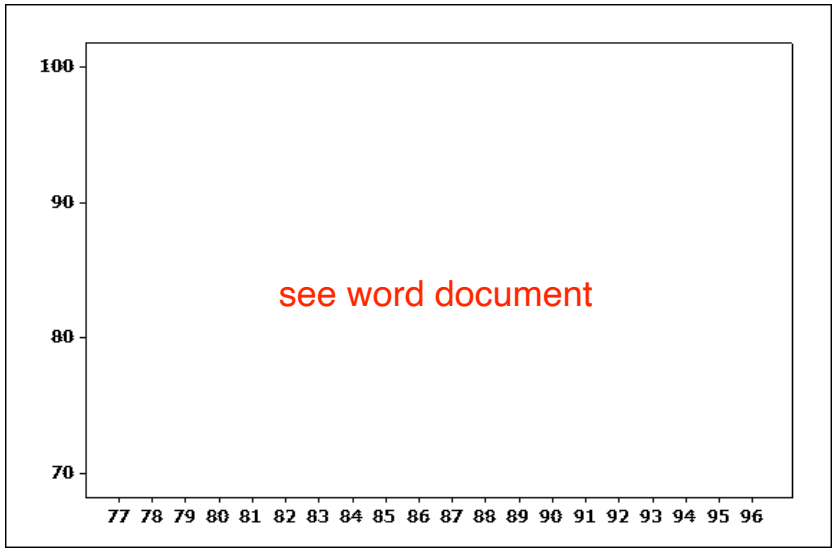
4. *Disc dog* is the more generic name for what is commonly called *Frisbee dog*. In *disc dog* competitions, dogs and their human *flying disc* throwers compete in events such as distance catching and somewhat choreographed freestyle catching. Scores differ among judges, even when the same performance is being evaluated. Khaldoun and Thomas have been selected to be judges in a local *disc dog* competition. The scores, reported by Khaldoun (J1) and Thomas (J2), for 11 competitors are located in a data file named **dogjudge.dat**,

```
file name: dogjudge.dat
dog             J1  J2
Abby            83  84
Cody            79  78
Dexter          77  73
Dusty           94  96
Kali            79  78
Korra           88  90
Lakota          95  90
MaxG            93  87
Mercedes        88  84
Shadow          96  95
MaxS            85  88
```

NOTE: The two lines of headings are included in the data file.

(i) Write the appropriate SAS code to read in and print out the data. Use the temporary SAS dataset name of **judging**. (DO NOT use do loops!) **(NOTE: DO NOT COPY AND PASTE THE DATA INTO YOUR SAS PROGRAM. READ THE DATA IN FROM THE EXTERNAL DATA FILE!)** (4 pts)

(ii) Fan and Rene would like a visual image of the data. Create a high resolution scatterplot of **J1** versus **J2**, using **J1** as the horizontal axis. Be sure to include a descriptive **title** on your plot. Sketch the plot in the space provided below. (Note: The plot below is a *generic* blank plot with numbers on the axes that may go beyond the numbers you get in your actual plot.) (4 pts)



5. Refer to Question 4. Suppose Abby and MaxS were caught taking performance enhancing dog treats and were therefore disqualified. Create a new temporary SAS data set that reads in and prints out the observations/rows for dogs (and their run times) that have **NOT** been disqualified. (Be sure to print out the results with a descriptive title!) (3 pts)

6. Suppose Ernesto and Natascha know the specification limits for a plastic bottle filling machine are 40.0 ± 2.50 ounces. In other words, the filling process is considered to be functioning at an appropriate level (functioning "in control") if the fill in the bottles is between 37.5 ounces and 42.5 ounces. Suppose the filling process is normally distributed with a mean of 40 ounces and a standard deviation of 1.75 ounces. They select a bottle at random from the assembly line. Let X = amount of fill in the bottle.

   (i) Write the appropriate SAS code to find the probability that a bottle is overfilled (in other words, the amount of fill is more than 42.5 ounces). (2 pts)

   ANSWER: 0.076564

   (ii) Write the appropriate SAS code to find the probability that the amount of fill is between 37 ounces and 41 ounces. (3 pts)

   ANSWER: 0.67291

   (iii) Write the appropriate SAS code to find the 96^{th} percentile (i.e., find x such that $P(X \leq x) = 0.96$). (1 pt)

   ANSWER: 43.0637

7. Sarah and Ziguo found that, according to a recent survey, 35% of current students prefer the change in class times that begins Fall Quarter 2020. Sarah and Ziguo were interested in examining this claim. A random sample of 24 students was selected.

Let X = # of students that prefer the change in class times. Use this information to complete the following.

(i) Write the appropriate SAS code to find the probability that exactly 10 of the students prefer the change in class times. (2 pts)

ANSWER: 0.13002

(ii) Write the appropriate SAS code to find the probability between 7 and 10, inclusively, of the students prefer the change in class times. (2 pts)

ANSWER: 0.17418

8. Let $y = e^{-\sqrt{2m-n}}$ for m = 1, 4, 7 and n = 1, 2. Use nested DO loops to generate the values of m and n and compute the values of y. Complete the following table with the appropriate values of y. (5 pts)

	m		
n	**1**	**4**	**7**
1	0.36788	0.07095	0.02717
2	1.00000	0.08634	0.03130

9. Consider the following SAS code, written for the new temporary SAS dataset `quest9`. Find and **correct** all the errors. (You may add this code right before the **quit** statement in your SAS program. Do not include the line numbers in your SAS program!) (3 pts)

```
Line 1:   data quest9;
Line 2:       input x y z
Line 3:       title5 'Question 9;
Line 4:       title6 ' ';
Line 5:   datalines;
Line 6:   1 6 4 8 3 9 2 7 3 5 5 2 4 3
Line 7:   ;
Line 8:   run;
Line 9:   proc print noobs data = quest9;
Line 10:  run;
Line 11:  proc sort data = diet_all;
Line 12:     by z;
Line 13   run;
Line 14:  proc print noobs
Line 15:  run;
Line 16:  proc gplot;
Line 17:      gplot y*x;
Line 18:  run;
```

List any corrections you have made in the space below. (Please include the Line number.)

CORRECTIONS:

See Word document

Turn in your completed exam paper as a PDF, your SAS program file, *e1su20_XX.sas* (where XX = your initials), and the Word/.docx document containing your plots. Prepare for R!

Luke & Ruihan