STAT 475 HOMEWORK 4 (*due Thursday, November 30, 2023*)

Exercises for Handout 24 material: Explain the reason for selecting the correct answer, as well as the reason for NOT selecting EACH of the wrong answers.

* Quiz for Chapter 5 in the advanced programming certification guide: Questions

1, 2, 5, 6, 7, 8, and 10.

Exercises for Handout 25 material: Explain the reason for selecting the correct answer, as well as the reason for NOT selecting EACH of the wrong answers.

* Quiz for Chapter 8 in the advanced programming certification guide: Questions 1 – 3.

Exercises for Handout 26 material: Explain the reason for selecting the correct answer, as well as the reason for NOT selecting EACH of the wrong answers.

* Quiz for Chapter 9 in the advanced programming certification guide: Questions 1-5.
* Quiz for Chapter 10 in the advanced programming certification guide: Questions 1-3.

And also do the following exercises:

Exercise 1. (Use proc sql) The files “grades550.xls”, “grades484.xls”, and “grades695.xls” contain student id’s and their final grades.

1. How many students took STAT550 and STAT484? What were their grades in these courses? List their id’s and respective grades.
2. How many students took STAT550 and STAT695? What were their grades in these courses? List their id’s and respective grades.
3. How many students took STAT484 and STAT695? What were their grades in these courses? List their id’s and respective grades.
4. How many students took all three courses? What were their grades in these courses? List their id’s and respective grades.

Exercise 2. (Use proc sql) File “glaucoma.xls” contains data on patients who underwent a glaucoma treatment in one of two eyes (Tx group), the other eye was kept as a control (Cx group). The other variables are number of medications (No of meds), time previously on medication (in years), gender, age, and type of glaucoma.

1. How many patients were in the study?
2. How many patients were currently on medication? How many were currently off medication or medication-naïve (never took medication)?
3. What were the mean, standard deviation, min, and max of the time previously on medication? Exclude medication-naïve patients.
4. How many males and how many females were in the study?
5. How many patients were in the study by type of glaucoma?
6. What was the mean age by gender and type of glaucoma?

Exercise 3. Write a SAS code to compute 1+1. Your code must contain the following keywords:

* title
* assignment is completed by
* &sysday
* &sysdate
* &systime
* Your full name
* %scan to extract your last name
* %eval
* %let
* %put
* %macro
* %mend

Hint: You might want to create a dummy dataset that you will be printing together with appropriate title(s). The titles might be

title1

This assignment is completed by Korosteleva on Saturday, 8APR20 at 22:08

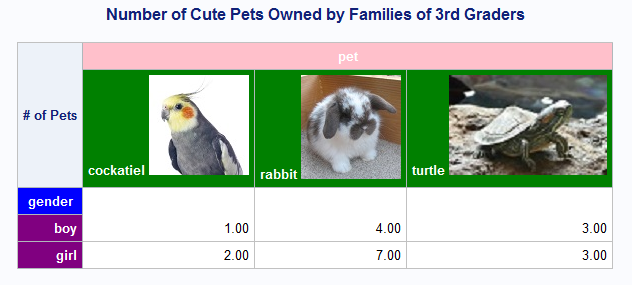
title2

1+1=2

Exercise 4. Run a macro loop to illustrate the Law of Large Numbers. This is what you need to do:

1. Generate 1, 2, 3, 4, 5, …, *n* Bernoulli(*p*) random variables and compute the means of 1, 2, 3, 4,5, …, *n* variables. Generate the variables and do the computations within the same loop (cf. the code we studied in class).
2. Plot the means to show that they converge to *p*.
3. Create a macro loop which has *n* and *p* as arguments.

Exercise 5. A teacher surveyed her 3rd graders in regard to what kind and how many pets their families own. The data in file “pets.dat” contain kids’ gender (boy/girl), name of the pet owned, and the number of pets owned. Use the magic of proc tabulate to reproduce the table below. Make sure to match colors and alignments. The pictures of the cute animals are in files “cockatiel.jpg”, “rabbit.jpg”, and “turtle.jpg”.



\

Exercise 6. Consider the dataset in file “products.dat”. The variables are Product Number, Product Name, Manufacturer Number, Product Type and Retail Unit Cost. Do the following:

1. Read instream data
2. Use PROC SQL to insert a new row

prodnum prodname manunum prodtype rtlcost

3480 Desktop Computer 780 Workstation $1,799

1. Use PROC SQL to modify the data to reflect a 20% price increase on all software products, and a 20% discount on all the other products.
2. Use PROC SQL to add title, footnote, labels and modify formats (if necessary) to obtain the following output (Note that the footnote contains today’s date).

|  |
| --- |
| Product Information |

| **Product Number** | **Product Name** | **Manufacturer Number** | **Product Type** | **Retail Unit Cost** |
| --- | --- | --- | --- | --- |
| 5009 | Dream Machine | 500 | Workstation | $2560.00 |
| 4506 | Business Machine | 450 | Workstation | $2676.00 |
| 2101 | Travel Laptop | 400 | Laptop | $2208.00 |
| 2212 | Analog Cell Phone | 230 | Phone | $28.00 |
| 4509 | Digital Cell Phone | 245 | Phone | $140.00 |
| 5003 | Office Phone | 560 | Phone | $116.00 |
| 1110 | Spreadsheet Software | 134 | Software | $360.00 |
| 1200 | Database Software | 113 | Software | $958.80 |
| 3409 | Statistical Software | 243 | Software | $2278.80 |
| 2102 | Wordprocessor Software | 245 | Software | $414.00 |
| 2200 | Graphics Software | 246 | Software | $718.80 |
| 3480 | Desktop Computer | 780 | Workstation | $1439.20 |

|  |
| --- |
| Updated on 12APR2022 |

Exercise 7. Consider five points and (7, 18). Note that these points lie on a perfectly straight line with the intercept and slope . The quantities and can be found as the solution of the linear regression equation, that is,

where and

Use PROC IML to find *a* and