

22S:166 Computing in Statistics
Fall 2014
Final Exam

You may submit your answers either as:

- a plain text file:
Number the questions; include all R and SAS code and R output, and put your name as the first line of the file.
- or, for 3 percentage points extra credit, as an .Rnw file and PDF file:
embed R code into the .Rnw file, and just copy and paste SAS code into a `verbatim` environment.

SAS output may either

- be included as part of your text file, or in a verbatim environment of your .Rnw file, or
- be submitted in a separate PDF file

Run R (and possibly \LaTeX) on the Linux machines. You may run SAS wherever is most convenient for you, including the Virtual Desktop.

1. SAS and R

- (a) One of the datasets that comes with R is called “CO2.” You can view it in R simply by entering
- ```
CO2
```

Write this dataset out to a text file called “CO2.txt.” Include column headings but no row numbers, and do not include quotation marks around character variables. The first four rows of your text file should look like this:

```
Plant Type Treatment conc uptake
Qn1 Quebec nonchilled 95 16
Qn1 Quebec nonchilled 175 30.4
Qn1 Quebec nonchilled 250 34.8
```

Include your R code in your answers file.

- (b) Write a SAS data step to read your “CO2.txt” file into SAS.  
Include your data step in your answers file. If you were not able to create the text file in R, you may use the “CO2.txt” file posted on the course web page under Handouts.

(c) Use `proc tabulate` to create the following table:

```

	Concentration									

	95	175	250	350	500					
	-----+-----									
	uptake	uptake	uptake	uptake	uptake					
	-----+-----									
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
-----+-----+-----+-----+-----+-----+-----+-----										
Treatment	Type									
-----+-----+-----+-----+-----+-----+-----+-----										
chilled	Mississippi	3	9.60	3	14.77	3	16.10	3	16.60	3
	-----+-----+-----+-----+-----+-----+-----+-----									
	Quebec	3	12.87	3	24.13	3	34.47	3	35.80	3
-----+-----+-----+-----+-----+-----+-----+-----										
nonchilled	Mississippi	3	11.30	3	20.20	3	27.53	3	29.90	3
	-----+-----+-----+-----+-----+-----+-----+-----									
	Quebec	3	15.27	3	30.03	3	37.40	3	40.37	3
-----+-----+-----+-----+-----+-----+-----+-----										
All		12	12.26	12	22.28	12	28.88	12	30.67	12
-----+-----+-----+-----+-----+-----+-----+-----										

```

(Continued)

```

	Concentration			

	675	1000		
	-----+-----			
	uptake	uptake		
	-----+-----			
	N	Mean	N	Mean
-----+-----+-----+-----+-----				
Treatment	Type			
-----+-----+-----+-----+-----				
chilled	Mississippi	3	18.27	3
	-----+-----+-----+-----+-----			
	Quebec	3	37.50	3
-----+-----+-----+-----+-----				
nonchilled	Mississippi	3	30.53	3
	-----+-----+-----+-----+-----			
	Quebec	3	41.50	3
-----+-----+-----+-----+-----				
All		12	31.95	12
-----+-----+-----+-----+-----				

```

Include your SAS procedure step in your answers file. Include the output either in your answers file or in a separate PDF file.

- (d) Write a SAS macro that will produce tables with the same columns as the table in the previous problem. However, your macro should let the user specify which variable(s) will define the rows. Call your macro to produce the two tables given below.

|            | Concentration |       |        |       |        |       |        |       |        |       |
|------------|---------------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
|            | 95            |       | 175    |       | 250    |       | 350    |       | 500    |       |
|            | uptake        |       | uptake |       | uptake |       | uptake |       | uptake |       |
|            | N             | Mean  | N      | Mean  | N      | Mean  | N      | Mean  | N      | Mean  |
| Treatment  |               |       |        |       |        |       |        |       |        |       |
| chilled    | 6             | 11.23 | 6      | 19.45 | 6      | 25.28 | 6      | 26.20 | 6      | 26.65 |
| nonchilled | 6             | 13.28 | 6      | 25.12 | 6      | 32.47 | 6      | 35.13 | 6      | 35.10 |

(Continued)

|            | Concentration |       |        |       |
|------------|---------------|-------|--------|-------|
|            | 675           |       | 1000   |       |
|            | uptake        |       | uptake |       |
|            | N             | Mean  | N      | Mean  |
| Treatment  |               |       |        |       |
| chilled    | 6             | 27.88 | 6      | 29.78 |
| nonchilled | 6             | 36.02 | 6      | 37.38 |

|            |             | Concentration |       |        |       |        |       |        |       |        |       |
|------------|-------------|---------------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
|            |             | 95            |       | 175    |       | 250    |       | 350    |       | 500    |       |
|            |             | uptake        |       | uptake |       | uptake |       | uptake |       | uptake |       |
|            |             | N             | Mean  | N      | Mean  | N      | Mean  | N      | Mean  | N      | Mean  |
| Treatment  | Type        |               |       |        |       |        |       |        |       |        |       |
| chilled    | Mississippi | 3             | 9.60  | 3      | 14.77 | 3      | 16.10 | 3      | 16.60 | 3      | 16.63 |
|            | Quebec      | 3             | 12.87 | 3      | 24.13 | 3      | 34.47 | 3      | 35.80 | 3      | 36.67 |
|            | All         | 6             | 11.23 | 6      | 19.45 | 6      | 25.28 | 6      | 26.20 | 6      | 26.65 |
| nonchilled | Type        |               |       |        |       |        |       |        |       |        |       |
|            | Mississippi | 3             | 11.30 | 3      | 20.20 | 3      | 27.53 | 3      | 29.90 | 3      | 30.60 |
|            | Quebec      | 3             | 15.27 | 3      | 30.03 | 3      | 37.40 | 3      | 40.37 | 3      | 39.60 |
| All        | 6           | 13.28         | 6     | 25.12  | 6     | 32.47  | 6     | 35.13  | 6     | 35.10  |       |
| All        | Type        |               |       |        |       |        |       |        |       |        |       |
|            | Mississippi | 6             | 10.45 | 6      | 17.48 | 6      | 21.82 | 6      | 23.25 | 6      | 23.62 |
|            | Quebec      | 6             | 14.07 | 6      | 27.08 | 6      | 35.93 | 6      | 38.08 | 6      | 38.13 |
| All        | 12          | 12.26         | 12    | 22.28  | 12    | 28.88  | 12    | 30.67  | 12    | 30.88  |       |

(Continued)

|            |             | Concentration |       |        |       |
|------------|-------------|---------------|-------|--------|-------|
|            |             | 675           |       | 1000   |       |
|            |             | uptake        |       | uptake |       |
|            |             | N             | Mean  | N      | Mean  |
| Treatment  | Type        |               |       |        |       |
| chilled    | Mississippi | 3             | 18.27 | 3      | 18.73 |
|            | Quebec      | 3             | 37.50 | 3      | 40.83 |
|            | All         | 6             | 27.88 | 6      | 29.78 |
| nonchilled | Type        |               |       |        |       |
|            | Mississippi | 3             | 30.53 | 3      | 31.60 |
|            | Quebec      | 3             | 41.50 | 3      | 43.17 |
|            | All         | 6             | 36.02 | 6      | 37.38 |
| All        | Type        |               |       |        |       |
|            | Mississippi | 6             | 24.40 | 6      | 25.17 |
|            | Quebec      | 6             | 39.50 | 6      | 42.00 |
|            | All         | 12            | 31.95 | 12     | 33.58 |

Include your SAS code for both the macro definition and the macro calls in your answers file. Include the SAS output either in your answers file or in a separate PDF file.

## 2. Relational database concepts

Xiao is the program chairperson for a statistics conference. She wishes to set up a relational database to organize and store information about the talks being given at the conference.

Talks are organized into sessions. Each session may have 1 to 6 talks. Any talk is in only one session.

Talks have authors. A single talk may have more than one author, and a single person may be the author of more than one talk.

The data that Xiao wishes to store are:

```
session date
session time
session location (room number)
talk title
which session talk is in
talk author names
talk author affiliations
talk author email addresses
```

- (a) The relationship between sessions and talks is (put the right answer in your answer file):
- i. one-to-one
  - ii. one-to-many
  - iii. many-to-one
  - iv. many-to-many
- (b) The relationship between authors and talks is (put the right answer in your answer file):
- i. one-to-one
  - ii. one-to-many
  - iii. many-to-one
  - iv. many-to-many
- (c) For these data, design a relational database structure in third normal form. Identify tables, fields, primary keys, and foreign keys.