22S:166 Computing in Statistics
PRACTICE PROBLEMS for 2006 FINAL
Fall 2005
Final Exam

Copy your code and output for each problem into a Microsoft Word document or a LaTeXdocument. **Include your name and a 4-letter or 4-digit code for me to use in posting your grade. The code must NOT be part of your Soc Sec number or student i.d**.

Email me either the Word document or a PDF file at

`kcowles@stat.uiowa.edu`

1. THE 2006 EXAM WILL *NOT* HAVE A SAS QUESTION LIKE THIS ONE. Download the dataset "blissymbols.dat" from the "Datasets" section of the course webpage. Read the description of this dataset in the "blissymbols.info" file.

   Use SAS to produce the following table using the Bliss symbols data. For this question, submit all your SAS code and the table you produce.

	Number correct	
	Mean	StdDev
Learning style		
Active	24.42	6.31
Passive	17.88	4.03
All	21.15	6.19

2. The secretary of the statistics department has asked you to help her design an efficient way of storing information about the faculty members and the committees they serve on. Each committee has more than one faculty member on it. In addition, the same faculty member can serve on more than one committee. Currently, the secretary is storing all the information in one file. The attributes in that file are:

```
Name of Committee
Committee type (ad hoc or permanent)
Name of chairperson of committee
telephone number of chairperson
email address of chairperson
name of committee member 1
telephone number of committee member 1
email address of committee member 1
name of committee member 2
telephone number of committee member 2
email address of committee member 2
.
.
.
name of committee member 6
telephone number of committee member 6
email address of committee member 6
```

   Help the secretary by recommending how to reorganize her data into a set of relational tables that are in 3rd normal form. Identify all primary keys and foreign keys.

   For this problem, type your answer and include it in the document you send to me.

3. This question requires the "caffeine.dat" dataset from the course web page "Datasets" section. Read the "caffeine.info" file to learn what the variables in the dataset are.

   (a) Download the dataset. Use the R `read.table` function to read it into an R matrix.

   (b) Use the `t.test` function in R to carry out a *paired* t-test of the null hypothesis that the population mean of depression scores is the same in caffeine-deprived people as in non-caffeine-deprived people. FOr this part of the problem, turn in your R code and a summary of the output.

   (c) Note that the `t.test` function returns a list, one item of which is the value of the t statistic (called "statistic"). Carry out the jackknife procedure to calculate an "unbiased" estimate of the t statistic for this hypothesis test and the standard error of the statistic. Turn in your R code and its output.