UPCOMING EXAM
Exam on February 17, 1999

▷ **Date of Midterm Exam 1: Wednesday, Feb 17, 1999**

▷ **Time of Midterm Exam 1: IN CLASS, 9:30–10:20am**

▷ **Place of Midterm Exam 1:**

- The exam will be held in our usual classroom, DH 2315
- You may be directed by the TA's to overflow seating in DH 2105

▷ **Ground rules:**

- *Closed book*, closed notes.
- Bring a *calculator*. You will not be able to share or borrow one during the exam.
- Bring *one 8.5" by 11" sheet of paper* with anything you want written on either side.
- *Show your work* for numerical calculations, and justify your qualitative answers. You will not get full credit unless you do. Also you will not get any credit for a wrong answer if you do not show work.

▷ **Review Session:**

- There will be a review session held by two TA's Monday Night after Supper.

- *Check with TA's in Lab for time and location.*

- **Bring Questions!** We will *not* have any prepared lecture! Everything we say will be in response to your questions.

▷ **Office Hours During Exam Week:**

- Monday 2:00–3:00, OSC 219, Iuliana Ianus
- Tuesday 2:00–4:00, Baker 232C, Brian Junker
- Tuesday 6:30–7:30, Poerter 226D, Alix Gitelman
- Office Hours Wed Feb 17 and and Thu Feb 18 are cancelled.

▷ **Labs and Homework**

- *Labs are cancelled Feb 18 and Feb 19.*

- Graded HW 4 will be inside Baker 132 on Tuesday for pickup.

▷ **All Class Materials are at:** `http://www.stat.cmu.edu/~brian/201/`

You are responsible for all class materials, reading, hw, and labs from Week 01 through Week 05, for the exam.

▷ **What you should know for Exam 1:**

- **Statistics—Exploratory Data Analysis**

  1. Difference between a quantitative (continuous) and qualitative (categorical) variable.
  2. Describing the distribution of a quantitative variable
     (a) Shape and features: symmetry, modes/gaps, outliers
     (b) Numerical summaries—know how to interpret
        – Central Location—mode, *mean**, *median**, *quartiles**.
        – Spread—*range**, *IQR**, standard deviation.
        – 50% rule for boxplots.
        – *68%–95%–99.7% rule** for symmetric, unimodal distributions; *Z-scores**.
     (c) Transformations to improve "normality": $\log_{10}(x)$, $\ln(x)$, x^p for $p < 1$ and $p > 1$.
  3. Methods for displaying the distribution of a quantitative variable
     (a) Histogram—How to read and interpret.
     (b) Stem-and-Leaf Plot—How to make one; how to interpret.
     (c) Boxplots—How to make one; How to read and interpret.
  4. Comparing Distributions
     (a) Back-to-back stem-and-leaf plots—How to interpret.
     (b) Side-by-side Boxplots on the same scale—How to interpret.
     (c) Scatter plot—Linear/Nonlinear, Trend, Clusters, Outliers, Unequal Variability—How to interpret.
     (d) Two-way contingency table.
     (e) Explanatory vs. response variable.
  5. Regression Analysis
     (a) Median trace plots ("boxplots within vertical strips") when a straight line is no good.
     (b) Straigh line analysis:
        – Association and *correlation** r; r^2 —How to interpret.
        – Least-squares line: $y = a + bx$, $b = r \cdot SD_y/SD_x$, $a = \bar{y} - b\bar{x}$ —How to use and interpret.
  6. Two-way Tables
     (a) Frequencies (counts) and Rates (proportions, percents)
     (b) *Row Percents**, *Column Percents**; Explanatory/Response Variable
     (c) Independence vs. Association
     (d) *Expected table**, *Standardized residuals** —How to interpret.
     *Know how to calculate quantities marked with * above by hand!*
     <u>*Use the methods on pp. 77–78 of Siegel and Morgan for* **median**, **quartiles**, *and* **IQR**.</u>

- **Minitab**

  – Know how to read and interpret the plots, summary statistics, and data tables you have created with MINITAB for homework assignments and labs.
  – Review readings, labs, lab and homework solutions, and class lecture notes; know how to use these plots and summaries, and how to write analyses based on them.