Course Level Assessment Report
MTH107 Statistical Concepts and Analysis

Derek H. Ogle

August 26, 2015

Course Information

Background

Statistical Concepts and Analysis (MTH107) is the entry-level statistics course at Northland College. Taught by instructors in the Mathematical Sciences program, the course is required by several majors in the Natural Resources, Environmental Sciences, and Social Responsibility departments and can also fulfill the "quantitative reasoning" requirement for the College's Natural Connections program.

I have taught MTH107 since 1996, including 13 sections of MTH107 since 2010, the period for which this assessment document covers. On average, the mean (standard deviation) enrollment was 29.8 (3.48) students, 7.3% (5.28) of the students dropped the course (after the initial add-drop period), and the mean course gpa was 2.33 (0.28). Over this time, there was no detectable linear trend in mean course GPA ($p = 0.3250$), percentage of students dropping the course ($p = 0.8107$), percentage of students that earned an A ($p = 0.5556$), or percentage of students that earned a D or F ($p = 0.2147$; Figure 1).

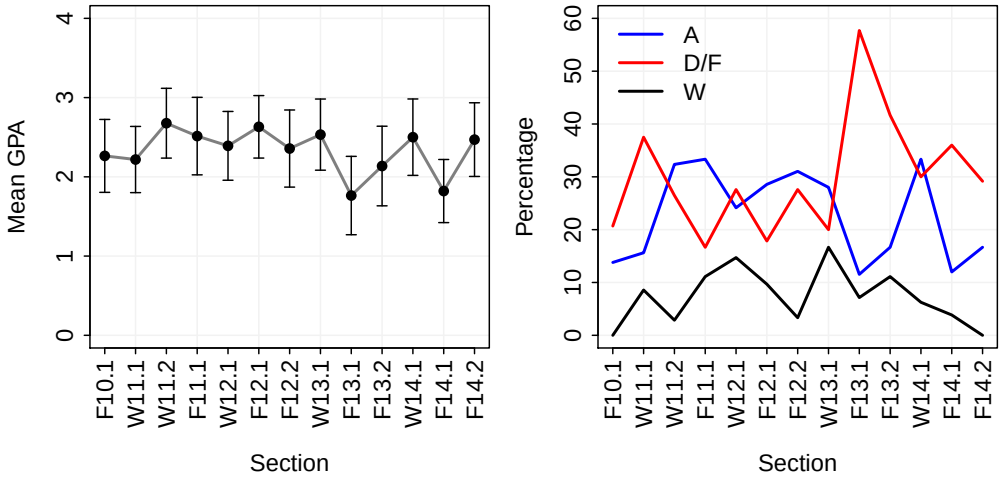

Figure 1. Mean course GPA (with 95% confidence intervals; Left) and the percentage of students (Right) that earned better than an A (labeled as an "A"), worse than a D+ ("D/F"), or withdrew ("W") from the course by section (abbreviated with semester (F=Fall, W=Winter), last two digits of the calendar year, and, following the period, the section number (1 or 2)).

Intended Learning Outcomes

I modified the intended learning outcomes (ILO) for MTH107 in preparation for teaching the course in the Fall of 2012. This list of ILOs has stayed the same (except for minor wording differences) since then. The

ILOs as they have appeared on the course syllabi are:

1. describe why statistics is central to scientific inquiry (& your field of interest);
2. define basic statistical words and symbols;
3. perform appropriate exploratory data analyses (univariate and bivariate);
4. identify the purposes of and conduct and interpret the results of a linear regression;
5. design simple experiments and sampling strategies;
6. construct and interpret confidence intervals for one and two-sample mean and proportion problems;
7. identify the appropriate hypothesis test to perform in one- and two-sample quantitative and categorical data situations;
8. construct and interpret the results from a hypothesis test for one- and two-sample quantitative and categorical data situations; and
9. communicate statistical results and ideas in a succinct and informative manner.

Some of these ILOs are "double-barreled" (i.e., represent more than one outcome) to save space on the syllabi. For example, ILO #8 represents at least four specific methods in statistics (1-sample t-test, 2-sample t-test, goodness-of-fit test, and chi-square test) for which I will assess student achievement. For assessment purposes, the "double-barreled" ILOs are unpacked to single outcomes for analysis.

Assessment Methodology

Beginning in Fall, 2012, the final exam for the course consisted of questions that were specific assessments of most of the course ILOs. Scores for individual questions were recorded separately for each student. Scores were summed across individual questions for each student to provide an overall grade on the exam for that student. However, scores on individual questions were summarized across all students to provide a measure of student achievement for each ILO. These summaries are used in this assessment report.

Raw scores for individual questions were converted to the proportion of the total points available for the question. These proportions were summarized for all students within a specific section of the course. For a specific ILO, the distribution (i.e., boxplot) and summary statistics of the proportions were computed. Results were compared to *a priori* goals for student achievement for that ILO.

Assessment of ILO #1

One question on the final exam asked the students to explain the importance of statistics to understanding science with explicity directions to highlight specific concepts from the course that would illustrate the student's argument. This question was meant to assess ILO #1.

I have two goals for student achievement of this ILO. First, given that this course is required for several majors and serves as a Natural Connections course, I hoped that most students would answer this question with some broad knowledge. Second, I also hoped that the "highest achieving" students in the course would provide a "very good" answer to this foundational question. I translated these hopes into the following two goals:

1. The first quartile (the lowest 25% of values; Q1) of proportional scores on this question would be 0.50 or better (i.e., at least half of the available points).
2. The third quartile (the top 25% of values; Q3) of the proportional scores on this question would be 0.90 or better.

Results

Student achievement on this question varied substantially among sections (Figure 2). The first goal was met in five of eight sections, including the last three sections. The second goal was not met for any section, though the "F12.2" and "W13.1" sections were close to having met the objective value.

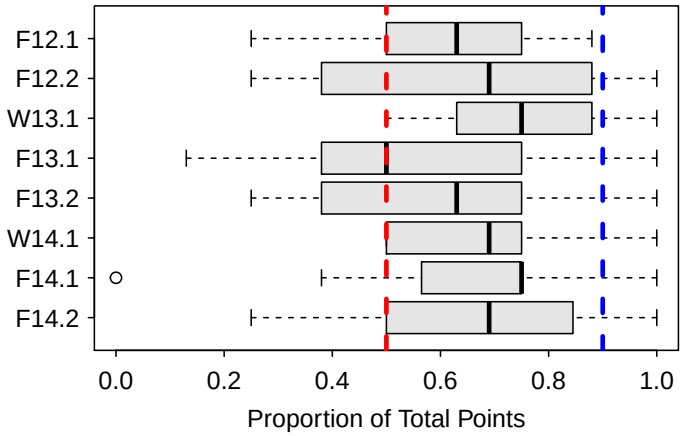

Figure 2. Boxplot (ends of the boxes are first and third quartiles, the solid line inside of the box is the median) of proportional scores on the 'Why Statistics is Important' question (ILO #1) by section (same abbreviations as in Figure 1). The red and blue vertical dashed lines are the goals for the Q1 and Q3 results, respectively.

Table 1. Summary statistics of proportional scores on the 'Why Statistics is Important' question (ILO #1) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	Q1	**Q3**	Q3	% zero	percZero
F12.1	28	28	0.63	0.18	0.25	0.50	0.63	0.75	0.88	0.0
F12.2	26	26	0.68	0.26	0.25	0.41	0.69	0.88	1.00	0.0
W13.1	21	21	0.75	0.15	0.50	0.63	0.75	0.88	1.00	0.0
F13.1	25	25	0.59	0.26	0.13	0.38	0.50	0.75	1.00	0.0
F13.2	23	23	0.61	0.22	0.25	0.38	0.63	0.75	1.00	0.0
W14.1	28	28	0.69	0.15	0.50	0.50	0.69	0.75	1.00	0.0
F14.1	23	23	0.67	0.21	0.00	0.56	0.75	0.75	1.00	4.3
F14.2	24	24	0.66	0.20	0.25	0.50	0.69	0.83	1.00	0.0

Actions

The course material related to this ILO was presented throughout each section, but was only explicitly discussed in the first chapter during the first week of each semester. To address this short-coming, I will develop a short mid-term assignment where students will re-visit the core principles related to this ILO by finding, summarizing, and reflecting on a "piece" (newspaper article, journal paper, etc.) that demonstrates the importance of statistics in the student's chosen field or to developing knowledge from data.

Assessment of ILO #8 (Two-Sample t-Test)

One question on the final exam required students to conduct a thorough analysis of data with a "Two-Sample t-Test". The question was graded with a rubric that listed the eleven steps of any hypothesis test (this rubric was supplied to the students before and during the exam). This question was intended to assess the "two-sample quantitative" portion of ILO #8.

I have two goals for student achievement of this ILO. First, I hoped that most students would have gained enough knowledge to at least provide a reasonable start to an answer for this question. Second, I hoped that the majority of students would provide a "very good" answer to this technical question. I translated these hopes into the following two goals:

1. The Q1 of proportional scores on this question would be 0.50 or better.
2. The median (the top 50% of values) of proportional scores on this question would be 0.80 or better.

Results

Student achievement on this question varied substantially among sections (Figure 3). The first goal was met by each section with the exception of "F13.1." The second goal was met by only two sections, though it was nearly met by three other sections.

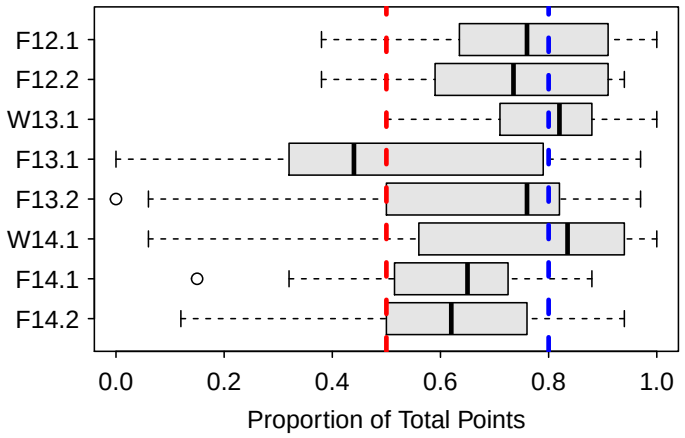

Figure 3. Boxplot of proportional scores on the 'Two-Sample t-Test' question (part of ILO #8) by section (same abbreviations as in Figure 1). The red and blue vertical dashed lines are the goals for the Q1 and median results, respectively.

Table 2. Summary statistics of proportional scores on the 'Two-Sample t-Test' question (ILO #1) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	**median**	median	Q3	% zero	percZero
F12.1	28	28	0.76	0.17	0.38	0.64	0.76	0.91	1.00	0.0
F12.2	26	26	0.73	0.18	0.38	0.60	0.74	0.90	0.94	0.0
W13.1	21	21	0.79	0.13	0.50	0.71	0.82	0.88	1.00	0.0
F13.1	25	25	0.52	0.31	0.00	0.32	0.44	0.79	0.97	12.0
F13.2	23	23	0.65	0.29	0.00	0.50	0.76	0.82	0.97	4.3
W14.1	28	28	0.73	0.27	0.06	0.56	0.84	0.94	1.00	0.0
F14.1	23	23	0.62	0.19	0.15	0.52	0.65	0.72	0.88	0.0
F14.2	24	24	0.63	0.22	0.12	0.53	0.62	0.76	0.94	0.0

Actions

This ILO was largely met. I will not make any drastic changes to the curriculum regarding this ILO.

Assessment of ILO #8 (Chi-Square Test)

One question on the final exam required students to conduct a thorough analysis of data with a "Chi-Square Test." The question was graded with a rubric that identified the eleven steps of any hypothesis test (this rubric was supplied to the students before and during the exam). This question was meant to assess the "two-sample categorical" portion of ILO #8.

My two goals for student achievement of this ILO are the same as those described above for the two-sample t-test. Specifically,

1. The Q1 of proportional scores on this question will be 0.50 or better.
2. The median of proportional scores on this question will be 0.80 or better.

Results

Student achievement on this question varied substantially among sections (Figure 4). The first goal was met by all sections except for the two sections in Fall, 2013 ("F13.1" and "F13.2"). The second goal was met in five of eight sections.

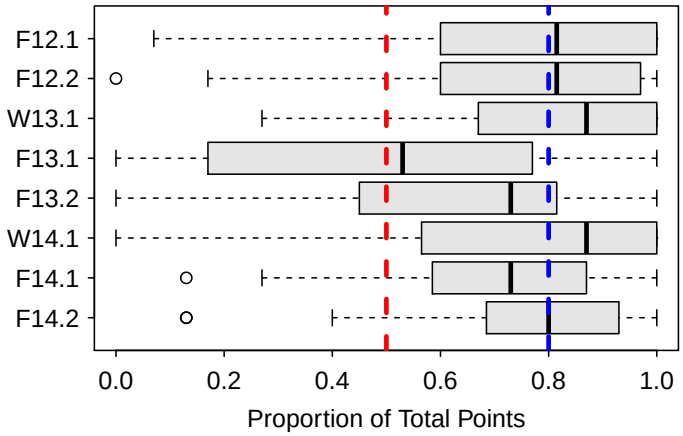

Figure 4. Boxplot of proportional scores on the 'Chi-Square Test' question (part of ILO #8) by section (same abbreviations as in Figure 1. The red and blue vertical dashed lines are the goals for the Q1 and median, respectively.

Table 3. Summary statistics of proportional scores on the 'Chi-Square Test' question (ILO #1) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	**median**	median	Q3	% zero	percZero
F12.1	28	28	0.76	0.17	0.38	0.64	0.76	0.91	1.00	0.0
F12.2	26	26	0.73	0.18	0.38	0.60	0.74	0.90	0.94	0.0
W13.1	21	21	0.79	0.13	0.50	0.71	0.82	0.88	1.00	0.0
F13.1	25	25	0.52	0.31	0.00	0.32	0.44	0.79	0.97	12.0
F13.2	23	23	0.65	0.29	0.00	0.50	0.76	0.82	0.97	4.3
W14.1	28	28	0.73	0.27	0.06	0.56	0.84	0.94	1.00	0.0
F14.1	23	23	0.62	0.19	0.15	0.52	0.65	0.72	0.88	0.0
F14.2	24	24	0.63	0.22	0.12	0.53	0.62	0.76	0.94	0.0

Actions

The objectives for this ILO were largely met. However, the large variability in proportional scores within a section suggests a wide range of achievement with regards to this ILO. This variability may be due to the lecture material related to this ILO being the last material covered in each semester and often rushed. Furthermore, I did not collect homework related to this ILO; thus, students did not receive any feedback on their performance with this material (though, several worked examples with an answer key were provided). In the Fall, 2014 sections I created more room in the lecture schedule for work on this topic and pushed the final exam back to the last possible date to allow students more time with this material. This may have narrowed the distribution of proportional scores. In future sections, I will consider methods to "force" students to do practice exercises on this topic.

Assessment of ILO #3 (Exploratory Data Analysis)

Three questions on the final exam required students to conduct exploratory data analyses (EDA). The analyses below assess their univariate and bivariate EDAs for quantitative data.

As this is fairly easy material that is foundational to all statistical analyses, I hoped that most student would perform well on these questions. I also expected most "high achieving" students to provide very strong answers to these questions. I translated these hopes into the following two goals (for each of the univariate and bivariate questions):

1. The Q1 of proportional scores on this question will be 0.7 or better.
2. The median of proportional scores on this question will be 0.9 or better.

Results

The first goal was met in only three sections, with less success recently (Figure 5). The second goal was met by four of eight sections. There is considerable variability within sections, especially recently, which suggests that a large number of students performed very poorly on this question.

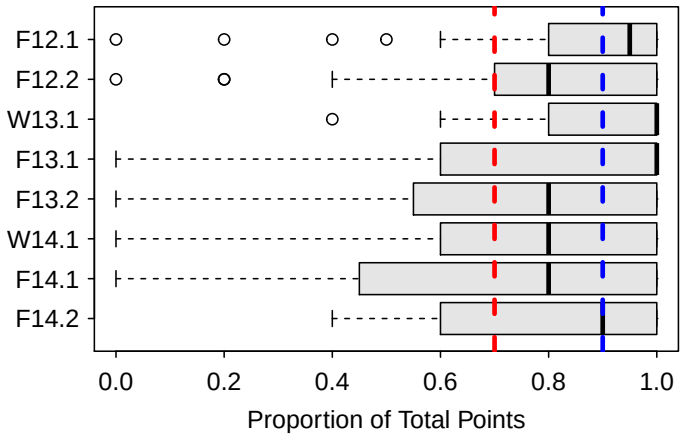

Figure 5. Boxplot of proportional scores on the 'univariate EDA' question (part of ILO #3) by section (same abbreviations as in Figure 1). The red and blue vertical dashed lines are the goals for the Q1 and median, respectively.

The first goal was met in only three sections, with less success recently (Figure 6). The second goal was met by only one section. There is considerable variability within sections, especially in recent sections, which suggests that a large number of students performed very poorly on this questions.

Table 4. Summary statistics of proportional scores on the 'univariate EDA' question (ILO #3) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	**median**	median	Q3	% zero	percZero
F12.1	28	28	0.82	0.27	0.00	0.80	0.95	1.00	1.00	3.6
F12.2	26	26	0.77	0.28	0.00	0.70	0.80	1.00	1.00	3.8
W13.1	21	21	0.86	0.19	0.40	0.80	1.00	1.00	1.00	0.0
F13.1	25	25	0.79	0.30	0.00	0.60	1.00	1.00	1.00	4.0
F13.2	23	23	0.74	0.28	0.00	0.55	0.80	1.00	1.00	4.3
W14.1	28	28	0.74	0.32	0.00	0.60	0.80	1.00	1.00	7.1
F14.1	23	23	0.69	0.33	0.00	0.45	0.80	1.00	1.00	8.7
F14.2	24	24	0.82	0.21	0.40	0.60	0.90	1.00	1.00	0.0

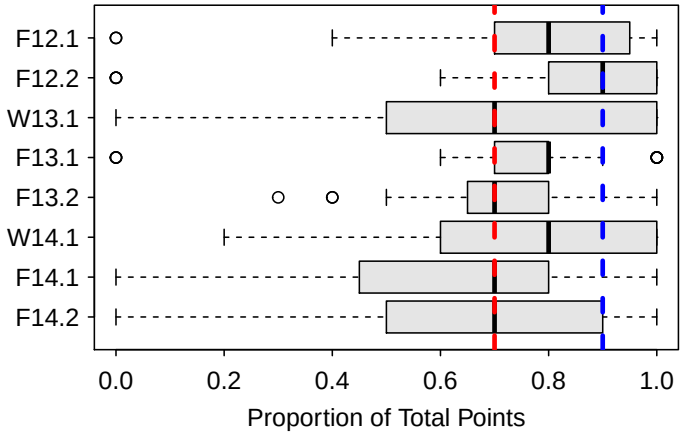

Figure 6. Boxplot of proportional scores on the 'bivariate EDA' question (part of ILO #3) by section (same abbreviations as in Figure 1). The red and blue vertical dashed lines are the goals for the Q1 and median, respectively.

Actions

The objectives for these two portions of this ILO were largely not met. The univariate EDA is one of the earliest topics in the course and students may be rusty with it by the end of the semester. Additionally, it is one of the easiest concepts in the course so students may not prepare for it as thoroughly, thinking that they will be able to perform it adequately on the final exam. Students generally perform worse on the bivariate EDA questions, so my objectives may be too high. However, it is clear that students need more practice late in the semester with both of these topics.

Assessment of ILO #4 (Linear Regression)

Three questions on the final exam required students to answer questions from linear regression results. Scores on these three questions were summed to provide an overall assessment of the student's ability to apply linear regression results.

Linear regression is a foundational statistical method in nearly all fields of study. Therefore, I hoped that most student would perform well on these questions. I also would like to see most "high achieving" students provide very strong answers to these questions. I translated these hopes into the following two goals:

1. The Q1 of proportional scores on this question will be 0.5 or better.
2. The median of proportional scores on this question will be 0.9 or better.

Table 5. Summary statistics of proportional scores on the 'bivariate EDA' question (ILO #3) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	**median**	median	Q3	% zero	percZero
F12.1	28	28	0.76	0.27	0.00	0.75	0.80	0.92	1.00	7.1
F12.2	26	26	0.80	0.31	0.00	0.80	0.90	1.00	1.00	11.5
W13.1	21	21	0.67	0.30	0.00	0.50	0.70	1.00	1.00	4.8
F13.1	25	25	0.72	0.30	0.00	0.70	0.80	0.80	1.00	12.0
F13.2	23	23	0.70	0.18	0.30	0.65	0.70	0.80	1.00	0.0
W14.1	28	28	0.78	0.22	0.20	0.60	0.80	1.00	1.00	0.0
F14.1	23	23	0.62	0.32	0.00	0.45	0.70	0.80	1.00	13.0
F14.2	24	24	0.67	0.28	0.00	0.50	0.70	0.90	1.00	4.2

Results

Student achievement on the univariate EDA question varied substantially (Figure 7). Neither goal was met by any section.

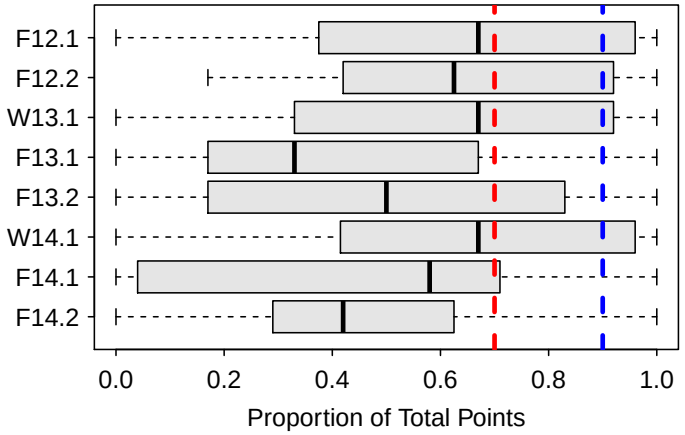

Figure 7. Boxplot of proportional scores on the 'linear regression' question (part of ILO #4) by section (same abbreviations as in Figure 1). The red and blue vertical dashed lines are the goals for the Q1 and median, respectively.

Table 6. Summary statistics of proportional scores on the 'linear regression' question (ILO #4) by section (same abbreviations as in Figure 1).

Section	n	nvalid	mean	sd	**Q1**	**median**	median	Q3	% zero	percZero
F12.1	28	28	0.63	0.30	0.00	0.40	0.67	0.94	1.00	7.1
F12.2	26	26	0.64	0.29	0.17	0.42	0.62	0.92	1.00	0.0
W13.1	21	21	0.60	0.29	0.00	0.33	0.67	0.92	1.00	4.8
F13.1	25	25	0.40	0.33	0.00	0.17	0.33	0.67	1.00	24.0
F13.2	23	23	0.50	0.38	0.00	0.17	0.50	0.83	1.00	21.7
W14.1	28	28	0.66	0.35	0.00	0.46	0.67	0.94	1.00	14.3
F14.1	23	23	0.46	0.36	0.00	0.04	0.58	0.71	1.00	26.1
F14.2	24	24	0.44	0.31	0.00	0.31	0.42	0.60	1.00	20.8

Actions

The assessment questions may not adequately address the objectives. Thus, I will reconsider the questions in future final exams. Regardless, it is apparent that students need more practice with linear regression methods and applications. I will explore methods to incorporate more practice into the class.

Future Assessments

Future assessments will continue to monitor student achievement on these five ILOs. In addition, I will develop analyses to assess student achievement on other ILOs.

```
Reproducibility Information
  Compiled Date: Wed Aug 26 2015
  Compiled Time: 8:39:23 PM
  Code Execution Time: 3.58 s

  R Version: R version 3.2.2 (2015-08-14)
  System: Windows, i386-w64-mingw32/i386 (32-bit)
  Base Packages: base, datasets, graphics, grDevices, methods, stats, utils
  Required Packages: FSA, car, xtable, magrittr, dplyr, knitr, plotrix, readxl
    and their dependencies (assertthat, DBI, digest, evaluate, formatR,
    graphics, grDevices, highr, lazyeval, markdown, MASS, methods, mgcv, nnet,
    pbkrtest, plyr, quantreg, R6, Rcpp, stats, stringr, tools, utils, yaml)
  Other Packages: car_2.0-26, dplyr_0.4.2, FSA_0.7.7, knitr_1.11, magrittr_1.5,
    plotrix_3.5-12, readxl_0.1.0, xtable_1.7-4
  Loaded-Only Packages: assertthat_0.1, DBI_0.3.1, evaluate_0.7.2, formatR_1.2,
    grid_3.2.2, lattice_0.20-33, lazyeval_0.1.10, lme4_1.1-9, MASS_7.3-43,
    Matrix_1.2-2, MatrixModels_0.4-0, mgcv_1.8-7, minqa_1.2.4, nlme_3.1-122,
    nloptr_1.0.4, nnet_7.3-10, parallel_3.2.2, pbkrtest_0.4-2, plyr_1.8.3,
    quantreg_5.18, R6_2.1.1, Rcpp_0.12.0, SparseM_1.7, splines_3.2.2,
    stringi_0.5-5, stringr_1.0.0, tools_3.2.2
```