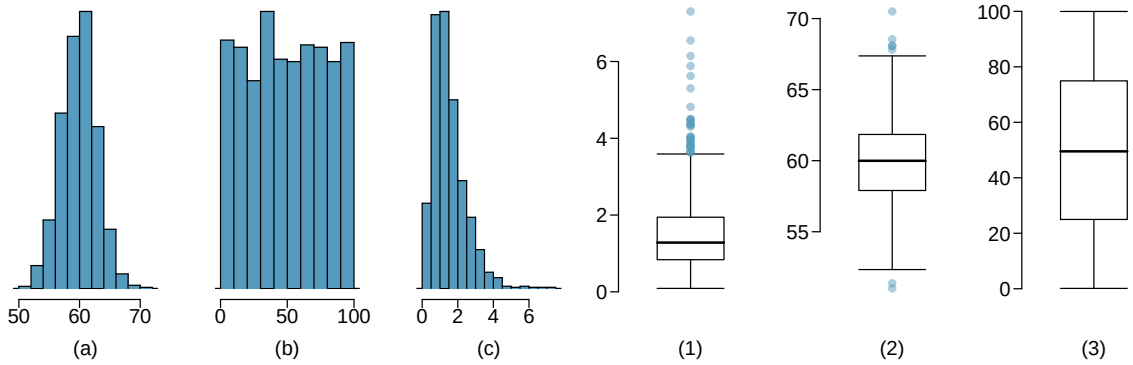
Homework №3

Math 107, Spring 2016 Due: April 8 by 4:00 pm

Problem 1

Describe the distribution in the histograms below and match them to the box plots.



Problem 2

In a class of 25 students, 24 of them took an exam in class and 1 student took a make-up exam the following day. The professor graded the first batch of 24 exams and found an average score of 74 points with a standard deviation of 8.9 points. The student who took the make-up the following day scored 64 points on the exam.

(a) Does the new student's score increase or decrease the average score?

(b) What is the new average?

(c) Does the new student's score increase or decrease the standard deviation of the scores?

Problem 3

Sophia who took the Graduate Record Examination (GRE) scored 160 on the Verbal Reasoning section and 157 on the quantitative reasoning section. The mean score for verbal reasoning section for all test takers was 151 with a standard deviation of 7, and the mean score for the Quantitative Reasoning was 153 with a standard deviation of 7.67. Suppose that both distributions are approximately symmetric and bell-shaped.

(a) What is Sophia's z-score on the verbal reasoning section?

(b) What is Sophia's z-score on the quantitative reasoning section?

(c) What do the Z-scores from parts (a) and (b) tell you?

(d) Approximately 95% of quantitative reasoning scores fall between what two values?

Problem 4

(This problem is inspired by 2.172 in the textbook.) The website TED.com offers free short presentations, called TED Talks, on a variety of interesting subjects. One of the talks is called "The Happy Planet Index," by Nic Marks.[1] Marks comments that we regularly measure and report economic data on countries, such as Gross National Product, when we really ought to be measuring the well-being of the people in the countries. He calls this measure Happiness, with larger numbers indicating greater happiness, health, and well-being.

You can find a tidy version of the 2012 Happy Planet Index (`hpi-tidy.csv`) on the course webpage (An untidy version can be downloaded from http://www.happyplanetindex.org/data/). In this homework problem you will explore the Happy Planet Index using R. Please copy all plots from RStudio into a typesetting program to reduce wasted paper.

A basic description of all of the variables included in the data set is given below:

Variable	Description
HPIRank	HPI rank for the country
Country	Name of country
LifeExpectancy	Average life expectancy (in years)
Wellbeing	"Ladder of Life" index from the Gallup World Poll (0 = worst possible life, 10 = best possible life)
HappyLifeYears	Index variable combining life expectancy and well-being
Footprint	Ecological footprint–a measure of the per capita ecological impact
HappyPlanetIndex	Happy Planet Index (0–100 scale)
Population	Population (in millions)
GDPcapita	Gross Domestic Product (per capita)
GovernanceRank	Governance ranking (1 = highest)
Region	Region of the world

(a) What are the cases?

(b) List each variable in the data set and classify it as either quantitative or categorical.

(c) Create a histogram of the Happy Planet Index scores and decribe the distribution, mentioning the number of modes, the shape, and the absence/presence of outliers. Remember that you should experiment with the bin-width until you find one that works well.

(d) Create a boxplot for the Happy Planet Index scores. What aspects of the distribution are easier to see using this boxplot than the histogram you created in part (c)? What aspects are harder/impossible to see?

(e) Create a density plot of the ecological footprint and decribe the distribution, mentioning the number of modes and the shape.

(f) Calculate the mean and standard deviation of the Happy Planet Index scores. Why do we prefer to report the mean and standard deviation in this situation?

(g) Calculate the five-number summary for ecological footprint. Why do we prefer to report the five-number summary in this situation?

[1] Marks, N. "The Happy Planet Index," www.TED.com/talks, August 29, 2010.