Tracking Your Life Project
The "Who" is You!
25 points
DUE: FRIDAY, DECEMBER 19th

The inspiration for this project came from Dan Meyer (blog.mrmeyer.com) & Nicholas Felton (feltron.com)

Pre-thinking task: Come up with at least 2 variables (at least 2 must be quantitative) regarding your life that you would want to track daily over the course of at least 3 weeks (21 days, or 15 school days). The variables do not need to be related, and you can interpret "in your life" rather loosely (e.g. you could record the temperature). Examples of variables for this project are: the number of minutes spent watching TV each day, how many hours of sleep you get each night, how many calories you consume on a daily basis, the number of text messages you send/receive every day, and how stressed out you feel (on a scale from 1-5, for example). Pick things that you are interested in and that you can collect data on fairly easily. In this day and age, technology makes data collection (and analysis) easier than ever, and in some cases, completely automatic. Therefore, you may want to use an app to help you. There are many, many options: *Datalove*, *Tally track*, *D.A.T.A.*, *MyFitnessPal*, *Lift*, *Way of Life*, etc.

Collection task: Once you receive approval from me, you can begin collecting your data. Using a notebook, memo pad, phone or other data collection instrument, keep track of the value of each of your variables on a daily basis. Think about what you are going to do if no data occurs on a particular date for a particular variable: will you record it as a zero, or leave it blank? Think about units and precision. Remember to note the date, as well as any other aspects of the day that you think might be pertinent. Based on past experience, I recommend recording whether or not it was a school day or night. In general, it's better to collect more data than less because you might need it later. Remember that collection is arguably the most important activity in Statistics, so be careful, diligent, consistent, organized, and neat in this regard. Remember that someone else has to be able to pick up your data collection notebook and understand the contents easily without you explaining it to them. Also, it is a good idea to back up your data periodically in case you lose your notebook or your phone, or if a technological glitch occurs. Taking photos of your notebook or storing your data in multiple locations are ways of reducing the probability of ultimate data loss. You will submit your raw data (a spreadsheet is okay).

Analysis task: Import the data into a spreadsheet, organize it in a clear fashion, and summarize and analyze it in the various ways that we explore in this course (see Ch. 3-5, and 7-9). Look deep into the data; look for patterns and relationships. Play around with different ways of displaying the data to see if you can uncover interesting things. Remember that the computer is a great tool for furthering your understanding of what's really going on. I hope to be able to show you some of the ways to analyze data with Microsoft Excel and Minitab. For quantitative variables, I expect you will provide a summary of each of the variables using all the descriptive statistics we learn about in Chapters 4 & 5 as well as time-plots, boxplots, and histograms. I expect at least one scatterplot showing the relationship (if any) between the quantitative variables. The scatterplot needs to show the regression equation as well as r and r^2. For the categorical variables (optional), I expect percentages, contingency tables, and bar or pie charts. You can also analyze the relationship between one or more of the quantitative variables and one or more of the categorical variables.

Interpretation task: After reviewing your summary statistics as well as your graphs, draw meaningful conclusions from them using the knowledge gained in this course. Talk about the meaning of the statistics. Describe the timeplots, distributions, and associations using the key focus items found in Ch. 4 and 7. Make sure that you express when something is a subjective conclusion versus something that is a hard fact. Remember, it's important to be clear and honest (to ourselves and to others) about what we know and do not know. And remember that correlation doesn't prove causation.

Presentation task: This part of the project is somewhat open-ended. Think about creative, clear, well-organized, well-labeled, and appropriate ways to illustrate your findings. Remember to include a description of the methods employed for collection and analysis, as well as your conclusions. In the past, students have done reports, PowerPoints, tri-folds, posters, and scrapbooks.

Project Grading Checklist

Collection

Raw data submitted? (spreadsheet okay)

Contents of data notebook and/or spreadsheet easy to comprehend?

Data recorded on a daily basis over the course of the specified study period?

Anomalies and other comments recorded in data notebook/spreadsheet?

No doubts regarding the integrity of the data?

Data collection methods explained in detail?

Errors in data collection mentioned in presentation?

Analysis

If applicable

Categorical variables analyzed in a contingency table?

Categorical variables analyzed using percentages?

Pie charts present?

Bar charts present?

Contingency table correct?

Summary statistics for quantitative variables present? (min, Q1, med, Q3, max, mean, IQR, standard deviation, range)

Scatterplot(s) present?

LSRL equation, r and r^2 present?

Timeplots present?

Boxplots present?

Histograms present?

Evidence of an in-depth analysis of the data?

Are outliers taken into account?

Any novel means of analysis?

Interpretation

Timeplots described?

Distributions described?

Associations described?

Regression equation interpreted correctly?

Are the descriptive statistics interpreted thoroughly and correctly?

Are the graphs interpreted thoroughly and correctly?

Are the conclusions sophisticated?

If there is potentially another explanation for results seen, does the author mention this?

Potential explanation for outliers given?

Presentation

Is the presentation interesting?

Is the presentation original?

Does the presentation evidence a substantial amount of time and effort devoted to this project?

Is the data, analysis, and interpretation presented in a way that is easy for the average Statistics student to understand?

Does the author/presenter avoid overwhelming the reader/audience with too many graphs and tables?