Project eCSite Unit: Mood Analysis

Content Areas: computer science, science & technology, health, mood, data analysis, statistics

Grade Level: Secondary (9 - 12)

Computational Thinking Connection: This unit teaches students about data analysis and basic statistical analysis and visualization in the context of analyzing data collected about their own health and mood over time.

Prerequisite Knowledge: None.

Time Required: 5 minutes per class for ~1 month, 1 entire class

Related Lessons / Activities:

1. Data is collected, either manually or with clickers, over a period of a month or more
2. Some preparation and analysis of the data is done, mainly involving getting the data into a spreadsheet and uploading it to Google Fusion Tables
3. Students analyze the data and fill out a worksheet while working through an in class activity

Contents:

1. Daily Mood Check Survey (1 page)
2. Instructions for Preparing & Uploading data (1 page)
3. Mood Analysis Activity (4 pages)

Other Notes: N/A

Contributors: Caleb Phillips and Thomas Kummer

Supporting Programs: NSF and GK-12 eCSite program
Daily Mood Check

Period #_____
Clicker #/Name ___

1) How do you feel?
   A - Bad
   B - Not so great
   C - OK
   D - Pretty good
   E - Great

2) How well did you sleep last night?
   A - Bad
   B - Not Great
   C - OK
   D - Pretty Good
   E - Great

3) Did you have a decent breakfast this morning?
   A - Yes
   B - No

4) On a scale of 1-5 what is your current stress level?
   A - 1 (Low)
   B - 2
   C - 3
   D - 4
   E - 5 (High)

5) How much exercise did you get yesterday?
   A - None
   B - 30 Minutes
   C - 60 Minutes
   D - 90 Minutes
   E - 120 Minutes or more

Instructions for Uploading to Google Fusion Tables
1) Look at the data included with the unit as an example (“Mood Data.csv”). Data in this format can be uploaded directly to Google Fusion Tables and used for the activity: [www.google.com/fusiontables](http://www.google.com/fusiontables)

2) If desired, daily temperature and snowfall data can be (manually or programmatically) extracted from NOAA data: [http://www.noaa.gov/](http://www.noaa.gov/)

3) Follow the instructions on fusion tables for uploading data.
Boulder High Mood-Analysis Activity  
April 27, 2012  

Due: Monday, April 30th  
Link: [http://swcurl.com/uXNK2](http://swcurl.com/uXNK2)  
<- Go here (Using the “Chrome” Browser)  

Introduction  

In the beginning of the semester, we used clickers to collect some data on your mood, and general wellness (sleep, exercise, eating habits, etc.). What we want to try to figure out is:  

*What affects your mood?*  

In this way, we can think about personal health, attitude and wellness as inter-related things that you have control over. In this activity, we’ll dig into that data which is 5 weeks of continuous data for 5 different health classes at Boulder High.  

We’ll begin this activity in class on Friday, but it’s **due the following Monday**. Please email the completed activity to your teacher. Make the subject look like: “Period 3: Esmeralda Lucas” (if you were in period 3 and your name was Esmeralda :P)  

Instructions & Questions  

We’re going to use something called “Fusion Tables” to look at the data. Pull up this link in a browser (the “Chrome” browser works best if you have it):  

[https://www.google.com/fusiontables/DataSource?docid=1xwkIjqK2_xx0fhm_EP49RbHpBBUApz1p1Ynge08](https://www.google.com/fusiontables/DataSource?docid=1xwkIjqK2_xx0fhm_EP49RbHpBBUApz1p1Ynge08)  

When the data pulls up you’ll be able to see it as a big table. Each row is a particular student’s entries on a particular day. Each column is a question or data item.  

Some of the columns are pretty clear, but some of them are less clear. Here’s a quick explanation for each:  

- date - the date the data is from, duh  
- period - which period the student is in Health class  
- clicker - their clicker number  
- mood - from 1-5 (feel bad...feel great)  
- sleep - from 1-5 (slept bad...slept great)  
- breakfast - 1 = yes, had good breakfast, 2 = no, didn’t  
- stress - stress level from 1-5 (low...high)  
- exercise - how much from 1-5 (none...120 minutes)  
- avgtemp - the average temperature that day  
- avgwind - the average windspeed that day
- snow - inches of snow that day (remember, we did this in January!)
- dow - day of week (1=monday, …, 5=friday)
- sid - unique id for each student

**Q1:** Do you expect that any of these **variables** might be correlated with one another? If so, why? If not, why not?

Now, click on “options”. Under the filter tab set “sid” = 310. Then click apply. This will just show the data for student 310 (it could be you!). Next click, “Visualize”, then “Line”. For the y-axis, choose “mood”. For the x-axis, try a few different variables.

**Q2:** Do there appear to be any correlation between mood and these other variables for this student? Do you think this student it **representative** (i.e., similar to how everyone else would look?). Why or why not?

One way to test whether a particular **variable** is affected by some number of **factors** (i.e., other variables that may influence it), is with a statistical test called an ANOVA. Pull up this link in a different window or tab:

http://smallwhitecube.com/ft/

This is a little application that uses the same data to perform an ANOVA using a statistical software package called ‘R’. You pick the variable you want and the factors you want to test and it will tell you if the factors (or a combination of those factors) is **significant**.

Try it with the variable set to mood and the other factors set to those you think might be correlated. If you select a bunch of factors, the output will be huge, so start with just one or two that you think might be correlated and could also be related to one another (do you think you sleep better after you’ve exercised?).

In the output, you can see which factors are significant because they will have three *** asterisks next to them. The Pr(>F) field is the “confidence”, and a smaller number is better.
**Q3:** Which factors did you find that were significant? Did this match your expectation? Why or why not?

Let’s take a look at sleep in particular. Go back to the fusion tables.

Go back to the filter and clear it. Now, click the “aggregate” tab. Under “show aggregate” pick the average mood. Then under aggregated by, choose “sleep”. For the plot, choose “sleep” for the x-axis and “AVERAGE mood” for the y-axis.

**Q4:** Does there appear to be a correlation between sleep and mood? What is the average mood level for people who say they slept great (i.e., sleep = 5)? What is the average mood of people who say they slept bad (i.e., sleep = 1)?

Take a look at this plot:

https://www.google.com/fusiontables/DataSource?snapid=S4715664t3D

This plot shows the average mood for each period for those students who said they slept badly (i.e., sleep = 1). Compare it to this plot, which is the same, but for folks who slept well (i.e., sleep = 5):

https://www.google.com/fusiontables/DataSource?snapid=S470492oNdI

**Q5:** Can you explain these plots? What seems to be going on?

**Q6:** Briefly reflect on this activity. Do you think the health choices you make affects your mood? Or, is mood seemly random? Would what you’ve seen here change the way you think about whether sleep, breakfast, or exercise are more important? What’s the best period to take a class if you want to learn the most?
If you have some time, play with other combinations of variables to see if you can find any other interesting relationships. For instance, do you expect that people in Colorado would be in a better or worse mood when it snows? What about when it’s windy?