Rob Acheson
racheson@fas.harvard.edu
CS-171 - Homework 1
2/7/2013

Find a visualization not discussed in class or used in a homework and answer the following questions pertaining to that visualization. Attach the visualization as a screenshot in your submission.

My chosen visualization is a screen from Ario Sevit's entry to the Hubway Data Visualization Challenge depicting the average speed of cyclists by time of day. Data is broken up between registered and unregistered cyclists, weekday and weekend speeds.

http://ariofsevit.com/hubway/images/daytimespeedtype.PNG

Consider Bertin's characterization of visual variables (position, size, shape, value, color, orientation, and texture). Pick 2 of Bertin's visual variables, and discuss them in relation to your visualization.

Two major visual variables discussed by Bertin in the chosen visualization are position and size. Speed is positionally encoded in the y-axis and time of day in the x-axis. From this, it is clear to see the variations in average speed over the course of a day. A less effective encoding method is the use of size to determine the number of riders at any given moment. No scale is provided, not that it would make much of a difference in determining accurate counts. What the size does convey is that few riders are on the road in the early morning and there are surges in numbers around 9 am and 6pm.

Munzner proposed a nested model for visualization design and validation. Discuss/validate your visualization with respect to domain problem characterization and data/operation abstraction design.

Munzner describes domain problem characterization as the first step in design. (Munzner p 921) Considering that this visualization was created for a Hubway sponsored contest, one could ascertain the purpose was for Hubway stakeholders to gain insight into their user habits and overall traffic patterns. Sevit's visualization set out to provide insight into large trends in traffic. Munzner describes the second step, data/operation abstraction design, as the step where data and vocabulary are evaluated and used to determine what tool is created. (Munzner p 922) Sevit was able to isolate a few important data types time, speed and number of users.

Based on Cleveland and McGill's results, does your visualization embody good practices (i.e. can people accurately perform the tasks based on the encodings?)

Sevit's graphs are easy to judge accurately at they encoded as a triple scatter plot. Its use of position on a common scale efficiently describes both the average speed and corresponding time of day. Adding area to these allows us to quickly determine which points represent more volume of bicycles on the road. By variations in color and saturation we can determine which type of user the areas represent and quickly notice that registered users ride faster than non-registered. Adding lines, similar to a "lowess curve" smooth the data and allow us to see in more detail the differences between user behavior on weekday and weekends.

Do you agree that visualization is a functional art? Explain.
When thinking of functional art, disciplines such as architecture and industrial design come to mind. Both processes involve creating something that serves a purpose, but is greatly benefitted by aesthetics and principles of traditional design. I would place visualization into this category as well.

Ask yourself what the designer is trying to convey and think of three to four possible tasks this visualization should help you with. Does the visualization achieve any of your tasks? (To view an example, see Albert Cairo, pages 26-28.)

Cairo states that visualizations are a tool that must present, compare, organize and correlate information. (Albert Cairo, pages 26-28.) Sevit's work successfully does all of these. User speeds are clearly presented and organized in the chart and we can see this data change over time. The use of color to separate registered and non-registered cyclists allows us to quickly see the differences between the casual user's speeds compared with the commuter. The size of the circles in the scatter plot allow us to see the relationship between the time of day and the number of users on the road.