Overall goal: To see how the application of known story curves to the presentation of (visualized) quantitative data can increase recall of information and data trends.

* Accentuate the story curve in text by using positive and negative sentiment words to describe changes in data.
* Prime user with explanation of story curve.

Data Domain: Raw quantitative data consists of measurements from Lake George taken between 1980 through 2009. Additionally, writing and figures from The State of the Lake report and the Lake George Chemical Monitoring Program site are used as human-level samples of writing.

Data and writing about three variables were chosen:

* Secchi Depth
* Total Phosphorous
* Chlorophyll A

Each variable’s data is grouped by Yearly, Seasonal (monthly), and/or Site (spatial) criteria. The variables were chosen due to their relationship. Phosphorous promotes the growth of Algae blooms, which raise the Chlorophyll A levels and which should reduced Secchi Depth. All three variables appear in both The State of the Lake report and on the Lake George Chemical Monitoring Program site, with human-written descriptions and visualizations of trends in the data gathered about each variable.

Test Materials:

For each variable, a data figure is chosen from one of the writing sources and it and its corresponding description are taken. A slightly edited version of this text is taken as the human writing source. Other variations contains some or all of the following story parts.

Story Parts (not necessarily in this order):

1. Story Curve Explanation - Explain the type of story curve that’s about to be used.

2. Sentiment Explanation – Explain what data changes are good and bad.

3. Domain Information – Give information about the measurement they’re about to look at.

4. Data Explanation – Tell them about the data itself.

Notes-to-self:  
- Language cannot be really separated from the task of portraying a story curve. A story curve is the theoretical curve of emotion/drama felt by the user; we cannot measure user sentiment/emotion as the story goes, but we can try to display appropriate amounts of positive or negative emotion using words with different sentiment. This is the only way we can really accentuate (or display at all) the story curve, since the agent doesn’t have compelling characters, a deep plot, or the ability to write deeper meanings into the text generated.

- Generated and trend-describing parts of the text, by themselves, are too short! It is trivial for users taking the tests to just repeat what the trends are. Use human-written text and explanations to lengthen, and chain several tests with different variations together.

- Just explaining the story curve before either type of explanation could have a priming effect, and could help the user remember just based on which story curve was explained.

- Are these stories too short, and the trends too trivial?

- If the direction of sentiment isn’t up-down, the story type explanations may cause confusion.

- Story curve explanations w/out the directional language.

* Human text vs. Story text only is not a balanced assessment, since human text includes a lot of extra explanation and domain-specific words. Difference between the two is too large, variation is not well controlled.
* Drama curves really only matter if the appropriate sentiment is displayed at the appropriate times.
  + Positive/negative needs to be portrayed.
  + Get lists of words from stock market for movement and such
* Generated visualizations don’t look good and human text does not discuss them
  + Try using existing visualizations from writing sources
* What variations should be used? User should be put through one “path” in a “grid” of variations
* What’s missing is language used in typical man-in-a-hole or rags-to-riches stories
* Have to let the user know what direction of data change their sentiment should be based on.

- Can’t separate text generation from the task of conveying dramatic curves

- Control sentiment generation at the text level

- Story-time doesn’t have to follow data-time. This allows for curves to be controlled regardless of the time axis in the underlying data.

- People need some explanation before they can begin to appreciate sentiment in the story

- what direction is good and bad

- Mentioning the story shape itself may be enough to help.