Research Questions & Hypotheses (part 2)

Today

• Discussion of Friday Crayon Exercise
• Assignment #3, Mini-Presentation volunteers?
• Paper for Friday: choice of paper on Parallel Coordinates Visualizations
Today’s Crayon Exercise

• Team up with someone from your side of the room, but in a different row
• Discuss your ideas for HW3: Big Data
• Pick one of the projects to use for this exercise...
• What is a specific Research Question that can be solved by studying this data? Write it down in 1-2 clear sentences.
• Brainstorm a specific (successful & interesting) visualization style for this data that will (attempt to) answer this question [this might require more than 1 week to complete, that’s ok, thing big/long term]
• Hypothesize 2 different specific detailed outputs of this data in this visualization style, and what conclusion will be drawn from this out as the answer to your research question (choose 2 of the following)
  – Your best, most likely prediction for what the data will look like
  – An example of a result that would confirm your hypothesis beyond a shadow-of-a-doubt
  – An example of what a surprising result would show
  – What failure of the visualization style to draw any conclusions from this data with respect to the research question

What is Science? from wikipedia & dictionary.com

• Science: the intellectual and practical activity encompassing the systematic study of the structure and behavior of the physical and natural world through observation and experiment
• Hypothesis: a supposition or proposed explanation made on the basis of limited evidence as a starting point for further investigation
What is Research? from wikipedia & dictionary.com

- Research question: Specifying the research question is one of the first methodological steps the investigator has to take when undertaking research. The research question must be accurately and clearly defined. Choosing a research question is the central element of both quantitative and qualitative research and in some cases it may precede construction of the conceptual framework of study. In all cases, it makes the theoretical assumptions in the framework more explicit, most of all it indicates what the researcher wants to know most and first.
  - To ask the NSF (National Science Foundation) for research money, we have to have a clear research question. Many computer scientist researchers fall in the trap of saying they will develop code, or implement something to solve a problem. Describing our work with this terms can doom a proposal, because these activities are not science (it’s not the National R&D Foundation!)

- Thesis Statement: A thesis statement is a short statement that summarizes the main point or claim of an essay, research paper, etc., and is developed, supported, and explained in the text by means of examples and evidence.

The Relationship Between the Research Question, Hypotheses, Specific Aims, and Long-Term Goals of the Project

Before you begin writing a grant proposal, take some time to map out your research strategy. A good first step is to formulate a research question.

A Research Question is a statement that identifies the phenomenon to be studied. For example, “What resources are helpful to new and minority drug abuse researchers?”

To develop a strong research question from your ideas, you should ask yourself these things:

- Do I know the field and its literature well?
- What are the important research questions in my field?
- What areas need further exploration?
- Could my study fill a gap? Lead to greater understanding?
- Has a great deal of research already been conducted in this topic area?
- Has this study been done before? If so, is there room for improvement?
- Is the timing right for this question to be answered? Is it a hot topic, or is it becoming obsolete?
- Would funding sources be interested?
- If you are proposing a service program, is the target community interested?
- Most importantly, will my study have a significant impact on the field?

http://www.theresearchassistant.com/tutorial/2-1.asp
A strong research idea should pass the "so what" test. Think about the potential impact of the research you are proposing. What is the benefit of answering your research question? Who will it help (and how)? If you cannot make a definitive statement about the purpose of your research, it is unlikely to be funded.

A research focus should be narrow, not broad-based. For example, “What can be done to prevent substance abuse?” is too large a question to answer. It would be better to begin with a more focused question such as “What is the relationship between specific early childhood experiences and subsequent substance-abusing behaviors?”

A well-thought-out and focused research question leads directly into your hypotheses. What predictions would you make about the phenomenon you are examining? This will be the foundation of your application.

Hypotheses are more specific predictions about the nature and direction of the relationship between two variables. For example, “Those researchers who utilize an online grant writing tutorial will have higher priority scores on their next grant application than those who do not.”

http://www.theresearchassistant.com/tutorial/2-1.asp

Strong hypotheses:

- Give insight into a research question;
- Are testable and measurable by the proposed experiments;
- Spring logically from the experience of the staff;

Normally, no more than three primary hypotheses should be proposed for a research study. A proposal that is hypothesis-driven is more likely to be funded than a “fishing expedition” or a primarily descriptive study.

Make sure you:

- Provide a rationale for your hypotheses—where did they come from, and why are they strong?
- Provide alternative possibilities for the hypotheses that could be tested—why did you choose the ones you did over others?

If you have good hypotheses, they will lead into your Specific Aims. Specific aims are the steps you are going to take to test your hypotheses and what you want to accomplish in the course of the grant period. Make sure:

- Your objectives are measurable and highly focused;
- Each hypothesis is matched with a specific aim;
- The aims are feasible, given the time and money you are requesting in the grant.

An example of a specific aim would be “Conduct a rigorous empirical evaluation of the online grant writing tutorial, comparing outcome and process measures from two groups—those with exposure to the tutorial and those without.”

http://www.theresearchassistant.com/tutorial/2-1.asp
Long-Term Goals:

- Why are you doing this research?
- What are the long-term implications?
- What will happen after the grant?
- What other avenues are open to explore?
- What is the ultimate application or use of the research?

These questions all relate to the long-term goal of your research, which should be an important undercurrent of the proposal. Again, they should be a logical extension of the research question, hypotheses, and specific aims.

It is also helpful to have a long-term plan for your own career development. Where would you like to see your career go in the next 5 years? How does the research you are proposing relate to that plan?

Now Write It Up...

http://www.theresearchassistant.com/tutorial/2-1.asp

Bolong & Theo

Research Question: How does the high temperature of a day indicate the total energy consumption of domestic households? To what certainty does the weather forecast predict total energy consumption of the grid for a given time period? Is AC the single biggest factor of variation in energy consumption?

Hypothesis: There is a direct correlation for energy rise in the summer due to air conditioning

Data Concerns: account for daylighting savings time, humans are unpredictable, need lots of data about many households, need smart grid before this data is available
Max & Shoom

Research Question: Are interactive/real-time improved audio simulation tools useful for architects during early design?

Experiment: Study on classroom/small meeting room/office acoustics

Hypothesis: Yes, it will be useful architecture students, on limited scope problems.

Why do we hypothesize this? Because our tool works for daylighting.

Concerns: 3D, materials, ...

Greg & Jaron

Research Question: What weapons are commonly taken together? Is there a dominant strategy? Or was design goal of many strategies successful?

[ Side Goal: find/fix bugs? Rebalance weapons ]

Hypothesis: Starting weapons strongly connected by thick lines. Defensive weapons commonly connected to everything. Offensive weapons no clear patterns. Late game weapons generally connected to each other in a relatively separated section.

Concerns: most people may just be using default weapons
Research Question: What cut of diamond sends most light out the top of the diamond (which is believed to be most attractive to consumers)? What properties differentiate popular diamond cuts from other cuts and other geometries?

Hypothesis: Teapots are better than ideal cut diamonds. Gemologists & heuristics are probably (mostly) right, will use computer simulation to prove, may find surprising other results.

Research Question: Is there a correlation between funding for elementary schools and scores on state reading tests?

Hypothesis: more funding = better results

Concerns: data processing – will be clumping on funding axis in funding per school, normalize/scale per school size
Caution: correlation is not causation! Are funding levels increased for schools with better scores last year?
Research Question: Is there structure to the first link on a wikipedia page?

Hypothesis: A very large % of pages will eventually lead to philosophy.
How to construct a page that breaks hypothesis: broken links, no links, cycles

Hypothesis: (failure case) there is an obscure subgroup of subjects that are connected to each other but not to philosophy
Research Question: Is color and effective visualization... for what application/datasets... What visualization is most effective for debugging details of meshing or simulation algorithm? Dealing with the limitations of 2D (What way do users find most intuitive? What visualization allows user to best understand actual 3D shape?)

Hypothesis: Using color to represent depth/zbuffer will be more effective that ____

---

Research Question: How does the class year at RPI affect the student’s major? How does a student’s major change over their 4 year timespan.

Hypothesis: Computer Science is on the rise? Younger students are more likely to be computer science. Certain majors have higher drop out rate, certain majors are common duals declared later.

[ Side goal: how does this compare to the Rensselaer plan? How does this correlate with admissions goals. ]

Concern: How do we get this data? Privacy? Properly dealing with students who are here < or > 8 semesters. Normalize for students who change majors and take a bit longer. Transfer students.
Today

- Discussion of Friday Crayon Exercise
- Assignment #3, Mini-Presentation volunteers?
- Paper for Friday: choice of paper on Parallel Coordinates Visualizations

Homework Assignment 3: due Monday @ 11:59pm
Big Data & Summarization

- Identify a “really big” dataset
  - That is research-related [Required for grad students]
  - Or is of personal interest [Undergrads only]
- Collect this dataset in a variety of sizes (start small)
- Parse/Process/Organize for a preliminary visualization
- Identify challenges/flaws: “Data is so big it…”
  - “… crashes my favorite visualization toolkit”
  - “… runs really, really slow”
  - “… font/data overlaps/is so small I can’t read it”
- As time permits, propose/implement automatic simplification/summarization of data
- Focus: Data Collection (primary) & Visualization Execution (secondary) & Visualization Design (secondary)

Team up with someone you haven’t worked with before [Optional]
Use a new visualization toolkit: Encouraged! But optional!
Today

• Discussion of Friday Crayon Exercise
• Assignment #3, Mini-Presentation volunteers?
• Paper for Friday: choice of paper on Parallel Coordinates Visualizations

Parallel Coordinates

http://eagereyes.org/techniques/parallel-coordinates
Parallel Coordinates

http://eagereyes.org/techniques/parallel-coordinates

“Uncovering Clusters in Crowded Parallel Coordinates Visualizations”, Artero, Ferreira de Oliveira, & Levkowitz, InfoVis 2004
"Angular Brushing of Extended Parallel Coordinates", Hauser, Ledermann, and Doleisch, InfoVis 2002