Lecture-1

Prof. Tulasi Prasad Sariki
SCSE, VIT, Chennai
www.learnersdesk.weebly.com

Outline

- Why have a human in the decision-making loop?
- Why have a computer in the loop?
- Why use an external representation?
- Why depend on vision?
- Why show the data in detail?
- Why use interactivity?
- Why is the vis idiom design space huge?

Outline

- Why focus on tasks?
- Why are most designs ineffective?
- Why care about effectiveness?
- Why is validation difficult?
- Why are there resource limitations?
- Why analyze vis?

Why Have a Human in the Loop?

- Vis allows people to analyze data when they don’t know exactly what questions they need to ask in advance.
- When people have well-defined questions to ask about data,
  - They can use purely computational techniques from fields such as stats and ML.
  - Some jobs that were once done by humans can now be completely automated.
  - If a fully automatic solution has been deemed to be acceptable, then no need for you to design a vis tool.

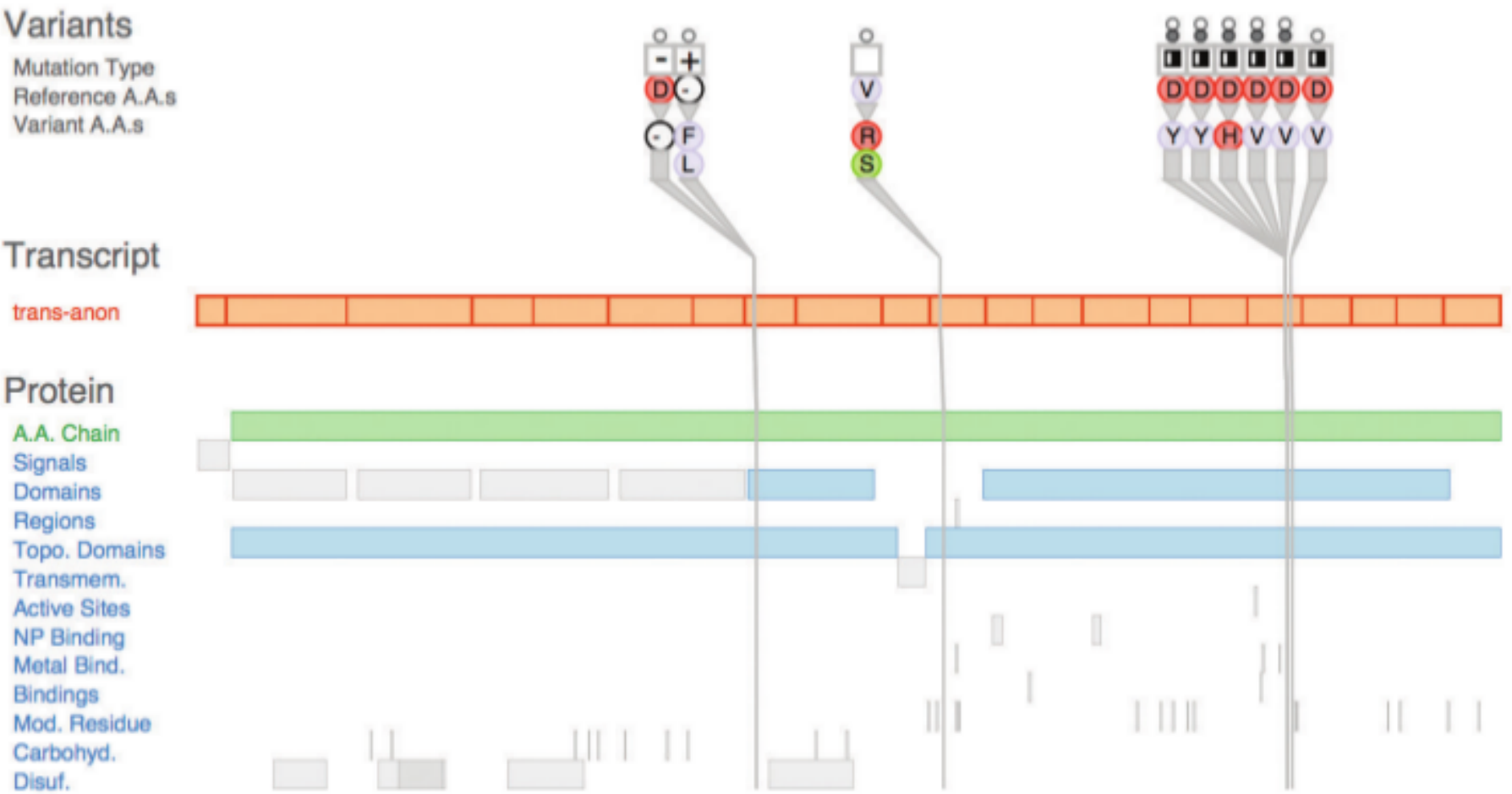
Why Have a Human in the Loop?

- Consider the **stock market trading**
  - Currently, there are many deployed systems for high frequency trading that make decisions about buying and selling stocks when certain market conditions hold, when a specific price is reached, for example, with no need at all for a time-consuming check from a human in the loop.
  - You would not want to design a vis tool to help a person make that check faster, because even an augmented human will not be able to reason about millions of stocks every second.

Why Have a Human in the Loop?

- Many analysis problems are ill specified: people don’t know how to approach the problem.
  - There are many possible questions to ask—anywhere from dozens to thousands or more—and people don’t know which of these many questions are the right ones in advance.
  - In such cases, the best path forward is an analysis process with a human in the loop, where you can exploit the powerful pattern detection properties of the human visual system in your design.
  - Vis systems are appropriate for use when your goal is to augment human capabilities, rather than completely replace the human in the loop.

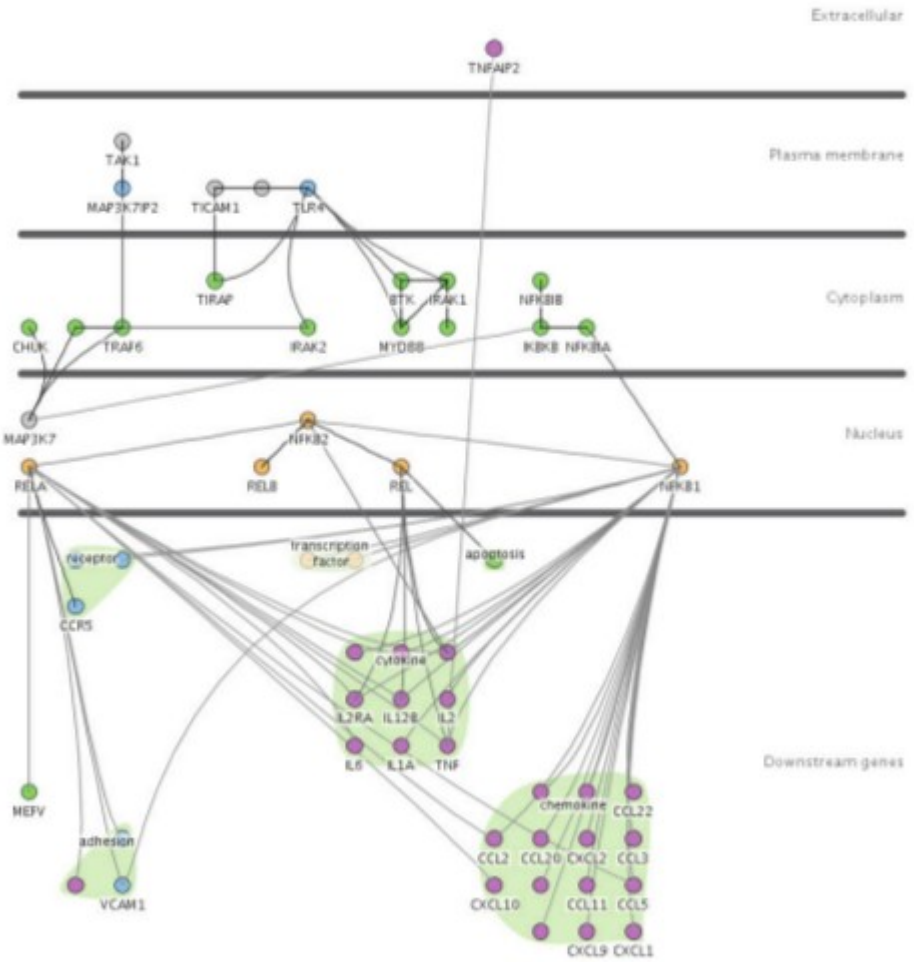
Why Have a Human in the Loop?



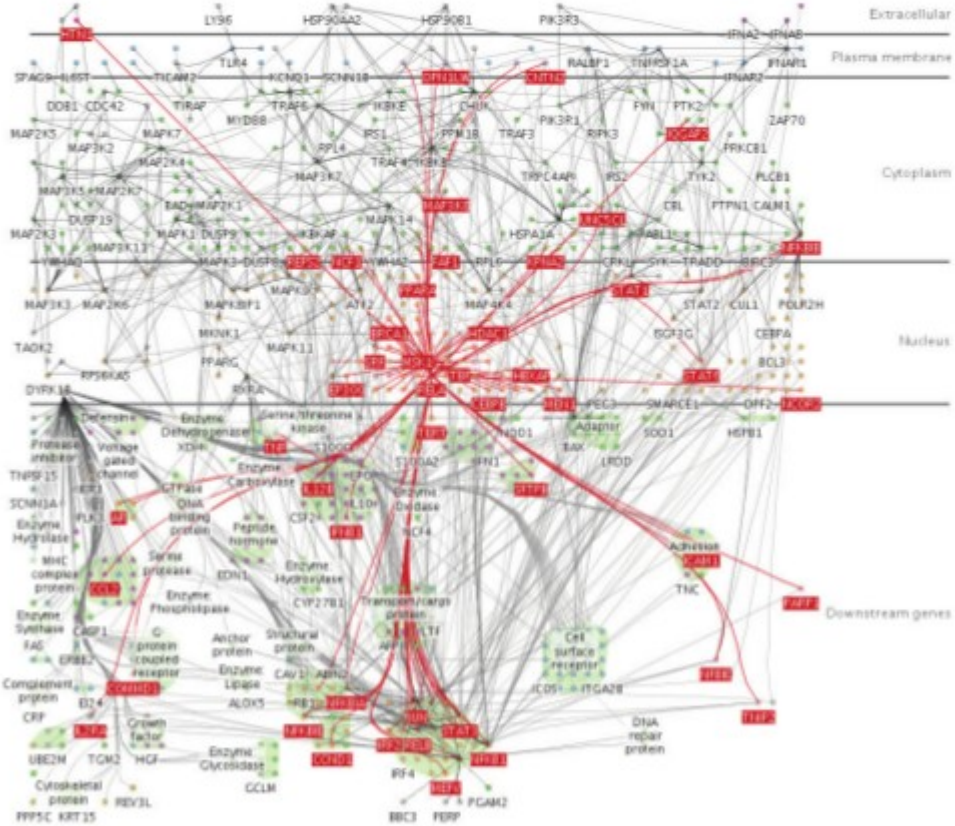
Tool supports biologists in assessing the impact of genetic variants by speeding up the exploratory analysis process.

Why Have a Computer in the Loop?

- Tools that allow people to explore or present large data sets that would be completely infeasible to draw by hand, thus opening up the possibility of seeing how data sets change over time.



(a)



(b)

(a) A small network of 57 nodes and 74 edges might be possible to lay out by hand with enough patience. (b) Automatic layout handles this large network of 760 nodes and 1269 edges and provides a substrate for interactive exploration

Why Use an External Representation?

- External representations augment human capacity by allowing us to surpass the limitations of our own internal cognition and memory.
- Information can be organized by spatial location, offering the possibility of accelerating both search and recognition.
  - Search can be sped up by grouping all the items needed for a specific problem-solving inference together at the same location.
  - Recognition can also be facilitated by grouping all the relevant information about one item in the same location, avoiding the need for matching remembered symbolic labels.

Why Depend on Vision?

- Visualization is based on exploiting the human visual system as a means of communication.
- The visual system provides a very high-bandwidth channel to our brains. A significant amount of visual information processing occurs in parallel at the preconscious level.

Why Show the Data in Detail?

- Vis tools help people in situations where seeing the dataset structure in detail is better than seeing only a brief summary of it.
- One of these situations occurs when exploring the data to find patterns, both to confirm expected ones and find unexpected ones.
- Another occurs when assessing the validity of a statistical model, to judge whether the model in fact fits the data.
- Statistical characterization of datasets is a very powerful approach, but it has the intrinsic limitation of losing information through summarization. Ex: illustrated by Anscombe's Quartet

Why Use Interactivity?

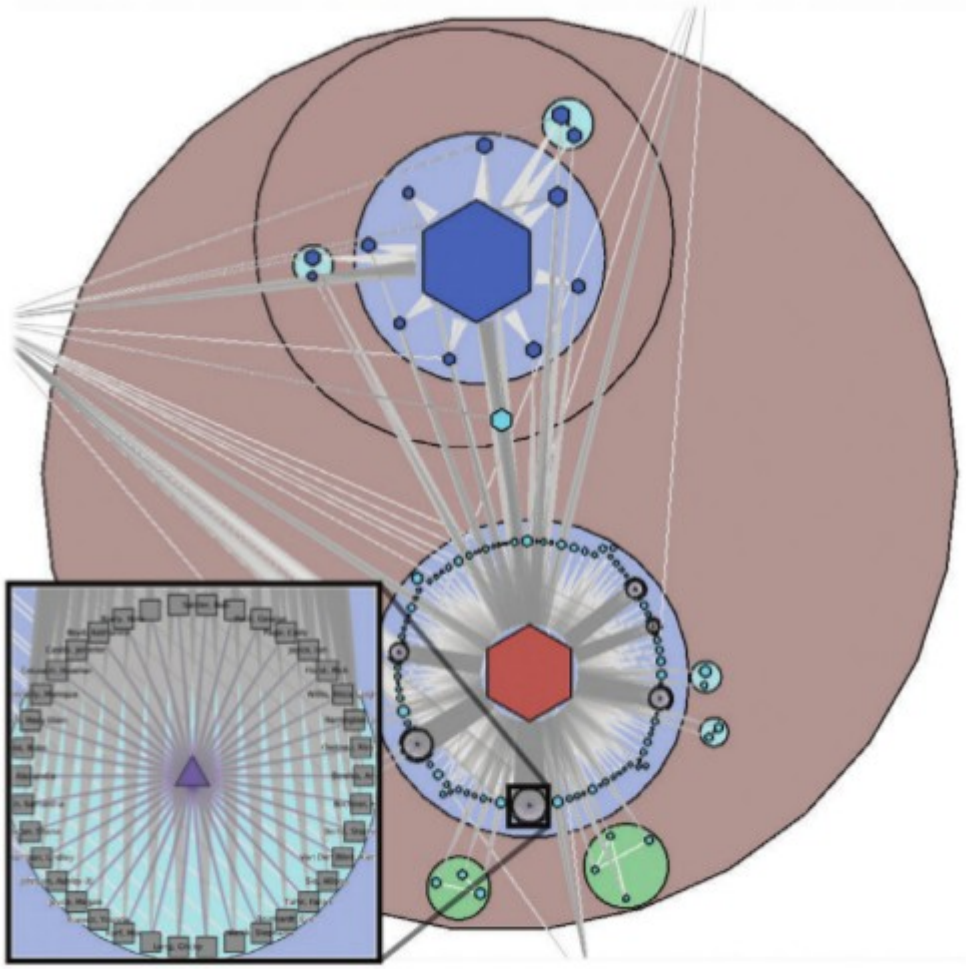
- Interactivity is crucial for building vis tools that handle complexity. When datasets are large enough, the limitations of both people and displays preclude just showing everything at once;
- A single static view can show only one aspect of a dataset. For some combinations of simple datasets and tasks, the user may only need to see a single visual encoding.
- In contrast, an interactively changing display supports many possible queries

Why Is the Vis Idiom Design Space Huge?

- A vis idiom is a distinct approach to creating and manipulating visual representations.
- There are many ways to create a visual encoding of data as a single picture.
- The design space of possibilities gets even bigger when you consider how to manipulate one or more of these pictures with interaction

Why Is the Vis Idiom Design Space Huge?

- Data from Internet Movie Data base showing all movies connected to Sharon Stone is shown, where actors are represented as gray square nodes and links between them



Why Focus on Tasks?

- A tool that serves well for one task can be poorly suited for another, for exactly the same dataset.
- The task of the users is an equally important constraint for a vis designer as the kind of data that the users have.
- Example, a vis tool can support presentation, or discovery, or enjoyment of information; it can also support producing more in formation for subsequent use.

Why Focus on Effectiveness?

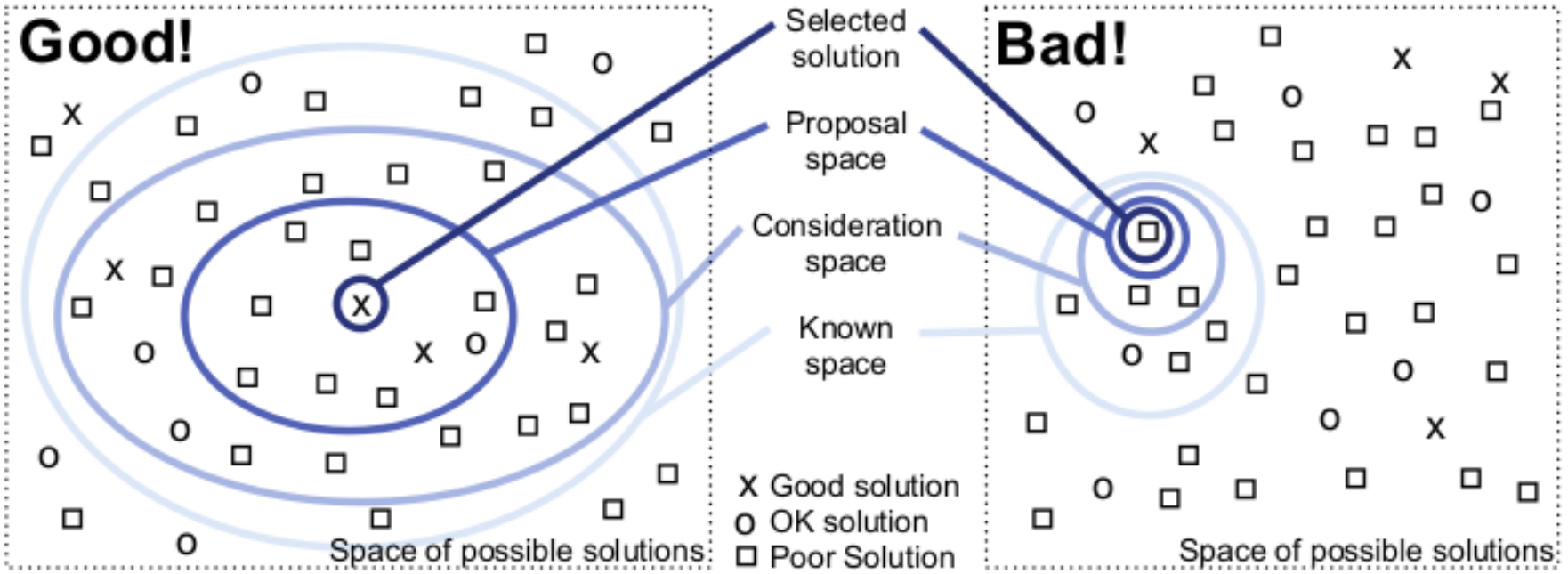
- The focus on effectiveness is a corollary of defining vis to have the goal of supporting user tasks.
- This goal leads to concerns about correctness, accuracy, and truth playing a very central role in vis.

Why Are Most Designs Ineffective?

- The most fundamental reason that vis design is a difficult enterprise is that the vast majority of the possibilities in the design space will be ineffective for any specific usage context.
- In some cases, a possible design is a poor match with the properties of the human perceptual and cognitive systems.
- In other cases, the design would be comprehensible by a human in some other setting, but it’s a bad match with the intended task.

Why Are Most Designs Ineffective?

- A search space metaphor for vis design.



Why Is Validation Difficult?

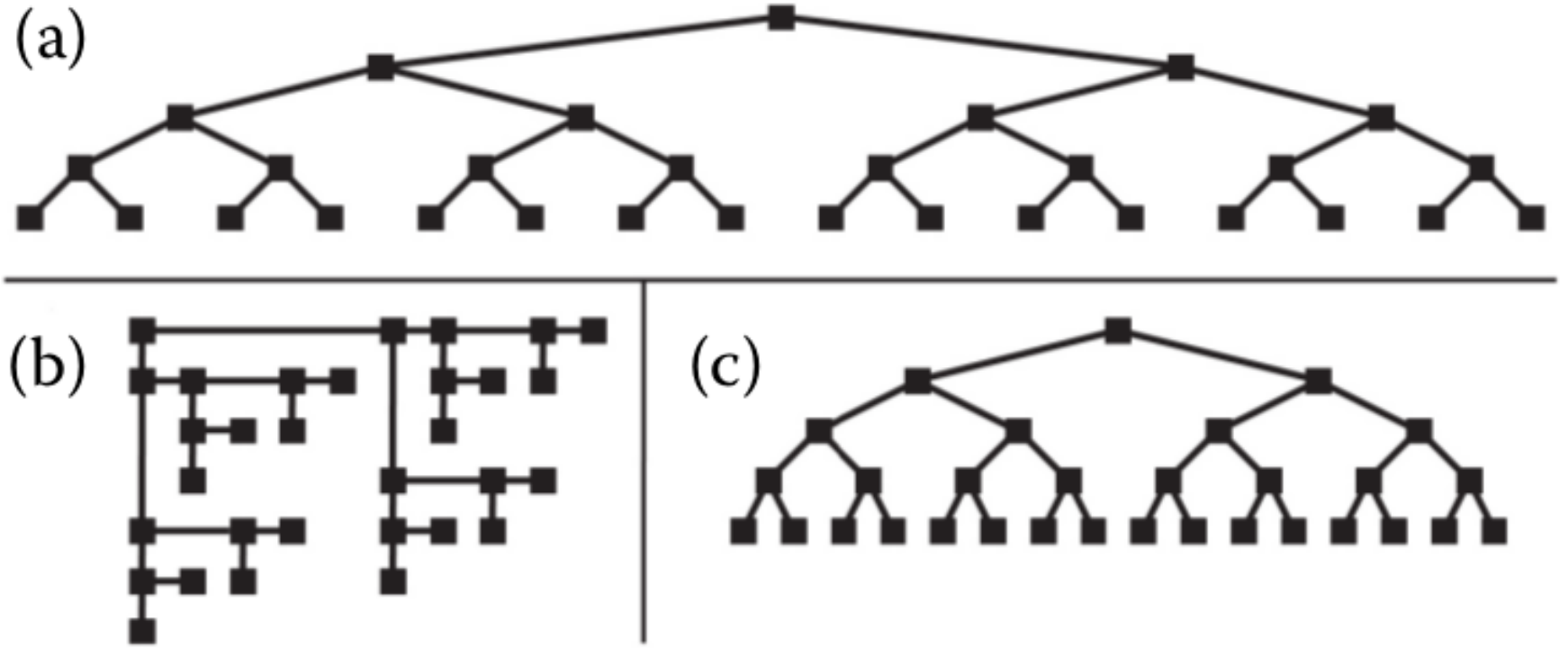
- The problem of validation for a vis design is difficult because there are so many questions that you could ask when considering whether a vis tool has met your design goals.
  - How do you argue that one design is better or worse than another for the intended users?
  - For one thing, what does better mean?
  - Do users get something done faster?
  - How do you decide what sort of benchmark data you should use when testing the system?

Why Are There Resource Limitations?

- Three different kinds of limitations:
- Computational capacity : Time, Space, Scale
- Human perceptual and cognitive capacity.
  - Memory and Attention
- Display capacity.
  - Vis designers often run out of pixels; that is, the resolution of the screen is not enough to show all desired information simultaneously.
  - The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Why Are There Resource Limitations?

- Display capacity

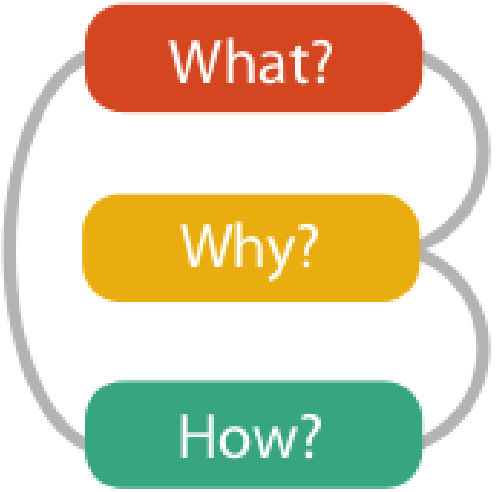


Why Are There Resource Limitations?

- Display capacity
- (a) encodes the depth from root to leaves in the tree with vertical spatial position. However, the information density is low.
- (b) uses nodes of the same size but is drawn more compactly, so it has higher information density. However, the depth cannot be easily read off from spatial position.
- (c) shows a very good alternative that combines the benefits of both previous approaches, with both high information density from a compact view and position coding for depth.

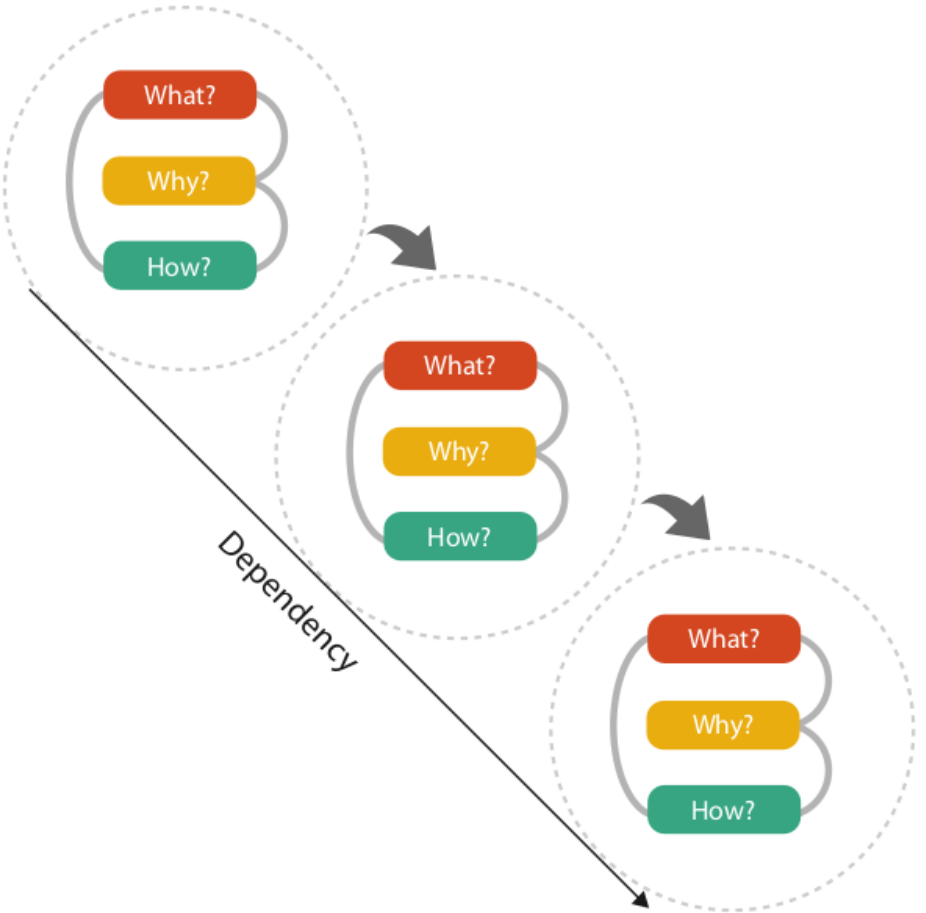
Why Analyze?

- The high-level framework for analyzing vis use according to three questions:
- What data the user sees.
- Why the user intends to use a vis tool.
- How the visual encoding and interaction idioms are constructed in terms of design choices.

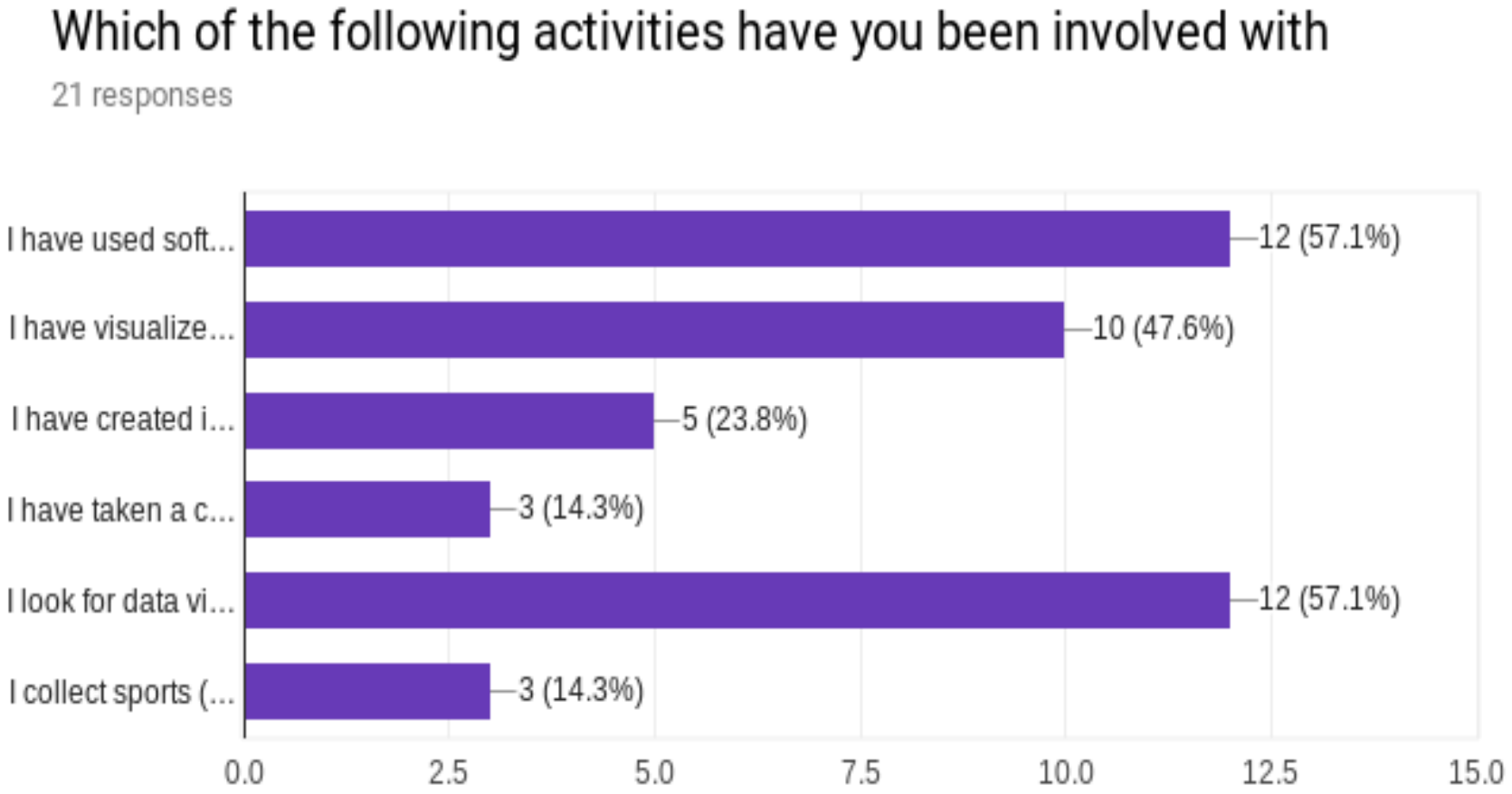


Why Analyze?

- Each three-fold what–why–how question has a corresponding data–task–idiom answer trio. One of these analysis trios is called an instance



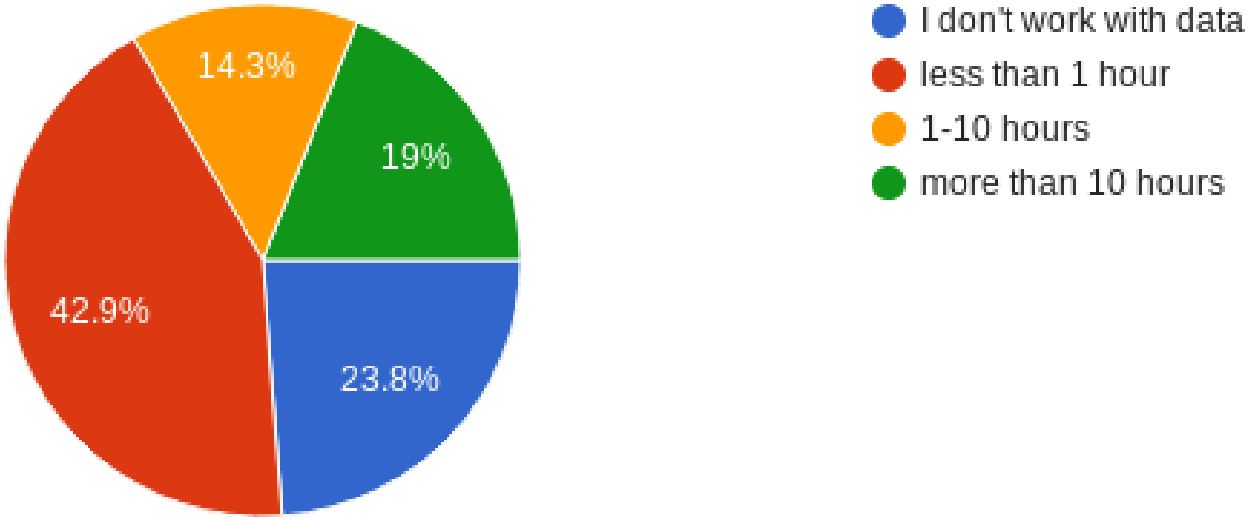
Survey Form Visuals



Survey Form Visuals

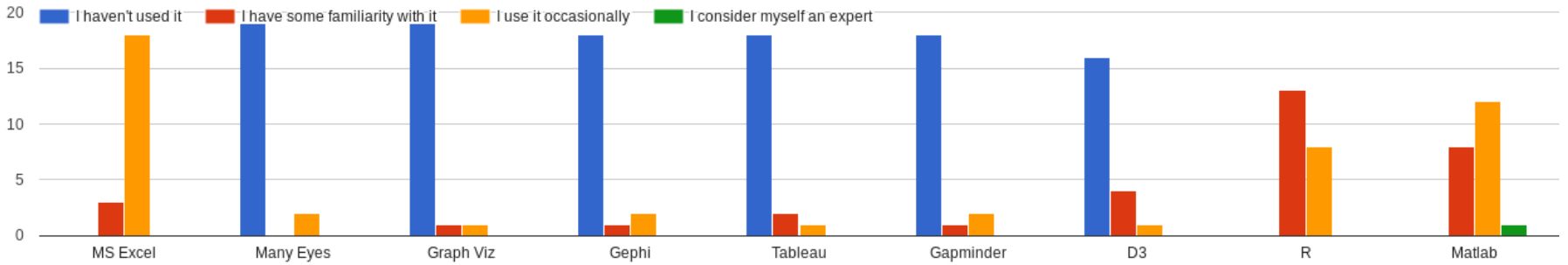
Approximately how many hours per week do you work with data?

21 responses

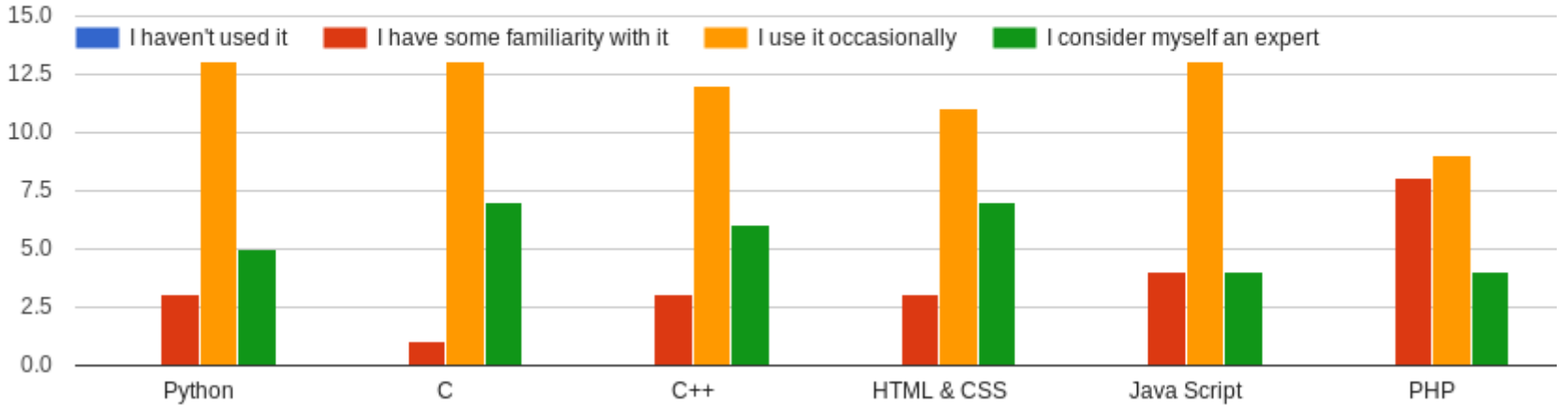


Survey Form Visuals

What is your proficiency with the following software?



What is your proficiency with the following programming languages?



Thank You