Low-Level Components of Analytic Activity in Information Visualization

Robert Amar, James Eagan, and John Stasko

College of Computing/GVU Center
Georgia Institute of Technology

ABSTRACT

Existing system-level taxonomies of visualization tasks are geared more towards the design of particular representations than the facilitation of user analytic activity. We present a set of ten low-level analysis tasks that largely capture people's activities while employing information visualization tools for understanding data. To help develop these tasks, we collected nearly 200 sample questions from students about how they would analyze five particular data sets from different domains. The questions, while not being totally comprehensive, illustrated the sheer variety of analytic questions typically posed by users when employing information visualization systems. We hope that the presented set of tasks is useful for information visualization system designers as a kind of common substrate to discuss the relative analytic capabilities of the systems. Further, the tasks may provide a form of checklist for system designers.

CR Categories and Subject Descriptors: H.5.0 [Information Interfaces and Presentation]: General; J.0 [Computer Applications]: General

Additional Keywords: Analytic activity, taxonomy, knowledge discovery, design, evaluation.

1   INTRODUCTION

Information visualization research, especially that dealing with the automatic generation of information presentations [10,15], has produced several taxonomies of system tasks that map visualization operations to user cognitive processes. In one sense, these taxonomies might be considered low-level task taxonomies or hierarchies, since they form part of a compositional language upon which automatic generation systems build higher-order externalizations of data.

However, considering these taxonomies as a basis upon which to build models of analytic activity is made difficult by their origins. While their elements can be algorithmically composed into presentations, the composition process itself is ad-hoc, relying on a designer's own insight and expressive capability within a particular tool. These taxonomies reflect this system-oriented approach, rather than providing ways to think about all the different analytic tasks a user may perform in a given space.

1.1   Representational Primacy and Task Focus

We have previously argued in [1] that these taxonomies typify thought under a paradigm we call "representational primacy", a data-centric view of information visualization that relies on user skill to generate insight. While effective representation is a prerequisite to useful visualizations, we feel that "analytic primacy", which can be described as a focus on more closely mapping visualization systems to user analytic goals, will increase the value and utility of information visualization.

Email: bob, eaganj, stasko@cc.gatech.edu

With the aim of generating an actionable means for supporting analytic activity, we wish to rethink some of the lower-level task taxonomies that focus on a generated presentation as an end result. In general, information visualization can benefit from understanding the tasks that users accomplish while doing actual analytic activity. Such understanding achieves two goals: first, it aids designers in creating novel presentations that amplify users' analytic abilities; second, it provides a common vocabulary for evaluating the abilities and affordances of information visualization systems with respect to user tasks.

We argue that a stronger focus on user tasks and analytic activities in information visualization is necessary as current tools do not seem to support analytic activity consistently. A 2004 study by Saraiya and North found that insights generated from tools used to visualize gene expression data were not generally valuable according to domain experts [11]. Systems such as IN-SPIRE [7] support analytic activities within the domain of document search but may not generalize across domains. Current tools may not even support representational activity very well; consider, for example, the Kobsa study showing only 68-75% accuracy on relatively simple tasks during commercial tool evaluation [8].

1.2   The Nature of Analytic Activity

User analysis questions and tasks as part of analytic activity typically range from broader, "high-level" goals to much more specific, "low-level" inquiries. For example, a person studying the history of motion picture films may have "high-level", uncertainty-tinged knowledge goals such as understanding trends in popularity over time or determining how to predict which movies will win Academy Awards. In the process of acquiring this knowledge, the person may generate more specific, low-level queries such as identifying the Academy Award-winning pictures of the past ten years and determining whether or not movie length correlates to the film's popularity.

It is this latter set of questions, more specific and focused in nature, on which we focus on in this article. In particular, we are interested in generating a relatively small set of question types that encompasses the set of user inquiries made while working with information visualization systems. While it seems unlikely that a small set of questions types would be complete and cover all user queries, we hope that a small core set that addresses the majority of user goals will emerge.

Efforts such as this invariably lead to some controversy. A taxonomic decomposition of such a broad and diverse domain, data analysis with information visualization, can never be perfect or settled, and we expect some healthy discussion of whether we have proposed the "right" set of tasks, whether other tasks need to be added, and so on. Our goal is to stimulate such consideration, however, as such discussion fosters an increased focus on analytic activities that we believe will be ultimately beneficial.

2   RELATED WORK

Many researchers have studied the problem of best facilitating the discovery of useful relationships within data sets. Approaches have evolved over the years, starting with tools and guidelines

based largely on the properties of the data itself. Taxonomic analyses of both system and user tasks have evolved as a way to manage the wider variety of representations that has come into standard use.

2.1 Data-Centric Approaches

Jacques Bertin, one of the earliest practitioners of data analysis and presentation, understood the deduction of relationships to be a matter of permutation [2]. Bertin proposed a synoptic that differentiated between ordered, reorderable (what we might today call "nominal"), and topographic data, established retinal qualities of marks that would allow viewers to differentiate between marks, and provided guidelines for representing data as arrays of marks, histograms, and curves based on its dimensionality. One can understand lower-level analytic activity as the organization of information gained from the differentiation between and the permutation of graphical marks, although such a framing does not always provide an understanding of how this activity is organized.

John Tukey developed several methods known collectively as exploratory data analysis [13]. Tukey was interested in using statistics to *extract* potentially useful hypotheses from data, as opposed to confirming existing proposed hypotheses. To accomplish these goals, he introduced quantitative methods to reduce the effect of outliers, such as resistant lines and median polish analysis, and visual techniques such as box plots, rootograms, and Pareto charts that emphasize summary statistics and enumerate potential root causes of phenomena.

2.2 Task-based and System-based Taxonomic Approaches

Wehrend and Lewis [14] propose a classic taxonomy of what could be called "cognitive tasks". They create a matrix of representation sub-problems that correspond to a particular combination of an object type, such as scalar or vector, and a user cognitive task, such as correlation, distribution, or point identification; the authors identify eleven such user tasks based on a literature search (identify, locate, distinguish, categorize, cluster, distribute, rank, compare within entities, compare between relations, associate, correlate). They then populate the matrix with representation techniques to create a mapping between techniques and problems. Finally, they show how techniques from this matrix can be used to generate visualizations of flow in irrigation systems.

Efforts in automatic presentation generation have produced a different perspective on low-level visualization system tasks.

The Roth and Mattis [10] taxonomy informs presentation design within the SAGE tool. While much of the taxonomy presented deals with static characteristics of the data, one of its dimensions deals explicitly with user information-seeking goals. Roth and Mattis use two characteristics to deal with such goals: display functions, which vary presentation of a data set based on whether users desire exact value lookup, comparison, or correlation; and distribution functions, which specify how to distribute sets of related information within the presentation.

Zhou and Feiner [15] examine techniques for automatically creating multimedia presentations in their tool, IMPROVISE, based on user goals. The authors group high-level presentation goals into two intents: "inform", which deals with elaboration and summarization of data; and "enable", which deals with data exploration and derivation of relationships. They then refine the Wehrend and Lewis operations into visual tasks organized by their visual accomplishments (low-level user or presenter goals) and visual implications (what visual capabilities are called upon in the attainment of the visual accomplishments). Each presentation intent maps to visual tasks that achieve it; for instance, the intent "enable-compute-sum" has correlate, locate, and rank.

Finally, other taxonomies take a more system capability-driven approach to characterizing visualization operations.

Shneiderman [12] posits a task-by-data-type taxonomy that crosses information-seeking visualization tasks (overview, zoom, filter, details-on-demand, relate, history, extract) with different types of data (1-D, 2-D, 3-D, multi-dimensional, time series, network, tree) and discusses both examples and missed opportunities for supporting the given tasks. The taxonomy assumes an implicit mapping between user goals and these visualization tasks.

Card presents a Visualization Reference Model [3] that emphasizes, among other things, the specific mappings of data tables into visual structures, and the iterative effects of human interactions with these mappings. Card uses the constraints presented by these mappings to organize information visualization techniques into three types of "visual structure" (simple, composed, and interactive) as well as focus + context abstractions. Card did not explicitly map user tasks to transformations within the reference model; more recently, though, Card, Pirolli and colleagues have done work in understanding analyst sensemaking techniques using cognitive task analysis techniques [4]. This work posits interlocked, bidirectional information foraging and sensemaking loops, and describes high-level tasks done in going both from theory to data as well as data to theory.

Chi [6] taxonomizes existing visualization techniques into several data categories (scientific visualization, geographic InfoVis, 2D, multi-dimensional, information landscapes, trees, networks, text, and web). He extends the Card reference model into a Data State Reference Model [5] in order to isolate common operational steps within each visualization type. Chi has the explicit aims of assisting implementers with choosing and deploying visualization techniques and broadening the visualization design space.

3 TOWARD CONCRETE TASKS FOR ANALYTIC ACTIVITY

Our original intent was to create low-level analytic tasks from experience and literature search. However, accomplishing the goal of depicting low-level concrete steps in analytic activity and providing operational guidance to designers requires a more grounded approach. We decided to draw upon a basis of existing document analytic activity to serve as a foundation.

3.1 Data Collection

To gather a corpus of analysis questions, we reviewed the work of students in our Spring 2004 Information Visualization course. As part of a series of assignments, students were asked to generate data analysis questions for provided data sets and then evaluate how well the questions could be answered using particular commercial visualization tools (in this case, Spotfire Pro 4.0, Eureka/Table Lens, InfoZoom and SeeIT). The students generated 196 valid analysis tasks that we used in the next phase.

The following table lists the data sets used, along with their dimensionality, cardinality, and the number of analysis questions students generated for each data set.

Data Set	Dimensionality	Cardinality	Questions Generated
Cereals	15	78	43
Mutual funds	14	987	14
Cars	10	407	53
Films	10	1742	47
Grocery surveys	8	5164	39

The directions to the students for generating questions read:

"Pick the two [data sets] that are most interesting to you to use in the assignment. Briefly scan the text of the files and familiarize yourself with the variables. Write down a few hypotheses to be considered, tasks to be performed, or questions to be asked about the data elements. Recall all the different kinds of analysis tasks that we discussed earlier in the course (identification, outliers, correlations, clusters, trends, associations, presentation, etc.)."

3.2   Analysis Approach

We used an affinity diagramming approach, grouping similar questions and iteratively refining the groups according to what we believed to be the core knowledge goal of the questions in each group. Clearly, a general concept such as "correlation" can involve subtasks or sub-operations; however, most of the sub-operations performed in the questions generally fell under tasks we already had isolated from other questions.

Our affinity diagramming approach led to ten component tasks across the analytic questions submitted by the students. Since we concentrated on user goals independent of particular visualization systems or paradigms (although the questions may have been influenced by the systems used in the assignment, as discussed later in this article), our list of tasks is free of system-specific operations such as "zoom". Tasks such as "filter" are offered in the spirit of users' analytic desires as opposed to that of low-level cognitive tasks or operators.

4   AN ANALYTIC TASK TAXONOMY

The ten tasks from the affinity diagramming analysis are:

- Retrieve Value
- Filter
- Compute Derived Value
- Find Extremum
- Sort
- Determine Range
- Characterize Distribution
- Find Anomalies
- Cluster
- Correlate

Each of the tasks is presented in the following sections, along with a *pro forma* abstract [9] and example questions that serve as general models and examples of the tasks. These tasks are not meant to be a normative picture of user analytic activity, but rather to provide a vocabulary for discussion.

In the task descriptions, we use the following terms:
- *data case*: an entity in the data set
- *attribute*: a value measured for all data cases in a data set
- *aggregation function*: a function that creates a numeric representation for a set of data cases (e.g. average, sum, count)

1. Retrieve Value

General Description: Given a set of specific cases, find attributes of those cases.

Pro Forma Abstract: What are the values of attributes {X, Y, Z, ...} in the data cases {A, B, C, ...}?

Examples:
- What is the mileage per gallon of the Audi TT?
- How long is the movie Gone with the Wind?

This task serves as a subtask for many other tasks; in particular, once a set of cases is known from another operation, such as finding a case with an extreme value of an attribute (see task 4, "Find Extremum"), this task is often used to read off relevant attributes. However, we only classify a question as a value retrieval if the particular cases to be examined are specified at the time of the question.

2. Filter

General Description: Given some concrete conditions on attribute values, find data cases satisfying those conditions.

Pro Forma Abstract: Which data cases satisfy conditions {A, B, C...}?

Examples:
- What Kellogg's cereals have high fiber?
- What comedies have won awards?
- Which funds underperformed the SP-500?

This is another task used as a subtask in many other questions. In particular, this type of question relies on classifying cases by a condition that can be measured independent of any other data case in the data set.

This leads to an interesting phenomenon for questions such as, "Which data cases have a high value for attribute X?" In this case, an operating definition for what entails a "high" value is necessary; once such a definition is established, the question becomes answerable in a concrete fashion. On the other hand, questions such as, "Which data case has the highest value of attribute X?" rely on properties of all other elements in the data set, and are thus not part of this category.

For example, in the question "What Kellogg's cereals have high fiber," there is an implicit definition of what it means for a cereal to be high in fiber, independent of the fiber values of the other cereals in the data set. Once the analyst makes that implicit definition explicit, this question becomes "What Kellogg's cereals have more than x grams of fiber" for some value of x. Questions of the form "What Kellogg's cereal has the highest fiber," by contrast, are only answerable relative to the other cereals in the data set.

3. Compute Derived Value

General Description: Given a set of data cases, compute an aggregate numeric representation of those data cases.

Pro Forma Abstract: What is the value of aggregation function F over a given set S of data cases?

Examples:
- What is the average calorie content of Post cereals?
- What is the gross income of all stores combined?
- How many manufacturers of cars are there?

Computing an aggregation (e.g. average, median, count) is a common task in data analysis. In particular, more complex aggregators such as "count-unique-values-of" can provide insights into the data itself. This task also appears as a subtask in other operations; in particular, some questions compare categories without a particular operating definition of what is being compared, such as: "Which cars are more fuel-efficient, Japanese cars or American cars?" These questions imply some sort of aggregator function without specifying exactly how that aggregation is calculated.

4. Find Extremum

General Description: Find data cases possessing an extreme value of an attribute over its range within the data set.

Pro Forma Abstract: What are the top/bottom N data cases with respect to attribute A?

Examples:
- What is the car with the highest MPG?
- What director/film has won the most awards?
- What Robin Williams film has the most recent release date?

Finding high or low values of an attribute was a very common operation across the student questions. Note that this task differs from "Sort" (task 5) since a complete sort is not always necessary to find an extreme value, and also differs from "Find Anomalies" (task 8) since anomalies are not always extreme values.

5. Sort

General Description: Given a set of data cases, rank them according to some ordinal metric.

Pro Forma Abstract: What is the sorted order of a set S of data cases according to their value of attribute A?

Examples:
- Order the cars by weight.
- Rank the cereals by calories.

Although this task is fairly self-explanatory, it appeared only infrequently as a task unto itself. Sorting is generally a substrate for extreme value finding, especially when searching for a number of values at the extreme and not just the most extreme value.

6. Determine Range

General Description: Given a set of data cases and an attribute of interest, find the span of values within the set.

Pro Forma Abstract: What is the range of values of attribute A in a set S of data cases?

Examples:
- What is the range of film lengths?
- What is the range of car horsepowers?
- What actresses are in the data set?

The range task is an important task for understanding the dynamics of data within a data set. Users can use range data to help decide the suitability of the data set for a particular analysis, or understand something about the general types of values found for a particular attribute. The "range" of a categorical attribute can be thought of as an enumeration of all its unique values in a set.

7. Characterize Distribution

General Description: Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set.

Pro Forma Abstract: What is the distribution of values of attribute A in a set S of data cases?

Examples:
- What is the distribution of carbohydrates in cereals?
- What is the age distribution of shoppers?

Distribution, like range, is another important task for characterizing data. Users can get a general sense of distribution to understand "normalcy" in data as opposed to anomaly (see task 8, "Find Anomalies"). Sometimes the distribution task is hidden; for example, "Compare Frosted Flakes' calories per serving to those of all other cereals" is really a question of location within a distribution.

8. Find Anomalies

General Description: Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers.

Pro Forma Abstract: Which data cases in a set S of data cases have unexpected/exceptional values?

Examples:
- Are there exceptions to the relationship between horsepower and acceleration?
- Are there any outliers in protein?

Anomalous values in a data set often provide a basis for further exploration. This task can be thought of as a complementary task to "distribution", although it is not always framed as such (e.g. sometimes a distribution is assumed, as in the case of a standard box-and-whisker plot).

9. Cluster

General Description: Given a set of data cases, find clusters of similar attribute values.

Pro Forma Abstract: Which data cases in a set S of data cases are similar in value for attributes {X, Y, Z, …}?

Examples:
- Are there groups of cereals w/ similar fat/calories/sugar?
- Is there a cluster of typical film lengths?

Users naturally group similar items together. This proximity can have a number of connotations depending on the clustering attributes; for example, similar products may be market competitors, members of a family of products, or simply represent the "normal" or expected case as opposed to outliers.

10. Correlate

General Description: Given a set of data cases and two attributes, determine useful relationships between the values of those attributes.

Pro Forma Abstract: What is the correlation between attributes X and Y over a given set S of data cases?

Examples:
- Is there a correlation between carbohydrates and fat?
- Is there a correlation between country of origin and MPG?
- Do different genders have a preferred payment method?
- Is there a trend of increasing film length over the years?

One of the most interesting observations about the corpus of student questions as a whole was how frequently students desired to "correlate" one or more non-numeric attributes. The semantics of such questions, in our interpretation, leaned more towards isolating "coincidences" of interest. Membership in a category may be predictive of certain attribute values, but does not predict those same attribute values in a different category; for example, a comparison of American, German, and Japanese cars' gas mileage does not allow you to predict the gas mileage of Korean cars. The semantics of "true" quantitative correlation questions deal with purely numeric variables and the ability to generate a mathematical predictive model relating the values of attributes within a data set.

Questions such as the fourth example question above involving trends over time were quite common in the corpus of questions. We interpret such questions simply as correlations with temporal variables.

5 DISCUSSION

It is difficult to construct a taxonomy that perfectly characterizes a domain so open to interpretation. In this section we discuss internal and external issues with the taxonomy and analysis methods.

5.1 Compound Tasks

Considering the set of tasks in the taxonomy to be analytic "primitives" allows us to examine some questions that do not cleanly fit into one category but rather appear to be compositions of primitive tasks. For instance, the task "Sort the cereal manufacturers by average fat content" involves a Compute Derived Value (average fat) primitive followed by a Sort primitive.

Further, users may simply be interested in particular attribute values from a set smaller than the entire data set; for example, on the movies data set, one might ask, "Which actors have co-starred with Julia Roberts?" One can complete this task by first finding all movies with Julia Roberts (Filter) and then enumerating the set of actors within those movies (Retrieve Value).

As another example, consider the question, "Who starred in the most films in 1978?" This task differs from the basic extremum task in that the domain of the extremum operation is no longer data cases, but aggregate relationships of those data cases; in specific, a user must enumerate each actor and actress in the relevant portion of the data set (Retrieve Value), assign each one a count of the number of data cases in which each actor/actress was present (Compute Derived Value), and finally determine which actor/actress has the highest count (Find Extremum).

5.2 Omissions from the Taxonomy

Even outside the context of combining analytic "primitives", there were a number of questions that still did not fit cleanly into the taxonomy. Such questions are marked either by a fundamentally mathematical or computational nature rather than an analytic one, or by uncertainty, either in the analytic process necessary to answer the question or user criteria employed during the knowledge-making process.

5.2.1 Low-level Mathematical and Cognitive Actions

In constructing the taxonomy, we abstracted away as low-level, and thus beyond the scope of the present work, some basic mathematical and cognitive operations, such as determining that a data case mathematically satisfies filtering criteria or conditions and computing aggregate values from a mathematical perspective. In particular, we explicitly acknowledge the existence of a low-level mathematical comparison operation, one in which a value is evaluated for being less than, greater than, or equal to another value or values.

This leads to the notion of questions whose overall goal is too "low-level" for our analytic task taxonomy. For instance, the following questions involve the aforementioned mathematical comparison operation:

- "Which cereal has more sugar, Cheerios or Special K?"
- "Compare the average MPG of American and Japanese cars."

These questions utilize Retrieve Value and Compute Derived Value primitives, respectively, followed by a mathematical comparison operation. We view this very low-level comparison as being a fundamental cognitive action taken by the person using a visualization tool, rather than as a primitive in our analytic task taxonomy.

5.2.2 Higher-level Questions

We have found that the proposed ten tasks cover the vast majority of the corpus of analytic questions we studied. Some questions, however, imply tasks not explicitly covered by our task set, but instead they can be thought of as guiding higher-level exploration in the data set. For example:

- "Do any variables correlate with fat?"
- "How do mutual funds get rated?"
- "Are there car aspects that Toyota has concentrated on?"

Much learning of a domain can occur in the use of a properly structured visualization, and discovering interesting relationships between domain parameters is one part of that learning. While the corpus of questions mainly limited such exploration to correlation (another factor to be discussed in the next section), a less-structured exploration is definitely possible.

5.2.3 Uncertain Criteria

Other questions in the corpus contained uncertain criteria, for example:

- "Do cereals (X, Y, Z…) sound tasty?"
- "What are the characteristics of the most valued customers?"
- "Are there any particular funds that are better than others?"

While these questions may be answered by supposing the existence of some black-box aggregation function, there may be other ways to answer the question, such as use of a distribution or clustering method. Fundamentally, each of these questions involves a value judgment that is beyond the proposed primitives.

Another style of question common in the set involves a comparison operation that is much higher in level and more abstract than the fundamental mathematical comparison operation discussed earlier in the section. For instance, consider the questions:

- "What other cereals are most similar to Trix?"
- "How does the Toyota RAV4 compare to the Honda CRV?"
- "Compare the distributions of values for sugar and fat in the cereals."

Each of these questions involves a more subjective evaluation of a data case, attribute, or derived value in comparison to others. We felt that the fundamental operation being performed in each of

these examples was not at the level of the ten primitives in our taxonomy.

6 OVERALL CONCERNS

6.1 Similarity to Existing Taxonomies

Of the taxonomies presented earlier, the closest in spirit to our efforts is that of Wehrend and Lewis [14]. While both taxonomies share many of the same tasks, several tasks differ either partially or entirely. We consider those tasks here.

We present four new analytic tasks not present in the Wehrend and Lewis framework: Compute Derived Value; Find Extremum; Find Anomalies; and Determine Range. The Wehrend and Lewis framework, similarly, presents operations not found in our taxonomy: associate; categorize; compare; and distinguish.

The four new tasks in our taxonomy frequently occurred in the corpus of analytic questions. The Compute Derived Value task frequently arose in performing visual analytic operations on attributes not directly present in the data set. Often, this task was implied as if the derived value were already a part of the data set. The Determine Range, Find Extremum and Find Anomalies tasks also occurred frequently enough in our corpus to warrant their own singleton tasks.

Wehrend and Lewis' associate, categorize, and distinguish tasks do not appear in our taxonomy. Questions relating to these tasks appeared rarely in our corpus. In each case, we decomposed that question into a combination of our primitive tasks. For example, distinguishing two cars involves Retrieve Value operations and the Comparison meta-operation.

Wehrend and Lewis include two compare operations, compare within relations and compare between relations. As discussed earlier, we view Compare as a higher-level meta-operation.

Finally, the Retrieve Value and Filter tasks relate to the identify and locate tasks of the Wehrend and Lewis taxonomy. Each involves acquisition of data cases and/or attribute values.

The composition of the tasks in our taxonomy may seem relatively unsurprising. While our approach comes from a different perspective, the tasks share much similarity with those from other taxonomies, as illustrated above. Furthermore, our tasks resemble operations that one would find in specialized tools such as Microsoft Excel or database languages such as SQL. While many of the data manipulation operations that one would perform are similar, we focus on analytic operations as they would apply to an information visualization system. Where languages such as SQL focus on data management, we focus on analysis.

6.2 Relationships to Higher-Level Tasks

A common theme in all of the taxonomies described in the Related Work section is that user goals are usually thought of as static and explicitly treated only as far as they map into low-level visual tasks. While such tasks are essential, they do not provide a firm basis for supporting the kinds of knowledge-making activities that people seek to perform every day. In [1], we provided examples of such activity:

- Complex decision-making under uncertainty
- Learning a domain
- Identification of trends
- Predicting the future

We further classified the gap between representation and analysis into two so-called "analytic gaps": the Worldview Gap, which concerns whether or not the right data and relationships are being shown; and the Rationale Gap, which asks whether the relationships inferred from a visualization can be trusted under uncertainty and used effectively to rationalize conclusions drawn from use of a visualization. We then posited six knowledge precepts for use in designing visualizations that bridge these gaps, as well as evaluated their use in existing systems. However, we have not examined how these knowledge precepts might map to lower-level concrete tasks for visualization systems.

Our analytic task taxonomy can help provide a basis for such a mapping. Consider again the task of learning a domain, as discussed briefly in Section 5.2.2. Our corpus of analytic questions clearly implies some meta-information tasks such as browsing for relationships (for example, through repeated Correlate and Characterize Distribution operations) and discovering the breadth of coverage in a data set (e.g. performing Find Extremum in tandem with Computed Derived Values, such as counts of unique values of a categorical attribute, and applying Determine Range of quantitative variables). In discovering a preponderance of underspecified criteria such as "high" or "low" in the questions asked, we assert that these criteria may be posed in the spirit of interactive discovery rather than operationally-defined pruning.

6.3 Methodological Concerns

The use of student questions as a basis for thought and evaluation obviously has limitations and caveats. Students were presented with the tools before coming up with questions, and may have fit their questions to match material presented in class as well as the tools themselves, as they had to answer those same questions using the tools as well as compare the tools' usability. For example, a preponderance of correlation-type questions existed even though the directions stated not to make all questions about correlation. We speculate that this might be due to the availability of scatterplots in Spotfire; similarly, ranking and sorting questions are well suited to Table Lens, and aggregations are a major analytical component of SeeIT.

Repeating our analysis and process, but instead using professional analysts such as drug discovery researchers or intelligence analysts to generate questions, may provide some new low-level tasks. It would be interesting to compare the task clusters emerging from the work of such domain experts to those found with the students. The tasks we identified tended toward deduction, while the efforts of domain experts often involve more exploratory analysis such as hypothesis formation and confirmation.

7 CONCLUSION

We present a set of ten primitive analysis task types, representative of the kinds of specific questions that a person may ask when working with a data set. The development of this set of primitives stemmed from ongoing research into the utility of information visualization systems. It also was driven by a large set of questions gathered from an assignment in which students generated queries about five sample data sets. We used an affinity diagramming approach to cluster the individual questions into sets of related items. Our set of analysis task types is similar to those generated earlier by other researchers, but a number of new primitives and ideas did emerge. Our focus here was more directly on the analysis primitives that one might expect people to be generating as they use an information visualization system to achieve some higher-level knowledge goals.

We believe that these primitives can serve as a form of common language or vocabulary when discussing the capabilities, advantages, and weaknesses of different information visualization systems. Researchers and system developers will be able to describe a system's support for performing these operations by identifying its particular representation, interaction, and interface

support for the operation. For example, how does the visualization presented by a system support characterizing distributions or finding anomalies? Does a system adequately facilitate the computation of derived values?

Furthermore, by identifying and enumerating these primitive analysis task types, we hope to foster a continued emphasis on the importance of analytic measures of information visualization systems. It is vital that information visualization system designers both develop innovative, new visualization techniques and clearly articulate the analytic qualities of those techniques. These primitive analysis task types also can serve as an informal check-list along which to assess and evaluate new systems and techniques.

REFERENCES

[1] Robert A. Amar and John T. Stasko, "Knowledge Precepts for Design and Evaluation of Information Visualizations", *IEEE Transactions on Visualization and Computer Graphics*, 11(4): 432-442, July/August 2005.

[2] Jacques Bertin. Graphics and Graphic Information Processing. Walter de Gruyter, New York, 1981.

[3] Stuart Card. "Information visualization." In Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. pp. 544 - 582, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 2002.

[4] Stuart Card and Peter Pirolli. "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis", Proc. International Conference on Intelligence Analysis 2005.

[5] Ed H. Chi and John T. Riedl, "An Operator Interaction Framework for Visualization Systems", Proc. InfoVis '98, pp. 63-70.

[6] Ed H. Chi, "A Taxonomy of Visualization Techniques using the Data State Reference Model", Proc. InfoVis '00, pp. 69-75.

[7] IN-SPIRE, http://www.in-spire.pnl.gov/, 2004.

[8] Alfred Kobsa, "An Empirical Comparison of Three Commercial Information Visualization Systems", Proc. InfoVis'01, pp. 123-130.

[9] William M. Newman, "A Preliminary Analysis of the Products of HCI Research, Based on Pro Forma Abstracts", Proc. CHI '94, pp. 278-284.

[10] Steven F. Roth and Joe Mattis, "Data characterization for intelligent graphics presentation," Proc. CHI '90, pp. 193-200.

[11] Purvi Saraiya, Chris North and Karen Duca, "An Evaluation of Microarray Visualization Tools for Biological Insight", Proc InfoVis'04, pp. 1-8.

[12] Ben Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations", Proc. IEEE Symposium on Visual Languages '96, p. 336.

[13] John Tukey. Exploratory Data Analysis. Addison-Wesley, Reading, MA, 1977.

[14] Stephen Wehrend and Clayton Lewis, "A Problem-Oriented Classification of Visualization Techniques", Proc. Vis '90, pp. 139-143.

[15] Michelle X. Zhou and Stephen K. Feiner, "Visual task characterization for automated visual discourse synthesis", Proc. CHI 1998, pp. 392-399.