

Augmenting Decision Making via Interactive What-If Analysis

Sneha Gathani*
Sigma Computing
San Francisco, USA
sneha@sigmacomputing.com

Madelon Hulsebos†
Sigma Computing
San Francisco, USA
madelon@sigmacomputing.com

James Gale
Sigma Computing
San Francisco, USA
jlg@sigmacomputing.com

Peter J. Haas
University of Massachusetts
Amherst, USA
phaas@cs.umass.edu

Çağatay Demiralp
Sigma Computing
San Francisco, USA
cagatay@sigmacomputing.com

Abstract

The fundamental goal of business data analysis is to improve business decisions using data. Business users often make decisions to achieve key performance indicators (KPIs) such as increasing customer retention or sales, or decreasing costs. To discover the relationship between data attributes hypothesized to be drivers and those corresponding to KPIs of interest, business users currently need to perform lengthy exploratory analyses. This involves considering multitudes of combinations and scenarios and performing slicing, dicing, and transformations on the data accordingly, e.g., analyzing customer retention across quarters of the year or suggesting optimal media channels across strata of customers. However, the increasing complexity of datasets combined with the cognitive limitations of humans makes it challenging to carry over multiple hypotheses, even for simple datasets. Therefore mentally performing such analyses is hard. Existing commercial tools either provide partial solutions whose effectiveness remains unclear or fail to cater to business users altogether.

Here we argue for four functionalities that we believe are necessary to enable business users to interactively learn and reason about the relationships (functions) between sets of data attributes thereby facilitating data-driven decision making. We implement these functionalities in SYSTEMD, an interactive visual data analysis system enabling business users to experiment with the data by asking *what-if* questions. We evaluate the system through three business use cases: marketing mix modeling, customer retention analysis, and deal closing analysis, and report on feedback from multiple business users. Overall, business users find the SYSTEMD functionalities highly useful for quick testing and validation of their hypotheses around their KPIs of interest, clearly addressing their unmet analysis needs. Their feedback also suggests that the user experience design can be enhanced to further improve the understandability of these functionalities.

*Also with University of Maryland.

†Also with University of Amsterdam.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution, provided that you attribute the original work to the authors and CIDR 2022, 12th Annual Conference on Innovative Data Systems Research (CIDR '22), January 9-12, 2022, Chaminade, USA.
CIDR'22, January 9-12, 2022, Chaminade CA, USA.

1 Introduction

Interactive visual data analysis systems targeting business users (e.g., sales, marketing, product, or operations managers) aim to help users make data-driven decisions. A basic yet overarching question related to this goal is what do business users essentially try to achieve with data analysis? Specifically, what should an enterprise data analysis system designed for business users be optimizing for? Answering this question is critical for operationalizing business data analysis.

The current view on interactive visual data analysis has been primarily shaped by John Tukey's emphasis on exploratory data analysis (EDA) [36]. Tukey, a prodigious figure with wide-ranging contributions [6] to statistics and beyond, considered data analysis in two stages: exploratory analysis, which he likened to the detective work of collecting evidence in an investigation, and confirmatory analysis (e.g., statistical hypothesis testing), analogous to the trial step of the investigation, where the validity and the strength of collected evidence need to be proved for a judge or jury. Tukey highlighted the importance of exploratory analysis and the use of graphics (visualizations) to that end, which had been overlooked by the statistical community of his time. The developments in the last two decades, including the wider adoption of interactive visualizations in data analysis and the success of commercial as well as open-source EDA tools, demonstrated the value of Tukey's perspective on data analysis. Paradoxically, this success also caused a tunnel vision that frequently turned exploratory analysis into the end itself by overly focusing on visual pattern exploration and qualitative hypothesis generation [20]. This skewed focus limits existing systems in helping business users to effectively complete their data analysis goals.

Consider these user questions: *how can I best use my \$200K marketing budget across advertisement channels? Which activities should I prioritize to increase my deal closing rate? How much do free trials help with acquiring customers? What drives the increase in revenue?* These decision-leading questions are, at best, difficult and time-consuming to answer through existing interactive exploratory analysis systems or data science tools by business users, who typically have no time for or background in coding, statistical analysis, or algorithmic modeling. These questions also hint at what we consider to be the purpose of business data analysis.

Purpose of Business Data Analysis. The fundamental goal of business data analysis is to improve business decisions. This is done by understanding the relationship between two sets of variables—input or independent variables which are hypothesized to be potential

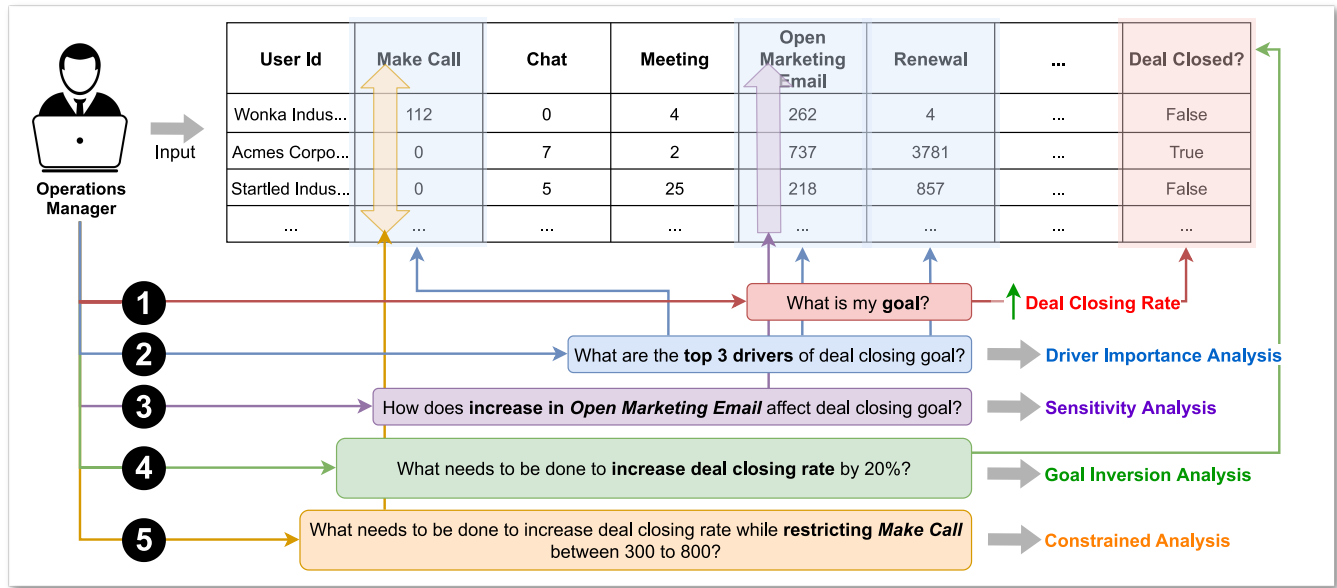


Figure 1: Four functionalities proposed by SYSTEMD to augment business users beyond exploratory analysis in decision making: Driver Importance Analysis, Sensitivity Analysis, Goal Inversion (Seeking) Analysis, and Constrained Analysis.

drivers and output or dependent variables (often a single variable) that are business key performance indicators (KPIs) and are hypothesized to be dependent on the input variables. For the remainder of the paper, we will refer to the input or independent variables as *drivers* and the output or dependent variables as *KPIs*.

Challenges. While interactive exploratory data analysis is useful, it is not sufficient for effectively carrying out the fundamental task above. There are four basic challenges.

Human Cognition. Limitations of human working memory and cognitive overload due to time pressures and data complexity limit the user’s ability to effectively run what-if scenarios, without getting help for rigorously following effective methods to generate, manage, and evaluate hypotheses. Confirmation bias [26], our tendency to fit the evidence to existing expectations, makes it hard to explore insights in an unbiased and rigorous manner [28, 37]. Thus, people often fail to focus on the most relevant evidence while sufficiently attending to hypotheses’ disconfirmation [26].

Interactive Exploratory Analysis. Interactive direct manipulation as a querying and data transformation paradigm does not scale well for learning relations (functions) between drivers and KPIs, often requiring a large number of transformations as well as slicing and dicing combinations.

Data Scale and Complexity. Increased data sizes and complexities exacerbate the two challenges above, easily turning the fundamental task of business data analysis into a daunting process if not impossible. Note that data in enterprise databases are constantly updated and appended, which cloud computing has been making easier and cheaper. Yesterday’s feasible choices or decisions informed by data can easily be suboptimal or infeasible today due to updates or the availability of new data. So, large complex dynamic data put additional pressure on human cognition and working memory.

Dead Data. Even if domain expert business users can build mental models between potential drivers and KPIs for a snapshot of data, there is no easy way for them to probe into, reason about, and run scenarios over to stress-test and utilize this mental model for decision making [18]. Mental models built out of exploratory analyses don’t lend themselves to simulation or scenario modeling based on hypothetical data, which is necessary for what-if analysis.

Elements of Business Data Analysis. What are the elements of the fundamental purpose of business data analysis introduced above? What does improving decisions mean? Based on our conversations with business users, we distill the following observations.

Improve Decisions. The goal of data analysis is to improve decisions based on data. An improved decision—an effective operationalization of insights—manifests itself differently in different domains and use cases. It can be increased sales, reduced cost, increased customer retention rate, reduced churn rate, reduced customer acquisition cost, and so on. Business users mean, well, business.

Understand Driver-KPI Relationship. Improving decisions requires users to learn, manually (mentally) or otherwise, the relationships (functions) between drivers in their data and KPIs on which their business objectives are based.

Use Data and Domain Expertise. Decision-making is an interplay between data and domain knowledge, including common sense. Neither data nor domain knowledge (expertise), which business users possess, is sufficient by itself for improved outcomes.

Value of BI Systems. The value of an enterprise data analysis or business intelligence (BI) system is its added value in effectively enabling improved decision-making using data and domain knowledge.

Desiderata for Systems. Based on the observations above and earlier work [9], we propose that enterprise data analysis systems integrate four interactive functionalities (Figure 1) to augment business users in decision making, going beyond exploratory analysis.

Driver Importance Analysis. Enables users to implicitly learn functions (models) allowing them to understand the relationships between drivers (input) and KPIs (output), along with the artifacts of these learned relationships such as the relative importance of various drivers and their interactions in predicting the KPI outcomes.

Sensitivity Analysis. Enables users to dynamically evaluate learned relationships for arbitrary driver values and observe the changes in KPI values. This helps users build their intuition about how their business works in a hands-on manner. To this end, systems should help users to experiment with the drivers by interactively perturbing (increasing or decreasing) their values and observing the effects on the KPI values.

Goal Inversion (Seeking) Analysis. Enables users to interactively set goals such as specific target values or optimization objectives (maximization or minimization) for the KPIs and observe multiple scenarios on how the driver values need to change to achieve the desired goals. For example, systems should provide recommendations for changes needed in driver values to achieve user-specified KPI goals such as doubling the revenue or minimizing the churn rate.

Constrained Analysis. Allows users to interactively set constraints or conditions over how the learned functions (models) are evaluated or inverted, enabling users to incorporate their domain knowledge such as business constraints and common sense to regulate these functions. This also enables users to quickly generate and evaluate multiple scenarios under various conditions specified by constraints. For instance, systems should enable users to set constraints (e.g., boundary, equality, or inequality) on one or more drivers and run the goal-seeking analysis to provide driver values satisfying user constraints. Note that there is rarely a single best context-free decision for improving a KPI goal; instead, there are often multiple feasible choices with dynamic costs and trade-offs bound to decision paths. Systems should enable rapid discovery as well as management and tracking of these choices (options), making them first-class citizens of data analysis.

We operationalize the desiderata above through an interactive visual data analysis system called SYSTEMD¹. The system enables users to experiment with their data and understand relationships between drivers (input) and KPIs of interest (output) through what-if scenarios. We evaluate SYSTEMD using three common business use cases: (1) marketing mix modeling, (2) customer retention analysis, and (3) deal closing analysis with real-world business users such as sales, marketing, operations, accounts, and campaign managers.

In summary we make the following contributions:

- We revisit the fundamental goal of business analysis to be making data-driven decisions via the ability to understand the relationship between drivers (input) and KPIs (output).
- We argue that the four functionalities we propose, *driver importance analysis*, *sensitivity analysis*, *goal inversion (seeking) analysis*, and *constrained analysis*, are necessary for effectively helping business users to make data-driven decisions.

- We implement SYSTEMD, an interactive visual data analysis system incorporating these functionalities, providing a coherent workflow and end-to-end solution for business users to improve their decisions.
- We evaluate SYSTEMD through three business use cases to test the applicability and perceived value of the system.

2 SystemD

SYSTEMD has a client-server architecture. The frontend client is responsible for rendering analysis views, managing user interactions, communicating with the backend server to fetch requested data, and (re-)rendering views accordingly. The backend server runs machine learning models to predict KPI objective values and packs them into efficient JSON data structures to send to the client in response to user interactions. We train two widely used models: linear regression models when the KPI objective is a continuous variable (e.g., sales) and classifiers when the KPI objective is a discrete variable (e.g., customer retained after 6 months or not, deal closed or not) to make predictions. Now, we briefly describe the views and functionalities of SYSTEMD as annotated in Figure 2.

(A) Use Case Selection. Users can select from three common business use cases supported currently. Figure 2 shows a snapshot of SYSTEMD during the deal closing analysis use case.

(B) Table View. On selection of the use case, the corresponding dataset is tabulated in this view. The columns in this view show the drivers and KPI of the dataset.

In the deal closing analysis use case dataset, every row represents a prospective customer and every column represents the counts of activities corresponding to the prospect such as *Chats*, *Meetings* attended, etc. Each prospect is also labeled as converted to a customer or not.

(C) KPI Selection View. Users can select one of the variables in the dataset to be the KPI objective using this view.

For the use case, users are interested in whether prospects close as customers, hence choose the KPI to be *Deal Closed?*.

(D) Driver List View. This view allows users to filter out drivers they aren't interested in analyzing.

Users want to understand the relationship between all activities of prospects and the average deal closing rate in the use case. Hence, they keep all variables selected and deselect only the *Account* textual variables to enable model training and prediction.

(E) Driver Importance View. This view characterizes the first functionality of SYSTEMD, the *driver importance analysis*. Users are interested in recognizing the drivers that drive the KPI. This view illustrates an interactive horizontal bar chart informing the users of the drivers (Y axis) of the KPI in decreasing order of importance (X axis) thereby guiding them to take actions in the appropriate directions.

We use Scikit-learn [27] to train machine learning models that predict KPI values: linear regression models for continuous KPIs and random forest classifiers for discrete KPIs. We choose the driver importance values to be the linear-regression coefficients and the random-forest feature importances because they are relatively easier for users to understand. However, we verify the importances using traditional measures such as Shapley, Pearson, and Spearman rank [3, 30, 39] to ensure that the model coefficients are not misleading. The

¹<https://tinyurl.com/SigmaSystemD>

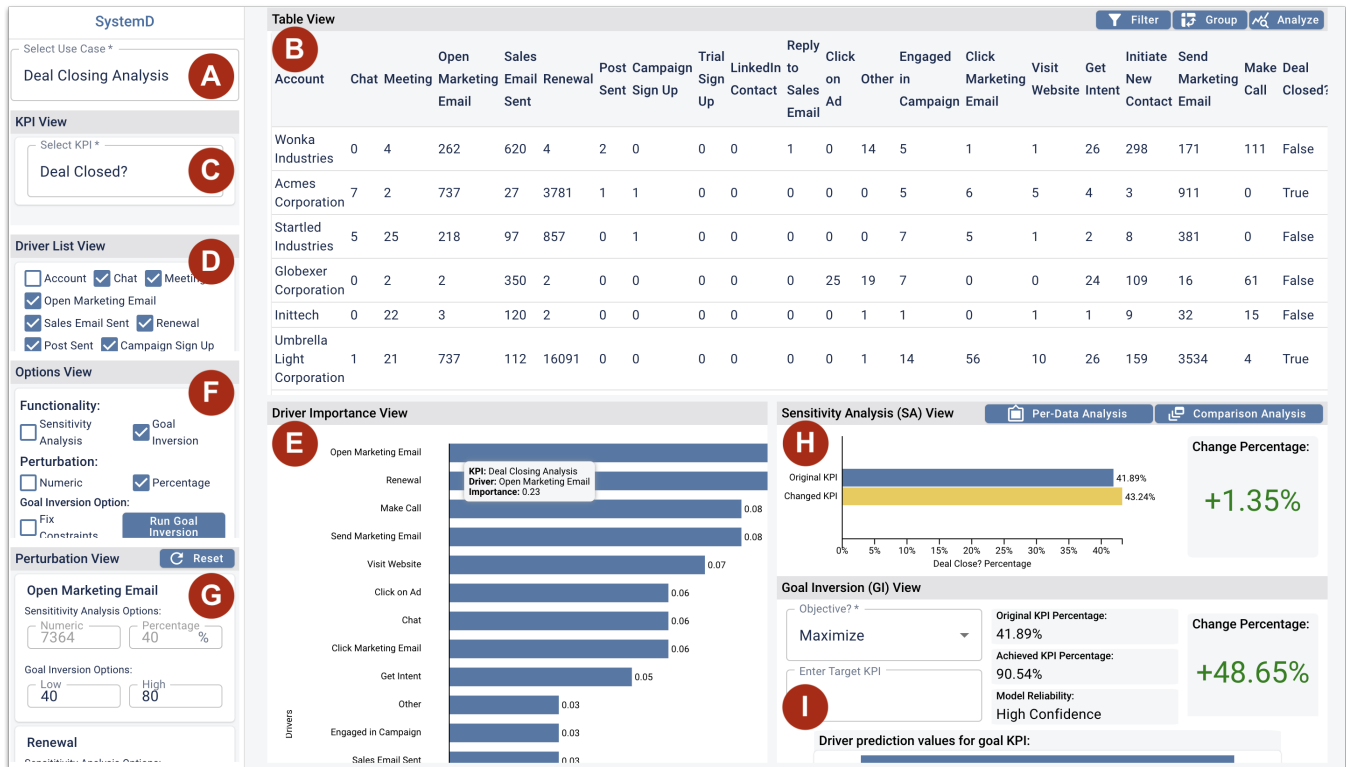


Figure 2: SYSTEMD user interface showing a snapshot from the deal closing analysis use case. It operationalizes four functionalities we deem necessary for augmenting decision making in data analysis systems. See Section 2 for details on each of the annotated views.

importance values range between -1 and 1 with extremes showing high negative and positive importance to the KPI respectively while closer to 0 shows low importance to the KPI.

Driver importance analysis for our use case shows that the three most important drivers of the deal-closing KPI are *Open Marketing Email*, *Renewal* and *Call* while the three least important drivers are *LinkedIn Contact*, *Initiate New Contact* and *Meeting*.

(F) Options View. This view provides options for types of analyses to perform and types of perturbations to apply to the original dataset. Users can check corresponding boxes to perform either sensitivity analysis or goal inversion analysis. The system supports two perturbation options: absolute and percentage values. The deal closing analysis use case illustrates outputs of both analyses using percentage perturbations.

(G) Perturbation View. Users can specify how much to perturb the original dataset using this view. For each of the drivers, they can set absolute or percentage perturbation magnitudes to perform sensitivity analysis and low and high constraints for goal inversion analysis.

Perturbations are made to each of the data points in the dataset; for example, a 40% increase on *Open Marketing Email* means increasing the marketing emails opened for every prospect by 40%.

(H) Sensitivity Analysis View. SYSTEMD helps users understand the relationship between drivers and the selected KPI using its second functionality of *sensitivity analysis*. Users can experiment with

the drivers by perturbing the original data — see (G) above — and observing the effects on the KPI. This view displays two bars: a static blue bar shows the KPI achieved on the original dataset and an interactive yellow bar shows the KPI achieved on the perturbed dataset. The change between KPI achieved on user perturbed data and the original data is explicitly shown to indicate the up-lift (positive, shown in green) and down-lift (negative, shown in red). Every perturbation re-runs the model prediction to re-calculate the KPI value.

To ease driver experimentation, SYSTEMD provides two additional sensitivity analysis features. To view sensitivity analysis in its entirety and compare KPI trends over all drivers, the *comparison analysis* feature shows the KPI achieved for every driver individually across a range of perturbations specified by the user. To allow users to drill down and experiment with individual data points, the *per-data analysis* feature lets users select a data point, perturb data of the selected data point and view effects on its KPI value.

The use case shows a 40% perturbation made to the *Open Marketing Email* predictor, which results in the deal-closing rate rising to 43.24% (yellow bar), an up-lift of 1.35% from that on original data.

(I) Goal Inversion View. The *goal inversion* functionality helps users to plan future actions that will help them achieve their KPI goals. Users can choose to freely optimize (minimize or maximize) the KPI or specify a desired KPI value they wish to achieve and run the model to provide values of the drivers that will achieve the goal KPI.

SYSTEMD uses Scikit-Optimize's Bayesian optimizer [23] to learn values of the drivers that attain the desired KPI value (maximum, minimum, or target). The view returns the best KPI attainable, the confidence of the model used, and a set (not necessarily unique) of driver values that achieve the user-specified KPI goal.

In practice, it is not always feasible for users to take the actions recommended by freely optimized goal inversion, such as when recommendations are not in line with domain knowledge or when optimal driver values exceed a budget, and so on. To obtain realistic recommendations, SYSTEMD provides the final functionality of *constrained analysis*. This functionality allows users to specify low and high constraints on one or more drivers and have the model recommend optimal KPI and driver values that are within the specified range.

The use case shows a *constrained analysis* in which users constrain the perturbation of the *Open Marketing Email* driver to be an increase between 40% to 80% as shown in (G). The resulting maximal deal-closing-rate KPI equals 90.54%, which is an up-lift of 48.65% from the KPI achieved on the original data.

3 Evaluation

We evaluate SYSTEMD through three business use cases to gather feedback on its interactive functionalities.

Participants and Protocol. We recruited five business users who play various roles at Sigma Computing and have a minimum experience level of four years. Three business users—a marketing manager, a campaign manager, and an account manager—participated in the marketing mix modeling use case. A product manager and a sales manager participated in the customer retention analysis and the deal closing analysis use case respectively. Although we recruited Sigma employees as our business users in order to facilitate our evaluation process, these employees are nonetheless fair representatives of the types of business users who daily perform analyses and make decisions.

Participants were briefed on the interview goals and structure. They were then asked to share analysis intent, workflows, current tools and challenges of their daily analysis in a pre-interview questionnaire, both to set the stage for the experiment and to let us understand in detail their context for decision-making using data. Then we shared our screen and demoed the system to the participants to explain the four functionalities. The three kinds of business users were each shown the system using their respective use case of interest so as to match their daily workflows. We demoed the system using certain interesting examples which we previously found while experimenting with the dataset, but we highly encouraged the participants to also guide us on the experiments they wanted to perform. For example, the product manager in the retention rate analysis use case explicitly asked us to remove an obvious predictor and perform the functionalities again. After the interview, participants were asked to provide feedback by completing a system usability questionnaire. Participants rated their agreement with system usability questions using a Likert scale of 1 (strongly disagree) to 5 (strongly agree). Additionally, we asked participants to provide us with more open-ended feedback on the features of, and concerns about, SYSTEMD. The set of questions from the pre-interview questionnaire, system usability questions and open-ended feedback questions are

listed in Table 1. Interviews typically lasted 45 minutes. We used a think-aloud protocol to encourage participants to share their thought process simultaneously as they got results from SYSTEMD. We recorded audio and captured screen for later review.

Use Cases. Here we briefly describe the three business use cases.

U1: Marketing Mix Modeling. Marketing Mix Modeling is a technique which helps quantify the impact of several marketing inputs on sales, market share, etc. With the rapid increase of new media, there has been an increased need for marketers to optimize their strategies that maximize KPIs such as return of investment, sales, profit, etc. across multiple channels.

To allow Sigma's marketing, campaign and account managers to use SYSTEMD for making decisions that optimize investments per media channel, we used a dataset describing investments made over a period of 6 months on 5 media channels (Internet, Facebook, YouTube, TV and Radio) and corresponding sales achieved per day. The participants guided the analysis for understanding the behavior between media channels and sales KPI and made decisions on which channel investments should increase or decrease to maximize sales.

U2: Customer Retention Analysis. Operations and product managers are interested in retaining existing customers because it is much easier, both in terms of effort and cost, to retain old customers than get new ones. To build strategies that retain customers, business users analyze large datasets of customer interactions with the product and their own activities with the customers that are collected by CRM and support systems. Examples of such activities are using help chat, opening new document, adding a visualization, etc. Additionally, business users often form hypothesis of additional features, or *formulas* that may drive customer retention such as customers attending 2+ demo meetings in the first two weeks, and so forth.

Sigma's product manager was interested in finding customer activities that would maximize retention after six months. Sigma's multi-touch attribution dataset was used for analysis; it consists of a customer's activities and product manager's hypothesis formulas such as pivoting on data, performing join operation, using 3+ formulas in two weeks, etc., during the last six months, along with a label indicating whether the customer was retained after six months. The product manager used SYSTEMD to optimize the six-month-retained KPI and understand the activities that lead to increased six month retention.

U3: Deal Closing Analysis. The approach to selling a product begins with initiating contact with a prospective customer and ends after a salesperson closes the deal. This process consists of multiple stages of back and forth communication between the prospects and the sales teams. Similar to the customer retention analysis use case, sales managers use a prospect's activities with the product, as well as the sales and marketing teams communication activities, to predict whether prospects will close to customers. Examples of such activities are prospects signing up for trials or demos, participating in campaigns, attending scheduled meetings, and so on.

The sales manager at Sigma wanted to analyze prospect data from the past few months similar to the data described in Section 2 to answer questions such as "what activities lead to the conversion of more prospects?" and "what is the ideal customer journey formula for Sigma?".

Table 1: Pre-study, system usability (elicited on a Likert scale of 5), and open-ended questions used for evaluating SYSTEMD.

Category	Questions
Pre-study	<p>Can you describe the kind of data you use? What is the intent of using the data? Given the data, what would you be most interested in analyzing? What is the purpose behind interest in the analysis of the data? Consider you are interested in sales (U1)/retention rate (U2)/deal closing rate (U3), can you describe what analysis would you perform to make decisions on investing in the right channels (U1)/increasing the retention rate (U2)/increasing deal closing rate (U3)? Which tools do you use typically to perform the analyses you described? How easy or hard would you say it is for you to analyze the data and make a decision? How much time would you approximately take to come up with a hypothesis and make a decision based on that? What strategies do you use to evaluate whether analyses results match your expected hypotheses (via your domain knowledge and/or experience)?</p>
System usability	<p>The functionalities of SystemD are useful in understanding the behavior of the data better. The functionalities of SystemD are useful in making optimal decisions. The interactions with SystemD are intuitive. Most users would learn to use SystemD very quickly. Various functionalities of SystemD are well-integrated. Compared to your process of analysis and current tools you use on a daily basis for making decisions (as described initially), how useful do you see SystemD helping you for the same tasks? Use SystemD in my daily work.</p>
Open-ended	<p>Compared to your process of analysis and current tools you use on a daily basis for making decisions (as described initially), how useful do you see SystemD helping you for the same tasks? Explain why. How useful is SystemD for making decisions that optimize interesting metrics (KPIs) in comparison to current tools? Explain why. List the most useful functionalities or features from most useful to least useful (Driver Importance Analysis, Sensitivity Analysis, Goal Inversion (Seeking) Analysis, Constrained Analysis). Which additional functionalities or features would become a more effective system to make decisions in SystemD? What would be your concerns with the SystemD?</p>

4 Results

We now discuss the qualitative feedback we collected from Sigma business users on SYSTEMD (Figure 3).

Understanding Driver-KPI Relationship. We found that participants were able to quickly understand the dynamics between drivers and their KPI objective and make optimal decisions using SYSTEMD. When describing her prior experiences with campaign analysis, the campaign manager said “[right now] we have to manually set accurate targets for our campaigns looking at the historical data and intuitively figure out the media channels [that] have the best [chance at] converting leads [of users] into actual customers” and noted that SYSTEMD is “definitely much more actionable!”. We found that all other participants shared similar sentiments. For example, while describing the challenges with current tools used such as reporting tools and spreadsheets, the marketing manager remarked that her “team consists of only marketers and not technical engineers or data scientists, making them really struggle ... to really see what is going on in spreadsheets and this [SYSTEMD would be] ... very exciting to perform ... [such] analyses”.

Value For Business Users. All five participants saw very strong value in using SYSTEMD as part of their daily workflows. Participants shared that they mostly used reporting tools such as Sigma, Microsoft Excel or Salesforce for their analysis. They also shared that they did not think any of these tools thrived in statistical analysis, making them unaware of any machine learning or predictive analytics functionalities in the tools. They repeatedly found the interactive statistical analysis of SYSTEMD beneficial as it could save them from the intuitive and time-consuming trial and error strategy of hypothesizing next steps, performing them and waiting for three to six months to see the results. The product manager liked the ability

of the interactive sensitivity analysis functionality to help her narrow this expensive strategy which “... is not something that [she] ... is easily able to do right now”.

Business users in two use cases were interested in working directly with the SYSTEMD team to fine-tune the system for their individual needs immediately. For example, after walking through the analysis demo, the accounts manager exclaimed that “[she] wanted to get access to [SYSTEMD] now!!!” and asked “whether we [her team] could access it ourselves [themselves]?”. This showed the participant’s interest in using SYSTEMD right away.

Usefulness and Completeness. When asked to rank the functionalities in terms of usefulness, we found that 3 of 5 participants ranked the driver importance analysis as being the most useful for their analysis while the other two participants ranked both sensitivity analysis and constraint-based analysis to be the most useful. All participants also ranked multiple functionalities second in usefulness and none being least useful. Thus, participants seemed to find all functionalities provided by SYSTEMD useful. Most participants also found the functionalities well integrated and connected to one another.

Suggestions for Improvements. Participants showed great interest in integrating SYSTEMD with Sigma Worksheet[15], a visual query-builder interface, so they could add calculations for additional features or drivers to the dataset. This could also serve any need to slice, dice and drill to obtain the required analysis data, such as per customer-cohort or prospect-stage analysis. Also, most participants needed clarification to understand the outputs of the functionalities, suggesting a need for better user experience workflows. We aim to resolve these concerns as we are starting to deploy SYSTEMD into Sigma.

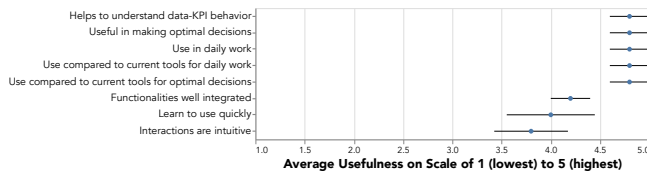


Figure 3: Feedback elicited through responses to usability questions on a Likert scale of 5. Participants found SYSTEMD to be highly useful for understanding data-KPI behavior, making optimal decisions, and adopting it in their daily work. On the other hand, they rated its interactions to be intuitive to a lesser degree.

5 Future Research Directions

Our study, along with our experience so far, confirms that helping business users make data-driven decisions implicitly using models is a difficult multifaceted problem with rich research and engineering challenges. We now discuss some of these challenges to indicate future research directions.

Communication and User Control. Decision making involves two constituents: domain expert user(s) and models implicitly created through user interactions, as shown in SYSTEMD. One challenge is how to best calculate and communicate the underlying model assumptions and confidences to users who have no background in statistics or data science? Another challenge is that some functionalities, such as goal inversion, may come across as “magic,” leading the users to form unrealistic expectations from the system. When the system fails to meet these expectations users may lose trust and become dissatisfied not just with the intelligent functionalities but also with the whole system. User trust is thus critical for automated methods, intelligent or otherwise [11]. Automated methods need to *just work* and when they fail, they should fail consistently, while giving users a sense of control. This leads to the future problems of both modeling as well as system and user experience design. For instance, how can a user give feedback if she thinks that what the system recommended is incorrect? And, how do we retrieve this feedback and update our models and systems? One solution is to use linear solvers [33] which can scale but face challenges in providing fast real-time response when the data is large or when robust models are used for prediction. There is ample scope for future research.

Robustness. The optimal solution from a given data-based model may be brittle: under small changes to the model or data, the solution may suddenly perform very poorly. This robustness problem is exacerbated by the fact that, for a given set of data, it is often the case that multiple models can reasonably explain the relationship between data and outcomes or objectives [4, 5]. These different models may yield different rankings of driver importance as well as different solutions to optimization and goal-seeking problems. Therefore, data analysis systems designed to augment decision-making need to take the robustness of solutions and multiplicity of explanatory models into account.

Interpretability vs Accuracy. It is well-known that some models are simpler and easier to interpret while others are more accurate but difficult to explain. It is essential that we study which models to pick

for our business users. Do we allow our users to have a say in this choice?

Specification and Reuse. In order to reuse the decision-making analyses process, it is important to have an editable specification of the experiments that SYSTEMD supports. It is crucial to identify the right grammar for specifying these data experiments and enabling their interoperability with, e.g., other data science languages or platforms such as computational notebooks. An interesting direction is to integrate these experiments with SQL as many data analysis systems [7, 8], including Sigma, compile the data analysis intent of users into SQL queries. Therefore, development of a declarative specification language for SYSTEMD is a potential future direction.

Conducting a Complete User Evaluation To rigorously learn the usefulness of SystemD beyond our initial evaluation, there is wide scope for conducting larger quantitative-based user studies. We want to understand various business user’s existing decision workflows compared to when they use SystemD, and to compare quantitative factors such as whether a decision can be made, the amount of time needed to make the decision, and so forth. Another potential direction is comparing the functionalities of SystemD to the functionalities provided by other existing commercial tools such as Salesforce’s Einstein Discovery [34] and Microsoft’s PowerBI [14].

6 Related Work

Model-based Analysis. The interactive visual analytics [22] literature has a plethora of prototype systems that incorporate models for data analysis. Typically these systems use unsupervised offline models to process data for feature extraction that subsequently augments interactive analysis. Less frequently, systems use online models to drive interactive analysis, for instance, the Praxis [9] forward projection and backward projection for dimensionality-reduction-based data analysis. SYSTEMD extends these two interactions to sensitivity analysis and goal inversion (seeking) analysis, respectively, for generalized model-driven relational data analysis for business users.

Model-driven Analysis. Prior work introduces interactive systems for understanding models. Some systems have interactions similar to sensitivity analysis (e.g., [2, 38, 40]). Others use counterfactual explanations (e.g., [10, 16, 17, 19]) on how to gain the desired prediction with minimal changes to the input, which is akin to our goal inversion (seeking) analysis. Unlike SYSTEMD, the purpose of these approaches is to understand and debug machine learning models, not to support model-driven analysis.

Data Farming. In the context of simulation, the data farming literature [31] emphasizes the value of a holistic understanding of a model (e.g., robustness properties, sensitivity) over simple optimization, which can yield high-reward but very brittle decisions. Unlike SYSTEMD’s hands-on approach, the emphasis in [31] is on non-interactive visualization, and goal-inversion (seeking) is not explicitly supported.

Spreadsheets. Spreadsheet applications generally enable optimization-based data analysis through add-ins and native functions. For example, Excel’s SOLVER [24] and GOAL SEEK [25] features allow solving for a desired output of a formula by changing its drivers. While these do not involve model building, they provide a level of functionality enabling what-if analysis, albeit with limited interactivity and expressivity.

BI Systems. Commercial business intelligence (BI) tools recognize the need to augment business user’s decision-making in data analysis systems. Both established players such as Power BI [14], Tableau [35], Salesforce [34], Facebook Robyn [29], Alteryx [21], SAS Visual Analytics [32] and newcomers such as c3AI [1], Einblick [13], and Sisu Data [12] aim to provide various predictive and prescriptive functionalities. However, these tools either provide partial solutions whose effectiveness and adoption remain unclear or fail to cater to business users. SYSTEMD specifically targets business users based on an intuitive set of interactive functions, providing a coherent workflow for data-driven business decision-making.

7 Conclusion

In this paper, we start by defining the fundamental goal of business data analysis: improving business decisions based on data. This goal requires understanding and probing into the relationships between data variables representing hypothesized drivers and those representing key performance indicators (KPIs) that business users are interested in. We then introduce four functionalities for interactively learning and reasoning about these relationships. The success of model-driven analysis for broader audiences with no background in coding or data science depends on understanding the users and their goals and designing the user experience accordingly. It is not a technology-first problem. We, therefore, focus on providing intuitive interactive functionalities and provide a preliminary use case driven evaluation here.

A close reading of Tukey’s writings suggests that his emphasis on EDA and graphical representations was a reaction to (or an antithesis of) dry, purely confirmatory approaches to data analysis of his time. It is however time to bring the pendulum of data analysis currently skewed towards exploratory analysis to a synthesis reflecting the needs of large swaths of users in performing data analysis. This is important because the purpose of enterprise data analysis in practice on the ground is neither pretty pictures nor exploratory insights but improved decisions.

References

- [1] C3 AI. 2021. C3 AI Ex Machina. (2021). <https://c3.ai/products/c3-ai-ex-machina/>
- [2] Saleema Amershi, Max Chickering, Steven M Drucker, Bongshin Lee, Patrice Simard, and Jina Suh. 2015. Modeltracker: Redesigning performance analysis tools for machine learning. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 337–346.
- [3] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise reduction in speech processing*. Springer, 1–4.
- [4] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.
- [5] Leo Breiman. 2001. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science* 16, 3 (2001), 199–231.
- [6] David R. Brillinger. 2002. John W. Tukey: His Life and Professional Contributions. *The Annals of Statistics* 30, 6 (2002), 1535–1575. <http://www.jstor.org/stable/1558729>
- [7] Matteo Brucato, Nishant Yadav, Azza Abouzied, Peter J Haas, and Alexandra Meliou. 2020. Stochastic package queries in probabilistic databases. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 269–283.
- [8] Zhuhua Cai, Zografoula Vagena, Luis Perez, Subramanian Arumugam, Peter J Haas, and Christopher Jermaine. 2013. Simulation of database-valued Markov chains using SimSQL. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. 637–648.
- [9] Marco Cavallo and Çağatay Demiralp. 2018. A visual interaction framework for dimensionality reduction based data exploration. In *CHI*. 1–13.
- [10] Furui Cheng, Yao Ming, and Huamin Qu. 2020. DECE: Decision Explorer with Counterfactual Explanations for Machine Learning Models. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1438–1447.
- [11] Jason Chuang, Daniel Ramage, Christopher Manning, and Jeffrey Heer. 2012. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 443–452.
- [12] Sisu Data. 2021. Sisu Datas Blog. (2021). <https://sisudata.com/>
- [13] Einblick. 2021. Einblick. (2021). <https://einblick.ai/>
- [14] Alberto Ferrari and Marco Russo. 2016. *Introducing Microsoft Power BI*. Microsoft Press.
- [15] James Gale, Max Seiden, Gretchen Atwood, Jason Frantz, Rob Woollen, and Çağatay Demiralp. 2021. Sigma Worksheet: Interactive Construction of OLAP Queries. arXiv:2012.00697 [cs.DB]
- [16] Sainyam Galhotra, Romila Pradhan, and Babak Salimi. 2021. Explaining black-box algorithms using probabilistic contrastive counterfactuals. In *Proceedings of the 2021 International Conference on Management of Data*. 577–590.
- [17] Oscar Gomez, Steffen Holter, Jun Yuan, and Enrico Bertini. 2020. ViCE: visual counterfactual explanations for machine learning models. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*. 531–535.
- [18] Peter J Haas, Paul P Maglio, Patricia G Selinger, and Wang-Chiew Tan. 2011. Data is dead... without what-if models. *Proceedings of the VLDB Endowment* 4, 12 (2011), 1486–1489.
- [19] Fred Hohman, Andrew Head, Rich Caruana, Robert DeLine, and Steven M Drucker. 2019. Gamut: A design probe to understand how data scientists understand machine learning models. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [20] Jessica Hullman and Andrew Gelman. 2021. Designing for interactive exploratory data analysis requires theories of graphical inference. (2021).
- [21] Carlisle Idoine, Peter Krensky, Erick Brethenoux, Jim Hare, Svetlana Sicular, and Shubhangi Vashisth. 2018. Magic Quadrant for data science and machine-learning platforms. *Gartner, Inc* (2018), 13.
- [22] Daniel Keim, Gennady Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. 2008. Visual analytics: Definition, process, and challenges. In *Information visualization*. Springer, 154–175.
- [23] Gilles Louppe. 2017. Bayesian optimisation with Scikit-Optimize. (2017).
- [24] Microsoft. 2021. Microsoft Excel Documentation. (2021). <https://support.microsoft.com/en-us/office/define-and-solve-a-problem-by-using-solver-5d1a388f-079d-43ac-a7eb-f63e45925040/>
- [25] Microsoft. 2021. Microsoft Excel Documentation. (2021). <https://support.microsoft.com/en-us/office/use-goal-seek-to-find-the-result-you-want-by-adjusting-an-input-value-320cb99e-f4a4-417f-b1c3-4f369d6e66c7/>
- [26] Raymond S Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology* (1998).
- [27] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12 (2011), 2825–2830.
- [28] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*.
- [29] Robyn. 2021. facebookexperimental. (2021). <https://facebookexperimental.github.io/Robyn/>
- [30] Alvin E Roth. 1988. *The Shapley value: essays in honor of Lloyd S. Shapley*. Cambridge University Press.
- [31] Susan M. Sanchez. 2020. Data Farming: Methods for the Present, Opportunities for the Future. *ACM Trans. Model. Comput. Simul.* 30, 4 (2020), 22:1–22:30.
- [32] SAS. 2021. SAS Visual Analytics. (2021). https://www.sas.com/en_us/software/visual-analytics.html
- [33] Zeyuan Shang, Emanuel Zraggen, Benedetto Buratti, Philipp Eichmann, Navid Karimeddini, Charlie Meyer, Wesley Runnels, and Tim Kraska. 2021. Davos: a system for interactive data-driven decision making. *Proceedings of the VLDB Endowment* 14, 12 (2021), 2893–2905.
- [34] Tableau. 2021. Einstein Discovery. (2021). <https://www.tableau.com/products/add-ons/einstein-discovery>
- [35] Tableau. 2021. What is Tableau Business Science? (2021). <https://www.tableau.com/about/blog/2021/3/what-is-tableau-business-science>
- [36] John W Tukey. 1962. The future of data analysis. *The annals of mathematical statistics* 33, 1 (1962), 1–67.
- [37] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* (1974), 1124–1131.
- [38] James Wexler, Mahima Pushkarna, Tolga Bolukbasi, Martin Wattenberg, Fernanda Viégas, and Jimbo Wilson. 2019. The what-if tool: Interactive probing of machine learning models. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 56–65.
- [39] Jerrold H Zar. 2005. Spearman rank correlation. *Encyclopedia of biostatistics* 7 (2005).
- [40] Xun Zhao, Yanhong Wu, Dik Lun Lee, and Weiwei Cui. 2018. iforest: Interpreting random forests via visual analytics. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 407–416.