International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# Vize: A Voice-Based Interface for Data Visualization and Summarisation with Recommendation System

**Anil Vasoya[1], Kamal Shah[2], Neha Gadiya[3], Palka Dhirawani[4], Pratik Kanani[5], Vedant Gandhi[6]**

**Submitted**: 03/10/2023 **Revised**: 23/11/2023 **Accepted**: 03/12/2023

**Abstract:** Data comprises information that may be used to improve business performance and keep a competitive edge. Extraction of knowledge and insights depends on data analytics and visualization technologies. The popularity of voice technology has skyrocketed recently. Users may communicate with technology using only their voices, which leverages speech recognition technology. A Voice User Interface (VUI) is intended to simulate dialogues that result in device-user interaction and facilitate users' ability to complete activities or perform information searches without requiring their hands or eyes. Thus, users can interact freely with their data, verify, and query it, and pursue their own lines of inquiry in a way that is both flexible and unfettered. Through a survey conducted, it was deduced that having a VUI could be a solution to streamline processes and give users an easier method to access data visualization tools. This research work aims to build a web-based tool that would enable data exploration and visualization through speech recognition. This equips users of all skill levels as well as those with sight impairments to easily upload their data and create visualizations with the use of straightforward voice commands. By simply instructing the software to do so, users may also request a summary of their data and customize the charts they create as they see fit. A Data Visualization Recommendation System (DVRS) was introduced that automatically created a dashboard that showcases the charts that can help users to understand their database quicker. It makes use of a simple feedforward neural network that provides insights into the underlying factors driving the predictions, resulting in charts with a high probability value. The charts with the highest probability values are chosen to be presented on the dashboard.

***Keywords:*** *Data Visualization, Voice user interface, Summarization, Accessibility, Speech Recognition, Recommendation System*

## 1. Introduction

There is a tremendous amount of data in the globe, and it keeps growing every second. In today's business world, as companies gather a huge volume of data, visualizing and analysing it effectively has become imperative. Data visualizations and dashboards enable decision-makers [1] to quickly identify trends, patterns, and insights that may otherwise go unnoticed. This, in turn, allows them to make better decisions, gain a competitive edge in their respective industries and take advantage of the opportunities presented by big data. In essence, data visualization transfers the responsibility of understanding numerical data from the viewer's numerical reasoning to their visual reasoning, making it easier to understand complex information [2].

On another note, the continuous advancement in technology has led to a growing interest in Voice User Interfaces. With more people relying on their smartphones for everyday tasks, VUIs offer a more convenient and efficient way to interact with technology. Voice-based systems enable people to communicate with digital devices by using voice detection and speech output. They can be basic components or full-fledged conversational bots, complete with cues, grammars, and discourse logic [3].

Moreover, voice systems have bridged the gap in making technology more inclusive as they help overcomes the challenges of accessing graphical interfaces faced by users with sight impairments. Text-to-speech technology enables individuals who rely on screen-readers to access the textual content in an audio format which allows for a hands-free experience. While voice interfaces in the form of virtual assistants, smart speakers, Interactive Voice Response (IVRs), etc., have been on the rise, [4] there are still numerous domains where currently there is a lack of voice-based accessible technology. One such domain is that of data analysis.

The motivation behind developing a voice-based data visualisation tool is to democratise information access while also advancing analytics and data literacy. Such a tool can assist people and organisations make more sound choices, advance numerous fields, and get around the drawbacks of conventional data visualisation techniques by offering a more simple and instinctive approach to engage with data.

Our contribution to the work aims to align the requirements of novice users to showcase complex data in the form simple visualizations using voice commands

*[1,2]Thakur College of Engineering and Technology, Mumbai, India.*
*[3,4,5,6] Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.*
*vasoyaanil32@gmail.com[1], kamal.shah@thakureducation.org[2], nehapgadiya@gmail.com[3], palkadhirwani8@gmail.com[4], pratikkanani123@gmail.com[5], gandhivedant01@gmail.com[6]*

without any coding required. The voice-based tool will aid blind users in performing a variety of data visualization tasks and get appropriate results in the form of dashboards which can be easily understood by screen readers. It also reduces the cognitive load of the user by providing interactive tutorials and a clean user interface. Another feature that would generate summaries for individual charts in the form of simplified text is implemented. The summary would display the common statistical measures necessary to extract insights from the chart. This would make it much easier to arrive at conclusions for inexperienced users and aid visually challenged screen-reader users in exploring the data.

## 2. Related Work

There has not been much research regarding creation and manipulation of data visualizations using voice commands. While data analysis tools like Tableau and Power BI are commonly used to perform data visualization, these tools have not explored the utility of integrating voice commands. Dundas partnered with Profunda Analytics [6] to enable users to build voice-controlled business intelligence systems. However, this tool is ideal for users that have extensive knowledge about data analytics and could be overwhelming for novice users. Along with this Dundas does not provide users with automated insights and fails to handle large data sets properly. It lacks the predictive analysis features and so a lot of trending data must be done in a roundabout fashion and scripts must be used in a number of places to make such complex data manipulations or visualizations thus only benefitting users who have adequate knowledge about data analysis.

Another plugin FlowSense [7], although easy to use because of its intuitive voice commands and its ability to create highly complex visualizations fails to provide a summary of the data and does not allow changing the UI of the charts, thus restricting any customizations to the graphs to make it more personal and user friendly.

Zierau et. al. carried out four extensive studies to ascertain how voice-based interfaces affect user experience. Through these experiments the authors found that [8] voice-based interfaces don’t necessarily have an advantage over text-based interfaces for a few varying reasons. According to research, text-based interfaces allow consumers to control the speed of processing information, which makes the data visualization process simpler. In contrast, voice activated interfaces result in higher levels of interface flow, leading to better experiential service perceptions, but more conversational turns which can be a disadvantage. Conversely, they concluded that voice-enabled interfaces lead to enhanced levels of interface flow, which increase experiential service perceptions but have a downside of having more conversational turns. There is a fine line between voice technology being beneficial and it becoming a nuisance, and the conditions under which the former would be true need to be examined in detail.

Murad et. al. identifies overlaps between existing GUI guidelines and the preliminary VUI heuristics proposed in their literature survey. In this study, the authors discuss the current guidelines for designing interactive graphical user interfaces (GUIs) and the emergence of voice user interfaces (VUIs). They conducted a thorough literature review of 21 articles, using specific search terms related to VUIs, [9] and found that issues commonly encountered in speech-based interactions, such as error recovery and system-to-real world mapping, have similarities with those addressed in GUI guidelines. However, the authors also highlighted the need for further consideration of how to develop design guidelines that can help designers create seamless, device-independent, conversational, and ubiquitous VUIs. The authors suggested a three-step plan to tackle this issue, which includes involving non-speech design experts in usability testing, verifying design heuristics through experiments, and evaluating the practical applications of these heuristics in realistic situations. The goal of this approach is to modify existing design principles for graphical user interfaces and apply them to voice user interfaces, resulting in a set of verified design guidelines for VUIs that can be used in practical applications.

Hua Guo et. al. conducted several experiments to understand how interactions with a visual analytics dashboard can lead to effective analytical insight generation. [10] They introduced a new analysis method involving curating 2 experiments: the first experiment quantitatively analysed the interaction features of the visualization tool to identify frequently performed actions and drew patterns which could then be segmented into interaction sequences. 4 primary patterns were finalized (orienting, locating, sampling, and elaborating) to carry out qualitative analysis of video segments and interaction logs to understand the correlation between these patterns and various insight metrics. They also observed that the search bar led to more elaborate actions as compared to using the query builder. Although various results were generated there did remain a few research gaps that could lead to more impactful results. The recruited participants should have been a mix of people having different levels of expertise of the system, leading to a much more insight driven visual analytical tool. Secondly, the order in which the interactions were presented was controlled. Finally, the temporal Aspects of Interaction History was not considered, i.e., how much time each candidate was given to come to conclusions and carry out interactions, thus some insights can be the ones that are generated after spending long hours while some may be incomplete.

Shen et. al. summaries V-NLP software as well as research articles and how the field of visual NLP is constantly evolving to help make better analysis tools to make visualization easier. They conducted a survey to summarize the findings of the different types of NLP toolkits, visualization types and characteristics of different types of V-NLP's. [11] The authors created a visualization pipeline of 7 stages (interpretation of queries, transformation of data, mapping of data into visual forms, transformation of views, human interaction, management of dialogue, and presentation of information) and discussed the improvements that could be made to each stage of the visualization process along with the tools or technologies used to perform each task. They also mentioned several research gaps that they would wish to fulfil. A promising solution to battle ambiguity when using V-NLPs was to track analytical tasks of the user and thus have a customized response to this ambiguity tailored to every user. Secondly, using advanced NLP models can help the system identify general grammatical structures from a large number of conversations related to visual analysis. This process will require utilizing existing high-quality textual datasets to create dependable NLP models for visual natural language processing (V-NLP), and then testing and training them.

Srinivasan et al. [12] provides a comprehensive overview of different techniques for assessing Natural Language Interfaces (NLIs) for data visualization, emphasizing their advantages and potential drawbacks, which can act as a benchmark for future studies. This study aims to be a useful reference for future researchers and professionals in choosing appropriate assessment methods depending on their specific research objectives and situations. In this study, the authors discuss the challenges of evaluating visualization NLIs, particularly with regards to training and task framing. To address these issues, they examine three popular task framing strategies for evaluating visualization NLIs: Jeopardy-style facts, open-ended tasks, and target replication tasks. The authors highlight the benefits of each strategy and provide guidelines for their use. The Jeopardy evaluation involves framing tasks as dataset statements or facts, which engages participants and provides a straightforward way to measure success. The use of open-ended tasks allows users to delve into data and assess the system's overall functionality and ease of use. Target replication tasks, on the other hand, provide specific criteria for users to replicate or achieve using the system and are useful for evaluating basic operations and command sequences with minimal risk of wording bias. Overall, the authors suggest that better visualization NLIs can be developed by overcoming the challenges of training and task framing, and by utilizing appropriate evaluation strategies. This can lead to improved usability and effectiveness of visualization NLIs.

Padmanandam et. al. explains the use of word clouds as a means of transforming tedious textual data into a visually appealing format that emphasizes significant textual content and quickly communicates important information. [13] Their proposed system generates word clouds for the input provided to it. This system takes various inputs (text from file, typed text, vocal input) thereby allowing effortless and efficient utilization by the user. However, this word cloud building system is limited to the English language exclusively. There is potential for further development to include multiple languages and real-time analysis of social media data to identify current trends in addition to analysing them.

Another research along similar lines is conducted by Marriott et al. [14], which studies the challenges faced by individuals with disabilities in accessing data visualizations. While most studies on this topic have focused on individuals with visual impairments, the authors emphasize the difficulties faced by individuals with cognitive and learning disabilities (CALD) as well as those with fine motor coordination issues. They highlight that such individuals may struggle with interpreting and operating interactive visualizations, which are often designed using abstract symbols and conventions that may not be accessible to them. To address this issue, the authors recommend adopting inclusive approaches to visualization design, such as providing information in multiple modalities and facilitating multimodal interaction, developing new visualization tools that enable individuals with disabilities to create and manipulate visualizations, and establishing evidence-based guidelines for accessible visualization design. They also stress the importance of understanding the specific disabilities that impact the use of data visualizations and identifying effective workarounds and accommodations to address them. These recommendations are aimed at promoting greater accessibility and inclusivity in data visualization design, enabling individuals with disabilities to better access and engage with visualized data.

As data visualizations are becoming more ubiquitous, Sharif et. al. [15] provides valuable insights into how voice-based visualizations can help people with visual impairments to analyse data and derive information. The first experiment they ran consisted of all participants having visual impairments. The authors aimed to understand how difficult it was for the users to derive insights from the tool that they have, and how similar is the visualization tool with respect to the software tools or the applications they engage with as part of their everyday activities. During the study, participants who utilized screen-readers showed dissatisfaction with the lack of direct access to the presented data points in the visualizations due to insufficient descriptions provided. They recommended that techniques such as sonification, summarization, and braille printouts could be useful in presenting data trends to individuals with visual impairments. Another experiment was conducted to compare the accuracy and speed of interaction between the system and screen-reader and non-screen reader users. It was found that non-screen readers had a 61% margin over the screen readers when it came to accurately drawing insights from the system. Thus, the authors realized the need to design systems that considers all types of users to accurately derive insights. They also highlighted the difficulties of using screen readers to understand visualizations, hence making the need for a voice-based data visualization tool more crucial and made suggestions in employing new techniques and strategies that would improve interaction. Even though the authors helped to identify a new target audience, they did not explore different types of screen readers that could eliminate the possibility of some of the highlighted issues occurring.

VoxLens, an open-source JavaScript plugin developed by Sharif et al, was inspired by calls to action to address the challenges faced by screen-reader users in accessing data visualizations. [16] This tool aims to enhance accessibility to online data visualizations and can be used with data visualization creation tools. VoxLens offers three modes of interaction: Question-and-Answer, Summary, and Sonification. In Question-and-Answer mode, users can extract information from visualizations using voice commands. Summary mode provides a summary of the information presented in the visualization, while Sonification mode maps data to a musical scale using Tone.js. Wizard-of-Oz studies for all three modes showed that participants appreciated clear instructions and responses, integration with their screen reader, and the ability to query specific terminologies. The research also showed that users preferred customized summaries for the visualizations. The researchers suggested incorporating an interactive guide to help users become more familiar with the tool, as well as a help section to identify which commands are supported, to enhance the user experience. VoxLens is only compatible with two-dimensional data visualizations that have a single series of data, and it is only compatible with bar charts, line graphs, and scatter plots.

**Table 1.** Literature Survey

| Paper Title | Advantages | Research Gaps |
|---|---|---|
| Voice bots on the frontline: Voice-based interfaces enhance flow-like consumer experiences & boost service outcomes [8] | Carried out 4 extensive studies to ascertain how voice-based interfaces affect user experience to come to conclusions that voice enabled interfaces lead to enhanced levels of interface flow which increase experiential service perceptions but have a downside of having more conversational turns. | Need to explore under which conditions voice technology can enhance consumers' experience versus when it becomes a nuisance or even detrimental. |
| Revolution or Evolution? Speech Interaction and HCI Design Guidelines [9] | VUI-specific (heuristic) design guidelines in the existing literature are in fact closely aligned with existing GUI guidelines. | Does not consider established GUI design heuristics as a starting point for the development of new design heuristics for VUIs. |
| A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights [10] | The typical insight-based evaluation approach does not specify a guiding framework for assessing the relationship between the tasks an analyst completes and the insights they produce. The authors employ a hybrid evaluation technique to explore how application design affects insight generation in order to close these gaps.<br>2 experiments conducted one after the another to identify patterns in carrying out analysis and thus drawing insights.<br>Both questions effectively answered. | Due to lack of time consideration, some insights are generated after spending long hours while some are incomplete.<br>The pattern extraction algorithm presented performs exact matching for top-level action sequences and is not able to capture more expressive patterns or ones with variable-length action sequences.<br>The order in which the interactions were presented was controlled whereas in general case scenario this might not the case and it may affect the results of the interaction and insight generation. |

| | | |
|---|---|---|
| | | Diverse group of participants not considered. |
| Towards Natural Language Interfaces for Data Visualization: A Survey [11] | A comprehensive overview of what characteristics are currently concerned and supported by V-NLI and scope for future work and improvements that can be made on the system.<br><br>Effectively breaks down the tasks involved in visualization into 9 steps and analyses each one in depth, the learning from the previous systems and what improvements can be made.<br><br>Provides a list of a wide variety of V-NLP tools already in market and their disadvantages as well as a number of toolkits for NLP that can be utilized at different stages of the visualisation pipeline.<br><br>Considers user experience along with effectiveness of particular V-NLPs<br><br>Compares different types of V-NLPs like text-based, voice based or selecting text from articles. | Rely on pre-existing toolkits for analysis, thus lacking adequate consideration of visualization element (e.g., mark, visual channel, and encoding property).<br><br>Up till now V-NLPs do not use models that provide robustness in the tasks they form and often lack performance. Applying more sophisticated NLP models to discover universal grammatical patterns from a huge corpus of talks in visual analysis is a promising path for building a more reliable system. The process will involve using the current high-quality text datasets to train and test reliable NLP models for V-NLI.<br><br>Does not consider Semantics of the dataset to build relations. Real-time prompts, richer interactive widgets and multimodal interaction can enhance the system. |
| How to Ask What to Say?: Strategies for Evaluating Natural Language Interfaces for Data Visualization [12] | Benefits of jeopardy-style facts:<br><br>i. Engaging participants in solving the task.<br><br>ii. Measuring success is straightforward since the facts are known beforehand and tasks involve true/false responses.<br><br>iii. Facts can mimic realistic analytical findings.<br><br>Benefits of open-ended or scenario-based tasks:<br><br>i. Relatively straightforward to devise tasks.<br><br>ii. Helps assess overall system usability. | This article motivates further research not only in developing new visualization NLIs but also methods for effectively evaluating these systems and tracking research progress. Similar to recent developments in visualization authoring system evaluation, we see the identification of concrete metrics for measuring contribution and success as an immediate area for research in the context of visualization NLIs. Eventually, these new metrics can not only help validate research progress but may also lead to the formulation of new evaluation methods and strategies that are best suited given the evaluation goals. |
| Inclusive data visualization for people with disabilities: a call to action [14] | Provides guidelines to keep in mind when working on the project and building visualisation tools.<br><br>Point no 4 is relevant to our topic and is what makes our project unique. | - |
| Understanding Screen-Reader Users' Experiences with Online Data Visualizations [15] | Involves participants with vision impairments, hence helps us identify a new target audience.<br><br>Also highlights difficulties of using screen readers to understand visualisations, hence making the need for a voice-based data visualisation tool more crucial.<br><br>Makes suggestions in employing new techniques and strategies that would improve interaction. | Makes use of screen readers the participants were already using and does not explore their different types. Some screen readers would not face the mentioned issues. |

| | | |
|---|---|---|
| VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In [16] | Incorporates voice input from user to extract info from the visualisation.<br>Provides a basic summary of all the important points from a graph.<br>Conducted good user research and interviews for every feature developed to build what the users actually want. | Voice commands are only used to extract info from the chart, and not to create the chart or modify it.<br>Lack of personalization for the chart summary and sonification feature.<br>Does not provide information about the overall trend through the Summary mode. Such information may be useful for screen-reader users in navigating line graphs. |

### 2.1 Research Gaps

Currently, numerous data visualization tools exist that allow analysts to create and manipulate charts as well as statistical graphs to derive underlying patterns and useful insights from their data. However, these tools involve using extensive dashboards with screen controls which demand a solid understanding and skill set to perform even basic operations. Any naive or visually impaired user may lack convenience as well as ease of use. Other tools that do provide voice commands do not let the user retrieve information accurately and are delimited because of their high pricing structure and complex user-interface [5]. Thus, a need for a data visualization and summarization tool with improved interaction using voice technology is identified.

Additionally, upon analysing the tools that do cater to the needs of the visually impaired, it was identified that the systems had very limited scope regarding the creating of graphs and the customizations that could be achieved. Voice commands could only be used to extract information from the charts but not for modifying it.

Hence, what was identified from the survey conducted above was the need for a data visualization software that not only creates plots and graphs, but also considers the needs of amateur and blind users by providing seamless user interaction.

## 3. Methodology

A structured approach is followed to design and develop VUI. It involves a set of principles, tools, and techniques used to create effective VUI experiences. This process typically includes steps such as user research, designing conversation flows, building speech recognition and synthesis models, testing, and refining the VUI based on performance. By following a methodology, VUI designers and developers can ensure that their systems are user-friendly, functional, and meet the needs of their intended audience ultimately leading to better user experiences.

### 3.1 Survey

In order to gain insights about how users interact with existing visualization tools and voice-based systems, a structural survey was conducted with 540 responses from participants that were either visually impaired or new to these tools. The data obtained showed that about 80% of the participants used data visualization tools at least once a week, out of which 30% used them every other day. The main issue identified was the complexity of the graphical user interfaces of these tools. They do seek help from the technical support provided or by asking an experienced acquaintance. However, it causes the users to lose interest in using the tool. Some specific challenges faced by the users that lead to their frustration are that they are sometimes unable to perform the desired operation in the correct manner or they do not get the desired output in the way they would like to have. Even though it is important to the users to customize the visualization and configure the tool to their specific needs, this process is not simple for them. When the idea of a voice-based interface for creating visualizations was introduced, a majority of the responses were positive, especially from the visually impaired participants. Even though they face some difficulties like not knowing how to use the voice system, what command to say, or the system not being able to understand the commands, 82% of the participants were comfortable with using voice commands. It was also deduced that receiving feedback and alerts on the interface, in the form of voice or text, for confirmatory messages or to display errors would be helpful to proceed while working.

### 3.2 System Architecture

The above responses informed the design process for developing Vize: a voice-based data visualization tool with a clean and simple UI. There are several approaches to building voice-based data visualization. One way is to use a voice assistant, such as Amazon Alexa or Google Assistant, to interpret voice commands and generate visualizations. Another approach is to use a custom-built voice interface that is integrated into a web or mobile application. Regardless of the approach, there are a few key concepts that are involved in building voice-based data visualization:

a. **Speech recognition**: This component listens for voice commands and converts them into text that can be understood by a computer.
b. **Natural language processing (NLP):** This component interprets the text generated by the speech recognition component and identifies the intent behind the command.
c. **Data retrieval and processing:** Once the intent of the command is identified, the system needs to retrieve the appropriate data and process it into a format that can be visualized.
d. **Visualization:** Finally, the system needs to generate a visual representation of the data based on the user's command.

After conducting the survey, a web-based frontend was set up. The frontend serves as the user interface where users can interact with the application, while the backend is responsible for processing data and providing relevant information to the frontend. The frontend interface was developed using ReactJS, a JavaScript library and ChartJS, an open-source library for creating interactive and customizable charts and graphs. The frontend and backend must be compatible and communicate effectively with each other to ensure seamless integration. The backend technologies comprise the Django REST framework and SQLite database.

Additionally, the voice assistant was integrated with the application, ensuring that it can process voice commands and provide relevant responses. The voice assistant was created using the inbuilt react libraries 'react-speech-kit and react-speech-recognition', which are popular open-source libraries that provide speech recognition and synthesis capabilities to React applications. The user can provide voice commands to create new dashboards, charts and even have the ability to customize or edit these components to their own likings. Keywords are extracted from the text and the required function is executed by necessary API calls. By ensuring that all components are connected properly and communicate effectively, Vize is developed as a working tool that provides users with a seamless experience.

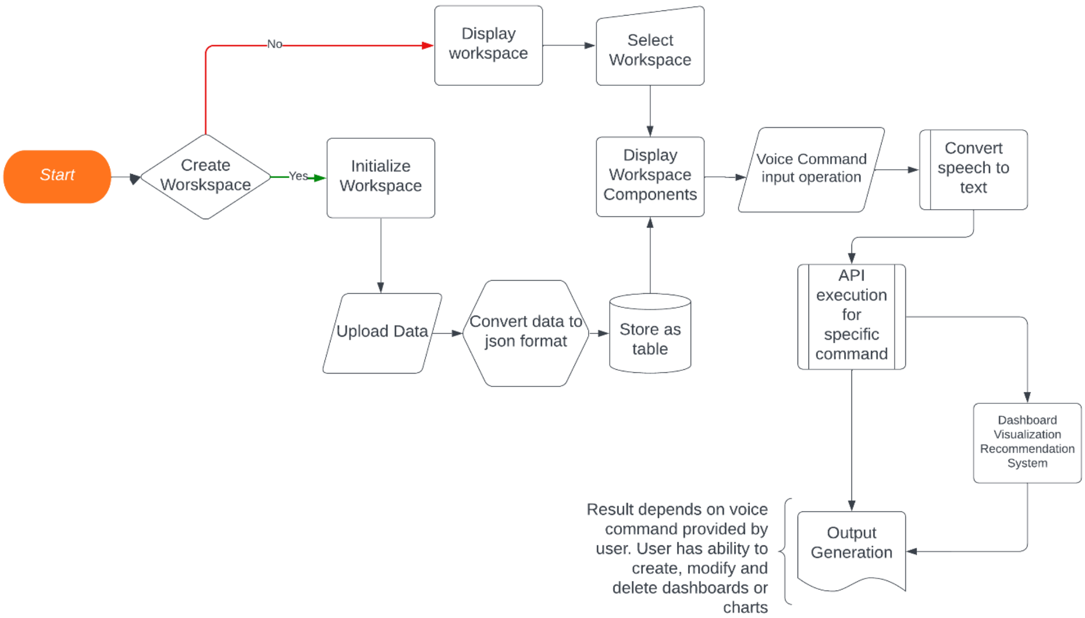


**Fig. 1.** System Architecture

***Fig 1*** shows how the system architecture was designed, thus explaining the internal functioning of the system and the steps that the system undertakes to generate the desired output. After logging into the VUI system the user can either create a new workspace or select an existing workspace. Upon creating a new workspace, the user is automatically directed to adding a CSV or JSON format database. The database is processed and converted to a standard JSON format using the python Pandas data processing library and saved in the backend.

Post creating a workspace, the system implements the Data Visualization Recommendation System (DVRS) which analyses the database to create automatically detect chart types based on input columns and proceeds to create a dashboard that showcases the visualizations that pass a set criterion [16]. The inner workings of the recommendation system have been explained in Fig. 3.

The system allows for voice-based input from the user, enabling the creation of new dashboards, charts, and customization of these components through natural language commands. The system extracts keywords from the spoken input using natural language processing techniques, which are then mapped to the appropriate functions through API calls to execute the desired actions. The resulting information is then used to generate the appropriate responses to the user's commands.

The user flow diagram shown in ***Fig 2*** below outlines the various functionalities of a voice-based tool. Upon registration, the user will be presented with an interactive tutorial to understand the tool's capabilities, which can also be accessed later from the workspace screen.

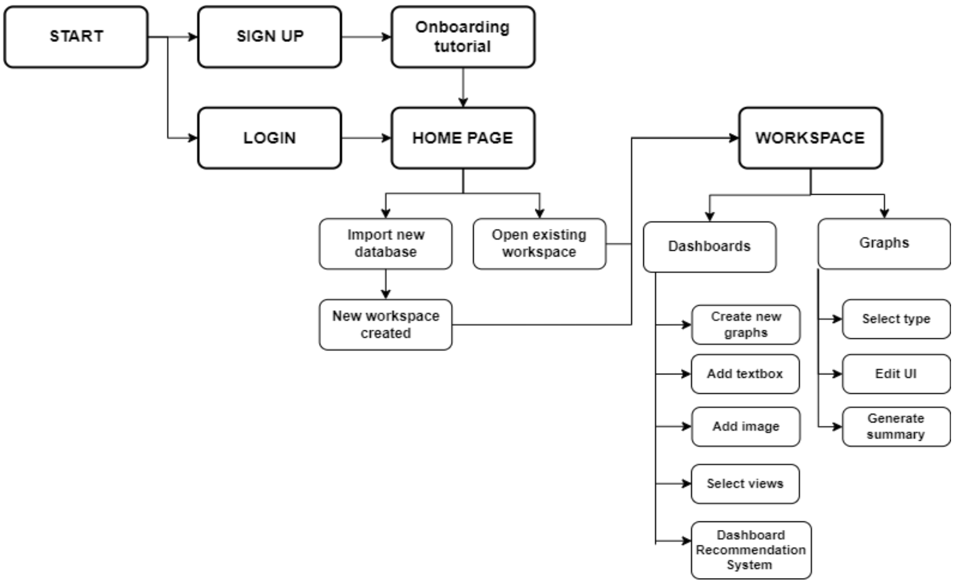


**Fig. 2.** User Flow and workspace operations through VUI

After registration, the user can upload data files to the workspace or access previously created ones from the homepage. In each workspace, the user can create multiple dashboards and graphs, which can be customized using voice commands to adjust properties. On uploading the database, the DVRS will run and suggest some charts to add in the dashboard which the user might find helpful. The user can decide to continue working with these charts or choose to delete them. Additionally, the tool can identify statistical insights from the graphs and generate summaries to enhance accessibility.

## 4. Implementation and Results

As technology continues to advance, there has been an increased interest in using voice-based interfaces to interact with any system or tool. In the field of data visualization, voice-based interfaces can provide a new and innovative way for users to interact with and explore data. By using natural language processing and speech recognition, users can now use their voice to query data, ask questions, and explore visualizations. This can provide a more intuitive and accessible approach to data analysis, making it easier for users to gain insights and make better decisions. In this context, the implementation of voice-based data visualization can be seen as a promising area of research and development, with potential applications in a wide range of fields including business intelligence, healthcare, and scientific research.

4.1 Data Visualization Recommendation System (DVRS)

4.1.1 Comparative Analysis of DVR Systems

The introduction of DVR Systems is comparatively new in the field of visualization and analysis. Hence, the existence of these systems and their workings have a limited scope. [19] SeeDB is an effective data visualization recommendation tool, whose approach to recommendation is based on a rigorous analysis of the data distribution and query workload, which makes it more reliable than other ad hoc approaches. Additionally, it makes use a metric called *deviation from reference* to judge the “interestingness” of a visualization and rank the visualizations accordingly. Therefore, SeeDB does not consider the interpretability or usefulness of the recommended visualizations, which could limit its usefulness in some scenarios where insights need to be communicated to non-technical stakeholders. Additionally, it does not support all types of visualizations, which could limit its applicability in some scenarios.

Another recommendation system, [17] VizML, proposes a machine learning-based approach for visualization recommendation, which leverages the power of data-driven techniques to provide personalized and context-aware recommendations. It takes into account various factors such as dataset characteristics, user preferences, and task requirements to generate visualization recommendations that align with users' needs and incorporates a deep learning model for learning the embeddings of visualizations and a recommendation model to predict the relevance of visualizations for users, providing a flexible and adaptable framework. Although, the effectiveness of this approach limited to specific datasets or domains, and its generalizability to diverse scenarios needs further investigation. Additionally, the usability of VizML in various real-world scenarios was not discussed appropriately.

Data2Vise is another tool for that generates visualization recommendations by using a Sequence-to-Sequence Recurrent Neural Network. [20] This particular tool has the ability to handle variable-length input sequences and generate diverse visualization designs that effectively represent the underlying data patterns. It enables users to provide high-level specifications and textual descriptions of desired visualizations, which are then automatically translated into complete and coherent visualization designs. What the tool lacks is its inability in handling ambiguous or incomplete input specifications and the resulting impact on the generated visualizations require further investigation. Moreover, the scalability of Data2Vis in terms of handling large datasets or complex visualization requirements needs further exploration.

**Table 2.** Comparative Analysis of DVR Systems

| Tool | Advantages | Limitations |
|---|---|---|
| SeeDB | Rigorous analysis of data distribution and query workload for reliable recommendations.<br>Uses the "deviation from reference" metric to judge interestingness and rank visualizations. | Does not consider interpretability or usefulness of recommended visualizations.<br>Does not support all types of visualizations.<br>Limited usefulness in scenarios requiring communication with non-technical stakeholders |
| VizML | Leverages data-driven techniques for personalized and context-aware recommendations.<br>Considers dataset characteristics, user preferences, and task requirements. Incorporates a deep learning model for visualization embeddings and a recommendation model. | Effectiveness limited to specific datasets or domains.<br>Generalizability to diverse scenarios needs further investigation.<br>Usability in various real-world scenarios not extensively discussed. |
| Data2Vis | Handles variable-length input sequences and generates diverse visualization designs.<br>Enables high-level specifications and automatic translation into coherent designs. | Difficulty in handling ambiguous or incomplete input specifications.<br>Scalability for large datasets or complex visualization requirements needs further exploration. |

The implementation shown below, depicts how the DVRS used by Vize learns from the above existing tool to create a custom recommendation system that works well for the use case and the target audience of our product.

4.1.2 Implementation

The Data Visualization Recommendation System is implemented after a workspace has been created by the user. The main idea behind the DVRS is to provide the user with an automated analytical dashboard to get a better understanding of the data at hand. The DVRS curates a dashboard by selecting all the charts that pass the specified criterion. ***Fig 3*** drawn below explains the internal functioning of the DVRS to arrive at a conclusion of which charts would be present on the automated Dashboard.

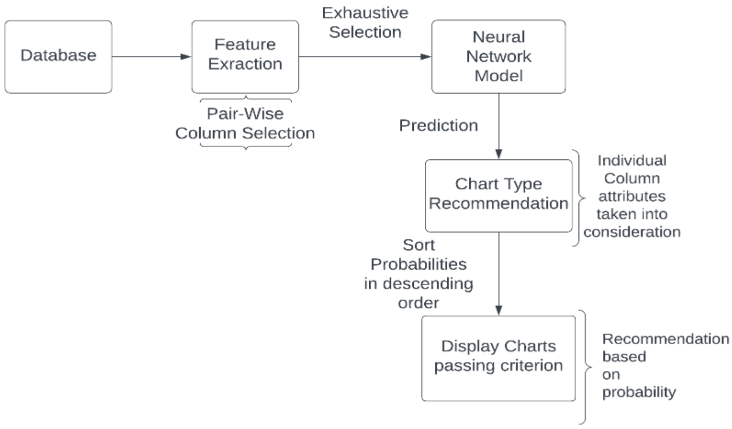


**Fig. 3.** Data Visualization Recommendation System

Once the user uploads the database a single column feature extraction procedure is conducted to eventually determine the pairwise column selections that can be fed as inputs to the neural network [16]. The neural network in use is a simple feed-forward neural network and makes use of the Sequential Model. The output of the neural network gives the chart type along with the probability of accuracy. After an exhaustive selection is carried out with all possible columns and chart types, the probabilities are sorted in descending order of accuracy. The charts which have a probability value above the passing criterion of 80% are chosen to be represented on the dashboard (in case the number of charts exceeds 6, then the top 6 are chosen).

Once the charts have been displayed on the dashboard the user has the flexibility to customize any chart or the dashboard itself or even delete any chart that they deem unfit. This allows the user to have maximum control over the analysis they want to conduct but save time in creating basic graphs to understand the data.

The DVRS is a powerful tool that can help users to make better decisions based on data. It has several advantages, including saving time and effort and providing a more comprehensive view of the data. With respect to our user, it bridges the gap between the user understanding the data and creating visualizations, enabling them to create other visualizations with the context that they have been provided with. However, the DVRS also has some disadvantages, including the possibility of creating inaccurate or incomplete dashboards, not being able to handle complex data sets, not being able to meet the specific needs of all users, and not being able to keep up with the latest data visualization trends. Overall, the DVRS can be helpful for users who are looking to make better decisions based on data, but it is important to be aware of its limitations before using it.

### 4.2 User Interface

The user interface has been designed in such a way as to make the user experience as easy and convenient as possible. The charts created are displayed in the form of a dashboard with soft, pastel colours and adequate white spacing. The design makes sure to provide all the features needed by the user without making the UI overly complex. Use of the Material UI library has been made.

#### 4.2.1 *User Workspaces*

The feature to turn on voice commands and access tutorials is provided on the home and workspace screens respectively. ***Fig 4*** displays the workspace interface where the user the ability to create a new workspace or view existing workspaces. Each workspace has 3 main components:

a. **Database**: Each workspace has one database. The user can filter data according to different columns and sort in the order they want. Additionally, the data can also be exported as csv.
b. **Dashboards:** Workspaces can have multiple dashboards that allow the end user to create a number of stories. Each dashboard has multiple charts and several widgets that can enhance its quality by adding texts, images or filter buttons.
c. **Charts:** Charts allows users to create the chart they want by selecting the chart type. Users can additionally customize each chart with respect to several options provided by the application to best visualize the data.

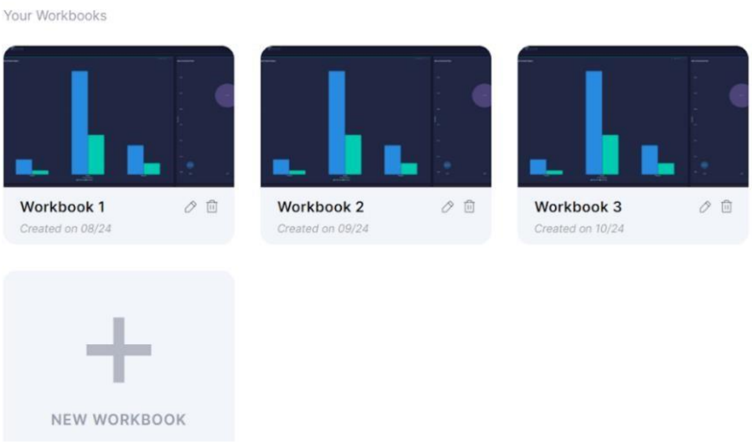


**Fig. 4.** Workspaces Created

4.2.2 Database

***Fig 5*** screen displays the database that has been imported on which the visualizations will be carried out. The data can be filtered according to the analysts' needs allowing them to generate distinct views of the database for visual understanding and analysis of data. The database displayed can be edited by filtering or selectively viewing columns. The density of the data displayed can also be changed and the edited database can be exported on the user's local machine.

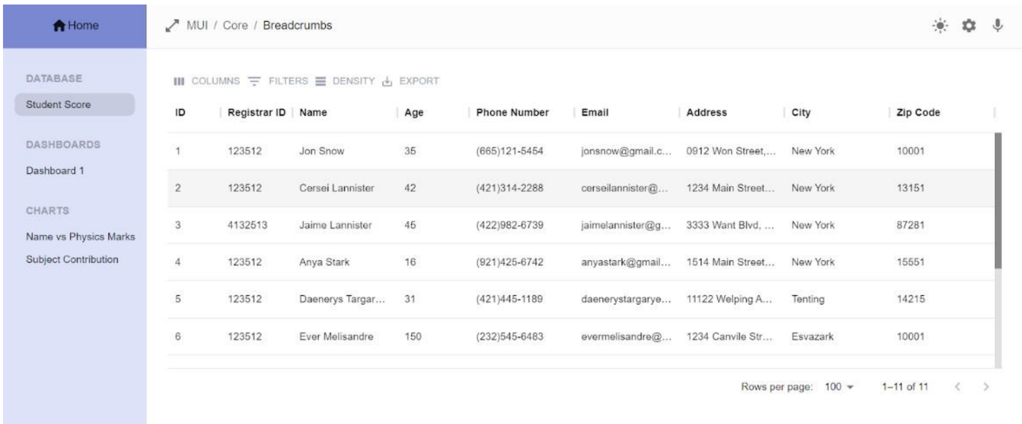


**Fig. 5.** Uploaded Database

4.2.3 *Dashboard*

Dashboards give analysts the ability to create stories that enable them to present and showcase their analysis in a creative, concise and apparent manner. The dashboard is automatically created by the DVRS when the user uploads the data file and the user receives a prompt as shown in ***Fig 6***. Suggested charts are added to the dashboard and an alert is prompted. These charts can be customized or deleted using voice commands if not needed. Along with the ability to add new charts on the dashboard, the user has the facility to add objects such as text boxes or images which can aid the visualizations for better analysis. ***Fig 7*** shows how all the voice commands are also listed to reduce memory load so that the user can refer to them whenever needed.

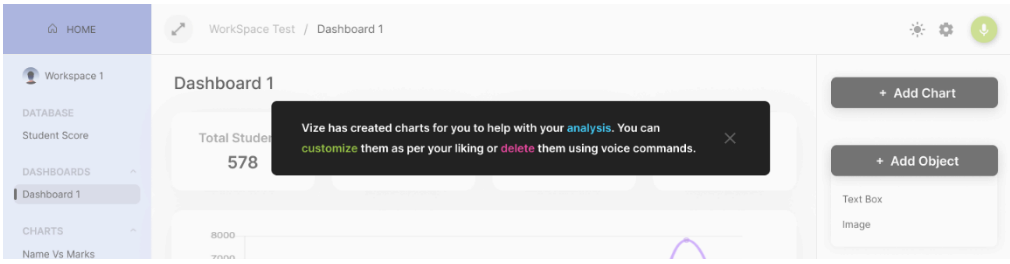


**Fig. 6.** Data Visualization Recommendation System Alert

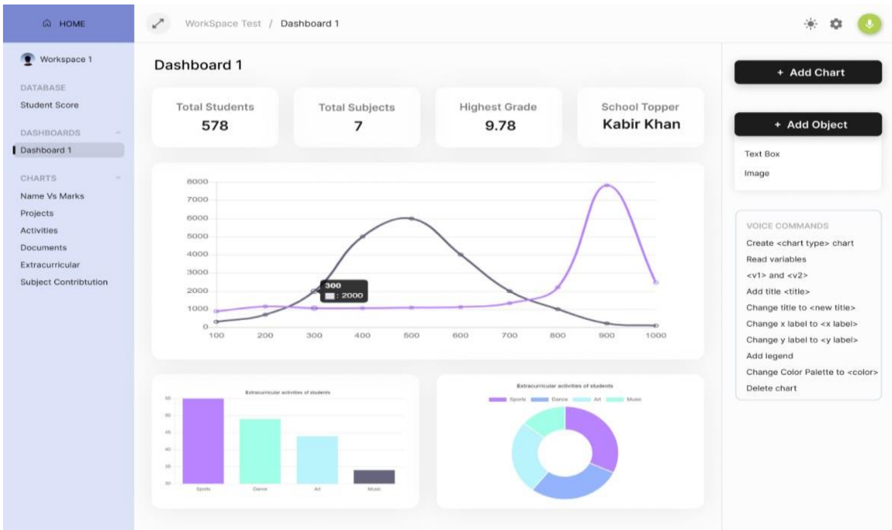


**Fig. 7.** Dashboard Screen

4.2.4 Charts

This is the screen where users can view each chart created individually. Each chart helps to provide a distinct analysis of the data that has been selected. The system allows users to select the chart they want to create and provide a list of data attributes that can be utilized to create that chart. A list of customizations that can be executed like adding titles or changing axes labels using voice commands is also provided. ***Fig 8*** also shows how the system curates and displays a summary to explain the statistical findings of the chart.

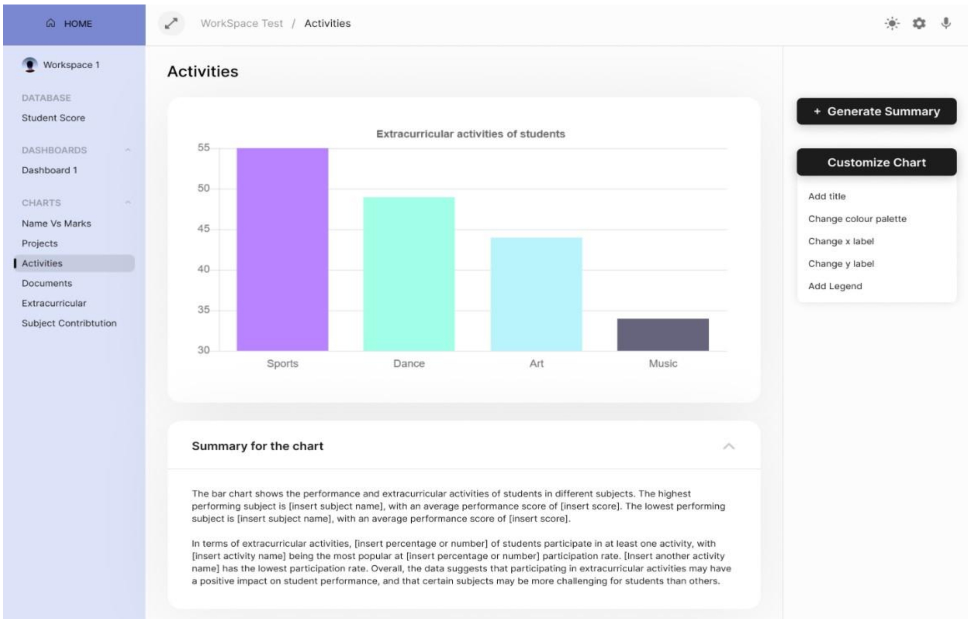


**Fig. 8. I**ndividual Chart Screen

### 4.3 Data Summarization

To aid users with visual impairments, Vize generates an accurate summarization of the chart in textual and numeric format which allows screen readers to provide a comprehensive voice output explaining the visualization. For numerical data, common summary statistics include measures of central tendency such as mean, median, and mode, as well as measures of variability such as variance and standard deviation. For categorical data, summary statistics include counts and proportions of each category in the chart or frequency tables displaying the distribution of the data. In addition, correlation coefficients is calculated to determine the strength and direction of the linear relationship between the two variables. Other information such as the maximum and minimum value and the range of data points is also mentioned.

### 4.4 Voice Commands

The user has the flexibility to perform numerous actions via voice commands as displayed in Table 1 below. Every action undertaken through clicking on the user interface can be directly translated to handsfree interaction by using any of the voice commands provided below, thus aiding users with visual impairments.

**Table 3.** Voice commands for Vize

| Command | Response | Action |
|---|---|---|
| Create <*chart type*> chart | Select variables for creating <chart type> chart | Opens a pop-up modal that has a list of all the variable names in the database |
| Read variables | Variables are: | Reads out all the variable names in the database |
| <*v1*> and <*v2*> | Creating <*chart type*> chart with variables <*v1*> and <*v2*> | Creates the desired chart and redirects to that individual chart page |
| Add title <*title*> | Adding title <*title*> | Adds desired title to the chart |
| Change title to <new title> | Changing title to <*new title*> | Changes the title of the chart |
| Change x label to <*new x label*> | Changing x-axis label to <new x label> | Changes the x axis label of the chart |
| Change y label to <*new y label*> | Changing y-axis label to <new y label> | Changes the y axis label of the chart |
| Add legend | Adding Legend to chart | Adds legend to chart |
| Change colour Palette to <colour> | Changes colour palette of chart to <colour> | Changes the colour palette of chart |
| Delete Chart <Chart Name> | Deletes chart <Chart Name> | Deletes desired chart |

### 4.5 Testing APIs

For this data visualization software, the Django REST framework has been used. API calls are used to retrieve data from an SQLite database and pass it to a visualization tool to create charts or dashboards using the JavaScript ChartJS library. The APIs allow the user to get the list of workspaces and its components (dashboards, charts, and database) and the flexibility to edit or create any new workspace or its components.

To make an API call to retrieve data for visualization, a client sends an HTTP GET request to the /database/ endpoint, The server would then retrieve the data from the database and return it to the client in a structured JSON format. To initiate an API call and retrieve data for visualization, the client sends an HTTP GET request to the /database/ endpoint. The server, which runs on the Django REST framework, receives the request and processes it. The server then accesses the SQLite database and retrieves the requested data, which is structured in a format that is suitable for data visualization. This data is typically returned to the client in a structured JSON format, which is a lightweight data interchange format that is easy for the client to parse and use. Once the client has received the data, it can be passed to the Chart.js library, which can be used to create charts or dashboards based on the data it received. The client might make additional API calls to retrieve additional data or update existing charts as needed to create more complex visualizations.

### 4.6 Usability Testing and Evaluation

Performance Testing was carried out to test the usability of the system. 42 participants were recruited to evaluate Vize, out of which 30 were novice users with no prior knowledge of data analysis and 12 were people with partial or complete blindness. Each participant was asked to perform a set of tasks as part of the experiment. These tasks included creating at least 3 new charts (other than the ones generated by the DVRS), performing subsets of the chart customization operations and generating a summary for at least 3 charts. These tasks were performed using the graphical user interface first and then by using voice commands. The participants with visual impairments made use of only the voice system. After the experiment, all the participants were asked to fill out feedback for each of the tasks based on a 5-point Likert scale ((1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree) [17]. A Customer Effort Score (CES) of 4 out of 5 was obtained, which denoted that users responded positively to Vize. 76% of the users preferred voice commands over the graphical interface to create and customize data visualizations. 92% of them agreed that the DVRS made valuable suggestions that sped up the process. When it came to data summarization, 10 out of the 12 visually challenged participants acknowledged that the summaries generated helped them analyze the charts better.

## 5. Conclusion and Future Scope

Dashboards have become a popular tool for businesses to analyse and share information visually with the abundance of data available. Thus, by presenting complex data in an intuitive format, dashboards help identify trends and insights quickly. Being able to create and extract information from data visualizations is becoming more and more important today. However, accessibility to data analytics should not be limited to specific demographics, and the integration of voice-based interaction with a simplistic user interface can expand the user base. To achieve this, Vize utilizes advanced technologies such as natural language processing and speech recognition, to enable users to interact with data visualizations through voice commands. Vize's user interface is designed with keeping simplicity and accessibility in mind, and by using a minimalist approach to make it easier for users to navigate and understand the data. The DVRS compliments the system's delivery to novice users by automatically curating a dashboard from the database to provide an initial analysis of the data thus making interpretation easier and saving time on creating basic charts. The summarization tool enables users with visual impairments to arrive at insights and draw up conclusions of the visualization. Overall, these advantages would enable Vize to expand the user base of data analytics by making it more accessible to any kind of user, regardless of technical proficiency or demographic. By integrating voice-based interaction and a simplistic user interface, Vize creates an intuitive and convenient data analytics tool that can be used by anyone, anywhere, anytime thus focusing on designing for inclusivity and accessibility.

Vize has the potential to offer several useful features in the future that could overcome the gaps the system currently has to mitigate the system's current limitations. One such feature is multilingual support, which would allow users to interact with the software in their preferred language using voice commands, making it easier for non-native speakers to use the software. The software could also provide additional chart customization options, allowing users to modify and personalize their charts to better meet their needs. Multiple database upload options could enable users to upload multiple databases and perform join operations, providing a more complete picture of the data. AI-based statistical summary generation could automatically analyse data and generate statistical summaries, which would be useful for users with large amounts of data. Additionally, a more accurate DVRS model would be able to capture more nuanced relationships between the variables in the dataset, which would provide users with a deeper understanding of the data. Finally, real-time data integration could allow users to connect their data sources to Vize and update their visualizations in real-time, keeping them up to date with the latest trends. These potential features could offer flexibility, control, and convenience to users, making data analysis and visualization easier and more efficient.

**Declarations**

**Conflict of interest**: The authors declare that they have no conflict of interest.

**Ethical approval**: This article does not contain any studies with human participants or animals performed by any of the authors.

**Availability of data and materials**: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

**Competing interests**: The authors declare that they have no competing interests.

**Funding**: This work has no funding resource.

**Author’s contributions**: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by all authors. The manuscript was written and revised by all authors on previous versions of the manuscript. All authors read and approved the final manuscript.

**Acknowledgement**: We thank the anonymous referees for their useful suggestions.

## References


[1] Diamond, M. and Mattia, A. (2017) Data Visualization: An exploratory study into the software tools used by businesses., Journal of Instructional Pedagogies. Academic and Business Research Institute. 147 Medjool Trail, Ponte Vedra, FL 32081. Tel: 904-435-4330; e-mail: editorial.staff@aabri.com; Web site: http://www.aabri.com. Available at: https://eric.ed.gov/?id=EJ1151731 (Accessed: February 28, 2023).

[2] Kaidi, Z. (no date) Data Visualization - University of Illinois Chicago. Available at: https://www.cs.uic.edu/~kzhao/Papers/00_course_Data_visualization.pdf (Accessed: February 28, 2023).

[3] Kendall, Linus & Chaudhuri, Bidisha & Bhalla, Apoorva. (2020). Understanding Technology as Situated Practice: Everyday use of Voice User Interfaces Among Diverse Groups of Users in Urban India. Information Systems Frontiers. 22. 585-605. 10.1007/s10796-020-10015-6.

[4] Shakila Cherise S Joyner, Amalia Riegelhuth, Kathleen Garrity, Yea-Seul Kim, and Nam Wook Kim. 2022.Visualization Accessibility in the Wild: Challenges Faced by Visualization Designers. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 83, 1–19. https://doi.org/10.1145/3491102.3517630.

[5] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA, Article 14, 1–16. https://doi.org/10.1145/3441852.3471202.

[6] Profunda Analytics + dundas bi (no date) Premium Reporting & Data Analytics - Dundas Data Visualization. Available at: https://www.dundas.com/resources/customer-stories/embedded/profunda-analytics-dundas-bi (Accessed: October 5, 2022).

[7] FlowSense (no date) VisFlow. Available at: https://visflow.org/flowsense/ (Accessed: October 5, 2022).

[8] Zierau, N., Hildebrand, C., Bergner, A. et al. Voice bots on the frontline: Voice-based interfaces enhance flowlike consumer experiences & boost service outcomes. J. of the Acad. Mark. Sci. (2022). https://doi.org/10.1007/s11747-022-00868-5.

[9] C. Murad, C. Munteanu, B. R. Cowan and L. Clark, "Revolution or Evolution? Speech Interaction and HCI Design Guidelines," in IEEE Pervasive Computing, vol. 18, no. 2, pp. 33-45, 1 April-June 2019, doi:10.1109/MPRV.2019.2906991.

[10] H. Guo, S. R. Gomez, C. Ziemkiewicz and D. H. Laidlaw, "A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights," in IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, pp. 51-60, 31 Jan. 2016, doi:10.1109/TVCG.2015.2467613.

[11] L. Shen et al., "Towards Natural Language Interfaces for Data Visualization: A Survey," in IEEE Transactions on Visualization and Computer Graphics, doi: 10.1109/TVCG.2022.3148007.

[12] Srinivasan, Arjun; Stasko, John; Keefe, Daniel F.; Tory, Melanie (2020). How to Ask What to Say?: Strategies for Evaluating Natural Language Interfaces for Data Visualization. IEEE Computer Graphics and Applications, 40(4), 96–103. doi:10.1109/MCG.2020.2986902.

[13] K. Padmanandam, S. P. V. D. S. Bheri, L. Vegesna and K. Sruthi, "A Speech Recognized Dynamic Word Cloud Visualization for Text Summarization," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 609-613, doi: 10.1109/ICICT50816.2021.9358693.

[14] Marriott, Kim, Bongshin Lee, Matthew Butler, Ed Cutrell, Kirsten Ellis, Cagatay Goncu, Marti Hearst, Kathleen McCoy, and Danielle Albers Szafir. "Inclusive data visualization for people with disabilities: a call to action." Interactions 28, no. 3 (2021): 47-51.

[15] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA, Article 14, 1–16. https://doi.org/10.1145/3441852.3471202

[16] Sharif, A., Wang, O.H., Muongchan, A.T., Reinecke, K. and Wobbrock, J.O., 2022, April. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In

CHI Conference on Human Factors in Computing Systems (pp. 1-19).

[17] Kevin Hu, Michiel A. Bakker, Stephen Li, Tim Kraska, and Cesar Hidalgo. 2019. VizML: A Machine Learning Approach to Visualization Recommendation. In *Proceedings of CHI Conference on Human Factors in Computing Systems (CHI '19), May 4–9, 2019, Glasgow, Scotland UK.* ACM, New York, NY, USA 21 Pages.

[18] 5-Point Likert Scale. In: Preedy, V.R., Watson, R.R. (eds) Handbook of Disease Burdens and Quality of Life Measures. Springer, New York, NY (2010). https://doi.org/10.1007/978-0-387-78665-0_6363.

[19] Vartak M, Rahman S, Madden S, Parameswaran A, Polyzotis N. SeeDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics. Proceedings VLDB Endowment. 2015 Sep;8(13):2182-2193. PMID: 26779379; PMCID: PMC4714568.

[20] V. Dibia and Ç. Demiralp, "Data2Vis: Automatic Generation of Data Visualizations Using Sequence-to-Sequence Recurrent Neural Networks," in IEEE Computer Graphics and Applications, vol. 39, no. 5, pp. 33-46, 1 Sept.-Oct. 2019, doi: 10.1109/MCG.2019.2924636.

[21] Anupong, W., Yi-Chia, L., Jagdish, M., Kumar, R., Selvam, P.D., Saravanakumar, R., Dhabliya, D. Hybrid distributed energy sources providing climate security to the agriculture environment and enhancing the yield (2022) Sustainable Energy Technologies and Assessments, 52, art. no. 102142

[22] Aoudni, Y., Donald, C., Farouk, A., Sahay, K.B., Babu, D.V., Tripathi, V., Dhabliya, D. Cloud security based attack detection using transductive learning integrated with Hidden Markov Model (2022) Pattern Recognition Letters, 157, pp. 16-26.