

# **Integrating Natural Language and Visualizations for Exploring Data on Smartwatch**

**Kaavya Varadarajan**

**A Thesis Submitted to the Faculty of Graduate Studies**

**In Partial Fulfillment of the Requirements**

**for the Degree of Master of Arts**

**Graduate Program in Information System and Technology**

**York University**

**Toronto, Ontario**

**April 2024**

**©Kaavya Varadarajan,2024**

## **Abstract**

Smartwatches are increasingly popular for collecting and exploring personal data, including health, stocks, and weather information. However, the use of micro-visualizations to present such data faces challenges due to limited screen size and interactivity. To address this problem, we propose integrating natural language (voice) with micro-visualizations (charts) to enhance user comprehension and insights. Leveraging a large language model like ChatGPT, we automatically summarize micro-visualizations and combine them with audio narrations and interactive visualizations to aid users in understanding the data. A user study with sixteen participants suggests that the combination of voice and charts results in superior accuracy, preference, and usefulness compared to presenting charts alone. This highlights the efficacy of integrating natural language with visualizations on smartwatches to improve user interaction and data comprehension.

## Acknowledgments

I want to thank god for his countless blessings.

I am indebted to my supervisor, Professor Enamul Hoque Prince, for his unwavering support and guidance throughout my graduate studies. His profound knowledge and expertise in human-computer interaction have been instrumental in shaping my thesis. His insightful feedback has been a cornerstone in completing my thesis report.

I want to express my deepest gratitude to my mother for bringing me into this world and providing abundant love, inspiration, and support. Words cannot express my love for you. I am also immensely grateful to my wonderful siblings, Preethi and Nandha, for always being by my side during good and bad times and motivating me to push harder. To my lifelines, Varnika, Inniyaah, and Shivanya, your constant cheering has strengthened me. I hope this thesis work inspires you to pursue a passion for education and research. I also want to thank Senthilnathan for his support.

I am grateful for my loving husband's encouragement and invaluable support, without which this could not be possible. Thank you for always being by my side and your unwavering commitment to everything you do, making our lives special. Your love, support, and dedication cannot be expressed in a few words.

Miss you appa.

Thank you all, Kaavya

# Contents

	<b>Page</b>
<b>Abstract</b>	<b>ii</b>
<b>Contents</b>	<b>iv</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Approach . . . . .	3
1.3 Scope . . . . .	3
1.4 Research Contribution . . . . .	4
1.5 Thesis Outline . . . . .	5
<b>2 Literature Review</b>	<b>6</b>
2.1 Data Visualization on Smartwatch . . . . .	6
2.2 Smartwatch Visualization Interaction . . . . .	7
2.3 Challenges in Smartwatch Visualization . . . . .	8
2.4 Smartwatch Visualization Types . . . . .	10
2.5 Chart Summarization . . . . .	13
2.6 Summary . . . . .	15
<b>3 Prototype Design</b>	<b>16</b>
3.1 Programming Language . . . . .	16
3.2 Integrated Development Environment . . . . .	18
3.3 Application Architecture and Design Component . . . . .	20

3.4	Application Integration with ChatGPT . . . . .	22
3.5	Chart Types . . . . .	25
3.5.1	Line Chart . . . . .	25
3.5.2	Bar Chart . . . . .	26
3.5.3	Multi-series Line Chart . . . . .	27
3.5.4	Donut Chart . . . . .	28
3.6	Interactive Features . . . . .	29
3.6.1	Text Highlight . . . . .	29
3.6.2	Tooltip . . . . .	31
3.6.3	Zoom . . . . .	32
3.7	Summary . . . . .	32
<b>4</b>	<b>User study</b>	<b>33</b>
4.1	Experiment Setup . . . . .	33
4.1.1	Training . . . . .	33
4.1.2	Experimental Trials . . . . .	33
4.1.3	Analysis . . . . .	34
4.1.4	Outliers and Exclusions . . . . .	34
4.1.5	Sample size . . . . .	34
4.1.6	Online User-study . . . . .	34
4.2	Method . . . . .	35
4.2.1	Study design . . . . .	35
4.2.2	Procedure . . . . .	36
4.3	Data Collection . . . . .	36
4.4	User Study implementation . . . . .	37
4.5	Summary . . . . .	39
<b>5</b>	<b>Analysis and Result</b>	<b>40</b>

5.1	Analysis of User Demographics . . . . .	40
5.2	Analysis of User Task . . . . .	43
5.2.1	Task Accuracy . . . . .	43
5.3	Statistical Analysis of Dependable Variables . . . . .	46
5.3.1	The Task Accuracy of the Design Components . . . . .	46
5.3.2	The Usefulness of the Components . . . . .	46
5.3.3	Descriptive data analysis of Usefulness and preferences . . . . .	47
5.3.4	Statistical Analysis of Usefulness . . . . .	47
5.4	Ranking Preference of the Components . . . . .	48
5.4.1	Statistical Analysis of Preferences . . . . .	48
5.4.2	User Feedback . . . . .	49
5.5	Summary . . . . .	50
<b>6</b>	<b>Conclusion And Future work</b>	<b>51</b>
6.1	Conclusion . . . . .	51
6.2	Design Consideration and Future Work . . . . .	52
6.2.1	Future Work . . . . .	53
	<b>Bibliography</b>	<b>54</b>
	<b>Appendices</b>	<b>61</b>
A	Appendix : Ethics approval . . . . .	61
B	Appendix : Consent form . . . . .	63
C	Appendix : Google form . . . . .	66

# List of Tables

- 1 Overview of studies related to visualization in smartwatch. . . . . 12
- 2 Sample SwiftUI code snippet . . . . . 17
- 3 Different types of questions and sample responses. . . . . 23
- 4 Text highlight implementation in array. . . . . 30
- 5 Overview of the study. . . . . 35
- 6 The Design components for the stock data . . . . . 38
- 7 Accuracy percentage . . . . . 44
- 8 Statistical test for Task Accuracy for Min/Max Values . . . . . 46
- 9 Statistical test for Task Accuracy for Trends . . . . . 46
- 10 Related-Samples Friedman’s Two-Way Analysis of Variance by Ranks . . . . . 47
- 11 Post-Hoc test for usefulness . . . . . 48
- 12 Statistical test for User Preference(One-Sample Chi-Square Test) . . . . . 49

# List of Figures

- 1 Examples of Smartwatch applications displaying data visualizations on weather, world clock, and sleep tracker . . . . . 2
- 2 Stock applications with text,number, icon, and chart components . . . . . 4
- 3 Smartwatch face with it’s representation types . . . . . 10
- 4 Smart watch running the WatchOS 9 . . . . . 16
- 5 WatchOS Simulator . . . . . 18
- 6 Application Architecture . . . . . 21
- 7 An example of generating chart summary for a given input chart . . . . . 21
- 8 The Menu page showing three alternative approaches . . . . . 22
- 9 Flow Diagram of the Request and Response of the ChatGPT Server . . . . . 24
- 10 Text to Voice Flow . . . . . 25
- 11 Line chart showing the trend of stock price . . . . . 26
- 12 Bar chart showing the Stock price of Apple . . . . . 27
- 13 Multi line Chart representing each stock data . . . . . 28
- 14 Donut chart representing each stock data . . . . . 29
- 15 Text highlights when Play button is pressed . . . . . 30
- 16 Tooltip in line chart which shows the accurate value . . . . . 31
- 17 How frequently the users use smartwatch . . . . . 41
- 18 Types of smartwatches used . . . . . 42
- 19 Applications used in Smartwatch . . . . . 43
- 20 Accuracy percentage related to trends . . . . . 45
- 21 Accuracy percentage related to min/max values . . . . . 45
- 22 Usefulness and Preference Average. Error bars indicate standard deviations. For usefulness, a higher is better. For preference, a lower score is better . . . . . 47

# 1 Introduction

## 1.1 Motivation

Smartwatches have surged in popularity as tools for gathering and exploring personal data, ranging from health metrics to stock updates and weather forecasts. They represent wrist-worn multi-functional devices that offer quick and convenient access to applications and information via Bluetooth/wireless connectivity [1]. Smartwatches have been seamlessly integrated into daily life, serving diverse purposes such as mobile payments, childcare and patient monitoring, schedule management, fitness tracking, reservations, and health activities such as step count and heartbeat monitoring [2].

Many smartwatch applications employ visualizations to present a wide array of data [3]. These visualizations enable users to effectively discern important patterns, trends, and outliers within complex datasets [4]. Common visualizations include charts (e.g., bars, lines), maps, and diagrams, offering valuable insights. Figure 1 illustrates the digital display of an Apple smartwatch, showcasing applications such as weather, world clock, stock monitor, and sleep tracker, all utilizing small-sized visualizations.

Despite potential numerous applications and benefits, Visualizing data on wearable devices such as smartwatches comes with several design challenges. A key challenge for visualizing data on smartwatches originates from the small display size [5, 6]. Traditional visualizations are tailored for larger computer displays, easily accommodating visual elements such as charts and graphs. In contrast, smartwatch displays are exceedingly small, making it challenging to display multiple data items effectively. For instance, in Figure 1 (right), a bar chart with many bars may make it difficult to discern individual values. Research indicates that display size significantly influences how information can be effectively presented [7], yet research on data visualization for small displays like smartwatches remains very limited.

Another challenge arises when the user wishes to interact with the visualization to explore, filter,



Figure 1: Examples of Smartwatch applications displaying data visualizations on weather, world clock, and sleep tracker

and zoom into datasets. Previous studies suggest that while interacting with data on the smartwatch, a key requirement is to answer users' questions on the go, which is currently not well-supported in smartwatches [8]. Mobile and wearable displays usually offer touch-based interaction, however, users commonly face the well-known 'fat-finger' problem, i.e., when selecting targets on a touch device with a finger and the targets are smaller than the size of the finger contact area, users do not know if they hit the desired target [9]. Though there are several ways to alleviate the fat finger problem on mobile devices, this could potentially become more challenging for users when the display sizes are even smaller, which is the case for smartwatches.

To address the above challenges, we aim to integrate natural language with small-scale visualizations to support users in exploring data on smartwatches. This thesis propose that integrating natural language with visualizations on smartwatches can improve user interaction and data comprehension. This is because natural language and visualizations are two powerful communication modalities of human communication and by using natural language (voice and text), we can facilitate effective comprehension of charts. Therefore, we focus on exploring the design spaces for integrating natural language and visualization on smartwatches.

## 1.2 Approach

To explore the design space of integrating natural language and visualizations, we address several research questions (RQ).

*RQ1: How can we automatically summarize and integrate natural language with small-scale visualizations?* To tackle this research question, we utilize automatic chart summarization techniques [10] to generate concise and informative captions. Specifically, we leverage a large language model (e.g., ChatGPT) to automatically summarize small-scale visualizations and integrate them with audio narrations and interactive visualizations to assist users in understanding the data. Detailed explanations of this technique are provided in Chapter 3.

*RQ2: How can we enhance interactions with small-scale visualizations?* To enrich user interaction with chart data, we integrated several interactive features into the iWatch App we developed, such as zoom in/out, selections, tooltips, etc. Users have the option to scroll through the text summary or listen to the audio narration of the chart summary to understand the context of the chart. Further details about these features are provided in Chapter 3.

*RQ3: How does the integration of Voice + Visualization affect task performance and user's subjective measures (e.g., usefulness, preference)?* To understand the usefulness of combining voice and small-scale visualizations, we conducted a user study in which smartwatch users utilized our iWatch app to explore stock data. Specifically, we aimed to determine if our approach enhances users' performance in tasks related to chart data accuracy. Additionally, we sought to understand if participants prefer such an approach of integrating voice and visualizations over using only voice or only visualization. We discuss the user study design in Chapter 4 and the analysis of results in Chapter 5.

## 1.3 Scope

In our research, we have chosen Apple smartwatches as the primary hardware device. The Apple smartwatch ecosystem stands out prominently due to the abundance of third-party apps available. Since September 2015, Apple's app store has boasted over 10,000 smartwatch apps [11], solidifying

Apple’s position as a leader in the smartwatch market. The accessibility of iOS development allows individuals with innovative ideas to publish applications for Apple smartwatches.

Our research focuses on developing a small application to visualize the historical data of select stocks. We have selected the stock domain due to its relevance and compelling use cases in smartwatch applications. Many users seek to access stock information on the go, especially as they manage their investment portfolios. Stock applications offer features such as real-time stock data, portfolio management, stock list management, and historical stock data analysis. These applications effectively visualize historical stock data using static charts, facilitating the understanding of trends. Specifically, our research targets applications that visualize historical stock data. Please refer to Figure 22 illustrating existing smartwatch stock applications.



Figure 2: Stock applications with text,number, icon, and chart components

## 1.4 Research Contribution

As a part of this thesis, we provide a new dimension to the research towards the visualization of smartwatch displays. We apply different data visualization techniques and integrate voice and text with the visualization. Adding the audio and visual cues, along with the visualization of the smartwatches, is one-of-a-kind research. In this thesis, we conduct a user study that focuses on finding the most effective and efficient design features that can be used for presenting data. This thesis will contribute to users’ understanding of the different visualization data displayed in the

charts, text, and voice formats. As a result, this thesis will offer a valuable contribution to both academic and practical development of wearable technology interfaces.

## **1.5 Thesis Outline**

The rest of the thesis is organized as follows. In Chapter 2, we present a literature review of data visualizations for mobile devices and natural language generation. Chapter 3 delves into a detailed discussion of the design, features, and functionalities of the application prototype. Next, in Chapter 4, we describe the details of the user study, including the different design spaces that we experiment with, the user characteristics that we test, and the tasks and procedures of the study. Then, Chapter 5 presents the results of the user study in detail, which helps us answer our crucial research questions. Finally, Chapter 6 concludes the thesis by summarizing the key findings of this thesis as well as several directions for future work.

## **2 Literature Review**

In this chapter, we review previous research on the visualization of smaller displays with the challenges and investigate the existing analysis with the smartwatch visualization. This chapter is organized to start with the smartwatch data visualization and different interactions of smartwatch visualization, challenges with the smartwatches, types of visualization inside the smartwatch and conclude this chapter.

### **2.1 Data Visualization on Smartwatch**

In the context of the smartwatch, visualization means the visual representation of data shown on the smartwatch screen. The smartwatch displays are typically around 3-4 cm wide or height. Some deviate from a standard rectangular form to a circular geometry, which is an exciting design constraint for visualization. Smartwatches are sufficiently large to convey information to individuals who are relatively close to the display using touch, buttons, or a digital crown on the devices, and sometimes through speech [12].

The smartwatch data visualization led to several use cases in the healthcare sector and the sports domain. The research by Van et al. explores the visualization of the vital signs data on smartwatches for health care monitoring [13]. An important feature of these Smartwatches is that they can be coupled with various other hardware or sensors easily and can be used to tackle complicated use case scenarios. Lu et al. discuss usages of smartwatches and wearable sensors in healthcare applications [14]. Regarding the sports activities related data visualization, a study was conducted to measure outdoor activity, and the data was visualized in the smartwatch. Alexander Schiewe and their team worked on real-time visualization for smartwatches that could assist runners in learning new running techniques [3]. The smartwatch data visualization is nascent, and each use case would address unexplored research topics. In our research, we focus on the domain of stock data and present visualization solutions for such data.

## 2.2 Smartwatch Visualization Interaction

The IOS for smartwatch allow interactions through touch, spatial, dial (Crown), gaze and voice. The interaction type in the smartwatches could be categorized into hand and hands-free. Touch, spatial and dial would require a physical touch of the user on the screen. The gaze and the voice do not require any physical touch. The touch interaction on the hardware was taken to the next level with the introduction of the iPhone in 2007. Most of the hardware devices required hardware devices as input devices before the inception of Apple's iPhone. The touch interaction can be categorized as simple or extended based on the amount of pressure you put on the screen. The extended pressure you put on the screen will be identified as the 3d touch. Programmatically, developers could develop use cases based on the touch types. Use cases have been based on touch pressure, twist angles, pan movements, and combinations to create touch interactions on smartwatches [15].

Wang et al. explored pressure-based touch interaction on the mobile devices for the 3D visualization, where a normal touch was used for axis transition and pressure touch for 3D rotation [16]. Another critical aspect of the touch interactions would be using multiple fingers or specific gestures that could trigger certain actions on the display. Multiple fingers could be used to zoom data on the displays or rotate any 3d object shown on the displays. A gesture such as tapping and double tapping could be leveraged to create an interactive user experience for the users. Touch Wave is an application that introduces the touch gesture in manipulating stacked area charts on tablets [17]. Touch Wave started establishing touch interactions such as tap, tap and hold, pinch, and swipe. The main advantage of the touch interaction lies in the versatility of gestures that can initiate different operations.

The complexity level can also be increased, and actions such as pulling to refresh or swiping can be achieved with touch interactions. Tangere systems were developed for interacting with various types of charts, including line charts, bar charts, parallel coordinates, and scatterplots [9]. One advantage of touch interactions is that they are not restricted to human fingers; special hardware such as a stylus or hardware pen can be used to perform touch interactions. The TouchPivot system combined a pen and touch interactions to support data exploration [18].

The crown on a smartwatch can be configured to perform specific actions, drawing inspiration from traditional watches where the crown can change the time and configure dates. Smartwatches enable the crown to perform actions such as zooming in and out and moving to the next page by rotating it clockwise or counterclockwise. The gaze interaction technique enables hands-free input and leverages eye movement [19]. Certain actions can be programmed into smartwatches based on eye movement, such as the display automatically waking when the user gazes at the screen.

Smartwatches can also interact based on voice input. The tiny microphone allows users to provide voice commands that translate into actions on the smartwatch, such as making phone calls and filtering data on the screen. The BezelGlide research explores a novel suite that uses bezel interaction on smartwatches to graph information [20]. Bezel interaction minimizes screen occlusion and reduces incorrect keyboard input, resulting in improved interaction techniques for smartwatches. The study's results helped develop the best interaction technique to minimize screen occlusion. Using the Bezel improved overall smartwatch interaction and mitigated issues like fat fingers and screen occlusion.

Although various interaction types exist for users to engage with smartwatches, they come with specific challenges, which we will discuss in the subsequent section. Voice interaction is another method users can interact with directly. InChorus is a multimodal interface that incorporates pen touch and speech to facilitate data exploration on tablet devices [21]. In this thesis, we focus on smartwatch as opposed to the larger tablet screen.

### **2.3 Challenges in Smartwatch Visualization**

The smartwatch has been incorporated into the user's day-to-day life. Use cases developed in healthcare, sports, etc., provide valuable insight into the users' activity. Once when technology becomes more accessible to the masses, it comes along with a set of challenges. In this section, we will illustrate the challenges related to smartwatches. The first challenge will be related to the smaller displays. The display ratio of the smartwatches will be five times smaller than the mobile

displays. Most software applications must be rewritten wholly or entirely refactored to fit on smaller displays such as a smartwatch. Another critical aspect would be to modify all the text and widgets developed for mobile devices.

The design principles of the larger displays are entirely different from those of, the smaller displays. But the challenges seen on the larger displays are the same for, the smaller ones. In the context of data visualization, the touch target size, spacing, and zoom are some of the problems a user could experience in using mobile devices. The same set of problems could be translated to the smartwatch as well. The ubiquitous nature of the smartwatch introduces challenges related to logistical, privacy, and ethical issues. Some of the logistical issues the users face are the user interfaces needing to be more friendly and configurable, battery utilization of the smartwatches, and connectivity and compatibility with smartphones.

The smartwatches come with watch faces that are easily configurable to the users. A person using an Apple smartwatch can choose between many watch faces that can depict up to eight different data types simultaneously [22]. If the user does not like the user interface, then the use of the smartwatches could be a great hurdle for the user, and he/she may abandon the product completely. The massive amount of data generated in modern times overwhelms and creates an information-overloaded nature. The increase in available information as well as the number of different channels and the often historical or contradictory information can help people to be more informed but also lead to information overload [23]. The aspect of the sensor and capabilities of smartwatches to collect personal data create privacy issues in a broad spectrum. Android framework supports 13 sensor types supported by the operating system [24]. These sensor types could be leveraged to collect personal data and provide self-reflection. The WatchOS also provides the same set of sensor types.

The exciting capabilities of watchOS let you bring the best of your app to the wrist with timely interactions that focus on the content users care about most. Take advantage of incredible performance, background modes, built-in altimeter capabilities, direct connections to accessories, and

more [25]. These sensors provide a new dimension to the data collection for the users. When it comes to wearable technology, most people are only worried about their privacy. This is because they do not want their data online; also, many people are concerned with what they put online for people to access [26].

Another major problem related to smartwatches is the accuracy of the health monitoring data. The heart rate monitoring with smartwatches might have discrepancies compared to the certified medical devices [27]. Though smartwatches are getting more accurate as time progresses, the sensors and the operating system in the smartwatches need to be appropriately calibrated. Thus, blindly believing these data might lead to some ethical concerns.

## 2.4 Smartwatch Visualization Types

A watch face is a user interface component that can be completely customized so that the user can show the data on the screen. When the watchOS boots up, the first screen appearing is the watch faces. There are a plethora of watch faces available on the market. This research paper surveyed whether the user added more data on the watch faces showing a significant amount or only a [3]. Examples of watch faces with various representation types such as charts, icons, and text are provided in Figure 3.

Figure 3: Smartwatch face with its representation types



Smartwatches can display various types of data including text, charts, and images. Text data types can convey information directly to users on the screen. Text can be manipulated to highlight critical

information, such as alerts, notifications, or health-related reports. Images can be presented as infographics to convey information visually. Combining text and images can enhance information delivery to users. Presentation mode significantly influences the quality and processing of information. Messages presented in both image and text formats were rated higher in quality [28].

Most health and fitness data on smartwatches are visualized using line charts. Visualizations of this data were generally preferred over purely text-based representations, and the preferred chart type for fitness bands and smartwatches was often the same. Each chart type was deployed for different uses and had other effects on the users. Bar charts had a very high effectiveness when the form factor of the display was wide [5]. The voice could be added as an extra add-on to the visualization.

Our research work is motivated by the existing research, which helps us understand data visualization in smartwatches. Table 1 provides an overview of the user study research using the smartwatch in data exploration. These studies are done in health and fitness, sports, sleep monitoring and heart rate monitoring.

Table 1: Overview of studies related to visualization in smartwatch.

Paper title	Research questions explored	Research outcomes
Investigating the Users' Preferences of Heart Rate Data Types and Visualizations on a Smartwatch.[29]	Q1: Which are the user-preferred heart rate data types (eight types) and visualization types (20 types) on smartwatches during tracking activities and Monitoring Body Condition(MBC) motivations?	A few insights: Text scores are more significant than an icon in average heart rate (AVG-HR) under MBC, and line chart scores are more significant than min-max under TA
Data-driven study of the preferences and effects of wearable device visualisation in sport [30]	Q1: Is plain text more helpful in understanding the data than data visualization? Q2:In general, which Visualization is more famous? Q3: Do smartwatches and fitness bands show any specific benefits in visual representation?	A1: To enhance the user's understanding, Visualization is more helpful A2: Area map is more popular among the participants A3: From the experiment perception, smartwatch visualizations are more proficient
Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches[31]	Q1: Which chart types are more preferred by the users: bar charts, donut charts, and radial bar charts with various data sizes such as 7, 12, and 24 data values?	A1: Donuts and bar charts are preferred when quickly grasping data.
Preferences and Effectiveness of Sleep Data Visualizations for Smartwatches and Fitness Bands[5]	Q1: To comprehend the interests of the users in sleep visualization? Q2: To find which sleep visualization is preferred and effective according to the users.	A1: The users are more interested in nightly sleep phase and weekly sleep period A2: From the results, the hypnogram is preferred for sleep phases, area charts are selected for monthly recaps and horizontal bars are used for the social comparison of data.
Data Representations for In-Situ Exploration of Health and Fitness Data[32]	Q1: would the users be interested in exploring the data in situ? Q2: What are the design requirements in in-situ fitness and health data exploration?	A1: The participants are interested in in-situ data exploration because it can help them achieve their fitness goals quickly. A2: From the findings, the heart rates, distance, speed, calories and step count are the most vital data types of in-situ experiments. The combinations of data types can also provide high performance in fitness data exploration.
Visualizing Information on Watch Faces: A Survey with Smartwatch Users[3]	Q1: What types of data do people use in their watch faces? Q2: Which form is data being represented? Q3: What can we visualize more in watch faces?	A1: The health and fitness category data are shown more in people-watch faces A2: Icon+Text is the most popular data type selected by the participants A3: Since the data were mismatched, this answer is left unexplored.

In contrast to the works discussed in this section, our work will examine different design spaces on smartwatches and analyze the effectiveness of these design spaces. The design components such as chart and voice and the combinations of these design components are used to find out which elements are effective in designing the smartwatch application in Apple Watch. We will provide extra add-on features such as tool tip and measure their usefulness along with the design components.

## **2.5 Chart Summarization**

Summarization is the method of compressing lengthy information, such as texts, documents, or articles. It helps to derive essential details, key points, and primary ideas from the text and displays them briefly. Thus, readers can grasp the details quickly without reading the full text. People often use visualizations such as bar charts, line charts, and pie charts to represent quantitative data. The charts help to make informed decisions based on the insights from the data. However, Gathering insight from the charts can take time and effort. To overcome the challenges, Kantharaj et al. [10] proposed an automatic chart summarization model that aims to understand the chart and summarise the vital takeaways from the charts in natural language.

In recent years, automated summarization tools have been produced in natural language processing using deep learning techniques such as transformer-based language models [33]. However, the most groundbreaking research in natural language processing was formed by combining the generative Pre-trained Transformers(GPT) with Large Language Models(LLM) called ChatGPT. ChatGPT contains a user-friendly interface with the conversational language model, which gained massive popularity by showing a remarkable ability to produce human-based answers in various conversational systems [34].

The initial communication with the chatbot has demonstrated its capability to answer the questions clearly and precisely [35]. Due to its advanced language generation dominion, Chat Generative Pre-trained Transformer(ChatGPT) has been used tremendously in real-world scenarios on a daily basis. Some of the applications which ChatGPT applied are as follows: Personal assistants, various

language translation, storytelling and creative writing, customer support and services, programming and code writing, generation of content, tutoring, and education [36].

In recent years, ChatGPT has been explored in various research areas related to summarization. Tariq et al. assessed ChatGPT's performance in text-based tasks, finding that it excelled in text simplification and summarization by delivering grammatically accurate content. The study highlighted ChatGPT's ability to succinctly summarize text while capturing key ideas and facts. However, ChatGPT's limitation lies in its need to comprehend context and make subjective decisions regarding essential information, sometimes resulting in irrelevant output due to differences between users and the machine learning model [37].

Tang et al. [38] investigated zero-shot prompt-based LLM usage limitations in medical summarization using ChatGPT models. Their research questioned whether LLMs produce better summaries for longer inputs. The findings revealed that a lower understanding of longer text negatively impacted ChatGPT's ability to generate relevant answers. Moreover, longer inputs increased the likelihood of producing misleading summaries and incorrect factual errors.

Savci and Das [33] compared human-generated summaries, ChatGPT-generated summaries, and BART-large model summaries, noting that ChatGPT-generated summaries contained more information, while human-generated summaries provided a more concise version. Yang et al. [39] found that ChatGPT's summarization capabilities surpassed those of traditional fine-tuning methods based on Rouge scores, as evaluated across various aspect and query-based text summarization benchmarks.

Tiwari et al. examined the integration of ChatGPT features into smartwatches, highlighting the benefits of leveraging ChatGPT in such devices. Voice command integration in smartwatches facilitates activities like weather monitoring, smart home device management, and appointment and reminder creation. Integrating voice commands and ChatGPT enhances personalized and quick user information retrieval for specific queries [40]. However, for optimal performance, interaction design and user interface considerations are crucial. The integration design should ensure that the chatbot seamlessly operates within the limited screen space of smartwatches to provide precise information.

Studies also indicate that voice-based interaction and chatbots enhance the practical usage of smartwatches, enabling users to efficiently communicate with chatbot functionalities. To achieve optimal design, chatbots should understand users' preferences and demands, employing suitable design techniques [40].

## **2.6 Summary**

In the literature section, the exploration begins with the visualization and evolution of the smartwatch. The subsequent section delves into smartwatch interactions, encompassing touch, spatial, dial (Crown), gaze, and voice interactions. The challenges encountered by smartwatches are identified, underscoring the necessity for additional solutions to mitigate these issues. Several researchers have conducted studies centered around visualization types in smartwatches, with emphasis placed on crucial design components. Lastly, the incorporation of large language models for summarization in smartwatches is reviewed.

### 3 Prototype Design

This Chapter presents the prototype we developed for this research project. We go over the wearable technology, hardware, and software used, along with the background information related to it.

The Apple watches come with a well-defined set of display sizes that would be easy to develop. Unlike the other wearable operating systems, the display type of the Apple wearable devices was well-defined and could be programmatically configured easily. The Apple smartwatches are powered with the help of the WatchOS operating system. WatchOS brings the best in app development and provides timely interaction with the user. The WatchOS 9 comes with all the latest functionality that would leverage the IOS functionality and provide an independent application experience on Apple Watch (see Figure 4) [41]. The smartwatch can efficiently process the actions such as filter, zoom, and button click. The Assistive touch feature would engage watchOS applications without having touched the watch.

Figure 4: Smart watch running the WatchOS 9



#### 3.1 Programming Language

Swift is a general-purpose programming language built using a modern approach to safety, performance, and software design patterns [42]. The Apple Ecosystem has used Swift as one of the official programming languages for development since 2014. The popularity of this language has skyrocketed in recent times. The Swift programming language offers simpler syntax and some

SwiftUI widgets	Code
Text	Text(“Hello World”)
Button	Button(“This is a button”)
Horizontal View	HStack
Vertical View	VStack

Table 2: Sample SwiftUI code snippet

constructions than common programming languages [12]. The most exciting aspect of the Swift language is the platform support. The programs could be cross-compiled and developed across different platforms. Apple made sure Swift would be completely open source and will be community-driven. Any enterprise or person with a feature or an idea could adapt Swift as a primary language and drive the development. Swift programming language is a statically typed language. The statically typed nature provides the advantage of catching as many errors as possible, enhancing reliability.

In our research, we used the Swift language to develop the components. We used the SwiftUI framework to develop our application on the watchOS. SwiftUI is an expressive UI framework development that could use less code and would follow a declarative way for the development of the application. The SwiftUI provides out of the box implementation of the storage, views, layouts, and Event handling. Table 2 shows the sample code snippet for the SwiftUI, and the code required for the implementation.

Each SwiftUI widget comes with a set of actions such as click, press, and long press. Layout types could be defined easily using the horizontal stack and the vertical stack options provided by the SwiftUI framework [43]. We will go over some of the basic widgets in this section.

**Text**– Text is the foremost basic view that could be used to show one or more lines on the screen. Along with the text the font, size, color, shades etc.

**TextField**- A text field is kind of an input box that accepts texts from the user.

**Layouts**– Three different types of layouts are currently supported. The horizontal, vertical, and the

zstack for the Z coordinate.

The key advantage of using the SwiftUI framework is the unified codebase. The components developed for the MacBook or the smartphone could be imported into the watchOS. The SwiftUI programming language gave us several sample libraries and implementations that had already implemented the visualization components inside their application.

### 3.2 Integrated Development Environment

Apple provides Xcode to developers to help them create applications for Macs, iPhones, and iPads (macOS and iOS). Xcode was used to create many of our favourite iOS and macOS applications. Arguably, without such a powerful, refined, integrated development environment (IDE), the thriving ecosystem that is the App Store would not exist as it does today [44]. The XCode comes with built-in simulators that can simulate different types of hardware, and the same code base can be deployed across multiple hardware (Refer to Figure 5). The essential advantage of using the XCode IDE would be the instant compilation of the code with an actual preview of the output. Unlike the other IDE, the code written could be previewed visually on the screen. In our research, we used different watchOS simulators and Xcode, ensuring that our developed charts could be deployed on all watchOS versions. The LLVM compiler built along with the XCode could tune the code to run quickly and in an optimized way on the hardware.

Figure 5: WatchOS Simulator



The XCode IDE provides robust support for the SwiftUI framework. XCode integrates well with the

SwiftUI framework, thus providing a smooth and enhanced development experience. The XCode has several unique features and tools tailored for the SwiftUI.

SwiftUI canvas: Xcode's tool will provide a live interactive preview of the SwiftUI code. This canvas feature will give real-time updates to visualize and iterate the UI. The real-time updates to the code change will allow developers to rapidly prototype your idea into code.

Code Assistance and Autocomplete: The Xcode comes with a built-in suggestion and an autocomplete feature that will drastically improve the coding efficiency.

SwiftUI Data Flow: The data flow in the SwiftUI framework is well supported with the XCode IDE. One could use the data flow with the help of `@State`, `@Binding`, `@ObservableObject` and `@EnvironmentalObject`.

SwiftUI App Life Cycle: The XCode supports the SwiftUI's app life cycle. The `@main` attribute makes the app's entry point into the application; thus, XCode will streamline the SwiftUI projects.

SwiftUI Animation: XCode supports the animation systems available in the SwiftUI framework. The animation components can create beautiful animations inside the SwiftUI applications.

SwiftUI Debugging: Xcode supports the robust feature of debugging the SwiftUI code. The breakpoints and inspect variables help resolve issues in the code.

Localization Support: XCode can provide the localization and internationalization for the SwiftUI projects. The localization tools are available inside the XCode IDE, which helps manage the localized content.

Library and Inspector: The Library and the Inspector feature inside the XCode IDE will provide a wide range of views, modifiers, and controls. The drag-and-drop components from the library can be pulled onto your code, and the inspector will allow you to modify the properties value for all the visual elements.

### 3.3 Application Architecture and Design Component

The design components that we show in our application include only Chart, Chart+Voice, and Chart+Voice. All the charts used in our research are the refactored version of the open-source charts developed by AppPear [45]. These charts were licensed under the MIT open-source license, and we made modifications to fit the watchOS.

We support a variety of chart types within our application. Along with the charts, our application supported text-to-voice and summary of the data with the help of the chatGPT. The text-to-voice feature is delivered with the use of the Text-to-language module that was built along with the WatchOS. The Text to Language module will input the text and then convert the content into a voice format. There are several ways to transform the voice's speed and modulation the hardware device generates. Refer figure [6] for the detailed module architecture. The chart module in the figure represents the chart information that needs to be visualised on the screen. The text/voice module help with the TTL logic, which is the build-in module. The HTTPS module helps in making API request to the back-end ChatGPT server and render the response.

The ChatGPT API provides a way to ask any question and fetch an answer for the question that we requested from the API. In our research, we visualize the stock data inside our watch. We developed a template for the question that will be asked in the ChatGPT. Based on the template question we requested, the appropriate server would generate the ChatGPT response. Refer to Figure 7 for the methodology which depicts the input of the template-based question and the response from the ChatGPT in the smartwatch application.

Our application has a menu-driven approach to navigate between different features and components. The first page of the application would prompt the user to pick design features or the design component. Refer to figure 8 for the menu page in the prototype.

We collected 5 days of historical data on popular stocks such as Microsoft, Meta, and Costco. This information is publicly available, and we display the price information for these stocks in the last 5 days. The application will store the stock information and dynamically fetch it inside the application.

Figure 6: Application Architecture

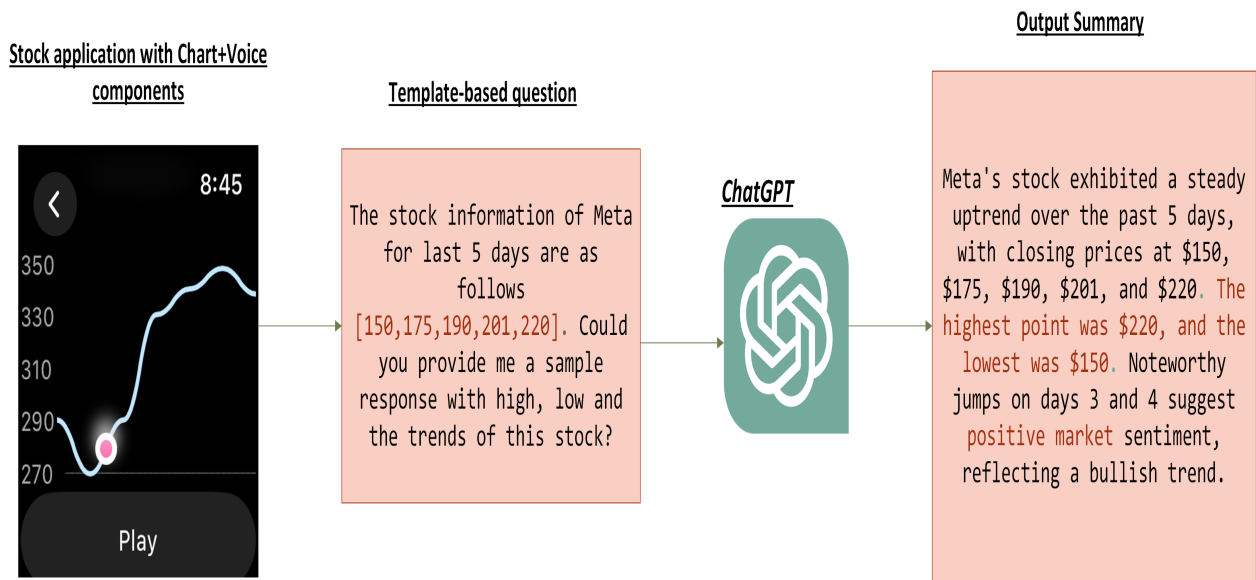
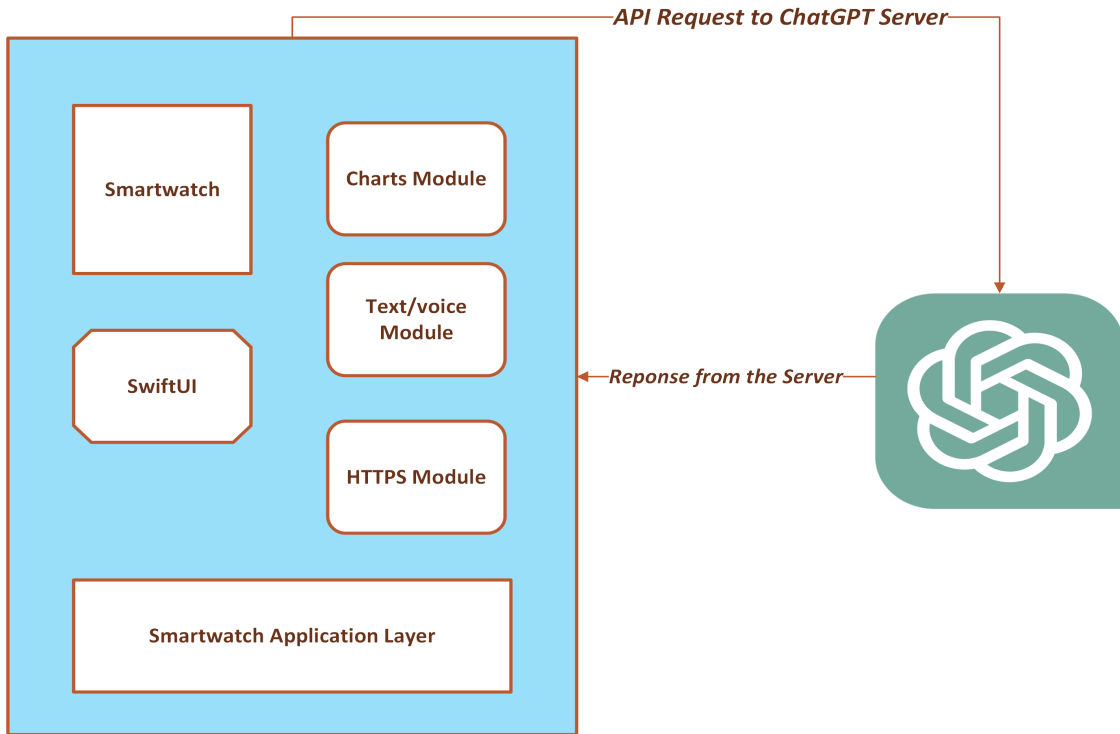
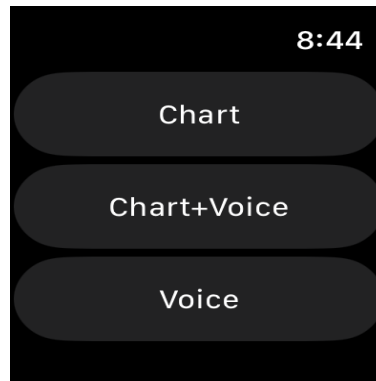


Figure 7: An example of generating chart summary for a given input chart

In the design feature, we have voice and graph. The user can pick the feature type to interpret the

Figure 8: The Menu page showing three alternative approaches



smart watch’s data. The same set of information would be passed to all the features. The user will be evaluated to rate which design feature was easy to understand. The line chart would be combined with the different types of features to create design components such as a chart with voice, only voice and only chart. The user will have the complete freedom to navigate between different design components. All the data used in the design components are stored in a central repository so that the pieces of information are uniform across various screens.

### **3.4 Application Integration with ChatGPT**

Our application integrates the ChatGPT API to perform text summarization. Based on our stock data information, we frame a template-based question and construct a question that could be requested via the API to the ChatGPT server. The template question that we designed will prompt the ChatGPT server for the stock’s high, low, and trends. We tried four different templates inside our Smartwatch application, and the response can be found in Table 3.

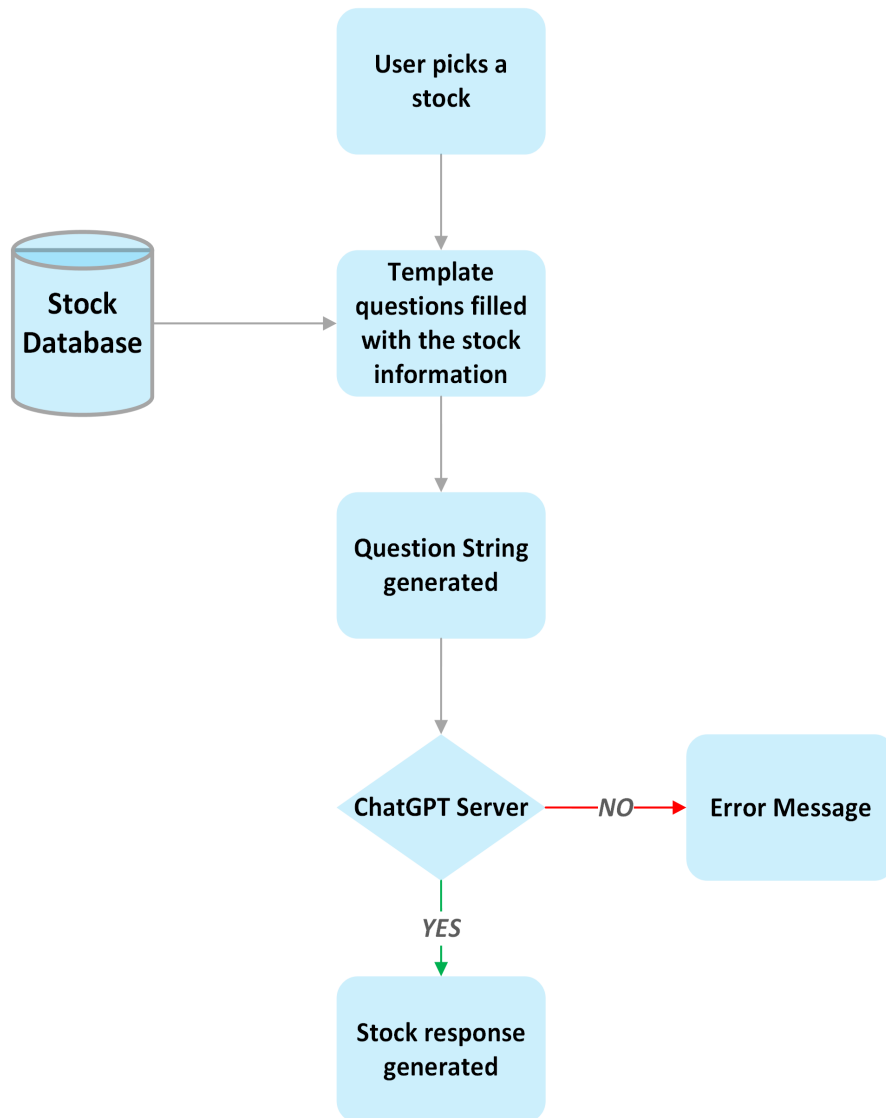
Given the display size constraint of smartwatch, we focus on generating short summaries with specific critical information for stocks such as extrema and trend. As such, we designed the following prompt: “The Stock Information of *Company Name* for the last five days are as follows [day1, day2, day3, day 4, day 5]. Could you please provide me a sample response with high, low and the trends of this stock”. The *Company name* and the stock data array will be text-processed to create the correct question string. Once the application processes the question string, the request

Table 3: Different types of questions and sample responses.

Questions	Sample response
<p>The stock information of Meta for last 5 days are as follows [150,175,190,201,220]. Could you provide me a sample response with high, low and the trends of this stock.</p>	<p>Meta's stock exhibited a steady uptrend over the past 5 days, with closing prices at \$150, \$175, \$190, \$201, and \$220. The highest point was \$220, and the lowest was \$150. Noteworthy jumps on days 3 and 4 suggest positive market sentiment, reflecting a bullish trend.</p>
<p>The stock information of Google for last 5 days are as follows [100,112,119,122,115]. Could you provide me a sample response with high, low and the trends of this stock.</p>	<p>Google's stock showed a mixed trend over the last 5 days, closing at \$100, \$112, \$119, \$122, and \$115. The highest point was \$122, and the lowest was \$100. A notable increase on day 4 was followed by a decline on the last day, indicating fluctuating market dynamics with an overall positive trend.</p>
<p>The stock information of Microsoft for last 5 days are as follows [350,375,390,401,405]. Could you provide me a sample response with high, low and the trends of this stock.</p>	<p>Microsoft's stock trended positively over the last 5 days, closing at \$350, \$375, \$390, \$401, and \$405. The highest point was \$405, and the lowest was \$350. With a steady increase in closing prices, the trend reflects strong market confidence and a bullish trajectory for Microsoft during this period.</p>

will be transferred to the ChatGPT server. The ChatGPT server will require a developer token, the question, and the API request. This request gets processed, and the responses are generated from the server. The sample responses can be found in Table 3. The sample responses are not unique and will vary slightly on each request. The entire request and response flow can be found in Figure 9.

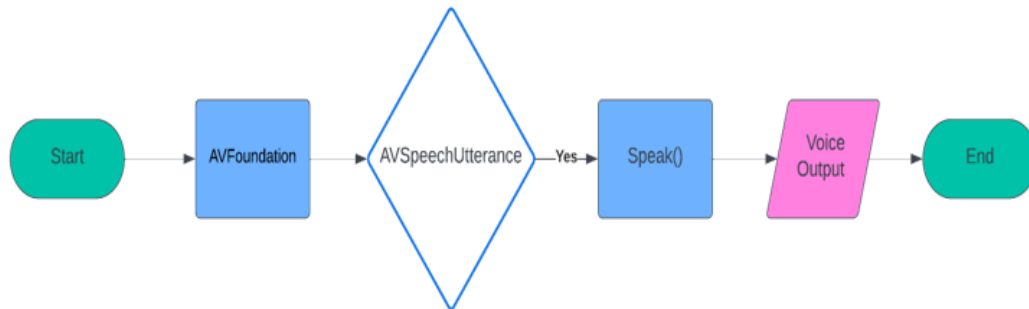
Figure 9: Flow Diagram of the Request and Response of the ChatGPT Server



The request and the response from the ChatGPT server will require the smartwatch application to be actively connected to the internet. When the request or response fails, a default error message will be shown on the screen. The text response that we received from the ChatGPT server can be displayed on the screen as output or played as a sound by the WatchOS. The text-to-voice in the WatchOS is supported with the help of the AVFoundation library. The AVFoundation library is part of the SwiftUI framework. The AVFoundation library will provide the AVSpeechUtterance class that controls the Voice, Speed, Pitch, Delay and Volume. An instance of the AVSpeechUtterance

is created inside the page that plays the voice, and the text information is provided as input. The language type, speed, pitch, and delay could be set as a setting. The AVSpeechUtterance has a function called speak. After all the required settings are provided, the speak function can be triggered, and the WatchOS will play the audio.

Figure 10: Text to Voice Flow



### 3.5 Chart Types

In this research study, we developed four different chart types: Line charts, Bar Charts, Multiline Charts, and Donut charts. The evaluation was done with the help of the line chart in this study. These charts were extended from an open-source implementation that was available for the iPhone and the Apple tablets, and modification was needed to fit these charts to accommodate the WatchOS. The chart library is publicly available on GitHub (<https://github.com/AppPear/ChartView>) and is released under the MIT license. The source code had to be slightly modified to support the WatchOS.

#### 3.5.1 Line Chart

Line charts can be used to visualize data that changes over time. This type of chart can be used to compare data across different categories if the categories are arranged in a logical manner. Some of the practical use cases where the line charts could be deployed are trend analysis, comparing data sets, performance tracking, and monitoring data. The Line charts are valued for their simplicity and clarity. The simplicity of the chart will make the user understand the complex information, identify

the trends, and make informed decisions.

Figure 11: Line chart showing the trend of stock price



We carefully designed the line charts to adjust for the smartwatch display by removing less important information that is not crucial to users. For example, for stock data x-axis labels and tick marks are not crucial as users already know that the x-axis represents time and therefore, we removed it from the design. The left of the x-axis represent Monday and the ends represent Friday which are hidden. Refer to Figure 11 for the line chart in the prototype application. In the user study, the line chart with the tooltip feature represents the chart components. The colour feature (red or green) can be utilized in the line to represent the negative or positive trends for the stock data.

### 3.5.2 Bar Chart

A bar chart is a graphical representation used to display and compare discrete categories of data through rectangular bars, where the length or height of each bar is proportional to the frequency or value of the corresponding type [46]. The bar charts can be oriented both horizontally and vertically. The length or height of the Bar chart will be proportional to the data value. Bar charts are mainly used to visualize sales, financial and healthcare data.

Figure 12: Bar chart showing the Stock price of Apple



We do not display x-axis and y-axis values to fit the chart in a small display. Refer to Figure 12 for the bar chart in the prototype application. Thus, the designed bar chart shows the corresponding values when the users hover over each bar, and the value is displayed above the chart. Depending on the bar sizes, the chart can display up to 12 to 14 bars on the screens, representing days or months depending on the application's requirement. Meanwhile, the x-axis is directly replaced by the values shown in the bar.

### 3.5.3 Multi-series Line Chart

Multi-series charts combine multiple series of data items in a single chart. A multi-series line graph shows the relationship between independent and dependent values of multiple sets of data [47]. The multi-series charts plot several lines, each representing a distinct data series against a common axis. In order to fit the data of multi-series line chart, we removed both the x and y axis and showed the line labels using the same color as the corresponding lines. Refer to Figure 13 for the multi-series line chart in the prototype application.

Figure 13: Multi line Chart representing each stock data



Multi-lines can be used in smartwatches to compare a maximum of three or four lines. Though the exact data values are difficult to map, comparing different lines representing different datasets can help understand the high-level data quickly and decrease the glance-ability time. For instance, the multi-line chart designed in our prototype represents each stock, such as Google, Meta, and Microsoft, with green representing Google, orange representing Microsoft, and purple representing Meta. The corresponding colours are represented in line charts, which help us simultaneously compare the stock price trends.

### 3.5.4 Donut Chart

Donut charts are a variation of Pie charts. The donut charts have a blank center that gives them the appearance of a donut. A donut chart is created simply by making a hole into a pie chart [48]. donut charts are used extensively in the media. The donut charts are visually appealing and have a higher advantage by enhancing the readability of the data. The circular design of the donut charts clearly explains the relationships inside the dataset. The circle is segmented into slices so that the viewers quickly understand the proportional value each category contributes to the dataset. The donut chart for the prototype code is inspired from [49] for the phone, which is changed accordingly to implement in the smartwatch.

Figure 14: Donut chart representing each stock data



The donut charts provide a wide range of customization compared to other charts. The ring, colour code and labels inside the donut charts can be changed and supply a wide variation in the charts. Refer to Figure 14 for the donut chart in the prototype application. The donut chart designed in our prototype shows the values of different stocks in percentages, where each arc represents each stock corresponding with the stock colours. The donut chart also provides the sum of the stocks selected in the centre. A Donut chart is the ideal option to create the user's portfolio in the smartwatch for different applications, which helps the user view the data on the go.

### 3.6 Interactive Features

In our prototype application, we have two interactive features to help the user interpret the chart. They are text highlights and tooltips. The text highlight feature is coupled with the voice and the tooltip and becomes visible when the user selects lines on the chart.

#### 3.6.1 Text Highlight

The text highlights feature is used to highlight each sentence when the voice narration is played for that sentence. The text highlight implementation is done with the help of the Swift UI text highlight.

Figure 15: Text highlights when Play button is pressed



Consider the sample text is “Microsoft’s stock trended positively over the last 5 days, closing at \$350, \$375, \$390, \$401, and \$405. The highest point was \$405, and the lowest was \$350. With a steady increase in closing prices, the trend reflects strong market confidence and a bullish trajectory for Microsoft during this period.”

Table 4: Text highlight implementation in array.

Array	Text
Element 1	If the sample text is “Microsoft’s stock trended positively over the last 5 days, closing at \$350, \$375, \$390, \$401, and \$405.
Element 2	The highest point was \$405, and the lowest was \$350. With a steady increase in closing prices, the trend reflects strong market confidence and a bullish trajectory for Microsoft during this period

The above sample text has two sentences; thus, the array length would be two. The Arrays in the prototype will be automatically resized based on the number of sentences in the text. When the user triggers the play button inside the prototype, the text colour of each sentence will be changed to a different colour and will have a direct sync with the Audio.

### 3.6.2 Tooltip

A tooltip is a simple add-on feature implemented for our prototype's line graph types. The tooltip would be triggered when the line graph would be clicked. The tooltip can show the appropriate data of the chart based on the area of the line we will click. The tooltip can provide additional information about a data point or segment when the user clicks or hovers over the graph with their finger. The main advantage of using a tooltip would be the detailed offering of the insights without cluttering the graph.

Figure 16: Tooltip in line chart which shows the accurate value



The advantage of this tooltip lies in its adaptability; it can be finely tuned in content, style, and visual appeal to meet the unique demands of the audience or to highlight particular data nuances. This customization ensures that the data is presented and communicated in a way that resonates with the viewer, making complex information accessible and understandable.

Tooltips can display precise values, percentages, or supplementary details about the data points on the line graph, such as exact sales figures, dates, percentages, or comparisons with previous data points. Tooltips can be customized in terms of content, format, and appearance. This allows the presentation of data to be tailored to the audience's specific needs or to emphasize aspects of the data.

### **3.6.3 Zoom**

We leveraged the Apple Watch's built-in zoom feature to perform the zoom operation on different screens. The zoom option for the Apple Watch defaults to false. The Zoom feature in the smartwatch can be enabled with the help of the settings application. The primary goal of Zoom is to provide the option to magnify the chart and see the information more clearly. The zoom actions zoom in or out, move around, and adjust magnification. The double tap on the screen will provide the zoom in and out of the screen. A single tap can pan around the page, which can be done with the help of the fingers or the crown.

## **3.7 Summary**

We developed the prototype application design to integrate the visualization and voice feature by adding ChatGPT to generate the summary. The overview of this prototype application, including the programming language and integrated development environment, is discussed in detail. The following two sections talk about the application architecture and the process of integrating the ChatGPT in the prototype. Lastly, the design components, such as chart types and special features, are discussed.

## **4 User study**

This section discusses the user study setup, methods, data collection, implementation and user feedback.

The main goal of the user study is to answer the three research questions:

- 1) Will voice + visualization be preferable to users compared to voice only or visualization only?
- 2) Will the voice +visualization more useful to users compared to voice only or visualization only?
- 3) What is the highest accuracy rate per trial for the design components from the user task?

The following chapter contains the user study implementation, methods and data collection.

### **4.1 Experiment Setup**

#### **4.1.1 Training**

The user study was conducted in person for all the participants. The users were asked to wear the watch when looking at the data. After exploring the application from the smartwatch, the Google form was sent to their respective email id to answer the questions. The answers were collected, statistical tests were performed to check the hypothesis, and results were drawn.

After exploring the application in the watch, participants could use the given laptop or their smart-phones to access the Google form. Two forms were created, namely form A and form B, which contain the same questions but in different order to counterbalance the study.

#### **4.1.2 Experimental Trials**

Participants answered ten questions using the given user task questions. We collected the answers for each question and the time taken to answer each question during this user study. The timer was used to calculate the time taken for the Participant to answer the question. After the Participant finished the user task, a few more questions are asked to find the usefulness and preference of the user in design components. Lastly, the post-study interview is done to ask additional questions for

future findings.

### **4.1.3 Analysis**

The study focused on finding the most effective and efficient design features that can be used for presenting data. Likert scale scoring measures the Participant's opinions from the post-study questionnaires. Along with this study, additional research was done to determine the user's understanding of the data displayed in chart and voice format by measuring the dependency variables.

### **4.1.4 Outliers and Exclusions**

The participants were excluded if they reported problems with the consent form or did not finish the complete experiment.

### **4.1.5 Sample size**

Sixteen users participated in the user study. Each Participant was asked to answer the questionnaire only once. Participants were recruited by advertising the user study, with a token of appreciation for the Participant 20\$

### **4.1.6 Online User-study**

The study was conducted online through the smartwatch simulator, and the users were asked to fill out the Google form simultaneously. There are three participants ( two males, and one female). All the users were frequent smartwatch users. The participants rated the chart+voice component more useful than just the chat and voice. From the user task, all the users answered the chart + voice questions correctly compared to the other tasks. Similarly, chart +voice components were preferred more when compared to the other two components. The informal study helped us to revise it for the final in-person user study. For example, we opted for multiple-choice questions instead of short answers as the short answers were not always relevant.

## 4.2 Method

### 4.2.1 Study design

The user study was a 3 x 2 within-subject design with the following independent variables. The data representation was chart-only, voice-only and chart +voice. The chart type is a line chart, and there are two tasks performed: 1) Extrema (min and max) and 2) Trends(directions). So overall, there were six trials per participant, and the counterbalancing was done with the randomization method. The overview of study variables and trials are as follows.

Table 5 The overview of study variables and trials are as follows

Task	User survey
Number of Participants	16
Design	Within Subject
Design Elements	Only Voice, Only Chart and Voice+Chart
Number of Trials	6

Table 5: Overview of the study.

The study was a within-subject because the same participants are subjected to different conditions. Compared to the between-subject design, the within-subject helps to increase the efficiency and statistical power. The efficiency is improved by having the same set of participants subjected to different conditions, which helps to save time and resources. In general, statistical power is more in within-subject design because each Participant acts as their own control. (i.e. fewer participants are required to find the significance).

The independent variable is the design components with the following conditions: Voice-only, Chart-only and Voice+Chart). The dependent variables are measured from the response to the user survey. The dependent variables are:

1. Task Accuracy per trial and task.

2. Subjective ranking on preference
3. Usefulness

#### **4.2.2 Procedure**

1. The consent form was explained to the Participant before the experimental session began.
2. The Google form containing the survey questions was sent to the users. The participants completed a pre-study background questionnaire, where demographic information (age range, gender), educational background, and experience with smartwatches were asked.
3. The demo was given to each Participant with the Apple smartwatch that explained the application's functionalities.
4. Afterward, the participants were given 5 minutes to become familiar with the smartwatch application, and if they had any questions, clarifications are provided instantly.
5. The participants were asked to perform a set of tasks using the smartwatch by exploring the different types of design components in the smartwatch application to complete the user task section in the Google form.
6. The post-study questionnaire was then completed based on their experience with the applications, and went through a brief interview with a few additional questions about the usefulness of the smartwatch application.

#### **4.3 Data Collection**

We collected 5 days of historical data on popular stocks such as Microsoft, Meta, and Costco. This information is publicly available, and we display the price information for these stocks in the last 5 days. The application will store the stock information and dynamically fetch it inside the application.

We conducted the user survey by asking the participants to complete the Google form containing the questions about our study. Later, the data from the survey was analyzed, and findings are made.

## 4.4 User Study implementation

The user study was implemented with four sections in each form. The following is a detailed explanation of each section.

**Section1:** In section one, the users can provide their valid email ID, which helps to validate the user's presence.

**Section 2:** In section two, the user demographic details are collected. These details include the user's age, gender, occupation, and which mobile OS they use. A few other questions that were asked in regards to their smartwatches are what type of smartwatches they are using, how often they use their smartwatch, how often they check the stock data, and which watch applications they use often in their smartwatch. These details help us to understand the user's knowledge and practical usage of smartwatches in their lives.

**Section 3 :** In section three, the user task questions are presented; with the help of design components, the users can answer the provided questions. The user task questions that are asked are as follows,

Chart-only:

Q1: Could you locate the minimum value of Costco's stock in this line chart?

Q2: What does the line chart reveal about the direction of Costco's stock?

Chart+voice:

Q3: Could you locate the peak value of Microsoft's stock and voice components in this line chart?

Q4: What do the line chart and Voice reveal about the direction of Microsoft's stock?

Voice-only:

Q5: Could you locate the peak value of Meta's stock from the voice components?

Q6: What does the voice feature reveal about the direction of Meta's stock?

The answers gathered from these questions help to find the accuracy of the design components. The Chart is the most efficient design component in finding the lowest value, but in finding the trends, it

is the least influential component (7 wrong answers). Chart + Voice effectively finds the min/max and trends where ten users find the answers correctly. Voice components have three wrong answers in trends and three in value findings.

No.	Design components	Screenshot
1	Chart	
2	Voice and chart	
3	Voice	

Table 6: The Design components for the stock data

**Section 4:** Section four contains the post-study questions to find the usefulness and user preferences of the design components. The questions in this section are as follows,

1. On a scale of 1(least useful) to 5 (most useful), how useful is the voice component in conveying the data?
2. On a scale of 1(least useful) to 5 (most useful), how useful are the voice and chart components combined in conveying the data?
3. On a scale of 1(least useful) to 5 (most useful), how useful is the chart component in conveying the data?
4. Rank the following interface on the application based on your preference, with one being the most preferred and three being the least preferred.

The answers to these questions are analyzed, and the results help to conclude our study.

**Section 5:** The Users were asked the following four questions after the survey:

- 1)Which component do they find it useful and preferred?
- 2)Is there any specific task for which they find these components effective?
- 3) How helpful is the tooltip in the Chart?
- 4)What can be added to these design components for future development?

## **4.5 Summary**

The user study's primary goal is to find the design components' usefulness, preferences and accuracy. Thus, The detailed version of the experimental setup, such as training, trials, sample size, and instruments, are discussed. The following section illustrates study design, procedure, and data collection, and finally, user study implementation and feedback data are collected.

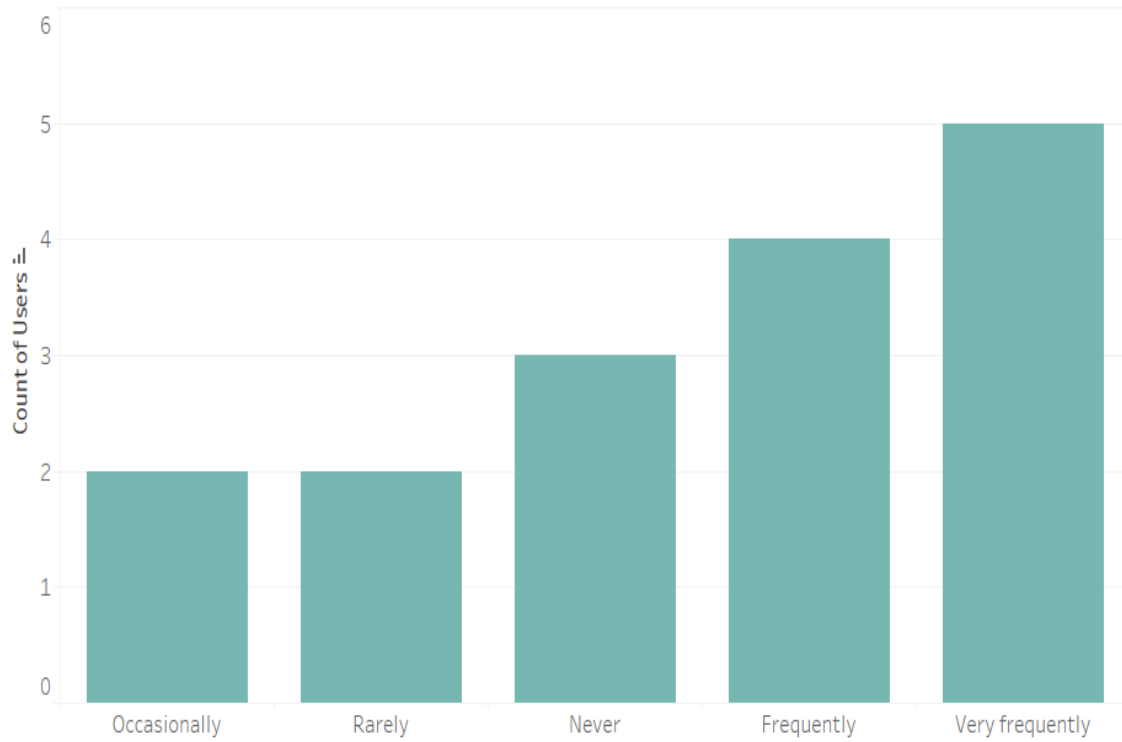
## **5 Analysis and Result**

In this section, the data collected from the user study is analyzed, and the results are presented. Three measures are taken into account to find which design components are most efficient and effective to understand the data. Those measures are Usefulness, preferences, and task accuracy. The measures taken to select the design components help us to understand what and how to use the design components in data exploration in smartwatches. We also analyzed the user summary and feedback from the data collected. The Tableau Software is used to create the dashboards, and the IBM SPSS Statistics is used for the statistical testing.

### **5.1 Analysis of User Demographics**

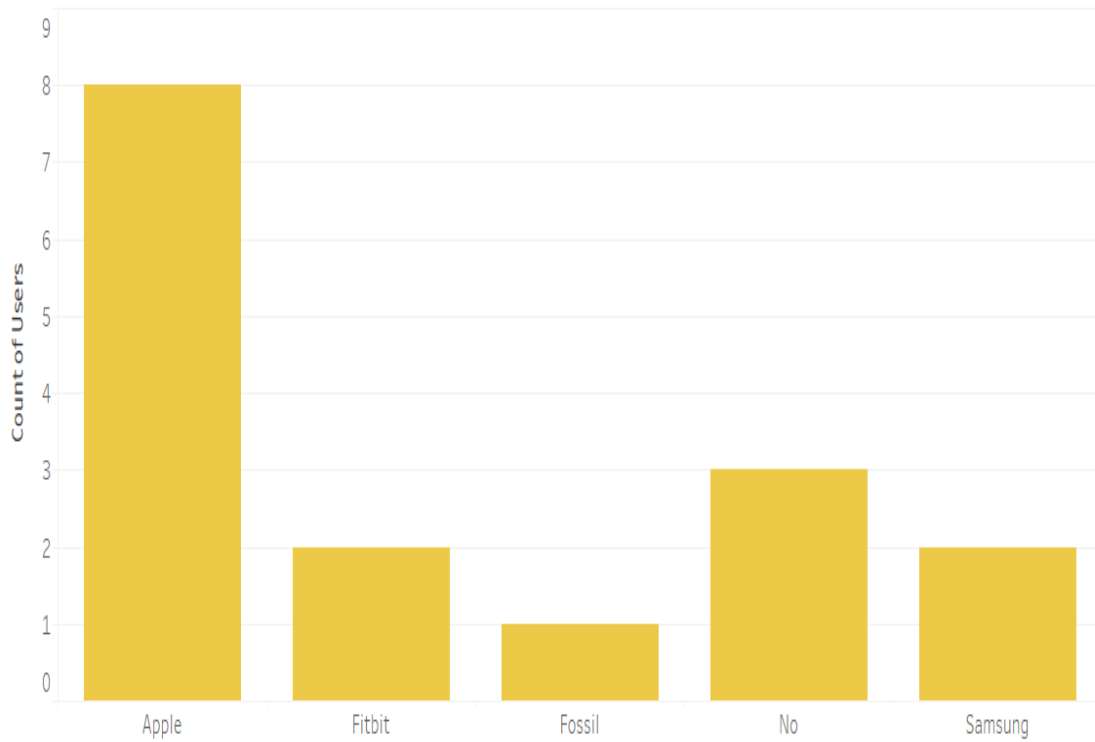
The user demographic details are collected to study more about the general characteristics of the small population in smartwatch usage. The survey takes place with eight female and eight male participants. The participants are mostly aged 31 to 40; a few are aged from 21 to 30, and one is above 50. Eight participants have IT-related occupations; two are students, three are unemployed, and the others are dentists, demand planners, and architects. Among these participants, four users have Android phones, and twelve use iPhones. Most participants actively use the smartwatches, while the three are non-smartwatch users. Out of 16, five people frequently check the stock-related data in their smartwatch. Refer to Figure [17] for the frequency usage.

Figure 17: How frequently the users use smartwatch



The types of smartwatches used by the user are also collected. Eight people use Apple smartwatches, two others use Samsung and Fitbit, and one uses Fossils. Refer figure [18] for the type of smartwatches used.

Figure 18: Types of smartwatches used



Users' most used applications in smartwatches are seven fitness apps. Refer to figure [19]: four use the weather app, five use the stock app, and two music and one message app. Many people use multiple apps in their smartwatches to gather their insights on data

Figure 19: Applications used in Smartwatch



## 5.2 Analysis of User Task

The user task is designed in our application prototype to measure the task accuracy when the user interacts with data using our application. For these purpose, three stock datasets were used: Costco, Microsoft, and Meta.

The two tasks used to assign by using these stocks are to find the trends and the peak or low values of the stock. Microsoft scored the highest number of correct answers in both the trends and the min/max values detection. Microsoft represents the chart and voice combined design components.

### 5.2.1 Task Accuracy

Accuracy is one of the critical usability measures that helps to find the efficiency of the design components. The average accuracy is calculated by dividing the number of correct responses by the total number of tasks[50].

We present our accuracy in percentages, so all the derived answers are converted to percentages. Thus, from the above, the calculated values for each design component are as follows,

Design components	Stock name	Accuracy %
Only Chart	Costco	Trends 37.5%
Only Chart	Costco	Value 75%
Voice + Chart	Microsoft	Trends 81.25%
Voice + Chart	Microsoft	Value 87.5%
Only Voice	Meta	Trends 62.5%
Only Voice	Meta	Value 75%

Table 7: Accuracy percentage

From the above Table, The highest accuracy percentage is executed in finding the value of Microsoft stock with the voice+chart design components(87.5%). The lowest accuracy percentage comes from the only chart component in finding the trends of the Costco stock(37.5%). From these findings, we observed that the Voice + Chart components have a higher accuracy rate in finding the answers correctly than the other design components.

Figure 20 and Figure 21 help us to understand the accuracy percentages from both trends and values more precisely and efficiently between the design components.

Figure 20: Accuracy percentage related to trends

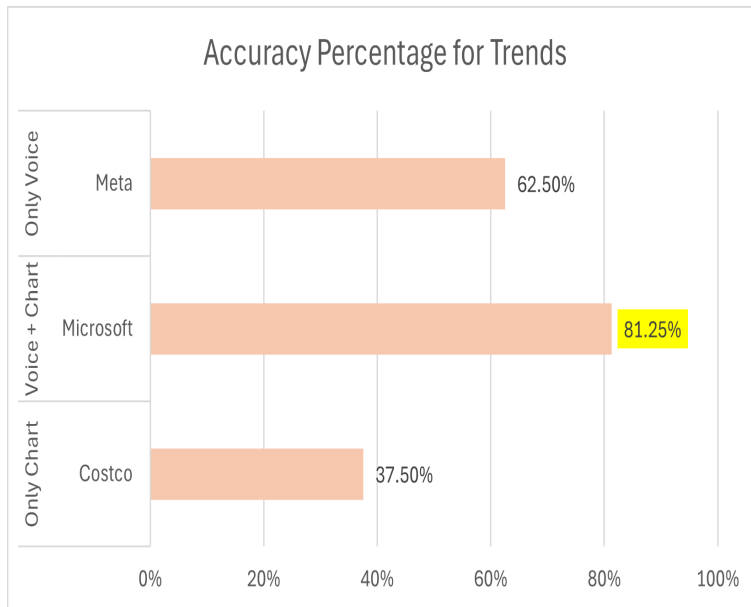
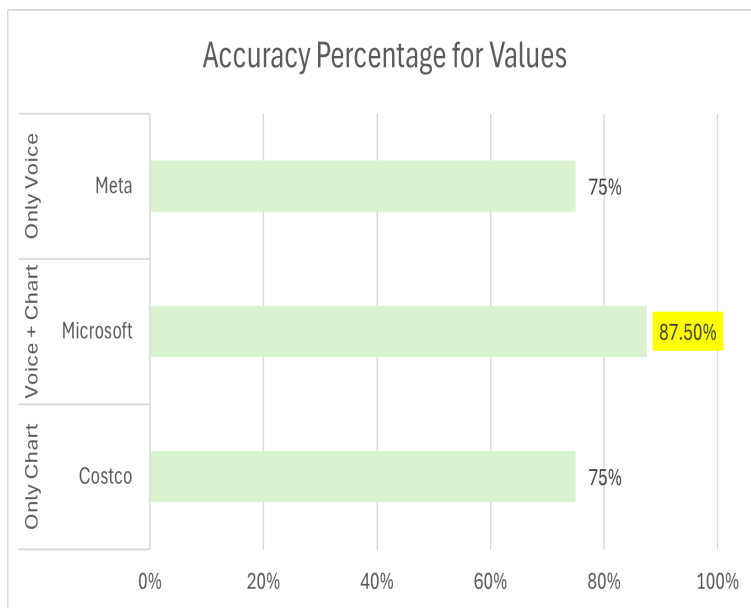


Figure 21: Accuracy percentage related to min/max values



### 5.3 Statistical Analysis of Dependable Variables

#### 5.3.1 The Task Accuracy of the Design Components

We converted the task accuracy data into categorical data, where the correct answers are named 'Correct,' and the incorrect answers are named 'Wrong.' For the categorical data, we ran the One-Sample Chi-Square Test. We ran the same test two times for the design components with the minimum or maximum values and trends.

Null Hypothesis(H0)	Significance (sig)
The categories of Chats occur with equal probabilities.	0.046(Reject)
The categories of Chat+Voice occur with equal probabilities.	0.003(Reject)
The categories of Voice occur with equal probabilities.	0.046(Reject)

Table 8: Statistical test for Task Accuracy for Min/Max Values

Null Hypothesis(H0)	Significance (sig)
The categories of only chart occur with equal probabilities.	0.317(Retain)
The categories of Voice+Chart occur with equal probabilities.	0.012(Reject)
The categories of only Voice occur with equal probabilities.	0.317(Retain)

Table 9: Statistical test for Task Accuracy for Trends

If the significance value is less than 0.05, the null hypothesis is rejected, and there is a significant difference in the expected and the observed frequencies.

#### 5.3.2 The Usefulness of the Components

In the post-study questionnaire, We asked the users to rate the least to most useful design components(1 being the least and five being the most). After the user task, the user selects the preferences after getting familiar with the application. Twelve users rated that voice + chart as the most useful,

followed by only chart and the least useful only voice. However, the chart and only voice are close in comparison.

### 5.3.3 Descriptive data analysis of Usefulness and preferences

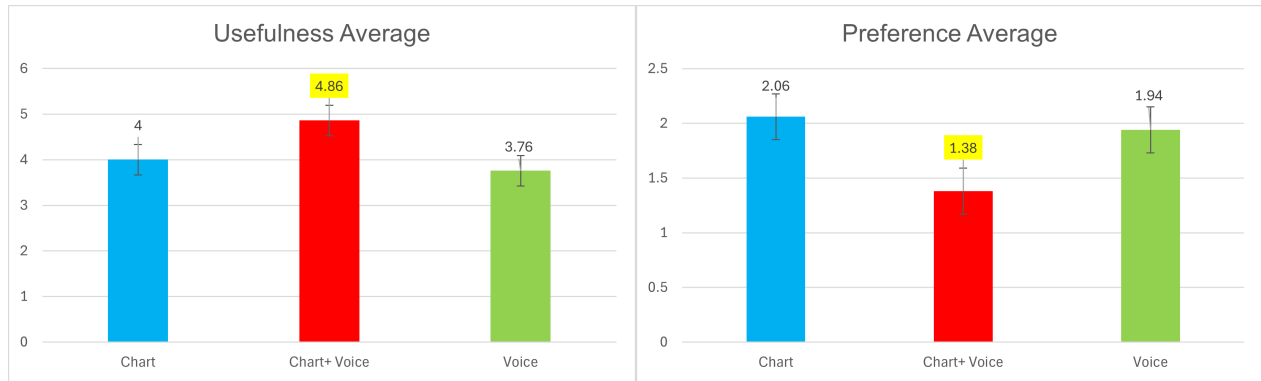


Figure 22: Usefulness and Preference Average. Error bars indicate standard deviations. For usefulness, a higher is better. For preference, a lower score is better

The chart and voice combined have a good average in usefulness and preferences. However, the average of chart-only and voice-only components in usefulness and preferences changes, with close values in comparison. Some might prefer voice only in contrast to the usefulness of only voice components.

### 5.3.4 Statistical Analysis of Usefulness

We ran the Friedman Test to do the statistical analysis. Since the data are non-parametric and taken from the same individuals, they fall under related samples. Thus, Friedman’s two-way analysis of variance by ranks test is performed to compare all three components.

Null Hypothesis(H0)	Significance (sig)
The distribution of voice,chart and voice+chart are same	0.004(Reject)

Table 10: Related-Samples Friedman’s Two-Way Analysis of Variance by Ranks

From the Table, the significance level of the distribution of all three components is 0.004, which is less than the significance level (0.05). Thus, it is confirmed that the distribution of usefulness rating in the only chart, only voice and chart+voice are not the same.

The post hoc test is also done to find which component shows significant differences and the pairwise comparison is done.

Design Components	Test Statistics	Std Error	Significance (sig)
Voice vs Chart	-0.124	0.354	0.724 (not significant)
Voice vs Voice+Chart	-0.906	0.354	0.010 (significant)
Chart vs Voice+Chart	0.781	0.354	0.027 (significant)

Table 11: Post-Hoc test for usefulness

The Table shows a pair-wise comparison of significant differences in the ranking of voice+chart vs only chart and voice+chart vs voice-only. The significance difference is because of the adjusted p-value, which is smaller than the significance level of 0.05. However, there is no significance between chart-only and voice-only. (Since the adjusted significance value  $> 0.05$ ).

## 5.4 Ranking Preference of the Components

In the post-study questionnaire, We asked the users to rank the most to least preferred design components (1 being the most preferred and three being the least preferred). After the user task, the user selects the preferences after getting familiar with the application. Twelve users ranked that voice + chart as the most preferred, followed by only voice and the least preferred only chart. However, only the chart and the voice are very close in comparison.

### 5.4.1 Statistical Analysis of Preferences

We converted the ranking into categorical data where one is named most preferred, two is second preferred, and three is least preferred. For the categorical data, the statistical test used is Pearson Chi-

Square. Chi-square is performed to observe whether the categorical variable frequencies significantly differ from equal probabilities

Null Hypothesis(H0)	Significance (sig)
The categories of only chart occur with equal probabilities.	0.646(Retain)
The categories of Voice+Chart occur with equal probabilities.	0.002(Reject)
The categories of only Voice occur with equal probabilities.	0.001(Reject)

Table 12: Statistical test for User Preference(One-Sample Chi-Square Test)

The Table shows the p-value of only voice and voice+chart, which is less than 0.05. The value suggests that the categories of these components do not appear with equal probabilities. However, The p-value is higher than 0.05 in only the chart category, which indicates that only the chart category appears with equal probabilities.

#### 5.4.2 User Feedback

The Users were asked the following four questions after the survey:

1) Which component do they find it useful and preferred?

Most users found that the chart+voice design is more useful and preferred compared to its counterparts. The users found this component useful because people who do not know Charts can use Voice simultaneously to understand the data better. Chart and Voice are useful because when you are not looking at the watch, you can hear; when you have a noisy background and can't hear, you can see. So, both components are used to counterbalance each other flaws. When the users multitask, you can see and listen to the data. Chart+voice can be used depending on the situation; when there is no disturbance, you can use the Voice, but when you can go outside and want to know the value, a chart can be helpful.

2)Is there any specific task for which they find these components effective?

Some users preferred Chart only because the voice component can not be used when there is noise. The chart component is also simple and helps to grasp the information quickly. Users found that the Chart depicts whatever it is as knowledge, and other users preferred the data to be seen visually for better understanding. One user found the Chart to be handy when the data is small.

3) How helpful is the tooltip in the Chart?

All the users found the tooltip practical and helpful in finding the accurate value and maneuvering the data.

4) What can be added to these design components for future development?

Many users suggested including the timeline in the x-axis to understand the stock data better. Some even said the timeline must be in the Chart to view the data according to their preferences. A user said to add the increase and decrease percentage of the stock and add more metrics like opening and closing values. Voice component placement can be in different places (say top corners) and should have the stop button. Can incorporate zoom In and out feature. With the voice component, the speed at which the Voice is delivered can also increase. We can improve the visual representation of the application by automatically scrolling the data and adding colourful icons beside the stock names. The simple grid line can be added to strengthen the Chart data's effectiveness.

## **5.5 Summary**

In this Chapter, we discussed the analysis of the results of the user study. In particular, we analyzed the demographic data of participants as well as their performance and subjective measures. From the analysis, we found that voice+chart was the most preferred and most useful component which also led to higher task accuracy compared to other design components.

## **6 Conclusion And Future work**

In this chapter, we provide our concluding remarks, including an overview of our findings and their significance. The design considerations and the possible future work for this research follow the subsection.

### **6.1 Conclusion**

This thesis aims to investigate data visualization on smartwatches and ascertain the significant utility of three design components: Chart, Voice, and Chart+Voice. These design components play a vital role in making the data in the smartwatch easy and quick to glance at and understand the data. In most research, the text is the standard natural language used to integrate with the chart type. In this research, the voice is used as natural language, and this new exploration of natural language paves the way for future research. The incorporation of LLM (ChatGPT) in our prototype is a distinctive feature that enables us to explore AI applications in smartwatches.

We can now revisit the research questions that need to be answered based on our findings from the user survey. Our initial hypothesis suggests that the combination of voice and visualization will be preferable to the user compared to the standalone mode with only the voice or visualization. The ranking rating for the design component data, on a rank case of 1 (being the most preferred) to 3 (being the least preferred), is collected from the users. From the analysis, we found that voice+chart was the most preferred.

Regarding our second question, the combination of voice and visualization proves to be more beneficial to users compared to using voice or visualization alone. Users rated the usefulness of design components on a scale of 1 to 5, with the combined voice and chart receiving higher ratings than standalone voice or chart visualizations.

The third question focuses on assessing user task accuracy, which involves identifying trends and min/max values of three stocks (Costco, Microsoft, Meta), each representing a design component. Analysis indicates that the accuracy rate for Voice+Chart tasks surpassed that of other tasks. The

user feedback also suggested that the former component helps answer the question efficiently and accurately.

In conclusion, our investigation suggests that combining different visualization types yields greater value and preference over standalone visualization methods. Stock data exploration remains an underdeveloped research area, with existing smartwatch applications offering static charts and text. There is potential for developing more interactive and user-friendly applications. This study lays the groundwork for exploring such opportunities in future research endeavors.

## 6.2 Design Consideration and Future Work

Based on our result, we summarize a few design considerations for the stock-related visualization on the IOS watch.

**Introducing Additional Features such as Tooltip:** The tooltip proves to be a valuable feature aiding users in accurately identifying values on the smartwatch interface. Application designers can consider integrating tooltips to mitigate the constraints posed by the small screen size of smartwatches. Moreover, tooltips can be linked with the crown function to facilitate seamless data exploration. Our research indicates that most users found this feature effective and user-friendly.

**Incorporating a Voice Component within the Application:** The Apple smartwatch comes equipped with an integrated voice feature that developers can leverage to enhance user interaction. Integrating a voice component into the application, along with a stop button, allows users to comprehend data without needing to glance at their watch. Many participants appreciated the voice feature for enabling multitasking capabilities, such as listening to application prompts while walking.

**Integrating the Large Language Model:** Our research presents a unique exploration of the LLM application prototype. By coupling ChatGPT with the voice component, data summarization, prediction, and generation become seamless processes. This integration presents numerous avenues for future research and development, offering abundant opportunities in mobile device applications.

**Exploring of Other Charts in Smartwatch Visualization:** While line, bar, radial, and other common chart types are prevalent in smartwatch data visualization, there remains room for exploring additional chart types, such as area charts, scatterplots and heatmaps. Smartwatches' size limitations necessitate the exploration of diverse chart formats to enhance data representation. However, the introduction of new charts alongside natural language interfaces must carefully consider maintaining the glanceability of smartwatch displays.

### **6.2.1 Future Work**

There are several possible ways for future enhancements of the prototype. The prototype application used in this research can be improved by adding the stop button to the voice component for future studies. A proper timeline or timestamp in the visualization will help us better understand the stock values. We can add more metrics like opening and closing values. Adding the icons beside the stock name and upwards or downward signs helps grasp the data quickly. Adding charts like sparklines, donuts, or bars helps to understand the data from different perspectives. The chart representation can be improved by adding the x-axis and Y-axis related to the data; even a simple grid line provides more accurate readability. We hope that further research in the area will help to develop more data accessibility by integrating natural language and visualization in smartwatches to overcome the challenges.

## Bibliography

- [1] M. E. Cecchinato, A. L. Cox, and J. Bird, “Smartwatches: The good, the bad and the ugly?,” in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’15, (New York, NY, USA), p. 2133–2138, Association for Computing Machinery, 2015.
- [2] M. Dehghani and K. J. Kim, “The effects of design, size, and uniqueness of smartwatches: Perspectives from current versus potential users,” *Behaviour and Information Technology*, vol. 38, pp. 1143–1153, 11 2019.
- [3] A. Islam, A. Bezerianos, B. Lee, T. Blascheck, and P. Isenberg, “Visualizing information on watch faces: A survey with smartwatch users,” in *2020 IEEE Visualization Conference (VIS)*, pp. 156–160, 2020.
- [4] “What is data visualization.” <https://www.tableau.com/learn/articles/data-visualization>.
- [5] A. Islam, R. Aravind, T. Blascheck, A. Bezerianos, and P. Isenberg, “Preferences and effectiveness of sleep data visualizations for smartwatches and fitness bands,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI ’22, (New York, NY, USA), Association for Computing Machinery, 2022.
- [6] J. Meyer, A. Kazakova, M. Büsing, and S. Boll, “Visualization of complex health data on mobile devices,” pp. 31–34, 10 2016.
- [7] K. J. Kim and S. S. Sundar, “Mobile Persuasion: Can Screen Size and Presentation Mode Make a Difference to Trust?,” *Human Communication Research*, vol. 42, pp. 45–70, 01 2016.
- [8] D. Dobbstein, G. Haas, and E. Rukzio, “The effects of mobility, encumbrance, and (non-)dominant hand on interaction with smartwatches,” ISWC ’17, (New York, NY, USA), p. 90–93, Association for Computing Machinery, 2017.

- [9] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 76. 2019.
- [10] S. Kantharaj, R. T. K. Leong, X. Lin, A. Masry, M. Thakkar, E. Hoque, and S. Joty, “Chart-to-text: A large-scale benchmark for chart summarization,” 2022.
- [11] J. Chauhan, S. Seneviratne, M. A. Kaafar, A. Mahanti, and A. Seneviratne, “Characterization of early smartwatch apps,” in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 1–6, 2016.
- [12] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 5. 2019.
- [13] H. M. T. Van, N. V. Hao, K. P. N. Quoc, H. B. Hai, L. D. V. Khoa, L. M. Yen, P. T. H. Nhat, H. T. H. Duong, D. B. Thuy, T. Zhu, H. Greeff, D. Clifton, and C. L. Thwaites, “Vital sign monitoring using wearable devices in a vietnamese intensive care unit,” *BMJ Innovations*, vol. 7, no. Suppl 1, pp. s7–s11, 2021.
- [14] T.-C. Lu, C.-M. Fu, M. Ma, C.-C. Fang, and A. Turner, “Healthcare applications of smart watches: A systematic review,” *Applied Clinical Informatics*, vol. 7, pp. 850–869, 09 2016.
- [15] H.-S. Yeo, J. Lee, A. Bianchi, and A. Quigley, “Watchmi: Pressure touch, twist and pan gesture input on unmodified smartwatches,” in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '16*, (New York, NY, USA), p. 394–399, Association for Computing Machinery, 2016.
- [16] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 80. 2019.
- [17] D. Baur, B. Lee, and S. Carpendale, “Touchwave: Kinetic multi-touch manipulation for hierarchical stacked graphs,” in *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces, ITS '12*, (New York, NY, USA), p. 255–264, Association for Computing Machinery, 2012.
- [18] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 77. 2019.

- [19] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, “Orbits: Gaze interaction for smart watches using smooth pursuit eye movements,” in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST ’15, (New York, NY, USA), p. 457–466, Association for Computing Machinery, 2015.
- [20] A. Neshati, B. Rey, A. S. Mohommed Faleel, S. Bardot, C. Latulipe, and P. Irani, “Bezelglide: Interacting with graphs on smartwatches with minimal screen occlusion,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI ’21, (New York, NY, USA), Association for Computing Machinery, 2021.
- [21] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 90. 2019.
- [22] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 213. 2019.
- [23] B. Lee and R. Dachsel, *Mobile Data Visualization*, p. 211. 2019.
- [24] “Sensors overview,” 2023. [https://developer.android.com/guide/topics/sensors/sensors\\_overview](https://developer.android.com/guide/topics/sensors/sensors_overview).
- [25] “Make the most of watchos,” 2023. <https://developer.apple.com/watchos/features/>.
- [26] Z. Alrababah, “Privacy and security of wearable devices,” 12 2020.
- [27] M.-C. M. Tsou, S.-C. C. Lung, and C.-H. Cheng, “Demonstrating the applicability of smartwatches in pm2.5 health impact assessment,” *Sensors*, vol. 21, no. 13, 2021.
- [28] K. J. Kim, “Shape and Size Matter for Smartwatches: Effects of Screen Shape, Screen Size, and Presentation Mode in Wearable Communication,” *Journal of Computer-Mediated Communication*, vol. 22, pp. 124–140, 04 2017.
- [29] M. Dong, “Investigating the users’ preferences of heart rate data types and visualizations on a smartwatch,” *International Journal of Human–Computer Interaction*, vol. 0, no. 0, pp. 1–22, 2023.

- [30] Z. Dong, L. Yi, W. Pan, A. Hu, and W. Ma, “Data-driven study of the preferences and effects of wearable device visualisation in sport,” in *Third International Conference on Computer Vision and Data Mining (ICCVDM 2022)* (T. Zhang and T. Yang, eds.), vol. 12511, p. 125110N, International Society for Optics and Photonics, SPIE, 2023.
- [31] T. Blascheck, L. Besancon, A. Bezerianos, B. Lee, and P. Isenberg, “Glanceable visualization: Studies of data comparison performance on smartwatches,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 630–640, 2019.
- [32] F. Amini, K. Hasan, A. Bunt, and P. Irani, “Data representations for in-situ exploration of health and fitness data,” in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth '17*, (New York, NY, USA), p. 163–172, Association for Computing Machinery, 2017.
- [33] P. Savci and B. Das, “Enhancing text summarization: Evaluating transformer-based models and the role of large language models like chatgpt,” pp. 1–4, 12 2023.
- [34] A. Nazir and Z. Wang, “A comprehensive survey of chatgpt: Advancements, applications, prospects, and challenges,” *Meta-Radiology*, vol. 1, no. 2, p. 100022, 2023.
- [35] J. Kocoń, I. Cichecki, O. Kaszyca, M. Kochanek, D. Szydło, J. Baran, J. Bielaniewicz, M. Gruza, A. Janz, K. Kanclerz, *et al.*, “Chatgpt: Jack of all trades, master of none,” *Information Fusion*, p. 101861, 2023.
- [36] A. Nazir and Z. Wang, “A comprehensive survey of chatgpt: Advancements, applications, prospects, and challenges,” *Meta-Radiology*, vol. 1, p. 100022, 10 2023.
- [37] R. Tariq, S. Malik, M. Roy, M. Z. Islam, U. Rasheed, J. Bian, K. Zheng, and R. Zhang, “Assessing chatgpt for text summarization, simplification and extraction tasks,” in *2023 IEEE 11th International Conference on Healthcare Informatics (ICHI)*, (Los Alamitos, CA, USA), pp. 746–749, IEEE Computer Society, jun 2023.

- [38] L. Tang, Z. Sun, B. Idnay, J. G. Nestor, A. Soroush, P. A. Elias, Z. Xu, Y. Ding, G. Durrett, J. F. Rousseau, C. Weng, and Y. Peng, “Evaluating large language models on medical evidence summarization,” *NPJ Digit. Med.*, vol. 6, p. 158, Aug. 2023.
- [39] X. Yang, Y. Li, X. Zhang, H. Chen, and W. Cheng, “Exploring the limits of chatgpt for query or aspect-based text summarization,” 2023.
- [40] M. Tiwari, M. Kumar, A. Srivastava, and A. Bala, “Inbuilt chat gpt feature in smartwatches,” pp. 1806–1813, 08 2023.
- [41] “watchos 9 is available today.” <https://www.apple.com/newsroom/2022/09/watchOS-9-is-available-today/>.
- [42] “Swift,” 2023. <https://www.swift.org/about/>.
- [43] “‘swiftui‘ framework learning and usage guide.,” 2020. <https://github.com/Jinxiansen/SwiftUI>.
- [44] M. Knott, *Welcome to Xcode*, pp. 3–12. 11 2016.
- [45] A. S. AppPear, “Chartview made in swiftui,” 2020. <https://github.com/AppPear/ChartView>.
- [46] “What is a bar chart?.” <https://www.jaspersoft.com/articles/what-is-a-bar-chart>.
- [47] “Multiple line graphs.” <https://flexbooks.ck12.org/cbook/ck-12-middle-school-math-concepts-grade-7/section/11.8/primary/lesson/multiple-line-graphs-msm7/>.
- [48] H. Siirtola, “Bars, pies, doughnuts tables – visualization of proportions,” 09 2014.
- [49] “build pie chart.”
- [50] “Accuracy.” <https://intercom.help/soma-technologies/en/articles/8058952-how-do-we-calculate-accuracy>.

- [51] “Infographic: Smartwatch market,” 2022. <https://www.counterpointresearch.com/infographic-smartwatch-market-q1-2022/>.
- [52] B. C. Thomas Gonzalez, “Mobile data visualization design,” 2019. <https://medium.com/swlh/mobile-data-visualization-design-18845fe6b660>.
- [53] A. Schiewe, A. Krekhov, F. Kerber, F. Daiber, and J. Krüger, “A study on real-time visualizations during sports activities on smartwatches,” pp. 18–31, 11 2020.
- [54] S. Pizza, B. Brown, D. McMillan, and A. Lampinen, “Smartwatch in vivo,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5456–5469, 2016.
- [55] S. Goodman, S. Kirchner, R. Guttman, D. Jain, J. Froehlich, and L. Findlater, “Evaluating Smartwatch-based Sound Feedback for Deaf and Hard-of-hearing Users Across Contexts,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, (Honolulu HI USA), pp. 1–13, ACM, Apr. 2020.
- [56] Gence, Y. A. Ekmekcciouglu, F. Balci, H. Ürey, and O. Özcan, “Howel: A Soft Wearable with Dynamic Textile Patterns as an Ambient Display for Cardio Training,” in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, (Glasgow Scotland Uk), pp. 1–6, ACM, May 2019.
- [57] V. Setlur and H. Chung, “Semantic Resizing of Charts Through Generalization: A Case Study with Line Charts,” 2021.
- [58] A. Turner, “Global smartwatch market share shipments,” 2023. <https://www.bankmycell.com/blog/global-smartwatch-market-share>.
- [59] E. Dellgren, “A case study on how the apple watch can benefit medical heart research,” 2017.
- [60] W. Playfair, “The commercial and political atlas, representing, by means of stained copper-plate charts, the progress of the commerce, revenues, expenditure, and debts of england, during the whole of the eighteenth century,” 1759.

[61] S. S. Gill and R. Kaur, "Chatgpt: Vision and challenges," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 262–271, 2023.

# Appendices

## A Appendix : Ethics approval

3/9/23, 7:49 AM

Ethics Approval



**OFFICE OF RESEARCH ETHICS (ORE)**  
5th Floor, York Research Tower,  
4700 Keele Street, Toronto ON  
Canada M3J 1P3  
Tel 416-736-5914, Fax 416-650-8197  
www.research.yorku.ca

### Memo

To: Enamul hoque Prince, School of Information Technology  
From: Alison M. Collins-Mrakas, Sr. Manager and Policy Advisor, Research Ethics  
(on behalf of the Chair, Human Participants Review Committee)  
Issue Date: Wed Mar 08 2023  
Expiry Date: Fri Mar 08 2024  
RE: **The User study on data Visualization of Apple smartwatch  
Certificate #: e2023-081**

I am writing to inform you that the Human Participants Review Sub-Committee has reviewed and approved the above project.

Should you have any questions, please feel free to contact me at: 416-736-5914 or via email at: [acollins@yorku.ca](mailto:acollins@yorku.ca).

Yours sincerely,

Alison M. Collins-Mrakas M.Sc., LLM  
Sr. Manager and Policy Advisor,  
Office of Research Ethics

### RESEARCH ETHICS: PROCEDURES to ENSURE ONGOING COMPLIANCE

Upon receipt of an ethics approval certificate, researchers are reminded that they are required to ensure that the following measures are undertaken so as to ensure on-going compliance with Senate and TCPS ethics guidelines:

1. **RENEWALS:** Research Ethics Approval certificates are subject to annual renewal.

**Failure to renew an ethics approval certificate or** (to notify ORE that no further research involving human participants will be undertaken) may result in the **closure of the protocol**. No further research activities may be undertaken until such time as a new protocol has been reviewed and approved. Further, **it may result in suspension of research cost fund and access to research funds may be suspended/withheld.**

<https://ore.research.yorku.ca/approvals/?c=3769>

1/2

2. **AMENDMENTS:** Amendments must be reviewed and approved **PRIOR** to undertaking/making the proposed amendments to an approved ethics protocol;
3. **END OF PROJECT:** ORE must be notified when a project is complete;
4. **ADVERSE EVENTS:** Adverse events must be reported to ORE as soon as possible;
5. **AUDIT:**
  - a. More than minimal risk research may be subject to an audit as per TCPS guidelines;
  - b. A spot sample of minimal risk research may be subject to an audit as per TCPS guidelines.

**FORMS:** As per the above, the following forms relating to on-going research ethics compliance are available on the Research website:

1. Renewal
2. Amendment
3. End of Project
4. Adverse Event

## B Appendix : Consent form

### Informed Consent Form Template

**Date:**

**Study Name:** Integrating Natural Language and Visualizations for Exploring Data on Smartwatch

**Researchers:** Enamul Hoque Prince  
Assistant Professor,  
School of Information Technology,  
York University,  
3052 Victor Phillip Dahdaleh Building (DB),  
Toronto, Ontario, M3J 1P3 Email: [enamulh@yorku.ca](mailto:enamulh@yorku.ca)

**Purpose of the Research:**

The primary goal of this study is to explore how to visualize data in a smartwatch effectively. In particular, we are interested in knowing how to integrate Charts (e.g., bar chart, line chart) with texts and/or speech to help users understand data. You will explore data with different designs of visualizations and then answer some questions about these visualizations which will help us to assess the effectiveness of these visualization designs.

**What You Will Be Asked to Do in the Research:**

The study will be conducted in person, and the session will last between 45-60 minutes. Throughout the session, you will access the smartwatch applications and answer the set of tasks. Upon your consent, the study session will be audio/video recorded and after the study, recordings will be transcribed to understand your feedback and your interaction patterns with the interfaces.

In this research you will be asked to perform the following tasks:

1. At first, the experimenter will explain the consent form to you before the experimental session begins.
2. You will be asked to complete a pre-study background questionnaire, where you will provide demographic information (e.g., age range, gender), educational background, and experience with smart watch.
3. You will be provided with a 10 minutes' demo that explains the functionalities of our smart watch application.
4. Afterward, you will be given about 5 minutes to get familiar with the smart watch application and if you have any questions, clarifications will be provided by the experimenter.
5. You will be asked to perform a set of tasks using the smart watch by exploring different type of visualization design in the smart watch application.
6. You will be asked to complete a post-study questionnaire and go through a brief interview with a few additional questions about the usefulness of the smart watch application.

As a token of appreciation for participating, we will offer you a \$20 cash. Please note that, you will be compensated the full amount (\$20) even if you withdraw without completing the experiment.

**Benefits of the Research and Benefits to You:** Taking part in this research study may not benefit you personally, but we may learn new things that could contribute new understandings valuable to the scientific community. Through this research, we aim to study more about the smartwatch display and its limitations in data visualization, which are highly desirable among visualization and smartwatch application developers and practitioners.

**Voluntary Participation and Withdrawal:** Your participation in the study is completely voluntary and you may choose to stop participating at any time. Your decision not to volunteer, to stop participating, or to refuse to answer particular questions will not influence the nature of the ongoing relationship you may have with the researchers and York University either now, or in the future. If you stop participating, you will still be eligible to receive the promised pay/compensation for agreeing to be in the project. In the event you withdraw from the study, all associated data collected will be immediately destroyed wherever possible.

**Risks and Discomforts:** We do not foresee any significant risks or discomfort from your participation in the research. However, if you feel any discomfort or risks, you may choose to stop participation. Please note that we are not testing you nor your abilities and your responses collected from the user study will be used to evaluate the designs of user interfaces.

**Confidentiality:** All personally identifying information you supply during the research – namely: your name and email address -- will be held in confidence and your name will not appear in any report or publication of the research. Your data will be collected through the use of Google Form (<https://docs.google.com/forms>). Your personally identifying data will be safely stored in a locked facility and a password-protected digital medium (USB key or DVD) and only research staff/research team members will have access to this information. The personally identifying data will be destroyed by the end of 2023. However, the non-identifiable response data will be retained at the researcher's website indefinitely for future research and for sharing with the research community. Confidentiality will be fully provided possible by law.

Your non-personally identifying responses -- including demographics (gender, age range, educational background), interaction log data, and post-study questionnaire responses -- may be published on the PI's websites (personal, lab) and shared with the research community as data files. The PI will screen the data and use his own judgment to identify accidentally or indirectly personally identifying features and will redact the parts in question or, if necessary, the entire response prior to availing it to third parties. The anonymized non-personally identifying data collected in this research project may also be used by members of the research team in subsequent research investigations exploring similar lines of inquiry. Such projects will still undergo ethics review by the HPRC, our institutional REB. Any secondary use of anonymized data by the research team will be treated with the same degree of confidentiality and anonymity as in the original research project.

The researcher(s) acknowledge that the host of the online survey (e.g., Google form) may automatically collect participant data without their knowledge (i.e., IP addresses.) Although this information may be provided or made accessible to the researchers, it will not be used or saved without the participant's consent on the researchers' system. Further, because this project employs e-based collection techniques, data may be subject to access by third parties as a result of various security legislation now in place in many countries and thus the confidentiality and privacy of data cannot be guaranteed during web-based transmission.

**Security limitations:** While York University researchers will not collect or use IP addresses or other information which could link your participation to your computer or electronic devices without informing you, there is a small risk with any platform such as this of data that is collected on external servers falling outside the control of the research team. If you are concerned about this, we would be

happy to make alternative arrangements (where possible) for you to participate, perhaps via telephone. Please contact Dr. Enamul Hoque Prince by e-mail ([enamulh@yorku.ca](mailto:enamulh@yorku.ca)) for further information.

Recordings (audio/video) will be saved in a password protected file to research team members' local computer, not the cloud-based service. Please note that it is the expectation that participants agree not to make any unauthorized recordings of the content of a meeting / data collection session.

**Questions About the Research?** If you have questions about the research in general or about your role in the study, please feel free to contact Dr. Enamul Hoque Prince by e-mail ([enamulh@yorku.ca](mailto:enamulh@yorku.ca)). This research has received ethics review and approval by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process, or about your rights as a participant in the study, please contact the Sr. Manager & Policy Advisor for the Office of Research Ethics, 5th Floor, Kaneff Tower, York University (telephone 416-736-5914 or e-mail [ore@yorku.ca](mailto:ore@yorku.ca)).

**Consent:**

By clicking "Agree" below, you confirm that you have read the consent form, are at least 18 years old, and agree to participate in the study "Integrating Natural Language and Visualizations for Exploring Data on Smartwatch".

- I agree to participate in the research
- I do NOT agree to participate in the research

**Consent for audio and or video recording:**

I understand that audio/video recording is voluntary and that I am free to request that these be stopped at any time during the recording. I also understand that if I should withdraw from the research, the audio and/or video will be destroyed. The procedures for protecting my identity and confidentiality include the destruction/deletion of the audio/video files after transcription and verification.

- I consent to the audio recording.
- I consent to the video recording.

Next >>

## C Appendix : Google form

2/19/24, 8:50 PM

The User Survey Form

### The User Survey Form

Pre study Questions

\* Indicates required question

---

1. Email \*

---

#### User Details

2. 1) Please indicate your age range \*

*Mark only one oval.*

- Below 20
- 21-30
- 31-40
- 41-50
- Above 50 years

3. 2) Please indicate your gender \*

*Mark only one oval.*

- Female
- Male
- Transgender Female
- Transgender Male
- Gender Variant/Non-Conforming
- Not listed
- Prefer not to answer

<https://docs.google.com/forms/d/1viQuxNX1vzRW0w1w2OfW6bRbK7D62NHahRvaYb5HycA/edit>

1/6

4. 3) Please indicate your occupation \*

\_\_\_\_\_

5. 4) What mobile platform are you using now? \*

*Mark only one oval.*

- iPhone OS
- Android
- Windows
- Other: \_\_\_\_\_

6. 5) Have you used a smartwatch? If yes, what type of smartwatch? \*

*Mark only one oval.*

- Apple
- Samsung
- Garmin
- Fitbit
- Other: \_\_\_\_\_

7. 6) How frequently are you using your smartwatch? \*

*Mark only one oval.*

- Never
- Rarely (several hours a year)
- Occasionally (several hours a month)
- Frequently (several hours a week)
- Very frequently (several hours a day)

8. 7) How often do you check the stock data from your smartwatch? \*

*Mark only one oval.*

- Never
- Rarely (several hours a year)
- Occasionally (several hours a month)
- Frequently (several hours a week)
- Very frequently (several hours a day)

9. 8) Do you use any specific types of watch applications? \*

*Check all that apply.*

- Fitness
- Weather
- Stocks
- Music
- Other: \_\_\_\_\_

**User Task**

10. Could you locate the minimum value of Costco's stock in this line chart? \*

*Mark only one oval.*

- 551
- 669
- 560

11. What does the line chart reveal about the trend of Costco's stock? \*

*Mark only one oval.*

- Upward Trend(Positive)
- Volatile Trend(Fluctuation)
- Downward Trend(Negative)

12. Could you locate the peak value of Microsoft's stock in this line chart and voice components? \*

*Mark only one oval.*

- 352
- 337
- 361

13. What does the line chart and voice reveal about the Trend of Microsoft's stock? \*

*Mark only one oval.*

- Upward Trend(Positive)
- Downward Trend(Negative)
- Volatile Trend(Fluctuation)

14. Could you locate the peak value of Meta's stock from the voice components? \*

*Mark only one oval.*

- 325
- 320
- 314

15. What does the voice feature reveal about the trend of Meta's stock? \*

Mark only one oval.

- Volatile Trend(Fluctuation)
- Downward Trend(Negative)
- Upward Trend(Positive)

Post Study Question

16. On the scale of 1(least useful) to 5 (most useful) ,how useful is the voice component in conveying the data? \*

Mark only one oval.

1 2 3 4 5

---

---

17. On the scale of 1(least useful) to 5 (most useful) ,how useful is the voice and chart components combined together in conveying the data? \*

Mark only one oval.

1 2 3 4 5

---

---

18. On the scale of 1(least useful) to 5 (most useful) ,how useful is the chart component in conveying the data? \*

Mark only one oval.

1 2 3 4 5

---

---

19. Rank the following interface on the application based on your preference, with 1 <sup>\*</sup> being the most preferred and 3 being the least preferred:

*Check all that apply.*

	1	2	3
<b>Only chart</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Only voice</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Chart + Voice</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

---

This content is neither created nor endorsed by Google.

Google Forms