DATA ANALYTICS AND VISUALIZATION FOR VIRTUAL SIMULATION

TABLE OF CONTENTS

List of Figures

ACKNOWLEDGMENT

My adviser, Dr. Thomas Wischgoll, has my sincere gratitude for letting me take part in the "Wright Life (Lifelike Immersion for Equity)" project. He continues to have faith in me, and I am appreciative of that. Dr. Thomas Wischgoll is a wonderful teacher, mentor, and thesis advisor, imparting good advice and encouragement in every step of the thesis. I'm proud and appreciative of the opportunity to work with him.

Then, I want to express my gratitude to Dr. Hershberger and Dr. Pei for their committee service as well as their helpful advice and counsel during this research. I am also indebted to my incredible computer science colleagues and themedical team. Thank you for all your help and support throughout the project. I will always remember our time together on this endeavor.

I would thank my parents for supporting my dream to study in a distant country, far from home. They always respected my opinions, supported my decisions, and made sacrifices for me to pursue this master's degree.

Finally, I want to show my thankfulness to the Ohio Department of Medicaid for sponsoring the research for my thesis, which is a component of The Medicaid Care Experience Simulation Project. The Ohio Department of Medicaid funds and the Ohio Colleges of Medicine Government Resource Center manage the Medicaid Care Experience Simulation (MCarES) initiative. The views expressed in this thesis are solely those of the author and do not represent the views of the state of Ohio or federal Medicaid programs.

1 INTRODUCTION

Healthcare professionals commit their entire life to helping and caring for others. Healthcare personnel has been fully committed to offering necessary treatment while demonstrating compassion and empathy. A variety of patients and caregivers are drawn to healthcare institutions. Healthcare workers engage with people from different racial, cultural, and socioeconomic backgrounds. When interacting with members of minority groups like those who identify as LGBTQIA+ (lesbian, gay, bisexual, transgender, queer, intersex, and asexual) or those who have autistic spectrum disorder (ASD), it is crucial to exhibit compassion and empathy to provide equal treatment.

Compassion is defined as the sympathetic consciousness of others' distress together with a desire to alleviate it. Healthcare is the effort made to maintain or restore physical, mental, or emotional well-being, especially by trained and licensed professionals [1]. The definition of empathy is "the capacity to comprehend and experience the sentiments of another." It is the ability to step into another person's situation, experience what they are going through, and share their emotions and sentiments. It entails acknowledging and validating a patient's fear, anxiety, discomfort, and worry. It is the capacity to comprehend the emotions of patients, enabling a more precise diagnosis and more compassionate care. Empathizing with patients does enhance humanism in healthcare. The *key ingredient* to improving the patient experience and patient interaction is exhibiting empathy in healthcare. In healthcare, empathy, and compassion both play crucial roles in the patient experience and are essential elements of the doctor-patient relationship [2].

Social and economic backgrounds influence the lives of people more than one can imagine, a person from a lower-income family faces more obstacles when it comes to

receiving proper healthcare. Socioeconomic status (SES) includes not just earnings but also the level of education, financial stability, and self-perceived social standing and class. A person's socioeconomic position might include aspects of their quality of life as well as the possibilities and privileges that are available to them in society. The pressures associated with poverty are numerous and include both physical and psychological ones [3]. U.S. Census Bureau data demonstrates the persistence of higher poverty rates for African Americans, Latinos, Asian Americans, children, single mothers, people with disabilities, and other groups, for example. An earlier Williams Institute study and other research showed that lesbian, gay, bisexual, Transgender, and Queer (LGBTQ) people were also more vulnerable to being poor [4].

Medicaid is a state and federally funded program that assists people with medical care and has been the main contributor to health insurance coverage for low-income people in the United States [5]. Every patient a practitioner encounters is unique, and it helps to understand the patient's history to manage the case more successfully. Through experience and exposure to various situations, we humans become more informed and adept at managing them. Therefore, it is advantageous for healthcare staff to receive training so they can manage situations where they must treat the patient with more understanding and compassion.

A simulation-based program for medical practitioners called Wright LIFE (Lifelike Immersion for Equity) was developed for this reason. Both the Apple Store and Google Play include this smartphone application. The development of clinician-patient relationships, the development of long-lasting compassion and empathy in healthcare workers for their actual patients, and the cultivation of mindfulness about social

determinants of health (SDoH) are all made possible by this engaging, interactive, and realistic application. The patient's willingness to be truthful and frank may be affected by this compassionate relationship, which may also help to lessen the consequences of any potential bias [6].

The "Wright Life (Lifelike Immersion for Equity)" project mimics comparable situations to aid healthcare professionals in becoming familiar with the cases through engaging game-like simulation. In addition to its emphasis on virtual simulation, "Wright Life" also gathers information from real-world professionals to examine and further analyze it to provide better results for dealing with and interacting with patients. This thesis's focus is also on data collection, analysis, and visualization with the goal to be better able to process and analyze the data to determine the effectiveness of the simulation and further improve the simulation where necessary. From the visualization, conclusions such as a change in user compassion, anxiety, and frustration can be drawn. Data can also be used to determine the trends based on demographic information like gender, profession, and age.

1.1 OVERVIEW OF THESIS

The thesis is broken up into chapters to concentrate on various project phases. The tools used throughout the project's duration are introduced in Chapter 2 of the thesis. The process of creating the animations for a character is briefly covered in chapter 3. Scene design and story development are the main topics of Chapter 4. The primary goal of chapter 5 is to cover data collection and data handling. Chapter 6 lists the problems and limitations. The paper is concluded in chapter 7 with the contributions and improvements, as well as a description of the scope of future work.

1.2 LITERATURE REVIEW

There have been many adaptations of Simulation-based technology used in the medical field but not many addresses Social Determinants of Health (SDoH) [7]. Also, there are not many tools available for increasing the awareness of Medicaid challenges and issues. “WRIGHT LIFE” is developed to address these challenges to improve the patient-provider relationship.

It is not very clear how health care organizations are developing tools to address the SDoH of a patient. Only a limited range of social determinants known to affect health outcomes are documented, and the SDoH data that are collected are rarely presented to care teams via automated structures designed to support care delivery, clinical decision making, and collaboration among multiple service organizations [7]. LaForge et al.[8] has described the methods used by some organizations to develop tools for handling SDoH and issues faced during those efforts. This article discusses how the organization is set up and developed screening tools for SDoH, EDH data collection in the form of electronic records, and data management and analysis. As the “WRIGHT LIFE” application will be available on mobile devices, in either the Apple App store or android google play store, this app will be one of the very few mobile applications, providing training to healthcare professionals to provide awareness.

The data traditionally collected through paper-based forms is not effective as it is difficult to acquire, consolidate and share. For collecting data in electronic form, surveys are one of the effective methods. David De Vaus [9] provides an idea of how to plan, conduct and analyze the surveys. Questionnaires are a widely used technique for the collection of information. The distinguishing functions of surveys are – the form of data

and the method of analysis [10]. Steps in the experimental analysis include – posing a question, conceptualizing the question, collecting the data, and analyzing the data. In the current project, the questions are formulated to understand the provider's mental state and behavioral responses in reference to the simulation. 5-point Likert scale that ranges from low level to high level. For example, "strongly disagree" to "strongly agree". Data Analysis starts with scoring the simple questionnaire and later combining the responses [11]. The data is not only quantitative but also qualitative. Studies that are quantitative focus on numerical or quantifiable data. Qualitative studies, on the other hand, focus on showing how individuals think or act in a certain social context [12]. To draw conclusions from the qualitative data, create a category system that enables the data to be categorized systematically by carefully reading the data [13].

It can be really hard to understand the raw data, especially for people who are not familiar with the data or what the data represent. Just like the saying "A Picture Is Worth a Thousand Words" suggests, visualizations can be effectively used to illustrate the impact of assets on health [14]. Showing the raw data may not make much sense but, if the data is transformed into a pictorial representation, the results can be conveyed easily. Using data analysis and visualization is an effective way to present the data. Multivariate data is defined as the "collection of data where each entry is a nominal or ordinal value corresponding to a dependent or independent variable". As the dataset increases with time, it becomes large and complex to analyze and interpret [15]. A large number of dimensions not only cause clutter, but it would also be hard to visualize the data [16]. In this project, the interactive visualization of the large dataset is achieved through the Parallel Coordinates technique [15] [16]. Parallel coordinates are a great choice as this visualization

technique can be extended to large multivariate datasets and have an advantage over orthogonal coordinates. Due to the orthogonal layout of the coordinate axes, scatterplots can only display a maximum of three dimensions; however, a parallel coordinate enables the depiction of many dataset properties as the parallel coordinate plot can handle arbitrarily many dimensions because the axes are drawn as parallel lines and then data points connected by line segments. Parallel coordinates are quite popular for the presentation and analysis of multivariate data, even though both techniques rely on projections of higher-dimensional geometry and are connected by a point-line duality [17].

2 TOOLS UTILIZED

In this simulation, there are multiple elements which include character design & development, animation, game/scene creation & narration, and finally, data analytics.

The transition from 2D to 3D is opening new areas and possibilities in the digital world. Recreating the world on a computer, just the way we see it, is important [18]. The best 3D modeling software unlocks your creativity, helping you bring your eye-catching and imaginative designs to life [19]. The character that is being created must be likable and relevant to the plot and scene in question. Users may only then relate to and get fully immersed in the game. Multiple tools are used for character creation, animating them, and adding vocals to enhance the user experience. Tools with the least amount of learning needs that can produce better outcomes are picked, considering the availability of resources like money and amount of time.

For 3D modeling, Autodesk 3DS Max, Blender, Autodesk Maya, and Realluminate were assessed. 3DS Max is especially popular with game developers, interior designers, and architects who need professional modeling, texturing, and meshing tools. Although it is perfectly capable of animations, experienced animators, for example, may be better served with Maya, from the same developer. Some users reported issues when using too many plugins, so they interact in unpredictable ways. Complex plugins that stress the PC can make 3DS Max unstable on occasion. Blender is an open-source 3D creation tool supported by a resolute community of developers and users, and free for personal and commercial use. Blender's independence and reliance on community support do mean it will not always have regular updates and bug fixes. Autodesk Maya is considered one of the best 3D character modeling software and it is often the software of choice for artists

and animators. Maya features seriously powerful tools for character creation and movement, and the simulation of natural elements such as water, fire, sandstorms, and explosions. However, Maya is a hugely advanced tool. And its interface is far less user-friendly [19].

iClone is the fastest real-time 3D animation software in the industry, helping users easily produce professional animations for films, previz, animation, video games, content development, education, and art. Integrated with groundbreaking real-time technologies, iClone simplifies the world of 3D Animation in a user-friendly production environment that blends character animation, scene design, and cinematic storytelling; quickly turning imagination into a reality [20]. Reallusion has provided multiple tools like iClone7, Crazy Talk, and 3DXchange which provided a full pipeline. As the tools can integrate, it made the choice easy to choose Reallusion tools.

Unity is a platform for games and ongoing virtual production. The game characters can be upgraded with a polygon reduction, LOD, UV, and material convergence for maximum efficiency using Reallusion's Game Tools for Unity. Reallusion's Character Creator computerized assistance for shader tasks, game regulation, and face exhibits enables users to import pre-made characters into game engines [21]. Unity 3D is chosen because we can import the created characters directly with animations and automatic shaders with reduced polygons. Due to this easy import option, we can save time by not doing the rework on the characters and project.

The data analytics and visualization software used in the project is Tableau. It is the market's top tool for data analytics and visualization (with a 33 percent market share, followed by Power-BI). Tableau has a very simple drag-and-drop user interface that makes

it simple to learn and allows you to work with practically any form of data. Because of this, it is a great option for data analysis [22]. Using Tableau users can create visuals, enhance the analysis, and perform data analytics, maps, and storytelling. Tableau is used to visualize the data of SDOH and Wright Life simulations. The visualization helps to present the data in a simple and impactful way for a better understanding of regular users as the raw data is hard to digest whereas data visualized as graphs and dashboards are easy to comprehend.

2.1 iClone7

Thanks to its extensive libraries and sophisticated character morphing capabilities, Reallusion's iClone 7 is a distinctive and multifaceted application that enables you to quickly get started in character animation without having to worry too much about modeling your characters or props from scratch [23].

The real-time 3D animation software in Reallusion's iClone allows users to animate an entire 3D scene. The tool’s libraries and marketplaces of pre-defined characters and props provide easy access to modify the animations in many ways as pleased. It is easy to render a character’s personality to allow the character to mimic the motions of talking and body movements. It can also emulate full body rigging and uses advanced tools for manipulating the curves of the animation. The rendering core of iClone has been redesigned by Reallusion in version 7 to be more compatible with widely used shader languages. This led to the development of a new Physically Based Rendering engine (PBR) that can produce an incredibly realistic appearance [23]. The program now features puppet tools, automatic lip-syncing, and Global Illumination, powered by NVIDIA VXGI (Voxel Cone Tracing). Specific to the “Write Life” project simulation, iClone is used only to produce animations based on resources, project scope, and skill levels.

The tool is easy to use and has many features, but the cost is one of the major drawbacks. Users can sign up for a free trial of iClone, but the full program must be purchased. Though the tool provides many pre-made character attributes and libraries, many of them are not available in the trial version and must be bought. The lack of resources such as tutorials available on the internet puts a toll on developers to learn the functionalities when the time is limited. When your time is limited, you may end up purchasing customized animations or templates by other individuals through the marketplace, which are costlier. Though iClone does not need a high-end configuration system, it requires specific system requirements which are not available on macOS.

2.2 Crazy Talk 8

Crazy talk is another software product from Reallusion. Crazy Talk can animate facial images using sound and text. By importing images, specifying the facial feature points, and recording your voice; it generates lip-syncs automatically to create 3D life-like, talking videos [24]. The tool has capabilities of producing complex facial expressions as well in addition to lip-syncing features.

Crazy Talk has excellent tools to fix geometry and animation, but the biggest benefit is the huge preset library. It comes with many faces, sounds, props, and backgrounds, which are added to if you register the software [25]. It comes preloaded with 2D and 3D actor models, as well as the ability to construct 3D models fromcharacter images and iAvatar files produced by other Reallusion products such as CC3 andiClone. Thereafter, these model heads can be used to create talking animations from voiceand text [6]. As the animations are created using iClone7, it became easy to integrate them with Crazy Talk due to compatibility. One of the most noteworthy features is the ability to make lip-syncs

utilizing auto-motions from a variety of scenarios, including Advice, Denial, Sympathy, Charge Attack, Mumbling, and many more. This feature has been used extensively in this project [6].

Crazy Talk comes with a set of limitations as well, one of them being its cost. It is a great product to make animations realistic but not realistic when it comes to pricing. Like any other, the free trial version does not include advanced features.

2.3 iClone7 3DXchange

3DXchange Pipeline is also a tool from Reallusion where users can narrate the characters or elements of a scene. The tool allows imports from various formats into iClone. This is extremely helpful when one needs to reuse characters. It specializes in the creation of virtual production, real-time cinematic animation, and motion capture solutions that expand the capabilities of professional 3D animation technologies. 3DXchange by Reallusion has import and export capabilities that extend to all types of 3D assets, including static objects, animated props, skin-bone rigged characters, and motion files. One can also repurpose 3D assets with output options tailored for Unity, Unreal, Maya, Blender, Cinema 4D, or Daz Studio [27].

As the tool is from the same vendor Reallusion, it integrates well with iClone. This program is used in this project to combine and export the characters and their animations in FBX format that is compatible with Unity. Unity 3D is another tool used for the project, which will be explained later in this chapter.

3DXchange has backward compatibility issues, therefore 3DXchange Pipeline 7 will only operate with iClone 7, not iClone 6 or earlier versions. However, it did not cause any issue in this project as iClone 7 is being used.

2.4 Kukarella

Most of us enjoy hearing things more than reading them. But producing speech content is typically expensive and time-consuming. Software for text-to-speech can change all of that [27]. Kukarella is an audio-to-text and text-to-speech online converter of the next generation. With effects and accents, it can produce 270+ realistic voices in more than 55 languages. It is easy to add small characteristics to computer-generated voices, such as whispers, softness, pauses, emphasis, and sighs, making it impossible to tell them apart from genuine voice actors [28].

It is also powered by Text to Speech (TTS) services from Google, Amazon, Microsoft, and IBM, which makes it a compelling option. In addition to monthly subscriptions, Kukarella also features a free plan with restrictions on usage. You can utilize the text-to-voice feature on the free plan for up to 2,000 characters per month and the audio-to-text option for around 5 minutes per month [27].

In this project, the characters need to interact and communicate with each other. Kukarella is used for voice generation in the scenes of the simulation. The tool helped with generating unique voices for each character instead of hiring expensive voice actors. It has a user-friendly interface and allows input texts to be organized as projects.

2.5 Unity 3D

A robust cross-platform 3D engine and user-friendly development environment are also features of Unity3D. Both the beginner and the expert may use it without difficulty. Anyone who wants to efficiently create 3D games and applications for mobile, desktop, the web, and consoles should be interested in Unity [29]. It can also be used to develop AR/VRapplications. Unity3D application community has exhaustive documentation which

provides complete description along with several examples including video and text tutorials and live training sessions to understand the engine. Apart from this, the growing community offers advice to help resolve any situations that may arise [30].

In this project, Unity3D is the game engine that has been used to bundle together everything by integrating the characters and animations all together into scenes, script the scenes, develop environments, and eventually the VR simulation app. The only disadvantage is that it is quite complicated and has a steep learning curve [6]. In terms of gaming engines, Unity 3D is quite huge. Due to the numerous bloats that come with utilizing an engine like Unity, even a blank project will frequently be bigger [31]. Unity generates temp and library files within the project folder, which accounts for the large project size. When an empty project is created it automatically imports the libraries which may not be needed for that project.

2.6 Tableau

Tableau is a visual analytics tool that is revolutionizing how we leverage data to address issues by enabling individuals and companies to maximize their data [32]. The Tableau platform, the market-leading option for modern business intelligence, is renowned for quickly and easily transforming any type of data from practically any system into meaningful insights. It only requires dragging and dropping [33].

Tableau's live visual analytics and interactive dashboard allow slicing & dicing datasets for generating relevant insights and exploring new opportunities. The choices from the interactive simulation are collected to analyze and provide insights to make better decisions. The predictive outcome in case this project is to determine the kind of responses by healthcare professionals that are best suited in different scenarios based on the case. The

tool allows to transform the data and generate analysis through varied functionalities and capabilities. It is easy to visualize the data as dashboards and reports using Tableau.

Tableau offers free one-year Tableau licenses to students at accredited academic institutions through their Tableau for Students program. Students can also receive access to the entire eLearning suite once verified. So, the tool fits the profile for data analysis and visualizations [34].

2.6.1 Excel vs. Tableau

Though Tableau is used for data analysis, the data is loaded into Tableau as an excel file. Excel provides capabilities for building dashboards as well but is limited. Excel can be used for reporting and quick visualization, whereas Tableau can be used to get insights from the data. The visualization built using Tableau is scalable, fully interactive, and reusable. The data source can be updated easily in case the source changes. The performance of Excel in terms of handling visualization is poor compared to Tableau. For creating visualization in Excel, one needs knowledge of custom functions and scripting (VBA) but for Tableau, no technical or coding knowledge is required. The charts generated in Excel are simple and limited. Tableau has a wide variety of built-in custom charts. Sharing the dashboards is very easy through Tableau. Though Excel has plotting capabilities, developing dashboards for unstructured data is difficult. The dashboards created in Excel will not be interactive like they are in Tableau [35]. As the Excel dashboards are not as powerful as Tableau, Tableau is chosen to create the visualization instead of Excel. One other major concern with Excel is collaboration. When we need to share dashboards /charts with others, raw data needs to be shared along as well. The survey data contains PII (Personal Identifiable Information) and PHI (Protected Health

Information), it violates the policies if the data reaches unauthorized persons. Through Tableau, we only need to share the final dashboards, which would improve the security vulnerabilities.

3 CHARACTER ANIMATION

Character animation gives a realistic view to the users. We can animate a character to show how living organisms behave and act in their environment. We can animate the characters by creating avatars and applying animations. Character animation is generally defined as the art of making a particular character move in a two- or three-dimensional context. It is a process central to the concept of animation. The idea of character animation has evolved through various types of animation techniques [36]. The 3d character animation technique is widely used to add movement to three-dimensional images. To create a perfect 3D character animation thoughts, personalities and emotions are added to the characters [37].

The game simulation in the scope of this thesis has 2 cases that mainly revolve around the clinical encounter of the characters and their experiences. Following is the summary of the simulations-

The LGBTQIA+ case: The is about a 60-year-old African American gay man Charles, who is well-educated and has a history of serious asthma and body weight management because of prescribed steroid use. Both Charles and his longtime companion Dani (now his spouse) have faced discrimination in employment and housing. Charles has been on antidepressant and anxiety medication for most of his adulthood, and he has also been a heavy smoker. Charles' visit to the community health center is due to growing knee pain, for which he has been informed that a surgeon will not perform a knee replacement unless his weight falls below a BMI of 35.

The character of Dani, a 55-year-old African American nonbinary gym instructor, and the receptionist, a 20-yar-old Asian female are developed and animated.

The ASD Case: This case is about an 18-year-old autistic woman, Ashley. Her mother's history of substance addiction caused her to be taken from her home. She was placed in foster care when she was five years old and has since lived with a variety of caregivers. She was bullied in primary school after being diagnosed with an intellectual impairment and autistic spectrum disorder. Ashley has been dating a 19-year-old lad, whom she met at high school, for almost 2 years now. Since she turned 18, Ashley has been seeking autonomy and freedom. Ashley came to the community health clinic to discuss birth control and to have a form signed for college so she could take tests in a quiet room. Because her normal provider has left the practice for a new position, Ashely sees a new provider at this visit.

In this case, the receptionist and Ashely's foster father Andy, a 50-year-old African American male are developed and animated.

While creating the characters Animation, considered 4 different aspects.

I. Character's idle pose – When the character is still, or not performing any actions while being passive in a conversation.

II. Character's sit pose – When the character is sitting. During the simulation, there are several instances where characters sit during the interactions.

III. Gestures – To depict the character's body movements, body gestures are included to make the characters appealing during the game narration. In Idle poses, though the character is still, added eye blinking and minor body movements to give a life-like feel.

IV. Facial Expressions – Facial Expressions are crucial to express the emotion of the characters.

3.1 LGBTQIA+ Case – Dani's Character animation

Cues – 55 years old African American, Non- binary, Gym Instructor, Married

Dani was Charles' partner for a long time and is now his spouse. Dani's preferred pronouns are they/them. They are Charles's support in battling depression and are extremely caring. Dani accompanies Charles on his visit and interacts with the provider while discussing the current treatment and surgery options. Dani should have a toned body as they are a gym instructor. Though They must have a toned athletic body, their face should look aged because they're a 55-year-old.

Figure 1: Dani's Idle Stand Pose

Figure 2: Dani's Idle Sit Pose

Figure 3: Dani when Stretching his Hands

Figure 4: Dani's Facial Expressions

3.2 LGBTQIA+ Case – Receptionist's Character animation

Cues – 20-year-old Asian Female, Friendly

The Receptionist plays a significant part in both the stories. She greets the main protagonists Charles and Ashley and appears generally polite and cheerful. To express more diversity in the simulation, she needed to be Asian.

Figure 5: Receptionist's Idle Stand Pose

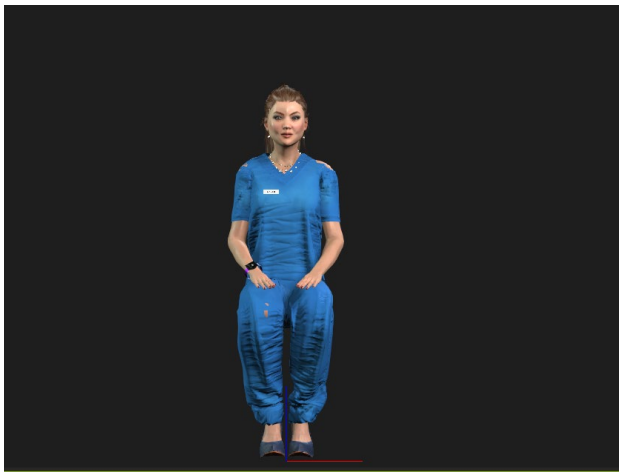

Figure 6: Receptionist's Idle Sit Pose

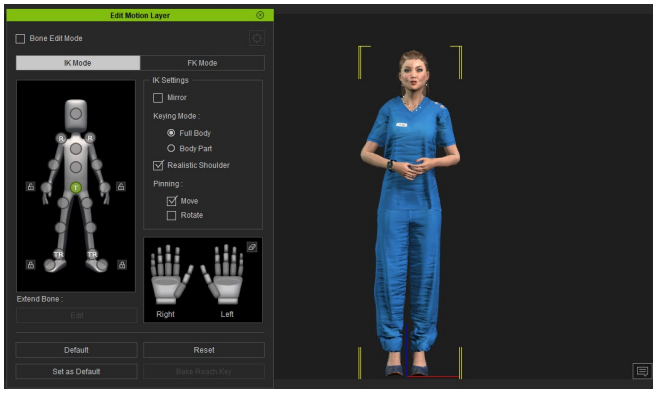

Figure 7: Receptionist's Holding Hands Pose

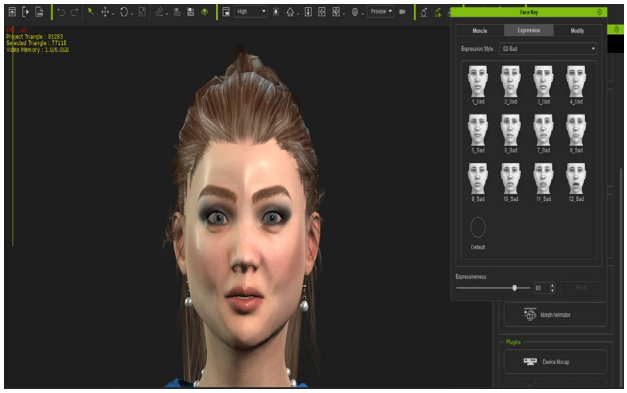

Figure 8: Receptionist's Facial Expressions

3.3 ASD Case – Andy's (Ashley's foster father) Character animation

Cues – 50-year-old African American Male

Andy is Ashley's foster father and her major supporter. He respects Ashley's choices and is supportive of her choice. The requirements for building this model's body were simple. The face needed a bit more detailing to make him look like someone with whom people will be more comfortable.

Figure 9: Foster Father standing Pose

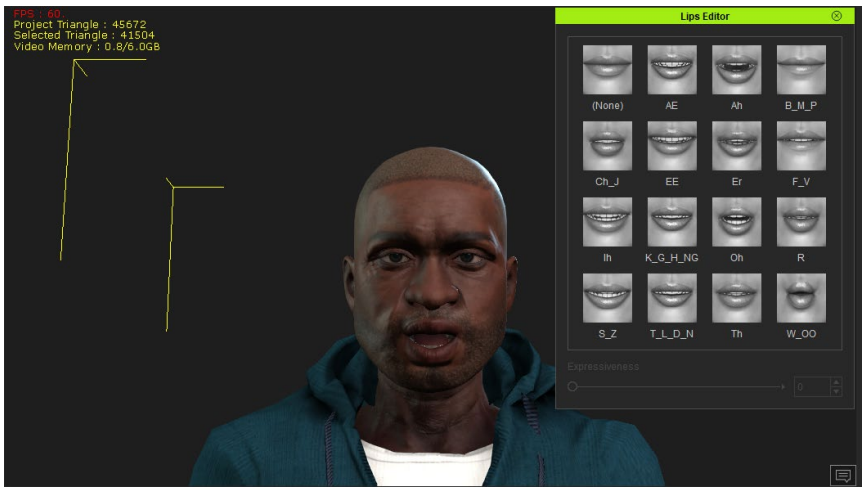

Figure 10: Foster Father Facial Expressions

4 DESIGN AND DEVELOPMENT

The scenes and storyline contribute to making the simulation interactive and interesting. So, the story and scenes are carefully designed and developed to capture the essence. For the story plot, extensively collaborated with the Medicaid team. Initial cues were provided by the team. Based on the inputs, the scenes were developed. Once the scene is developed, it was reviewed by the Medicaid team again to ensure the story requirements are translated into the simulation correctly. This simulation centers around Charles, a 60-year-old gay man with an appointment at the community health center for Knee pain. The story is developed to depict virtual scenes portraying Charles's commute to the community health center, interaction with the Front Desk receptionist, exchange with the medical assistant, conversation with the provider, and developed demographic and survey scenes.

The elements of the design and development are divided into 4 parts to better elaborate.

4.1 Background Story

The main character in the story is Charles, a 60-year-old African American gay man. Charles faced many hardships from friends and family for being gay. Charles's parents were not supportive of him and ridiculed him. He was bullied and beat up at school by his classmates. Charles was thrown out of the house when he came out as gay. While struggling with depression and social stigma, Charles graduated from High school and took a job to work on getting his apartment. He began college but dropped out for a year due to being bullied by anti-gay students. He later graduated with a bachelor's degree in History, which

is when he met Dani, his current husband. After 7 years in graduate school, Charles completed his Ph.D. in History.

Charles applied for a faculty position in a local private university, but he was not offered the job due to his sexual orientation assumed from his published papers on the history of the Gay rights movements. Charles and Dani were denied a spot for an apartment in a more desired location due to anti-LGBTQIA+ bias. Facing these scenarios, he continued to work at his low-paying job. Years passed, and his PCP (Primary Care Physician) prescribed him new anti-depressant paroxetine, which contributed to additional weight gain. Charles and Dani were able to adopt Dani's 17-year-old niece. They were able to get married when same-sex marriage was legalized.

Charles was denied opportunities and faced constant discrimination throughout his life. Charles and Dani struggle a lot financially as well due to their limited income from low-paying jobs.

4.2 Storyline implementation

The current story is designed to focus on Charles's visit to his newly established care with the Primary Care Physician (PCP) at the local community health center. He changed his previous provider as the PCP viewed him as "lazy" and did not take his ongoing struggle with depression and knee pain seriously. Being perceived as lazy due to his weight and persecuted for his sexual orientation, Charles became more depressed and self-conscious.

The story starts with Charles starting his journey to the health care Community Centre. The journey starts with the transit by bus for which he reaches the bus stop and waits for more than 15 mins. Because of the waiting, he becomes late for his appointment

and tries calling the clinic but due to insufficient balance, he could not make the call. This indicates that Charles could not afford a proper mode of transit and communication plan.

Once he reaches the clinic, he observes that there are no chairs for him to sit. This made him think that the clinic did not take into consideration people with high BMI (Body Mass Index). Although his knee hurts, he will have to stand till he is taken to the exam room. After waiting for some time, he wants to use the restroom and realizes that there are only restrooms for men and women. It made him think about his husband Dani who identifies himself as non-binary. The scenario describes the day-to-day struggle faced by people like Charles. He also interacts with the front desk receptionist regarding the appointment and his profile.

The next scene describes Charles's interaction with a medical assistant, which includes a general discussion about his visit and regular medical check-up. During the discussion, the medical assistant asks Charles about his emergency contact Dani and gets to know that Charles is gay. The MA acknowledges their pronouns and speaks a little about a friend of hers who is also gay to make Charles comfortable. During the check-up, the MA examines Charles and checks his weight and blood pressure which are very high. Meanwhile, Dani reaches the health care community center.

The MA introduces Charles and Dani to Dr. Jones. He is not Charles's primary physician, but he agreed to meet him as Charles's physician Dr. Avery is on vacation and his knee pain has worsened. After the introduction, Charles and Dani discuss Charles's struggles with weight management and recent gain due to a change in medication. The provider suggests some lifestyle changes and discusses different options.

4.3 Story Scene Creation

4.3.1 LGBTQIA+ Bus stop scene

In this scene, Charles reached the bus stop from where he needs to take a bus to reach the health care center. The scene is developed to contain the bus stop in a metropolitan city and some Non-Playing Characters (NPC). In the background, the scene contains trees and buildings to make it look realistic.

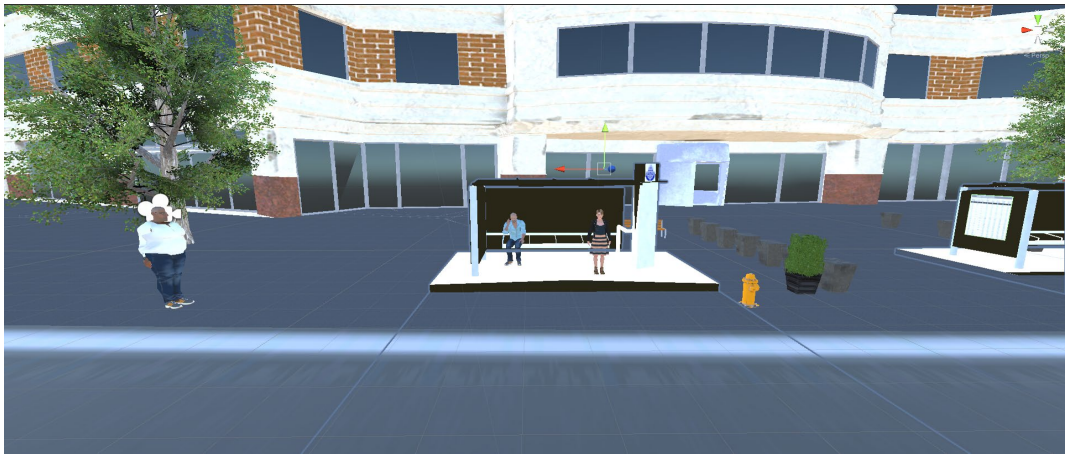

Figure 11: Bus Stop

Character's thoughts are denoted using "Thought bubbles" which can be seen in Fig – 12 and Fig – 13. The bus is late, so Charles thinks to call the clinic.

Figure 12: Charles's thought Bubble

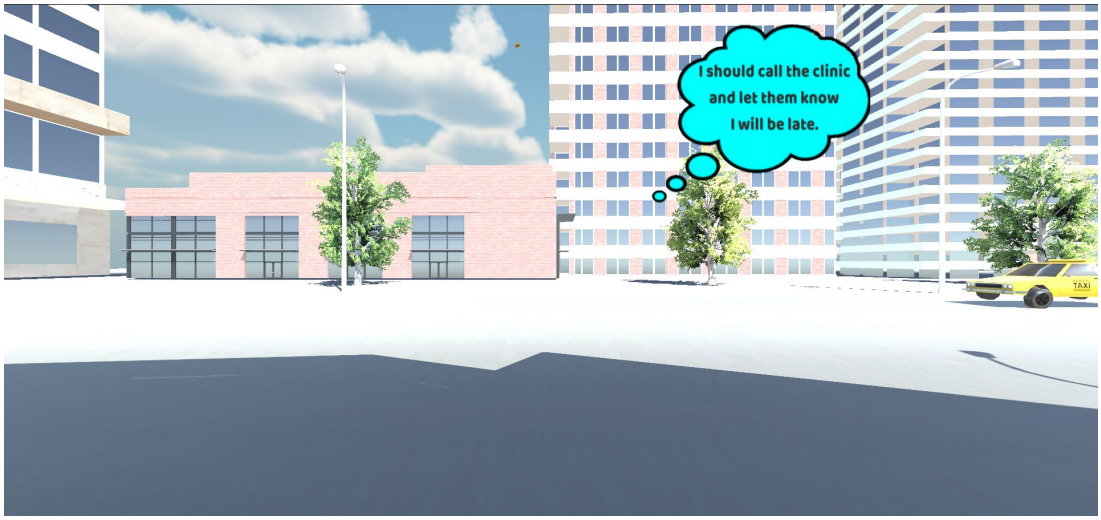

Figure 13: Thought Bubble with message

Unfortunately, his mobile service plan doesn't have sufficient balance and the call could not be connected.

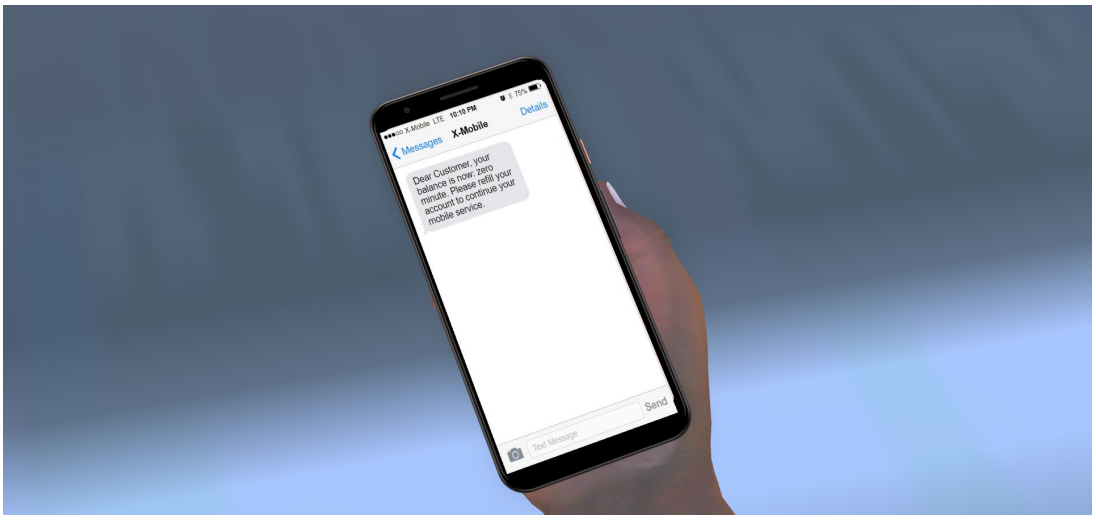

Figure 14: Charles's Mobile screen with the Text message

After some time, the bus arrives, and Charles boards the bus. There are some more NPCs added here as well to show the co-passengers.

Figure 15: Charles onboard the bus

4.3.2 Clinic Reception Scene

The scene starts with the reception of the clinic. The scene contains NPCs as patients waiting in the room. In the background, modern art pieces are added for ambiance.

Figure 16: NPCs in the waiting room

Charles couldn't find a seat where he can sit. Though his knee is hurting he decides to stand and wait.

Figure 17: Charles's thought bubble 1

He notices restrooms for men and women and the absence of restrooms for non-binary people like Dani.

Figure 18: Charles's thought bubble 2

Charles reaches out to the front desk receptionist, a 20-year-old Asian to enquire about his appointment. The front desk is designed with 2 receptionists – one of which is developed as part of this project and the other is an NPC. The background contains the hospital front desk cubical and records and file cabinets in the back. During the interaction,

the conversation is displayed in the text format using “Caption Box”. Caption Boxes are used to narrate the story in text for people with hearing impairment.

Figure 19: Receptionist and Charles’s conservation (Caption Box)

4.3.3 MA Hallway Scene

The Medical Assistant Barb, a 50-year-old white female comes out of the exam room into Hallway and calls out Charles’s name for an appointment who is waiting.

Figure 20: MA Hallway Scene

The MA hands out medical forms to fill out for Charles and conducts routine preliminary checkups. MA checks Charles's Blood Pressure, and weight and gets shocked by the numbers.

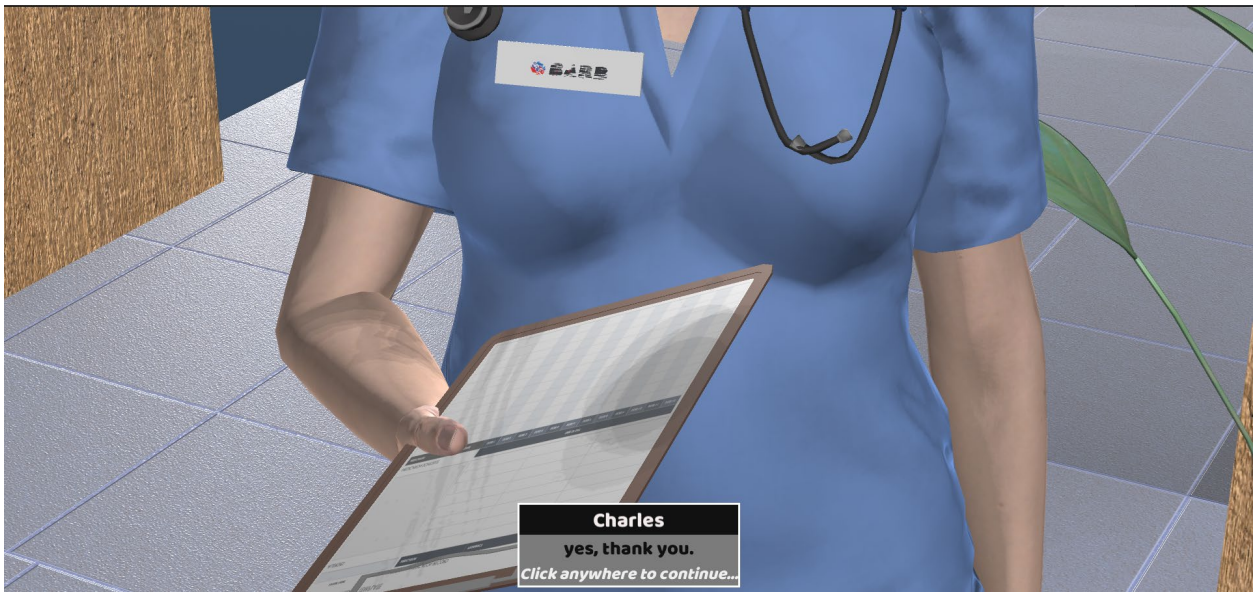

Figure 21: MA Holding a Clipboard

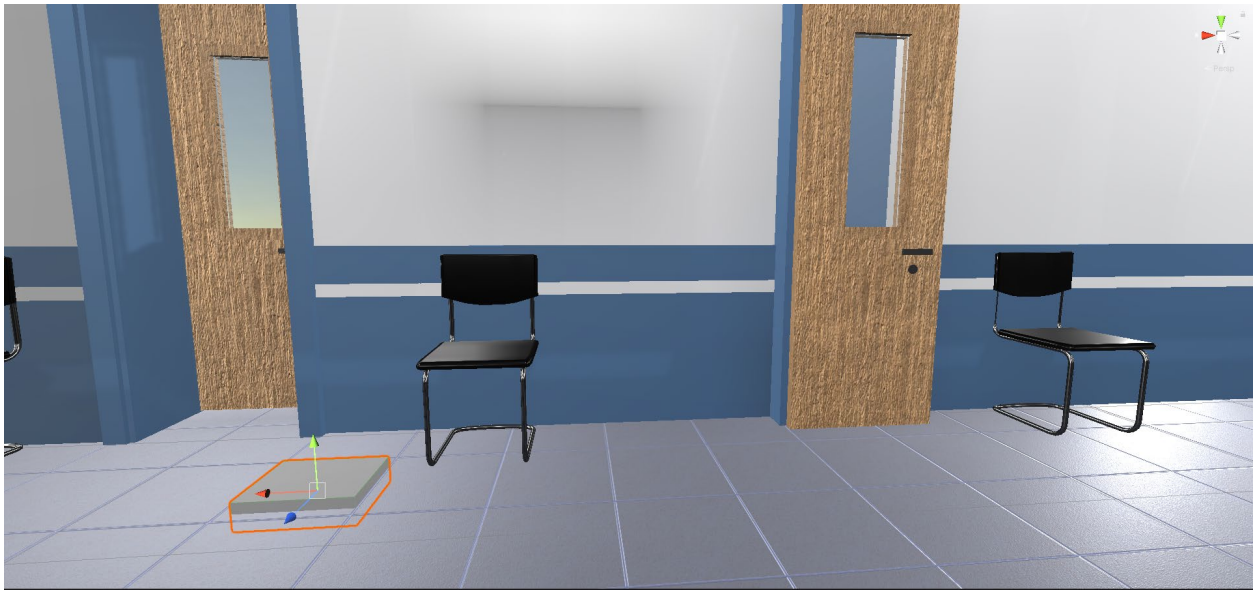

Figure 22: Weighing Machine

Figure 23: MA Checking Charles's Weight

4.3.4 Patient-Provider Conversation Scene

Charles and Dani meet with the physician Dr. Jones as Dr. Avery is on vacation that week.

Figure 24: Charles, Dani, and Provider interaction Scene with Caption Boxes

Charles introduces his spouse Dani and discusses the reason for their visit. They spoke to the doctor regarding the issues with Charles's weight, struggle with depression, and surgical options to manage his knee pain.

4.4 Demographics and Surveys

The objective of the simulation is not only to train the healthcare professionals but also to collect the data from Wright LIFE (Lifelike Immersion for Equity) to analyze it further. The data is collected through Survey Scenes incorporated into the gameplay. These scenes are developed to collect the participant's opinions about how they treat patients with biases like ASD (autism spectrum disorder) and LGBTQIA+. The provider's responses can be used to prepare for future training sessions and dissemination to the providers. In the case of both Charles and Ashley, the survey questionnaire and the options are the same. Only the reference would change.

There are three survey scenes. For these scenes, a different aspect ratio is set up to render the screen open on any device. The First Assessment starts after MA introduces Charles to the provider. The second Assessment starts after the completion of Charles and the provider's Interaction in the exam room. And the Final Assessment starts after the completion of the life Journey of Charles.

4.4.1 First and Second Assessment

The questionnaire starts with collecting providers' responses on how much *anxiety, frustration, and compassion* the providers feel when they are working with patients with biases on a scale of low to high.

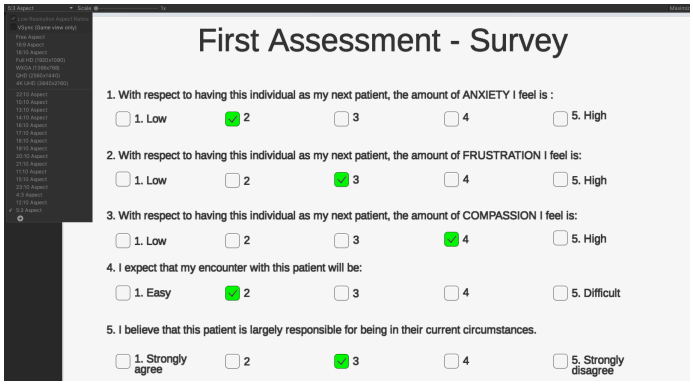

Figure 25: First assessment Survey scene 1

The Users participating in this survey need to answer all the questions without skipping any. If all the questions on the screen are not answered, the "Next" button will be disabled so that the users cannot move to the next page.

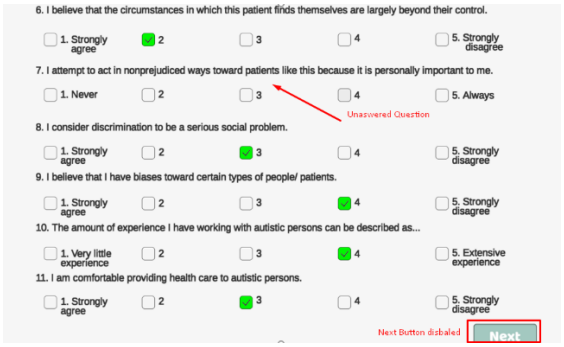

Figure 26: Survey scene with "Next" Button Disabled

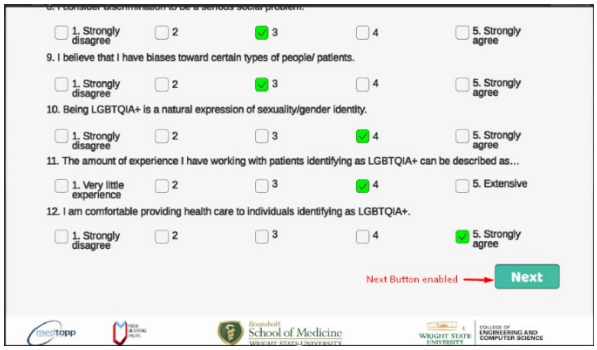

Figure 27: Survey scene with "Next" Button Enabled

The next part of the questionnaire is about the participant's encounter with the patient and their experience on a scale of easy to difficult. The questions touch up on the provider's opinion about the patient, discrimination, social problems, circumstances, comfortness, and bias towards them on a 5-point Likert scale of Strongly disagree to Strongly agree. Next, the questions are intended to assess patients' lack of prejudice on a scale of Never to Always. Finally, the questions inquire about the provider's previous experiences that they have acquired from collaboration with similar biases (ASD and LGBTQIA+) on a scale of Very Little Experience to Extensive.

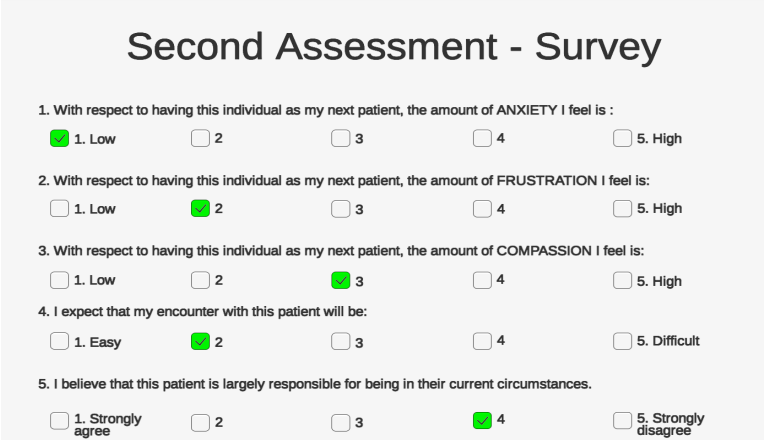

Figure 28: Second assessment Survey scene

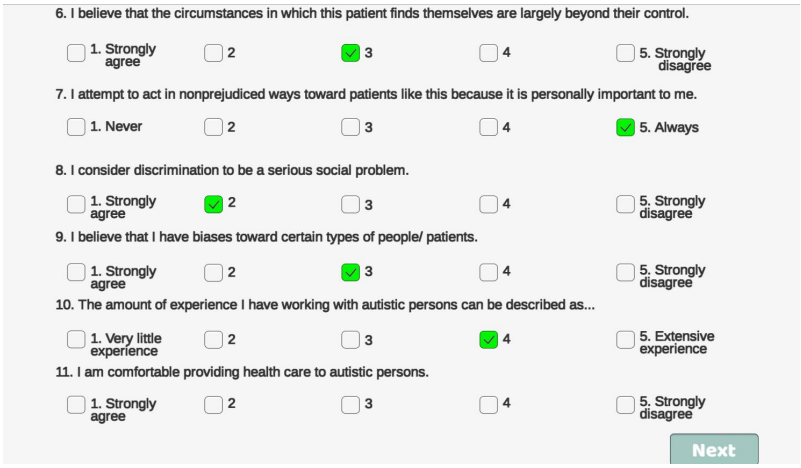

Figure 29: Second assessment Survey scene with Button disabled

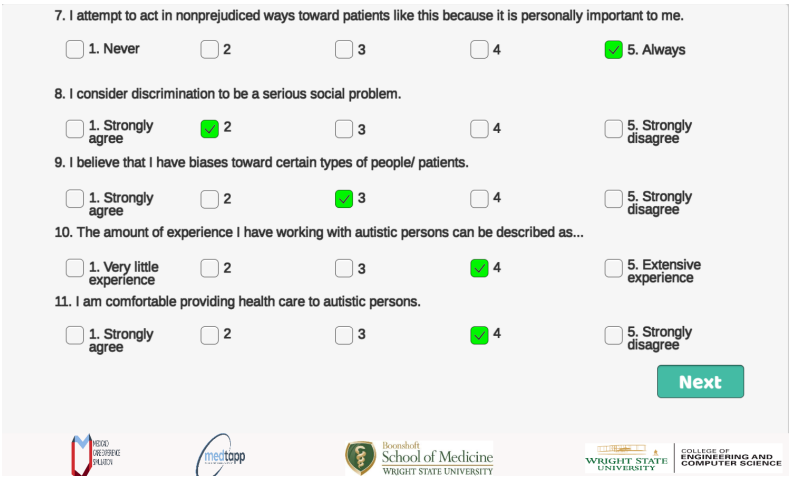

Figure 30: Second assessment Survey scene with Button Enabled

4.4.2 Final Assessment

In the final assessment, one-half of the questions are similar to the first and second assessments. The latter half of the questions are about providers' experience with the simulation and how they will apply the knowledge gained from the Wright LIFE simulation.

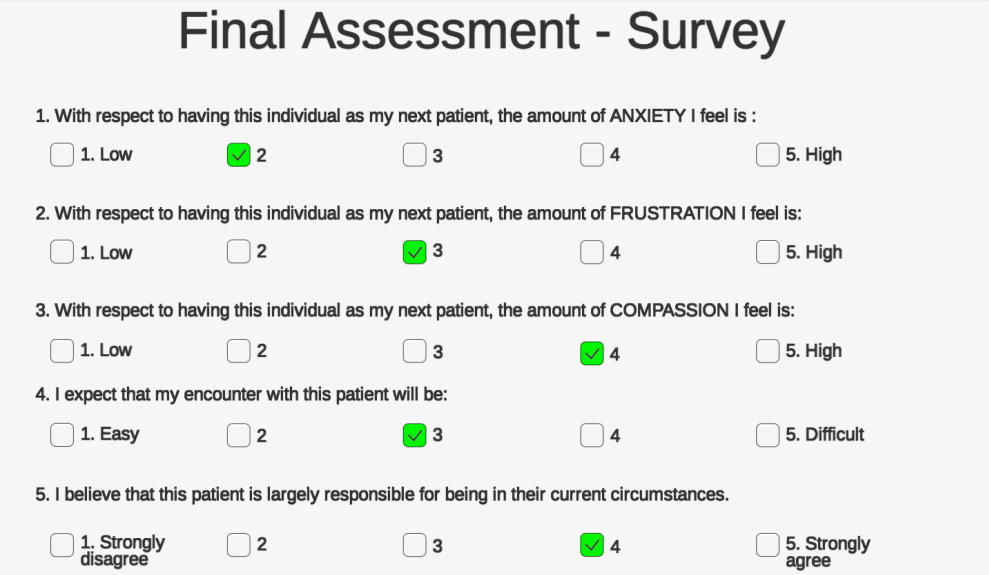

Figure 31: Final assessment Survey scene part 1

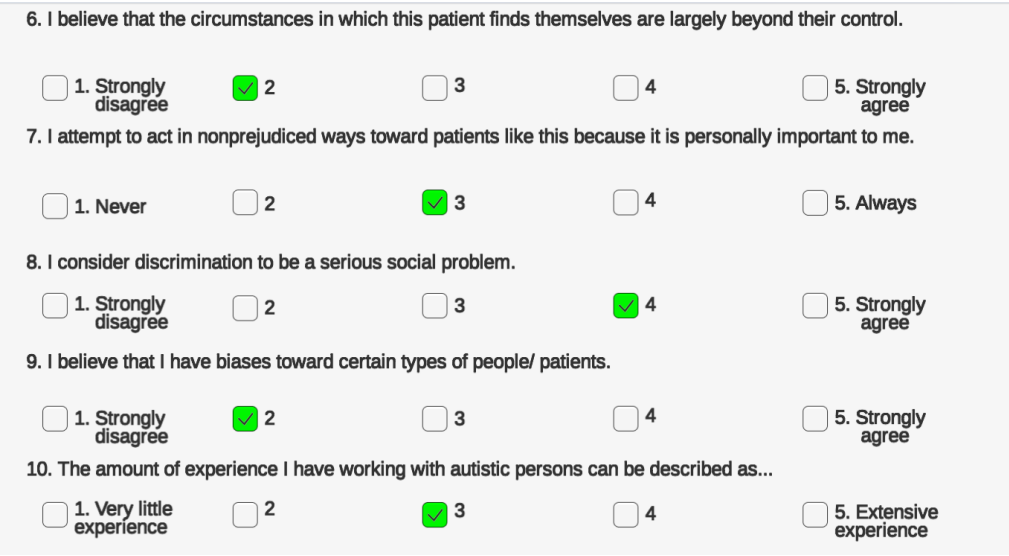

Figure 32: Final assessment Survey scene part 2

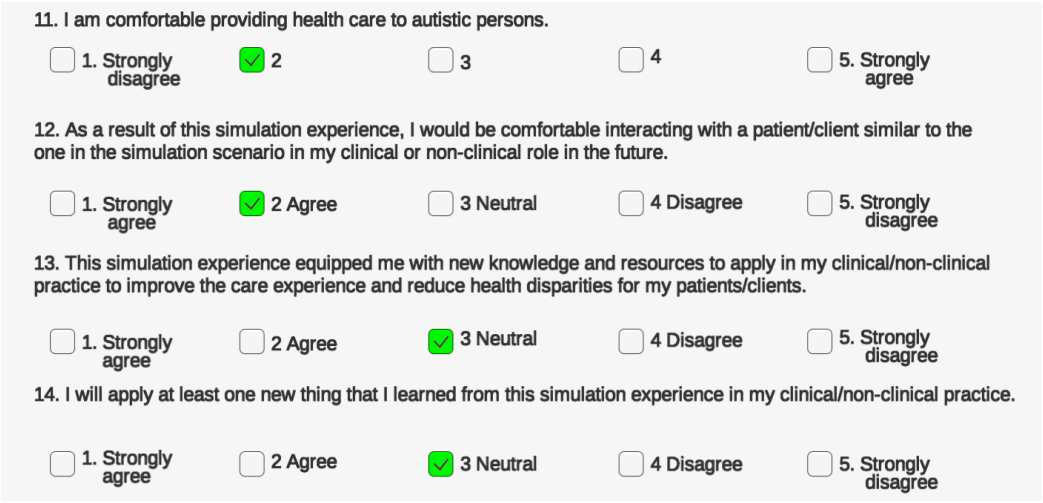

Figure 33: Final assessment Survey scene part 3

One of the questions is about the learning objective(s) listed in the scope before the questionnaire. Here, if they choose options other than strongly agree or agree the providers can give their opinion to answer which objective is not adequately addressed.

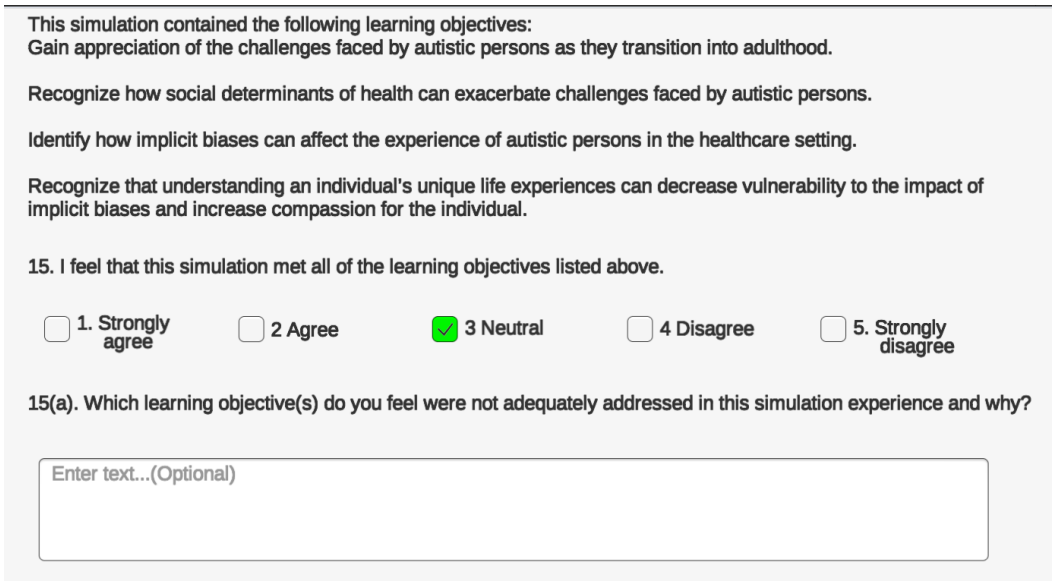

Figure 34: Final assessment Survey scene part 4

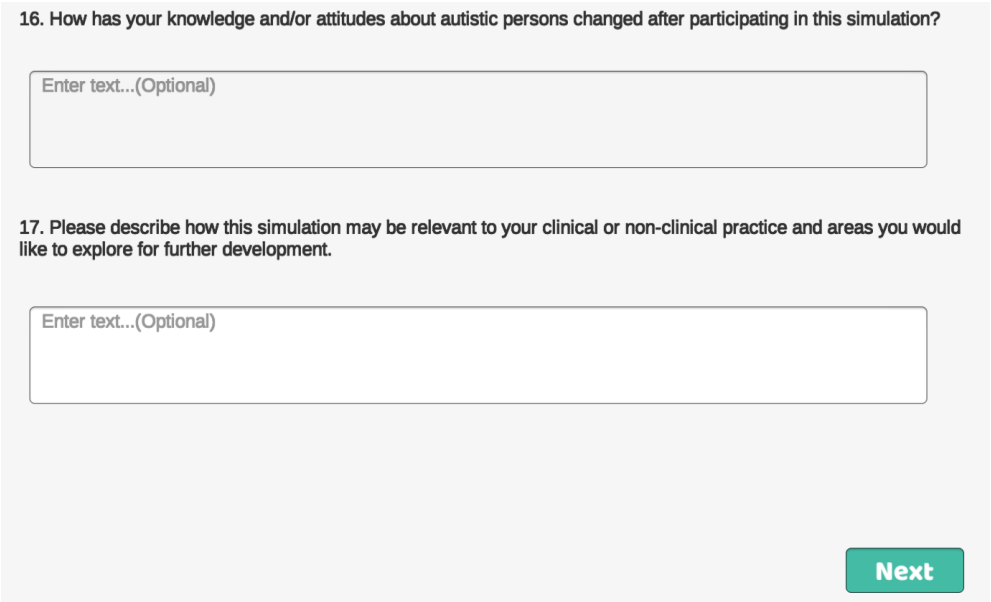

Figure 35: Survey scene Final assessment part 5

4.4.3 Thank you note and References

After completion of all the surveys, a Thank you note is placed to express gratitude for completing the simulation and providing their valuable opinions and feedback. The references are provided via a link to gain more knowledge about the simulation along with the sponsor information.

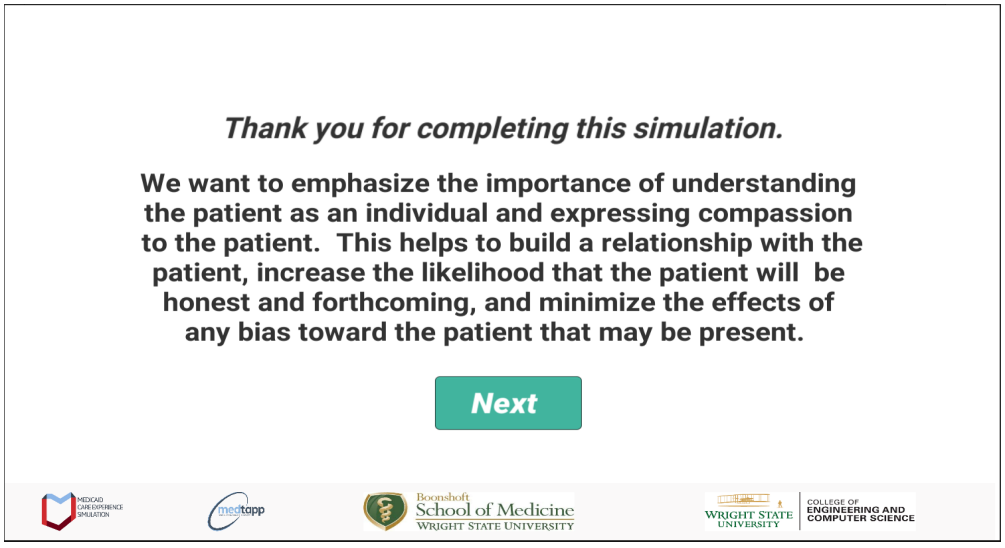

Figure 36: Survey scene - Thank you Note

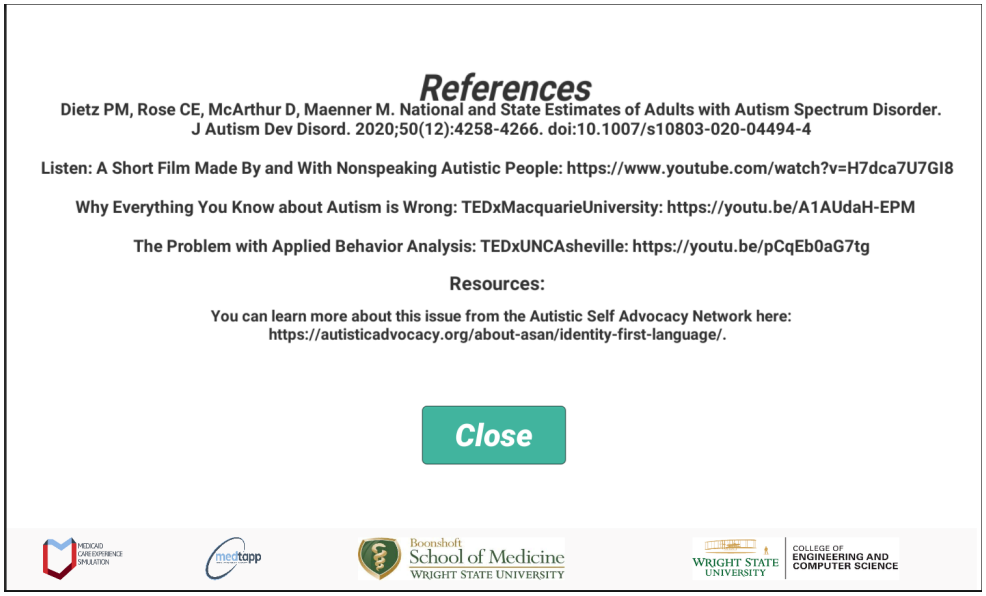

Figure 37: Survey scene - References

4.5 Demographic Scene

The demographic scene starts with an overview and objective of the simulation. Instructions are provided explaining the data collection policies to request the user's consent. The demographic information is needed to further analyze the data based on different categories.

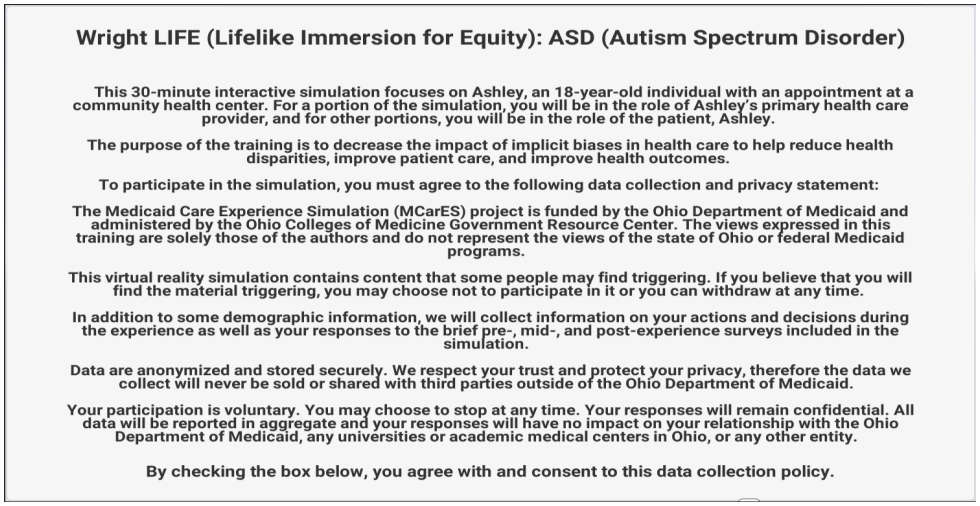

Figure 38: Demographic Scene - Summary Panel

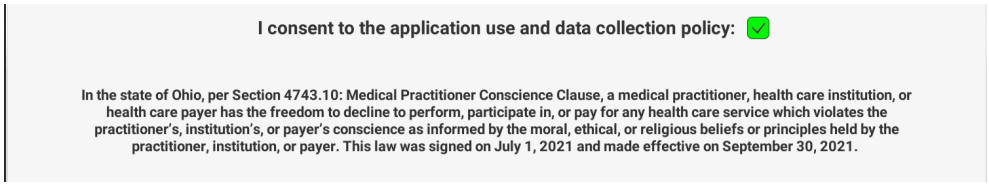

Figure 39: Survey Scene – Data Consent

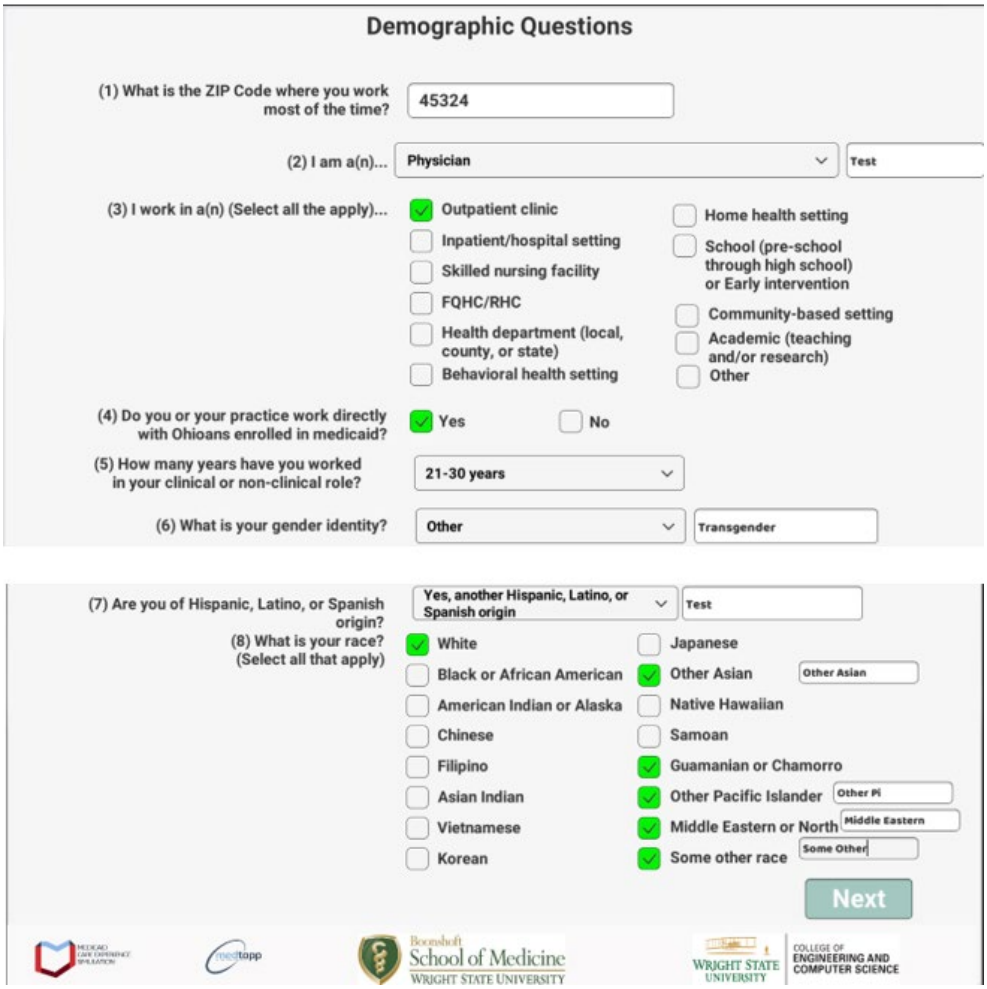

Figure 40: Survey Scene - Demographic questionnaire

5 DATA ANALYTICS AND VISUALIZATION

5.1 Data Collection

Data collection is the process of acquiring and processing information on relevant variables in a planned, methodical way that makes it possible to respond to specific research questions, test hypotheses, and assess results [39]. During the simulation, the application (SDOH) collects demographic information from users. The surveys are used to collect user responses. The survey responses are collected from pre- and post-clinical encounters. Finally, data are gathered through a questionnaire at the end of the simulation. All these responses which have been collected from users will be stored in file-based storage. The following screenshots show the data format of sample data collected from the simulation (SDOH).

session	epoch	age	gender	ethnicity	ethnicityOther	profession	professionOther	experience
	1567191152	27	male	asian_pacific_islander		other	Test	<1
1	1568236748	18	male	black_african_american		Nurse		1
1	1569896600	19	male	white_caucasian		Medical Assistant		<1
171	1558552143	27	male	white_caucasian		Physician		<1
171	1558552255	29	female	asian_pacific_islander		Physician		1
173	1558552189	30	female	white_caucasian		Physician		2
173	1558552931	30	female	white_caucasian		Physician		2
174	1558552245	28	female	white_caucasian		Physician		2
174	1558552260	38	male	latino_hispanic		Physician		12
175	1558552266	67	male	white_caucasian		Physician		38
177	1558552219	33	male	white_caucasian		Physician		3
178	1558552248	27	female	white_caucasian		Physician		<1
179	1558552242	28	female	white_caucasian		Physician		<1
179	1558552366	28	female	white_caucasian		Physician		<1
181	1558552220	34	male	white_caucasian		Physician		6
181	1558553081	34	male	white_caucasian		Physician		6
182	1558552243	26	male	asian_pacific_islander		Physician		1
182	1559828027	24	male	latino_hispanic		Pharmacist		7
184	1558552212	32	female	white_caucasian		Physician		1
185	1558552219	28	female	white_caucasian		Physician		2
186	1558552279	40	female	black_african_american		Physician		<1
188	1558552403	39	male	white_caucasian		Physician		1
189	1558552303	27	female	asian_pacific_islander		Physician		1
190	1558552354	34	female	white_caucasian		Physician		8
191	1558552402	27	male	white_caucasian		Physician		<1
294	1560894124	34	male	asian_pacific_islander		Physician		1
379	1561062470	22	male	asian_pacific_islander		Pharmacist		5
381	1561062627	24	male	asian_pacific_islander		Pharmacist		6
449	1561127491	48	female	white_caucasian		Other		25
450	1561127073	56	female	white_caucasian		Social Worker		32
452	1561128642	28	male	white_caucasian		Other		5
454	1561133143	23	female	latino_hispanic		Social Worker		4
457	1561130130	59	female	white_caucasian		Other		<1
458	1561130484	40	male	asian_pacific_islander		Social Worker		8

Figure 41: SDOH Simulation Dataset Part-1

Patient_look_time	MARTII_look_time	EHR_look_time	Total_look_time	pre1	pre2	pre3	pre4	pre5	pre6	pre7	pre8	pre9
115.626587	0	44.72064	370.926331	5	5	5	5	5	5	5	5	5
203.683075	0	11.71612	237.712036	3	3	3	3	3	3	3	3	3
257.9999	0	33.2199249	336.614746	4	3	2	4	5	3	3	4	4
133.1004	1.1337918	12.9073887	239.556625	3	3	4	4	3	2	4	4	4
229.387177	0	38.55028	286.614563	2	2	4	3	4	1	5	3	5
224.355881	0.333477974	0	276.419159	3	3	4	4	2	1	5	2	5
214.665085	0	21.1513042	332.3742	3	3	5	2	4	2	4	4	5
243.086029	0	0	312.401123	3	2	4	2	3	2	3	4	5
203.4117	0	0	275.356628	1	1	4	1	4	2	4	3	3
150.247284	0	12.5404072	263.210724	4	3	3	5	3	1	5	4	4
241.291885	0	0	304.260437	2	2	5	2	4	1	4	4	5
210.191025	0	0	285.8038	3	2	4	3	4	1	2	3	5
120.736534	0.133323	21.4117012	290.590027	1	1	5	1	3	1	3	3	5
196.730286	0	2.46725726	248.169067	1	2	4	1	2	1	5	4	5
235.919891	72.61338	0	328.899323	3	3	4	3	4	2	2	3	4
241.929077	0	11.2093782	264.302765	4	4	3	4	5	1	5	4	5
71.32839	13.3631277	0	321.5962	1	2	5	1	3	1	5	5	5
192.368484	0.063881	46.1729355	326.9492	3	1	5	1	3	1	4	3	5
240.990326	1.23404992	19.6426277	304.5855	2	3	3	4	3	3	3	3	4
250.192429	0	18.9292145	373.2903	3	2	4	3	4	1	4	4	4
255.644318	0	0	315.9622	4	4	2	5	5	2	3	5	5
211.915787	0	0	328.336	5	5	5	1	5	1	5	5	5
224.386429	0	20.1614971	273.5765	1	1	1	1	1	1	1	1	1
235.240128	0	21.3189945	288.3152	4	4	4	4	4	4	4	4	4
289.2013	0	18.5003757	322.1081	5	5	5	5	5	5	5	5	5
126.946304	0	0	281.721741	2	1	4	2	3	2	3	2	4
208.675613	0	19.56993	286.311	3	1	5	1	4	1	5	2	5
233.34375	0	11.7236071	260.919281	4	5	4	5	5	3	3	4	5
282.34436	0	19.8536	336.725433	2	1	5	1	3	1	4	5	5
157.087189	0	16.15974	284.183319	1	1	5	1	2	1	4	5	5
234.797134	0	0	305.668427	1	1	4	1	3	1	5	2	5
268.542	0	16.129221	294.9805	4	4	4	1	3	1	3	1	4
193.162	0.051288	0	267.708557	2	2	4	1	3	1	5	5	5
233.870346	0	12.1858883	256.683136	3	1	3	1	4	1	5	2	4
194.75296	0	19.3373337	291.299622	3	2	3	3	3	3	3	3	5
221.705246	0	0	278.9493	3	2	4	1	3	1	4	2	4

Figure 42: SDOH Simulation Dataset Part-2

post1	post2	post3	post4	post5	post6	post7	post8	post9	ageGroupMedPat3	post10	post11	post12	genderOther	LicensedMedicaidProviderOhio	scenario	FirstName	LastName	OfficeZipCode
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					Yes		Testing	Testing	45435
3	3	3	3	3	3	3	3	3	I do not see Medicaid patients					No		Jarhal	Duncan	45431
5	5	5	4	1	5	5	5	5	I do not see Medicaid patients					Yes		Guguf	Guyfyufu	34567
5	5	5	5	5	5	5	3	3	I do not see Medicaid patients					No		Vhvgh	Chjvj	45678
5	4	4	4	4	4	4	4	4	I do not see Medicaid patients					Yes		Mit	Vas	43567
4	4	4	4	5	5	5	5	5	I do not see Medicaid patients					No		Fes	Smith	95014
1	1	1	1	1	1	1	1	1	I do not see Medicaid patients					No		Igigi	Yiyi	45345
3	3	3	3	3	3	3	3	3	I do not see Medicaid patients					No		Hhh	Gugg	54545
1	2	2	3	3	4	4	4	4	I do not see Medicaid patients					No		Test	Test	45324
1	2	2	2	3	3	3	4	5	I do not see Medicaid patients					No		Test	Test	45435
4	3	4	4	3	2	1	1	5	I do not see Medicaid patients					No		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Test	Test	45435
1	2	2	3	3	2	1	2	3	I do not see Medicaid patients					No		Hhh	Jjj	55667
1	1	2	2	3	2	3	2	3	I do not see Medicaid patients					No		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					Yes		Fhfyh	Gufu	54321
4	4	4	4	4	4	4	4	4	I do not see Medicaid patients					No		Test	Test	45435
1	2	1	2	3	2	1	3	4	I do not see Medicaid patients					No		Jkbfakjb	Safjka	32452
1	2	3	4	5	4	3	3	4	I do not see Medicaid patients					No		Jvjbj	Hknkj	56432
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		T	E	45435
5	5	5	5	5	1	1	1	1	I do not see Medicaid patients					No		Ytr	Ulo	32456
2	1	5	3	4	1	5	4	5	I do not see Medicaid patients					Yes		Fhfufyhfu	Uugug	45646
4	1	5	3	2	1	4	2	5	I do not see Medicaid patients					No		T	T	45435
2	2	4	3	4	2	3	5	5	I do not see Medicaid patients					No		Angela	Clements	45415
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Dfg	Jkl	76543
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Tester	Tester	45435
1	1	1	1	1	1	1	1	1	I do not see Medicaid patients					Yes		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		Test	Test	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		T	H	45434
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					Yes		Hghhg	Fhfhf	45324
1	1	1	1	1	1	1	1	1	I do not see Medicaid patients					Yes		Test	Test	54632
5	1	5	5	5	5	4	5	5	I do not see Medicaid patients					No		Test	Test	54545
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		T	T	45435
5	5	5	5	5	5	5	5	5	I do not see Medicaid patients					No		T	H	45435
2	1	5	1	3	1	5	5	5	I do not see Medicaid patients					No		Test	Test	12345

Figure 43: SDOH Simulation Dataset Part-3

workInOther	perMedPatServe	ageGroupMedPat0	ageGroupMedPat1	ageGroupMedPat2	ethnicity1	ethnicity2	ethnicity3	ethnicity4	ethnicity5	ethnicity6
Temp2	My practice does not see Medicaid patients	Children	Adults	Older Adults	White (non-hispanic)	Black	Hispanic Latino	American Indian or Alaska Native	Asian Indian	Chinese
Test	Greater than 30%	Children	Adults	Older Adults	White (non-hispanic)	Black	Hispanic Latino	American Indian or Alaska Native	Asian Indian	Chinese
Testing	Less than or equal to 30%	Children	Adults	Older Adults	White (non-hispanic)					
Testing	Less than or equal to 30%	Children	Adults	Older Adults	White (non-hispanic)	Black	Hispanic Latino	American Indian or Alaska Native	Asian Indian	Chinese
Abc	Less than or equal to 30%	Children	Adults	Older Adults	White (non-hispanic)	Black	Hispanic Latino	American Indian or Alaska Native	Asian Indian	Chinese
Testing	My practice does not see Medicaid patients					Black				
Academic	Less than or equal to 30%		Adults		White (non-Hispanic)					
University	Less than or equal to 30%									Chinese
Development	My practice does not see Medicaid patients					Black				
Student	Greater than 30%		Adults		White (non-Hispanic)					
University	My practice does not see Medicaid patients									Chinese
University	My practice does not see Medicaid patients				White (non-Hispanic)					
Managed Care	Greater than 30%	Children	Adults	Older Adults	White (non-Hispanic)					
Test	My practice does not see Medicaid patients								Asian Indian	
Test	My practice does not see Medicaid patients		Adults							
Test	My practice does not see Medicaid patients		Adults							
Test	My practice does not see Medicaid patients	Children						American Indian or Alaska Native		
Test	Less than or equal to 30%		Adults							Chinese
Test	Greater than 30%	Children	Adults	Older Adults						
Community Clinics	Less than or equal to 30%		Adults		White (non-Hispanic)					
University	My practice does not see Medicaid patients				White (non-Hispanic)					
Dod	Less than or equal to 30%		Adults			Black				
State Agency	Greater than 30%				White (non-Hispanic)					
Ohio Department Of Medicai	Greater than 30%				White (non-Hispanic)					
Primary Care Association	My practice does not see Medicaid patients				White (non-Hispanic)					
University	My practice does not see Medicaid patients				White (non-Hispanic)					
t Academic Research Facility	My practice does not see Medicaid patients				White (non-Hispanic)					
State Medicaid Dept	My practice does not see Medicaid patients				White (non-Hispanic)					
University	Greater than 30%	Children	Adults	Older Adults	White (non-Hispanic)					
Oneil Center	My practice does not see Medicaid patients								Asian Indian	
Vision Dept	Less than or equal to 30%	Children	Adults		White (non-Hispanic)					
Lab	Less than or equal to 30%	Children	Adults	Older Adults	White (non-Hispanic)					
Lab	Less than or equal to 30%	Children	Adults	Older Adults	White (non-Hispanic)					
Mytesting	Greater than 30%				White (non-Hispanic)	Black	Hispanic Latino	American Indian or Alaska Native	Asian Indian	Chinese
Pharmacy	Greater than 30%	Children	Adults	Older Adults	White (non-Hispanic)					
University	My practice does not see Medicaid patients				White (non-Hispanic)					

Figure 44: SDOH Simulation Dataset Part-4

ethnicity7	ethnicity8	ethnicity9	ethnicity10	ethnicity11	ethnicity12	ethnicity13	ethnicity14	otherPacificText	ethnicity15	otherAsianText	ethnicity16	someOtherText
Filipino	Japanese	Korean	Vietnamese	Native Hawaiian	Guamanian or Chamoro	Samoan	Other Pacific Islander	Temp3	Other Asian	Temp4	Some other race	Temp5
Filipino	Japanese	Korean	Vietnamese	Native Hawaiian	Guamanian or Chamoro	Samoan	Other Pacific Islander	Test	Other Asian	Test	Some other race	Test
							Other Pacific Islander	Testing	Other Asian	Testing	Some other race	Testing
Filipino	Japanese	Korean	Vietnamese	Native Hawaiian	Guamanian or Chamoro	Samoan	Other Pacific Islander	Testing	Other Asian	Testing	Some other race	Testing
Filipino	Japanese	Korean	Vietnamese	Native Hawaiian	Guamanian or Chamoro	Samoan						
		Korean					Other Pacific Islander		Other Asian		Some other race	
						Samoan	Other Pacific Islander		Other Asian		Some other race	
						Samoan			Other Asian	Test		
Filipino	Japanese	Korean	Vietnamese		Guamanian or Chamoro	Samoan	Other Pacific Islander	Mytesting	Other Asian	Mytesting	Some other race	Mytesting

Figure 45: SDOH Simulation Dataset Part-5

The data from the application storage is accessed through secured web login. Once logged in, the system provides the data in a .csv (Comma separated values) formatted file.

The simulation data is directly imported into the Microsoft Excel sheet using the in-built data access from the Web method of the MS Excel application. This Excel file can be imported to Tableau. This eliminated the manual login to the server and downloading of the file. Apart from importing the data from the Web into Excel, it also allows to set up data refresh. The data is imported automatically and get updated based on the configured Refresh options. In this simulation, the data will get refreshed every 60 minutes.

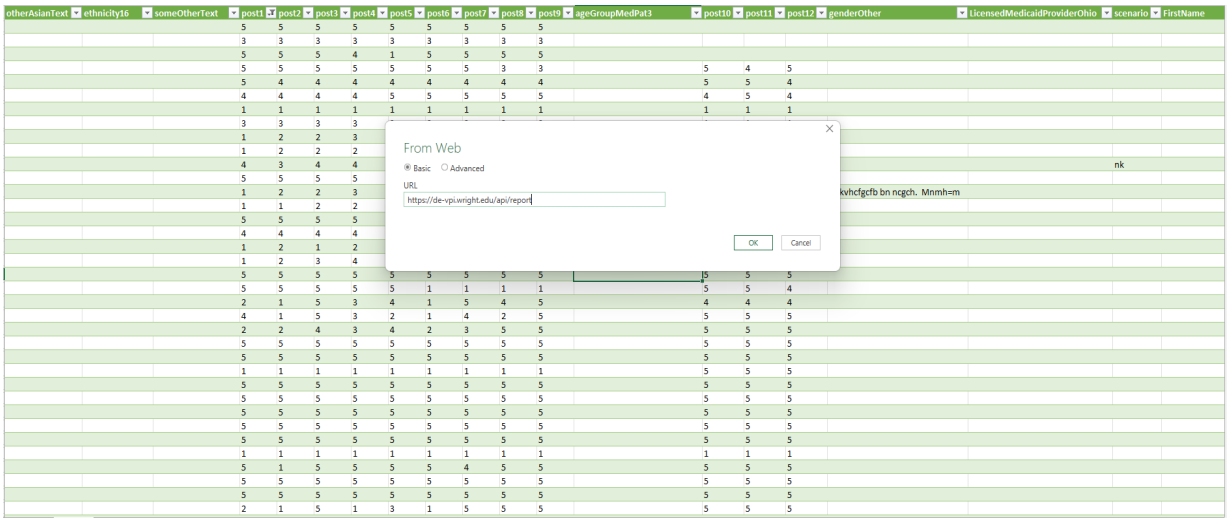

Figure 46: SDOH web data access in Excel

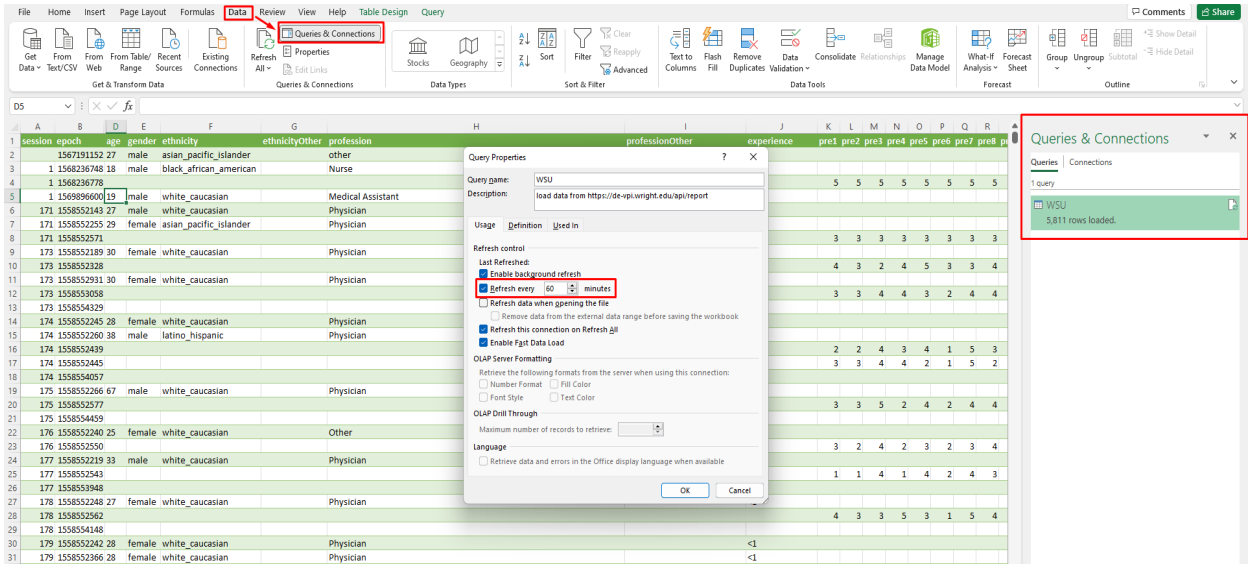

Figure 47: Data Refreshing properties for the connection

5.2 Data Analysis

Data analysis is the methodical application of logical and/or statistical tools to interpret and display, summarize, and assess data [40]. Analyzing the Pre- and post-questionnaire responses is the key objective for the analysis performed on the data. Before visualizing the data, the data is formatted and filtered to produce better graphs. The data consisted of several rows with raw, unprocessed data which is difficult to generate visualizations. So, data is processed using filters and formulas.

"COUNTIF" function is used to calculate the count of each option of every question.

Example: Option Count C(x) = COUNTIF (report! K: K, 1). Here K: K is the range of rows in column K in the excel worksheet '*report'*. The function counts the occurrence of (option) value 1 in row K. This can be noted in mathematical representation using the Kronecker delta (δij), defined to be 1 if i=j and 0 if i≠j,

$$C(x)= \sum_{i=1}^{n} \delta(x, Si)$$

x: option (1,2,3,4,5)
n: number of user responses
Si: user response value

The SUM function is used to calculate the total count of responses (from option1 to option 5) for each question.

$$\text{Total responses } T(x) = \sum_{x=1}^{n=5} C(x)$$

Here the function T(x) gives the sum of values of Option count for all the questions.

The mathematical formula (Column1/Column2) *100 is used to calculate the percentage.

Example: Option 1 percentage = (Option Count / Total responses) * 100For each question in pre- and post-questionnaires, applied calculations in Excel to summarize the total count of responses. Refer to Fig: 48. Options here represent the choices provided for each

question, such as low to high, Strongly Disagree to Strongly Agree, Never to Always and Easy to Difficult. The total number of responses received for each question (Ex: Pre-1, Pre-2, Post-1, and post-2, etc.) are calculated as shown. Based on the count, it is observed that the total count of responses for pre- and post- are not the same. Users took a pre-survey but may not have participated in the post-survey. To compare the responses, calculated the percentage (%) of each option for all questions individually.

Questions	OPTION1	OPTION2	OPTION3	OPTION4	OPTION5	Total responses	Option 1 percentage	Option 2 percentage	Option 3 percentage	Option 4 percentage	Option 5 percentage
pre-1(Anxiety)	348	119	175	105	375	1122	31.01604278	10.60606061	15.59714795	9.35828877	33.42245989
pre-2(Frustation)	348	205	104	78	387	1122	31.01604278	18.27094474	9.26916221	6.951871658	34.49197861
pre-3(Compassion)	189	79	141	186	527	1122	16.84491979	7.040998217	12.56684492	16.57754011	46.96969697
pre-4(Encounter)	285	193	171	108	365	1122	25.40106952	17.20142602	15.24064171	9.625668449	32.5311943
pre-5(Currrent Circumstances)	199	165	224	134	400	1122	17.73618538	14.70588235	19.96434938	11.942959	35.65062389
pre-6(Circumstances)	312	194	151	110	355	1122	27.80748663	17.29055258	13.45811052	9.803921569	31.6399287
pre-7(biases)	152	97	241	196	436	1122	13.54723708	8.645276292	21.47950089	17.4688057	38.85918004
pre-8(negative thoughts)	212	124	124	146	516	1122	18.89483066	11.0516934	11.0516934	13.01247772	45.98930481
pre-9(nonprejudiced)	150	81	98	124	669	1122	13.36898396	7.219251337	8.734402852	11.0516934	59.62566845
post-1(Anxiety)	118	62	56	28	152	416	28.36538462	14.90384615	13.46153846	6.730769231	36.53846154
post-2(Frustation)	99	89	51	23	154	416	23.79807692	21.39423077	12.25961538	5.528846154	37.01923077
post-3(Compassion)	34	11	35	59	277	416	8.173076923	2.644230769	8.413461538	14.18269231	66.58653846
post-4(Encounter)	108	66	46	46	150	416	25.96153846	15.86538462	11.05769231	11.05769231	36.05769231
post-5(Currrent Circumstances)	50	66	81	51	168	416	12.01923077	15.86538462	19.47115385	12.25961538	40.38461538
post-6(Circumstances)	125	72	33	38	148	416	30.04807692	17.30769231	7.932692308	9.134615385	35.57692308
post-7(biases)	44	23	60	60	229	416	10.57692308	5.528846154	14.42307692	14.42307692	55.04807692
post-8(negative thoughts)	43	41	37	57	238	416	10.33653846	9.855769231	8.894230769	13.70192308	57.21153846
post-9(nonprejudiced)	49	11	19	33	304	416	11.77884615	2.644230769	4.567307692	7.932692308	73.07692308
POST-10(training experience)	47	14	35	76	241	413	11.38014528	3.389830508	8.474576271	18.40193705	58.3535109
POST-11(effective learning platform)	27	30	35	82	239	413	6.537530266	7.263922518	8.474576271	19.85472155	57.86924939
POST-12(comfortable)	31	23	61	71	227	413	7.506053269	5.569007264	14.76997579	17.19128329	54.96368039
POST-13(feedback)	3	1	24	41	43	112	2.678571429	0.892857143	21.42857143	36.60714286	38.39285714

Figure 48: Response Summary

After analyzing the questionnaire data, started analyzing with more attributes based on demographics such as profession, gender, and experience. By analyzing the data based on these attributes, conclusions regarding the participants like gender group, experience level, and profession can be drawn. This helps to identify the user base who contributes to the survey more. Experience levels are not categorized. If experience is used directly in the analysis, it produces too many results ranging for each experience level. To make better sense of the data, created a calculated field to categorize the user experience level. These calculations can be changed as per the requirement. In Fig 49, users with experience of less than 1 year are considered “Beginners”, users with experience of 2 to 14 years are grouped as “Associates” and users with 15 years or more experience are categorized as “Analysts”.

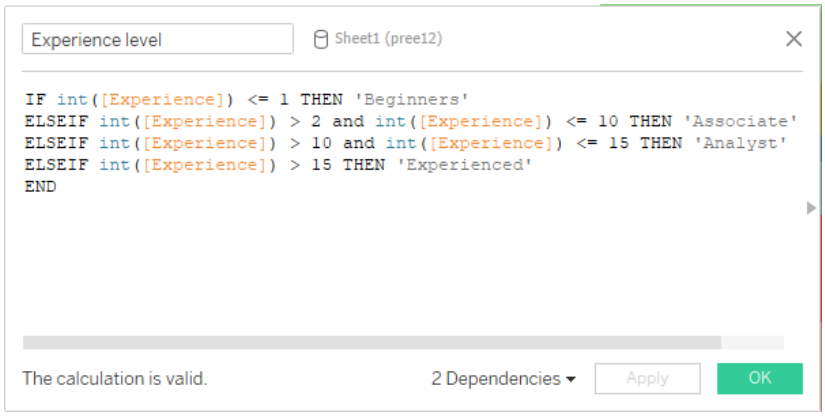

Figure 49: Field for Calculating Experience Level

While analyzing the data, Parallel Coordinate graphs are used. As the dataset increases with time, it becomes large and complex to analyze and interpret. Large number of dimensions not only cause clutter, but it would also be hard to visualize the data. The major advantage of using Parallel Coordinates is the ease of plotting high-dimensional data as 2-D visualization. Each row of the data table is shown as a line or profile in a parallel coordinate plot. This allows for making a comparison between individual observations. this visualization technique can be extended to large multivariate datasets and have an advantage over orthogonal coordinates. Due to the orthogonal layout of the coordinate axes, scatterplots can only display a maximum of three dimensions; however, a parallel coordinate enables the depiction of many dataset properties as the parallel coordinate plot can handle arbitrarily many dimensions because the axes are drawn as parallel lines and then data points connected by line segments. A data point, i.e. a row in the database, is represented as a line with each data element being plotted on its axis. These axes are visualized as parallel lines which is why this visualization is named parallel coordinates. Due to this, parallel coordinate plots resemble line charts in appearance, but the process of converting data into a plot is very different [41]. The scale on the axes of a parallel

coordinate plot is normalized. Since the data is presented here as a line, it is simple to see the trend that the data entries indicate. Parallel Coordinate plots are recommended to represent high-dimensional data [42]. To examine the trend between two separate axes, the axes can be positioned accordingly by changing the order. Below in Fig: 50 is an example of such a Parallel Coordinates Plot.

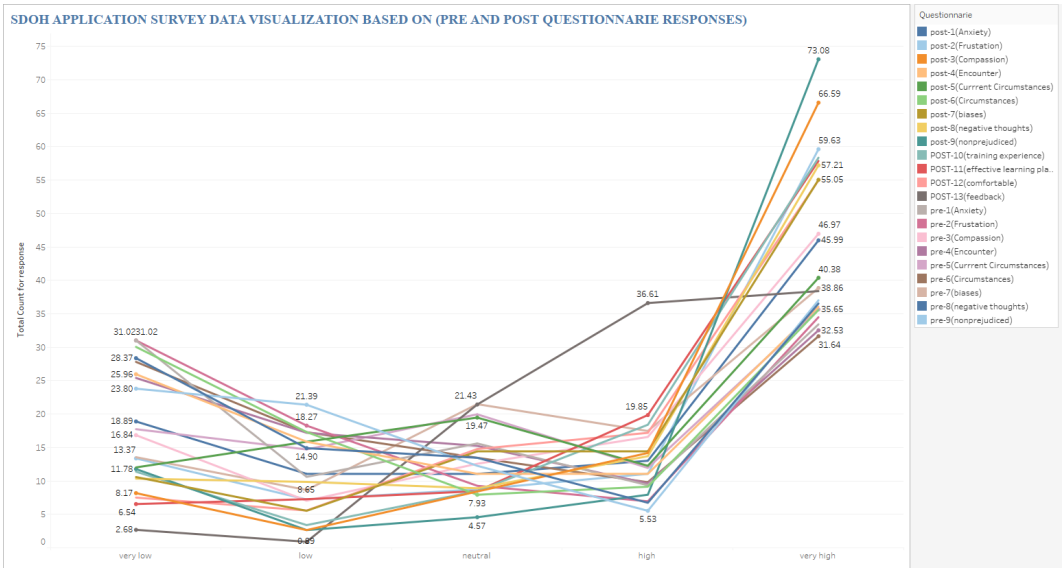

Figure 50: SDOH Application Questionnaire Visualisation - parallel coordinates

For analyzing the questionnaire data, using a Parallel coordinate plot is a better choice compared to a bar graph. As all the questionnaire trends needed to be visualized in a single visualization, it was possible only with a parallel coordinates plot. As shown in Fig: 51, it was not possible to generate a single graph that includes all the questions from the questionnaires whereas the parallel coordinate plot is capable of handling the number of dimensions.

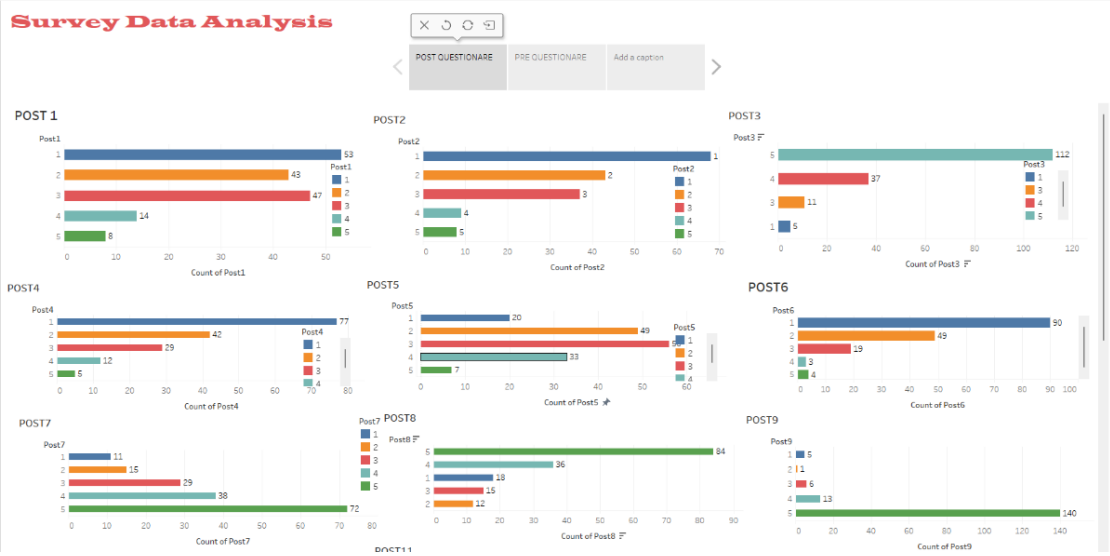

Figure 51: SDOH Application questionaries visualization - Bar Graph

5.3 Data Visualization

Data visualization is the representation of data using common graphics, such as charts, plots, infographics, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand [43]. The analyzed data is visualized by using the reporting tool called Tableau. Tableau has easy options to import the file-based data to develop the visualization.

5.3.1 SDOH Application Survey Data Visualization based on Questionnaire

In the SDOH visualization data is visualized based on the user's responses to pre- and post-questionnaires. The visualizations are built also based on gender, profession, and experience as attributes. Starting with visualizing SDOH survey data using a parallel coordinate plot, the participants' opinion choices are plotted to understand how the users responded. For each question in the pre-and post-questionnaire, the plot is generated to show whether users' responses contributed to higher or lower levels of parameters. Here, the parameters are "the intended quality or behavior" we are trying to understand through

the question – such as compassion, anxiety, and frustration. By using parallel coordinates, responses to all the questions are visualized in a single plot to provide an idea of how the data is trending.

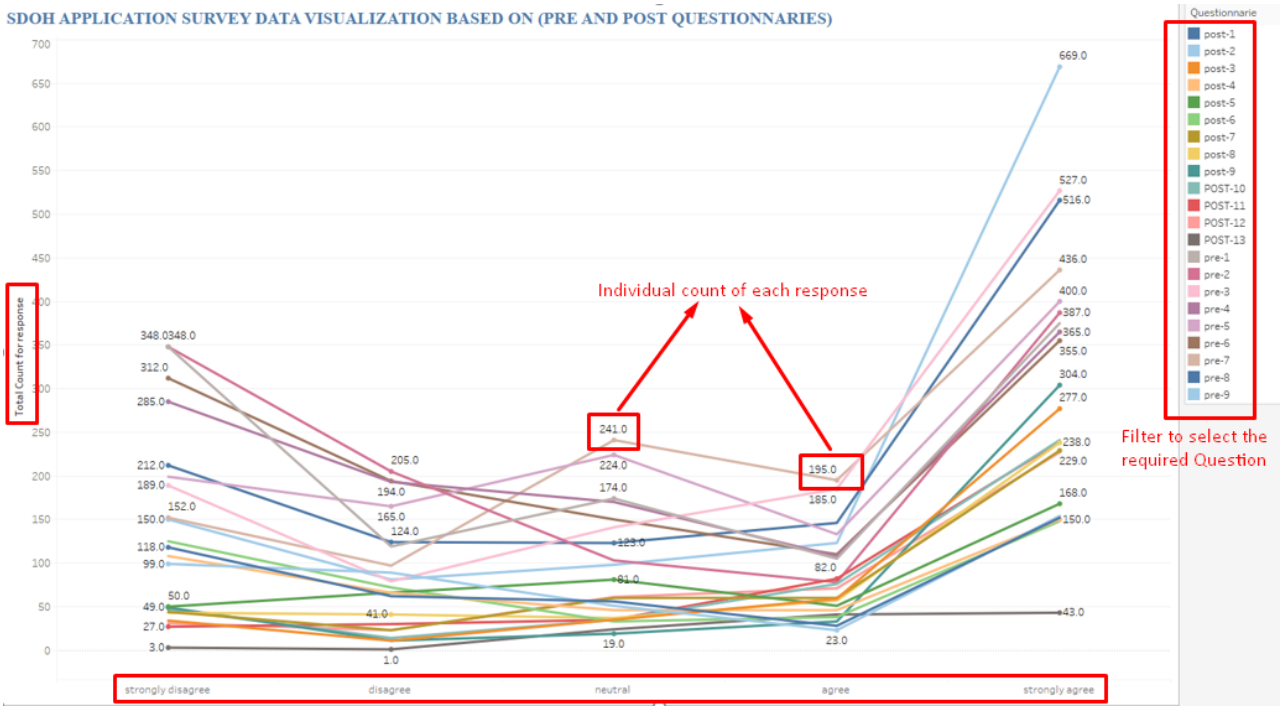

Figure 52: SDOH Application Dataset Visualization for Questionnaires

The above Fig 52 shows the parallel coordinates plot for SDOH data, visualized for questionnaire and responses. In the plot, the 5 options of the questions (Low to high, Strongly Disagree to Strongly Agree) are shown on X-axis and the count of each response is mapped on Y-axis. Each line represents an individual question, and each data point shows the count for the total number of times each option is selected. The graph shows the trend for all the questions in a single chart to provide a comparison.

If the user wants to see questionnaire responses individually, they can see them by simply clicking on that questionnaire. As shown in Fig 53., the right-hand side filter in the

graph shows the selected question and the plot shows a single line chart showing the count for each option.

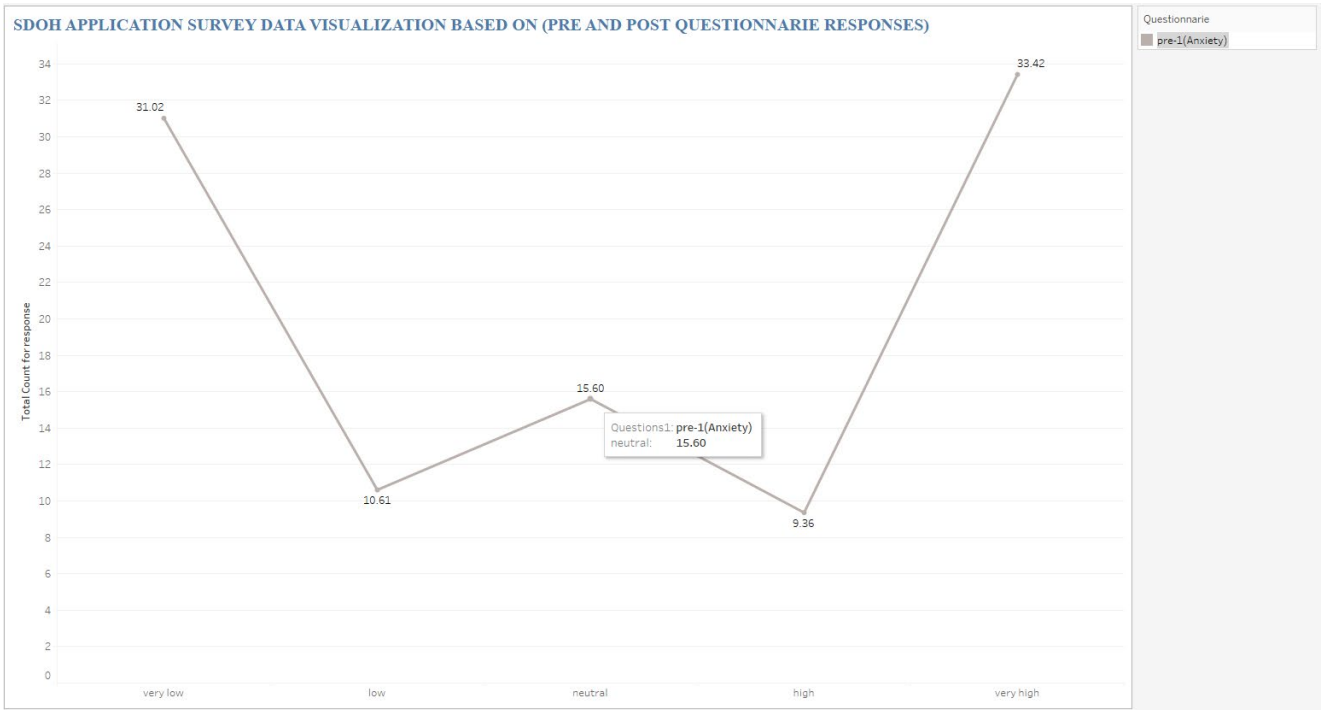

Figure 53: SDOH Application Dataset Visualization for questionnaire with Filters

5.3.2 SDOH Data Visualization Based on Experience and Age

After the Questionnaire data, visualized the data based on the user's experience and age. Among the providers who participated in the simulation and survey, there is a wide variety of group combinations based on gender, age, profession, and experience. Each generation responds to technology differently. When we plot age, it could help us segregate the data for each generation. By knowing the experience level, we can estimate who showed more interest in the learning program. For the Experience attribute, created a calculated field based on the user's experience level (Beginner, Associate, Analyst, and Experienced) as explained in chapter 5.2.

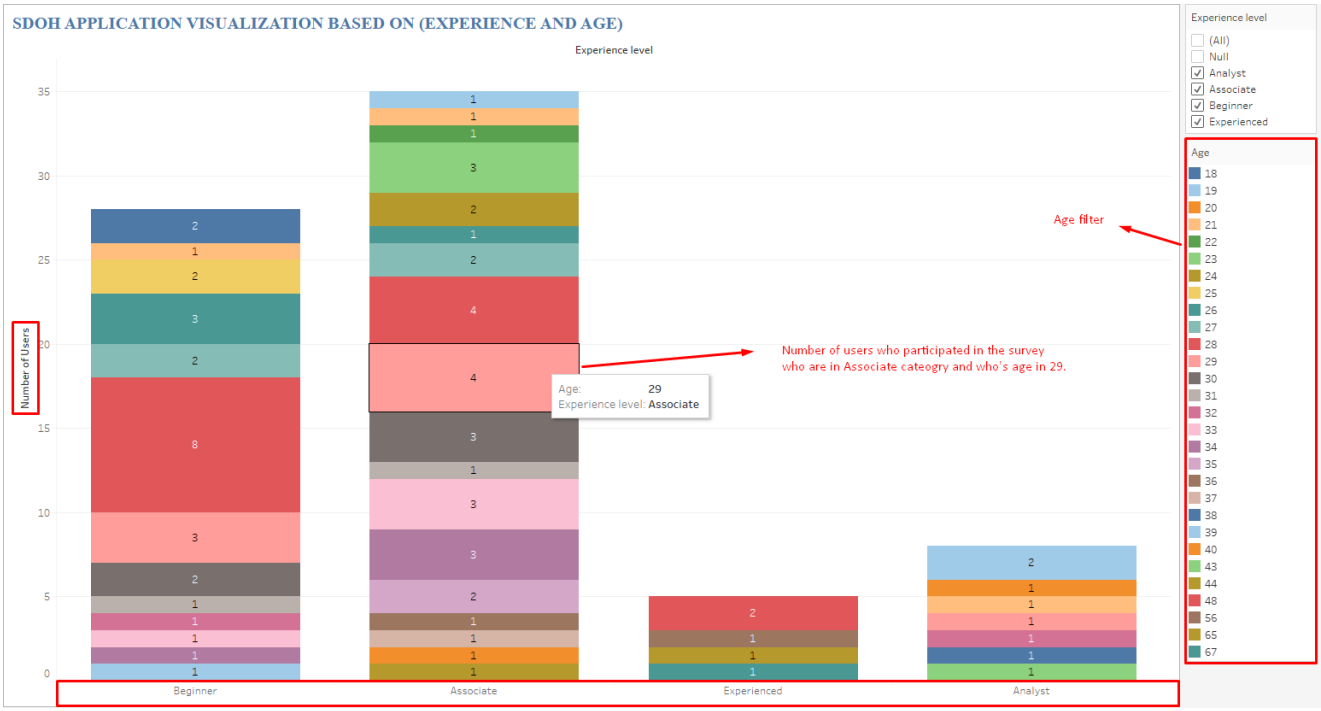

Figure 54: SDOH Application Dataset Visualization with Experience Vs Age

The plot shows the experience level on the x-axis and the number of users on the Y-axis. This represents the number of users who fall under the respective experience range for each age group. The age can be seen on the right-hand side as a filter as shown in Fig 54. Currently, the graph shows each age group separately. But if needed, even age can be categorized into different ranges. From the Demographic data, age and experience are some of the quantifiable fields. Using these fields, one can determine participants' experience level to understand how the experience can be a factor contributing to the simulation data. The bar graph is chosen to present the data in the above visualization. Using bar graphs, it is easier to show the users from each experience level as bar graphs can be used to display categorical data with bars that can show the values in terms of length.

If the user wants to see the responses individually, they can see them by clicking on the age filter. As we can see below Fig 55. The graph shows the number of users from each

experience level for age 56. There is only one user who is of age 56 and falls under experience level “Experienced”. We can say that there are not many experienced providers participated in the simulation. Here, age is not grouped like experience level. In the visualization, the focus was to find the trend of the count of users who participated in the survey at each experience level. If needed, age can be a group but the results would be the same.

Figure 55: SDOH Application Dataset Visualization with age filter

5.3.3 SDOH Data Visualization Based Gender and Profession

One more visualization is developed, considering the User’s profession and gender. Gender and profession also play an important role in this simulation to assess which group is users actively participated in the simulations and surveys. So, the graphs are developed to verify the trends. For this visualization, the “Packed Bubbles” graph is used. Data is displayed using bubbles or circles of varied sizes and colors on a bubble chart. A bubble chart's design can enable it to display numerous variables. Individual bubbles are used to

indicate dimension field values, which measure field values that determine the bubble's size and color. Thus, a plot with at least three variables, including one dimension and two measure fields, can be examined. A bubble chart or packed bubble chart in Tableau is a straightforward but comprehensive and illustrative picture [43]. As profession and gender have a relatively small number of groups, a bubble chart is chosen to display the information in a presentable graph. As per Fig:56 Profession and Gender are taken as dimensional fields to form the bubbles. The size of the bubble represents the number of users in each bubble corresponding to their respective profession and gender.

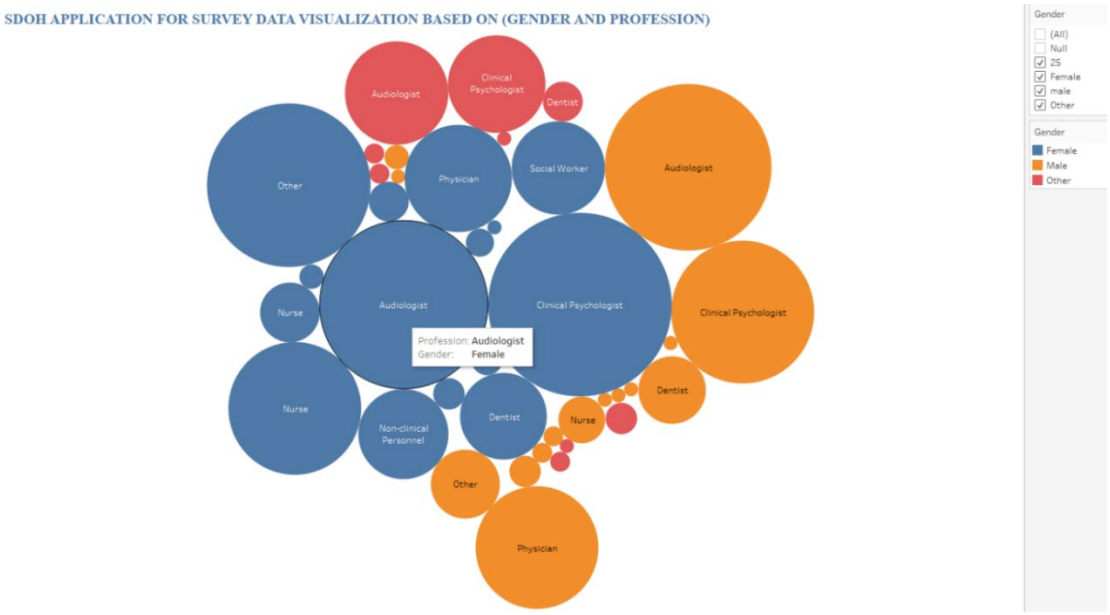

Figure 56: SDOH Application Dataset Visualization with Gender Vs Profession

If the user wants to see responses individually, they can see them by simply clicking on the gender filter. As we can see, Fig 57 displays the data for male users based on their profession.

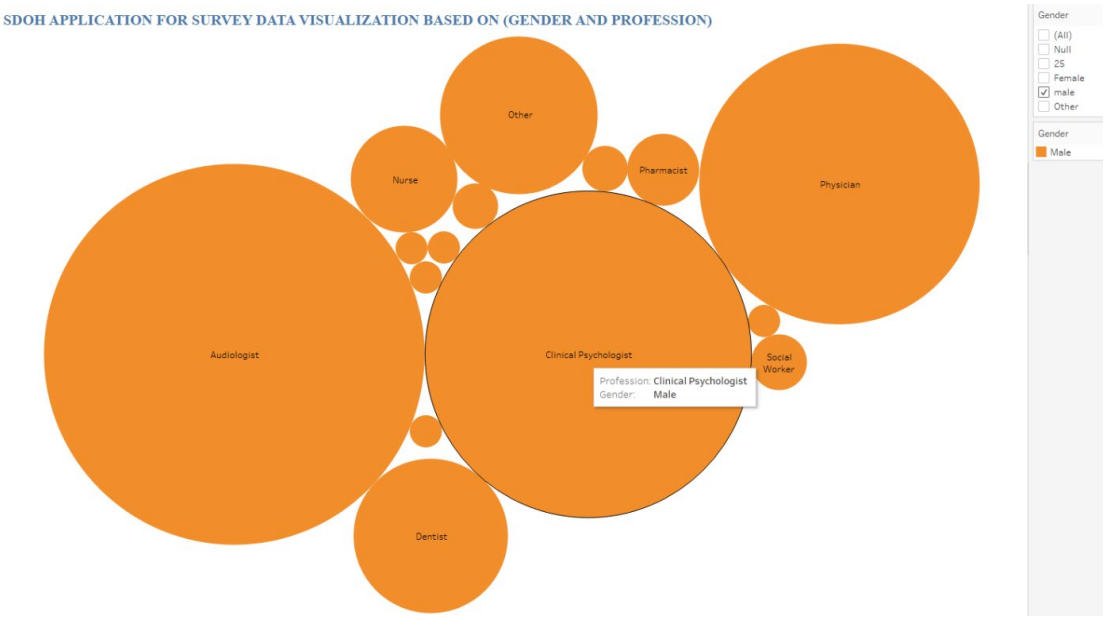

Figure 57: SDOH Application Dataset Visualization with Gender Filter

5.4 Data Analysis results

In the pre-and post-questionaries, most of the users have chosen Neutral, Disagree, and Strongly agree to options for the questionnaires. In Fig 52, the results for all the questions before and after the simulation with respect to user responses. This graph shows the trends for pre- and post-questionnaires. Fig 58 and Fig 59 show the trend for pre- and post-questionnaires respectively. When these figures are compared, the users have reported decreased *anxiety* and *frustration*, whereas their *compassion* levels have increased.

Unfortunately, the number of users who participated in both surveys has decreased. For starting the game/simulation, it is mandatory to fill pre-questionnaire, but the post-questionnaire is not mandatory. This shows that a lot of users have dropped from the simulation after the game is finished, without participating in the survey.

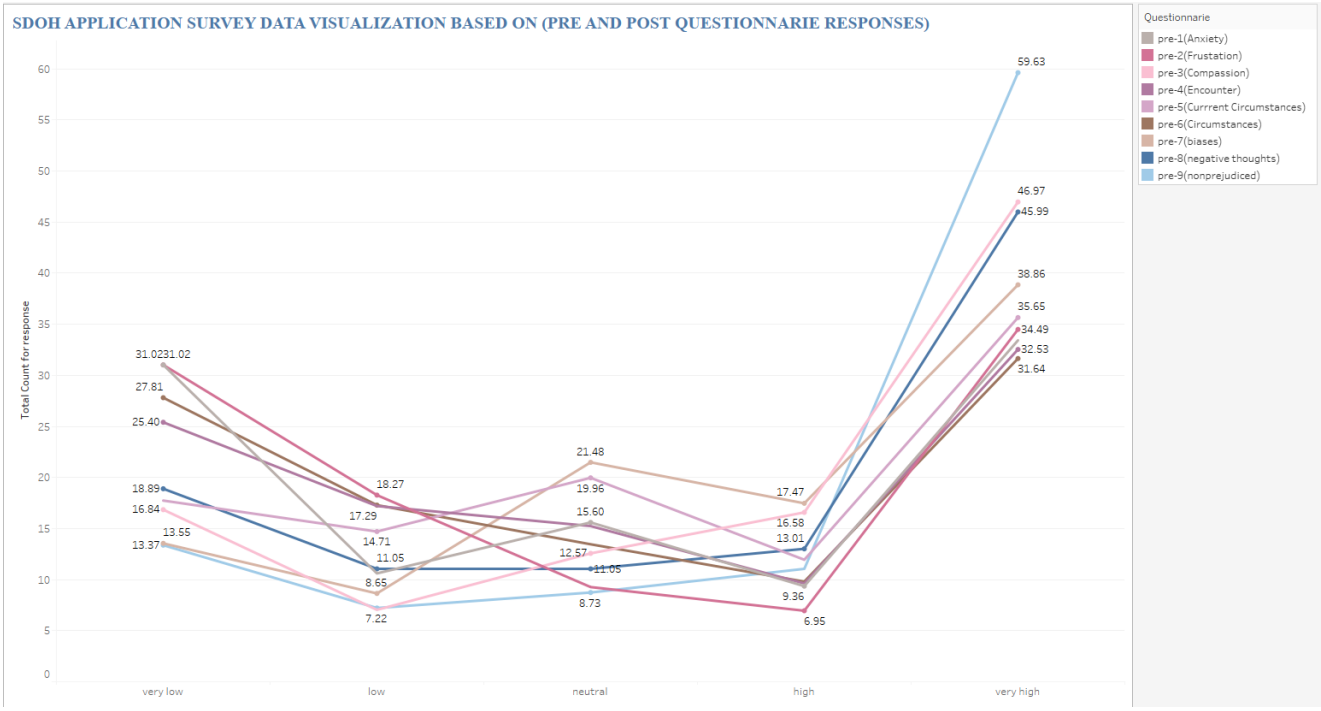

Figure 58: Pre-Questionnaires

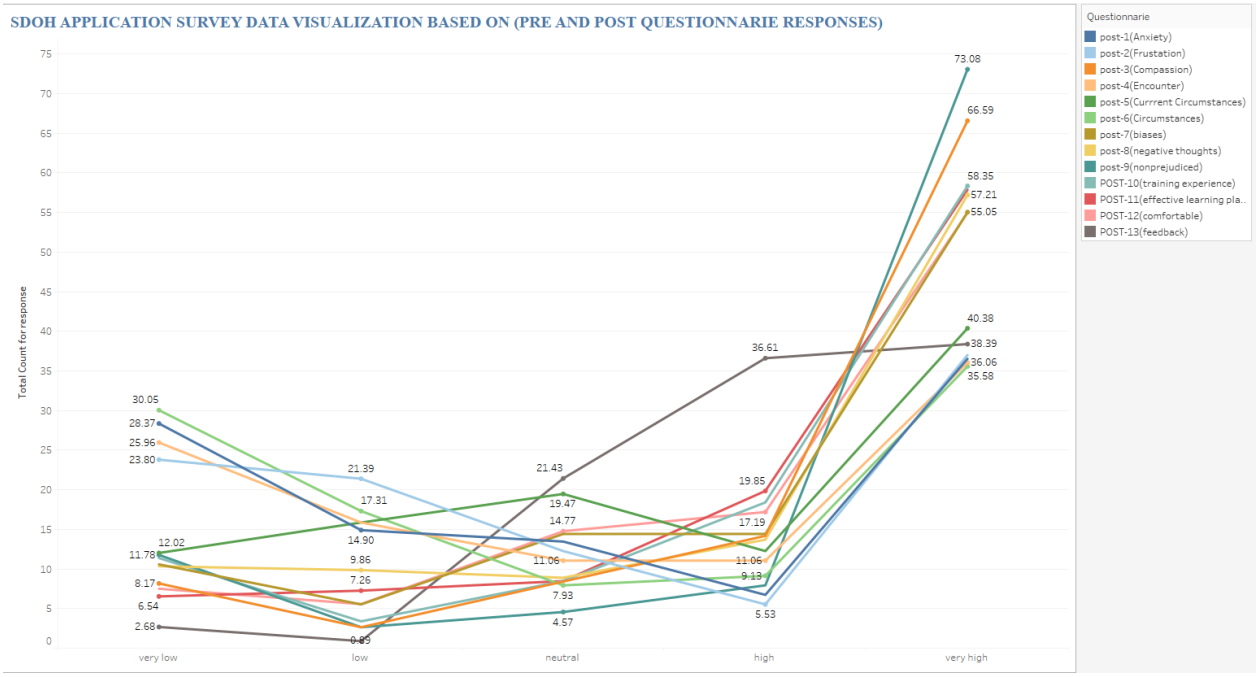

Figure 59: Post-Questionnaires

When drilling down on an individual question, the user response has improved in terms of lessened anxiety when treating or interacting with patients with biases. The Y-

axis in the graph denotes percentage of total count of responses for each question. If the percentage of total count of the responses from participants with very low levels of anxiety increases after simulation, it would mean the anxiety of the users decreased. Similarly, if the percentage of total count of responses of users with very high anxiety decreases post simulation means the user response has improved as the anxiety decreased. Refer to Fig 60 for individual question responses of *anxiety*. The percentage of participants who showed high levels of anxiety before the simulation reduced whereas the percentage of users showing low levels of anxiety increased after the simulation. We can see similar observations for *frustration*, and *compassion*.

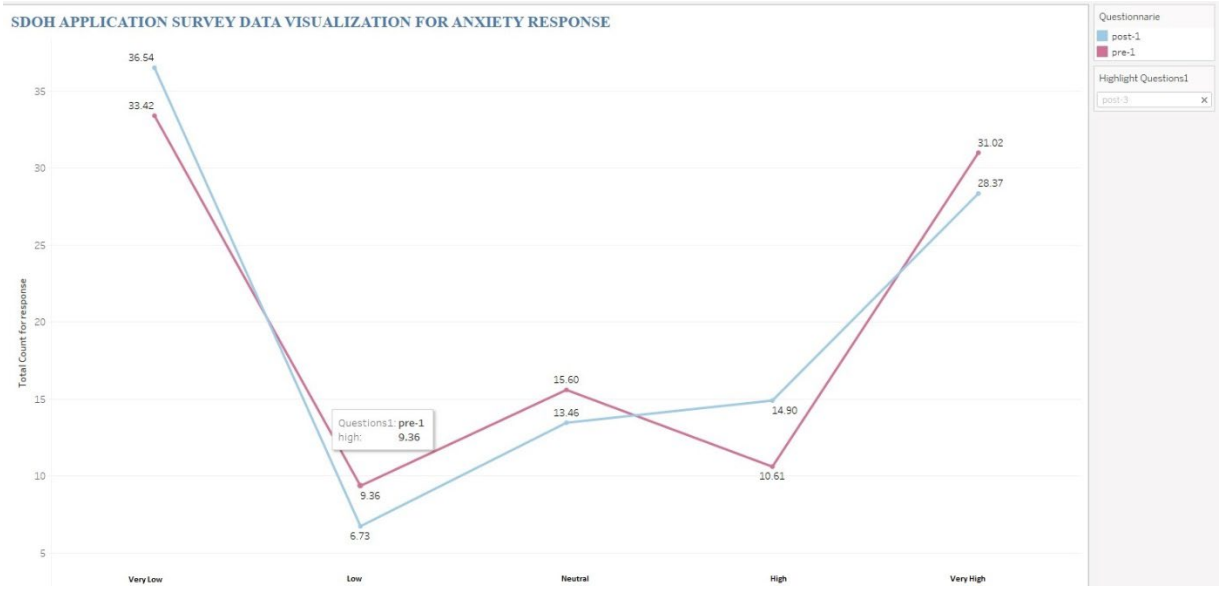

Figure 60: Pre1 and Post1 Questionnaire

Based on the gender-profession analysis, we can see that people identifying as "Females" actively participated in the survey when compared with other genders. Please see Fig: 62 for reference.

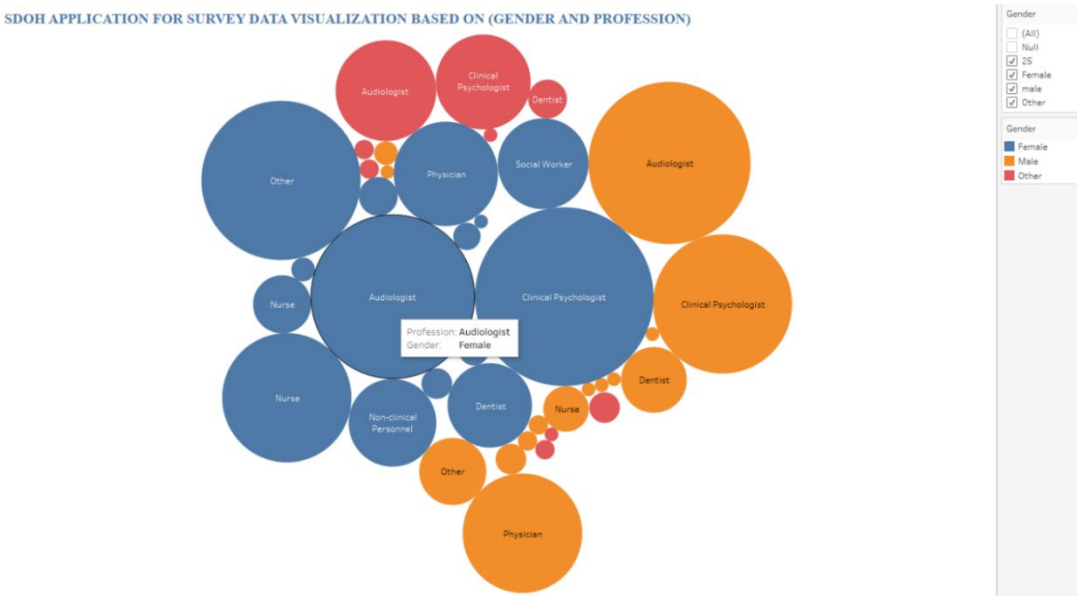

Figure 61: Simulation Visualization - Profession

The users whose profession is “Audiologist” and “Clinical Psychologist” have participated in the survey in large numbers. The users whose experience level is “associates” and “beginners” actively participated in the survey. Fig: 54 shows the trend. Also, the number of participants who participated in the simulation and survey are from “associates” and “beginners” groups. People from the age group 25-30 participated actively in the survey.

6 CHALLENGES

During the life cycle of a project, it is common to have issues and deviations. Some of these difficulties are caused due to lack of resources, expertise, and time. Such challenges faced in the project are discussed in this chapter.

6.1 Character Sliding- During the gameplay, users have noticed patchy spots on the screen when characters move around. This is especially noticeable when a character is in focus. The issue seems to be caused due to characters' motion. We have added gestures to the characters to make the game more appealing to the users. As the character's gesture speed is not constant. Based on the movement, the speed shall change. So, when the character's movements are fast, it leaves some patches behind. Also, there is a mention in the community that the issue could arise when Text comments or annotations are added to the screen. The issue is fixed by adjusting the speed of character movements. The "Break" feature is used to pause the character flow to provide retrieval time for the screen.

6.2 Diversity for black characters- As part of the assignment, 2 black characters are developed – One of them is Charles's husband Dani, a 55 years-old non-binary, and the other is Ashley's foster father Andy, 50 years-old male. The avatars and accessories available in the free version of Reallusion are limited. When it comes to black characters, the choices are minimal. Due to lacking diversity, the characters are developed using available resources. Another challenge was finding suitable hairstyles for this group of people.

6.3 Data Accuracy: One of the major issues faced by any simulation-based research is the accuracy of the data. It is difficult to be certain whether the users are providing

suitable options and are trying to avoid the questions by simply choosing an option. Such inconsistent data can make the outcomes inconclusive.

6.4 Data Accessibility and Importing- Tableau is chosen to visualize the data from the SDOH simulation. Tableau has provided a free license for students which is one of the main factors for developing the visualization in that tool. The visualizations are developed in parallel to simulation based on sample data. But when the time came to integrate the data from the application and Tableau, there are a couple of challenges faced.

- **Data storage**: The data from the simulations are stored in the application web server. As the application server is on-premises, we did not have direct access to the data. The data access is provided through secure login using APIs (Application Programming Interface). The data is stored in .csv file format.

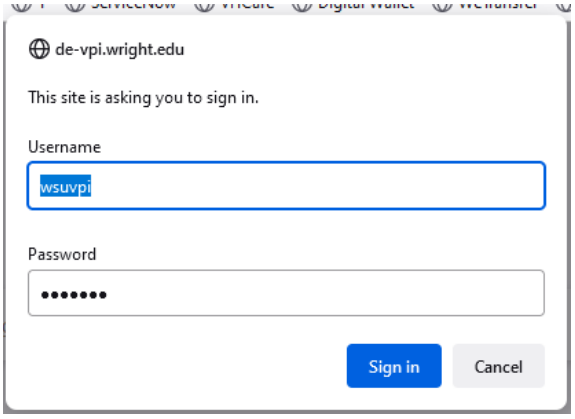

Figure 62: Webserver Login Page

- **Browser Compatibility**: While accessing the data using web browsers, there are compatibility issues. While Firefox and Edge had no issues with accessing the data, the Chrome browser had authentication issues. Chrome did not allow the authentication window popup.

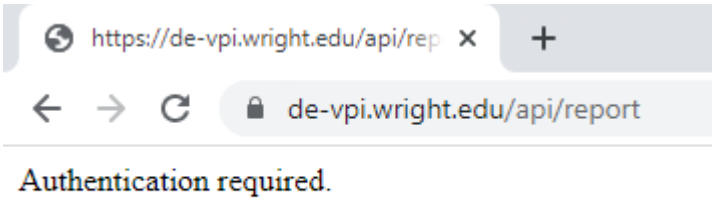

Figure 63: Google Chrome Authentication issue

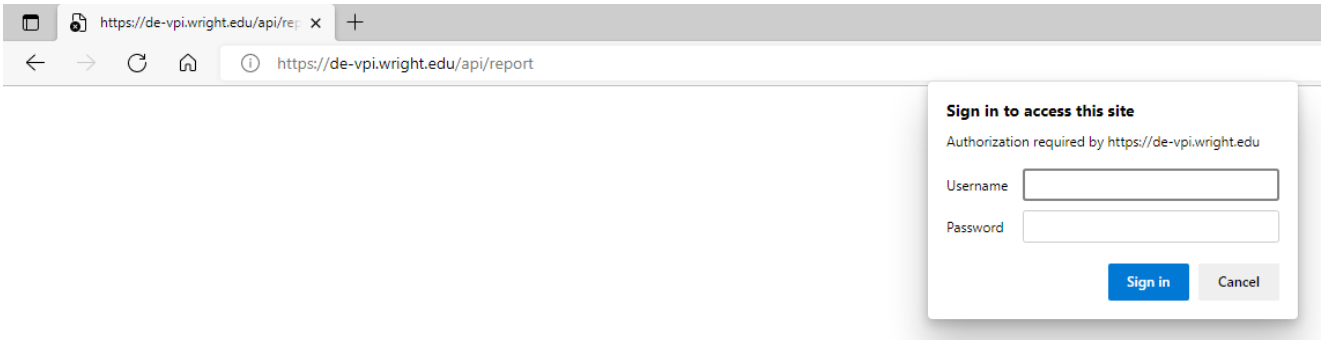

Figure 64: Edge Web Browser Authentication window

- **Tableau compatibility with Web-based data** – Though Tableau has many options for the input data source, it cannot access the data from the Web, particularly where the data is stored in a file. Available connectors were not compatible to import the data directly into Tableau. As an alternative, Microsoft Excel is used to eliminate some of the manual steps.
- **Data Import** – As the data is available on the Web, we needed to automate the process of updating the data. Comma Separated Value file (.csv) format doesn't provide many capabilities for data transformation. Whereas Microsoft Excel has advanced features which allow data management. One of such features is "Get Data from Web". Using this option, we can eliminate the process of manually logging into the Web server and downloading the data file. Excel also has data

refresh options, using which one can schedule the data to refresh to sync the data at regular intervals.

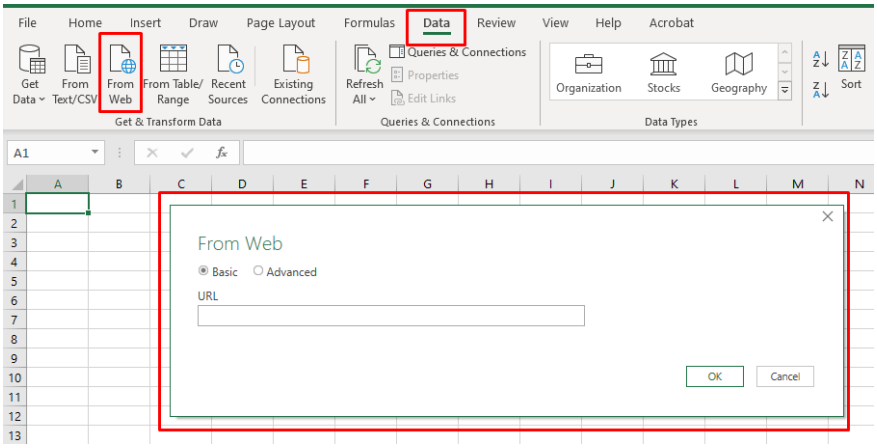

Figure 65: Getting Data through Microsoft Excel

Microsoft Power BI is another Data analysis & Visualization tool which can be used as an alternative to Tableau. Power BI can support direct integration with data. But due to licensing and expertise issues, Tableau is used in the current project.

7 CONCLUSION

It is crucial to foster awareness of SDoH as well as long-term compassion and empathy in healthcare professionals for their patients to improve the patient care experience and outcome. Because they immerse players in a variety of situations, serious role-playing games (RPGs) are among the most captivating types of games. They are widely acknowledged as a useful and successful instrument for educating, training, and cognitive therapy. A simulation-based SDOH training software called Wright LIFE (Lifelike Immersion for Equity) was created especially for mobile platforms. Better human character models that appear and feel more lifelike and relatable to the users are essential as gaming characters are crucial for player engagement and boosting their game experience.

The iterative character-development method has been covered in the thesis. The advantages can be extended to various company sectors to offer cutting-edge digital training and e-learning programs, aside from the healthcare industry. This can be made more realistic by using actual voice actors to capture audio for the simulation. Alternative strategies for enhancing the current technique of communicating with the interdisciplinary team can also be investigated. The length of the feedback cycle can be shortened by working in parallel with the Medicaid team.

Proper checks should be enforced to make sure the data from the simulations is not corrupt or invalid. Currently, user responses are not validated properly ensuring the data accuracy. The surveys are voluntary but setting up rewards such as training certification or vouchers for completing the survey can boost user participation in the program.

Interpreting the raw data is very difficult, especially when your audience is not familiar with the type of data or what data represent. This is where Visualization through data analysis can help bring life to the data and present an easy and understandable format such as graphs. Without analyzing and visualizing the data it would be challenging to convey the results and draw conclusions.

From the outcome of the data analysis, we were able to identify the groups and categories of people who have participated actively. These analyses can be used to focus on the groups who showed less interest and develop programs to improve their participation. Also, we have determined that the "WRIGHT LIFE" applications and the simulations designed have improved the medical professionals' responses to handling patients with biases. This could be revolutionary when simulation-based learning applications are promoted to educate healthcare workers.

The data analysis and visualization approach described in this document enabled us to determine how well the simulation worked in achieving the goal of, for example, reducing anxiety and increasing empathy. By providing a wholistic visualization, the user can drill down on specific groups of providers, such as different age groups, to determine the efficacy of the simulation based on those distinguishing factors.