



Evaluation for an Effective Data Visualization Method in Safety Radar System Demonstrator

Narra Vivek Reddy
Yashaswini Julapally

This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfilment of the requirements for the degree of Master of Science in Software Engineering. The thesis is equivalent to 20 weeks of full time studies.

The authors declare that they are the sole authors of this thesis and that they have not used any sources other than those listed in the bibliography and identified as references. They further declare that they have not submitted this thesis at any other institution to obtain a degree.

Contact Information:

Author(s):

Narra Vivek Reddy

E-mail: vinb18@student.bth.se

Yashaswini Julapally

E-mail: yaju18@student.bth.se

University advisor:

Dr.Deepika Badampudi

Department of Software Engineering

Faculty of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

Abstract

Background. Evaluation of Data Visualization methods is a major challenge within the software and other industries. With complex data and requirements, often organizations require effective visualization methods which impact business decisions and convince stakeholders. This is a similar challenge in the development of a software demonstrator for the innovative safety radar system at ABB Jokab Safety whose aim is to improve the detection reliability using multiple radar sensors and requires an effective visualization method which will satisfy all the requirements.

Objectives. The main objective of this study is to explore different data visualization methods involved in illustrating the raw data and with the help of developers, and other team members feedback with reference to existing literature and filter them with respect to the system functionalities. Establish evaluation criteria with relevant metrics to perform analytic evaluations on the visualization methods to determine an effective method.

Methods. A Case Study which includes a multivocal literature review, is conducted at ABB Jokab Safety. Initially, to gather information on the subject, both formal and grey literature are explored and documented to filter our appropriate data visualization methods for this system. A task-based evaluation using semi-structured interview is conducted on 14 participants to determine an effective visualization method followed by statistical analysis to derive proper validation of the findings. The Response time, Ease of understanding, Confidence and Accuracy of the visualization methods are evaluated with feedback from the participants.

Results. The Multivocal literature review filtered 16 primary articles which encouraged the use of 4 data visualization methods used in the safety radar system with distinct functionalities. A coordinate transformation engine to combine the data sets was also developed for the safety radar system contributing to the overall improvement of detection reliability. The evaluation including both quantitative and qualitative results validate each other's findings through statistical tests like Kruskal Wallis and Bonferroni post hoc followed by narrative analysis resulting a heat temperature plot to be more effective in the visualization of the radar data from multiple sensors.

Conclusions. The results from this research provide insight into how data visualization evaluations can be implemented for real-time industrial problems and furnish validation process to determine an effective data visualization method. This study helps object detection using similar radar technologies visualize their data in an effective way and provides a scientific approach for evaluating similar data visualization problems.

Keywords: Data Visualization, Evaluation, Millimeter wave, Radar detection

Acknowledgments

We want to express our heartfelt gratitude to our university supervisor Dr Deepika Badampudi for her constant and motivating guidance who despite being busy with her duties has given us valuable time during this thesis work and managed to put us in the right direction always.

We also thank our Industrial Supervisor Peter J Hessling firstly for selecting us and introducing some of the elite professionals at work and for his support at ABB Jokab Safety whose suggestions and feedback helped in a successful execution of our research. We are grateful to the continuous cooperation and support form all employees involved in this thesis who not only took part in interviews but also provided interesting insights during initial phases and helped contribute towards the research and development of an innovative product.

Lastly, We would like to acknowledge the fact that this thesis would not be possible without the love and support from our parents, families, and friends who showered their never-ending support, dedication and guidance towards a fruitful completion of the thesis.

We also share our prayers and well wishes to all the University officials, thesis workers and their families for good health and happiness during the Pandemic *Covid-19* who have achieved so much despite all the challenges.

Thank you all for your unconditional support and encouragement!

Contents

Abstract	i
Acknowledgments	iii
1 Introduction	1
1.1 Research Problem	3
1.2 Aims and Objectives	4
1.3 Thesis Outline	4
2 Background	5
2.1 Safety Radar System	6
2.1.1 Dataset	7
3 Co-ordinate transformation and System Architecture	8
4 Related Work	10
4.1 On Information Visualization Evaluation	10
4.2 On Radar Data	11
5 Method	13
5.1 Research Questions	13
5.1.1 Method Selection	14
5.2 Research Design	15
5.3 Multi Vocal - Grey Literature	15
5.3.1 Data Collection	16
5.3.2 Data Analysis	18
5.4 Case Study	20
5.4.1 Case Description	21
5.4.2 User Evaluation Criteria and Metrics	22
5.4.3 Data Collection	23
5.4.4 Task Based Evaluation	26
5.4.5 Data Analysis	28
6 Results and Analysis	32
6.1 RQ I	32
6.1.1 Multi-vocal Literature Review Results	32
6.1.2 Thematic Analysis Results	34
6.2 RQ II	38

6.2.1	Evaluation over Response Time	43
6.2.2	Evaluation over Ease of Understanding	46
6.2.3	Evaluation over Confidence	49
6.2.4	Evaluation over Accuracy	52
6.2.5	Interview Summary & Data Triangulation	55
7	Threats to Validity	59
7.1	Reliability	59
7.2	Internal Validity	59
7.3	External Validity	60
7.4	Construct Validity	60
7.5	Limitations	61
8	Discussion	62
9	Conclusions and Future Work	65
A	Supplemental Information	71
A.1	Interview demographics and questionnaire	71
A.2	Data visualization methods	74
A.3	Data Distribution : Quantitative Data Analysis	76

List of Figures

1.1	Thesis Document Structure	4
2.1	Safety Radar System Structure	7
2.2	System Setup at the Company	7
3.1	Demonstrator Software Architecture Overview	8
3.2	Code Logic for Coordinate Transformation of the data points	9
5.1	Research Process	15
5.2	Literature Search Process results	18
5.3	Data Extraction from the Literature search	19
5.4	Case Study Protocol	21
5.5	Unit of Analysis	22
5.6	Interview Protocol	24
5.7	Task evaluation design	27
6.1	Process for thematic analysis	35
6.2	Thematic Network	36
6.3	Age of the participants in the study	38
6.4	Participants Gender	38
6.5	Educational Qualification level of the Participants	39
6.6	Educational Field of study of the participants	39
6.7	Experience in the Research field	40
6.8	Experience with data visualization methods and their evaluation	40
6.9	Previous encounter with data visualization in either course work or projects	41
6.10	Familiarity of the participants with the concept of object detection	41
6.11	Familiarity of the participants with the concept of machine safety	42
6.12	Familiarity with the usage or influence of radar technology with object detection	42
6.13	Feedback regarding Data visualization and its impact on understanding complex data	43
6.14	Kruskal-Wallis Test Result: Response Time	44
6.15	Bonferroni Results : Response time	45
6.16	Response Time : Box plot	45
6.17	Kruskal-Wallis Test Result : Ease of Understanding	47
6.18	Bonferroni Results : Ease of Understanding	48
6.19	Ease of Understanding : Box plot	48

6.20	Kruskal-Wallis Test Result : Confidence	50
6.21	Bonferroni Results : Confidence	51
6.22	Confidence : Box plot	51
6.23	Kruskal-Wallis Test Result : Accuracy	53
6.24	Bonferroni Results : Accuracy	53
6.25	Accuracy : Box plot	54
6.26	Overall Selection of an Effective Data Visualization by the participants	57
A.1	V1 : Time Domain plot	74
A.2	V2 : Scatter Plot	74
A.3	V3 : Contour Plot	75
A.4	V4 : Heat Temperature Plot	75
A.5	Normal Q-Q Plot : Response time	76
A.6	Normal Q-Q Plot : Ease of Understanding	76
A.7	Normal Q-Q Plot : Confidence	77
A.8	Normal Q-Q Plot : Accuracy	77

List of Tables

5.1	Multivocal Literature review : Search String Formulation	16
5.2	Exclusion Criteria for the Literature	17
5.3	Inclusion Criteria for the Literature	17
5.4	Literature Quality Assessment Criteria	18
5.5	Participants in the Interviews	25
5.6	Task Evaluation Table	26
6.1	Multi Vocal Literature Review Results	34
6.2	Codes from thematic analysis	35
6.3	Theme Identification	36
6.4	Thematic Analysis: Results from the Literature	37
6.5	Summary - Descriptive Statistics for evaluation over response time . .	43
6.6	Quantitative Variables : Response Time	44
6.7	Summary - Descriptive Statistics for evaluation over Ease of Under- standing	46
6.8	Quantitative Variables : Ease of Understanding	46
6.9	Summary - Descriptive Statistics for evaluation over Confidence of the participants	49
6.10	Quantitative Variables : Ease of Understanding	49
6.11	Summary - Descriptive Statistics for evaluation over Accuracy of the participants	52
6.12	Quantitative Variables : Accuracy	52
6.13	Analysis Summary	55
A.1	Interview Questionnaire	71
A.2	Demographics Questionnaire	73

List of Visualization Methods

- V1 Time Domain Plot
- V2 Scatter Plot
- V3 Contour Plot
- V4 Temperature Heat-map Plot

In the present day, the Software industry is striving to deliver meaningful, usable and efficient products to the customers with real-world problems and applications. Software-intensive systems and software embedded hardware systems often undergo developmental changes and are introduced to various new functions to improve and create a new product. To illustrate this, organizations utilize different data visualization methods which aids in convincing stakeholders, users and other practitioners in making decisions appropriately. Knowing which visualization method gives an effective and accurate intended outcome is inevitable. Data Visualization has become a process of converting raw data into understandable information and knowledge through graphical depiction. It is often the primary goal of an organization to meet the customer requirements with a high level of satisfaction, and it is not so surprising that an effective visualization tool, framework or method helps present the accurate comprehension of the data structures or trends. This mainly has an effect on the decision-making skills of a user and convincing the stakeholders. For example, a limited amount of data in the form of a verbal or written format is useful. Still, there is a significant difficulty in understanding the data when there is a huge amount of it being generated and even harder on a larger scale[1]. Generally, the graphical illustration or depiction shall allow the users to interpret the data in complex scenarios as it does in ideal conditions.

Data Visualization is a powerful concept through which complex and difficult problems are illustrated in a simple format. Data is either in a structured or unstructured form of information which is used to make decisions or take actions based on the requirements of a given problem. Having a surplus amount of data and no proper medium of presenting it is just considered to be unresourceful and leading to the failure of a product. Often, many industries require attractive and understandable charts, graphs or pictorial representation of the information which give meaning to the data[2][3][4].

Over the past decade, many industries have developed various products and made discoveries which surround us and our daily works, where every system, software and functionality of a machine is evaluated based on the reliable results. Reliability and understandability are key factors to many real-time applications and helps measure the correctness and consistency in the system. Understanding the need to measure or determine reliability ensures good efficiency, performance and accuracy of such applications[3]. These applications, for example, could be using a smartphone,

travelling in an autonomous vehicle, production equipment, Heavy payload manufacturing machines and many more which are dependent on different safety factors which help them perform better. If these products or developments were to fail or show some signs of incorrectness or faults, then it would pose a great threat to humans and their surroundings.

The development of such machines and products undergo many years of research which contribute to all the performance and efficiency of such a system. Often these products require a software infrastructure to support all the features and operations. And synthesizing data from multiple team members involved in developing takes multiple iterations with significant challenges. One of the significant challenges which are the synthesis of data and visualization when different team members use other visualization methods, but require to show collective results to the stakeholders or customers. User satisfaction and understandability of data is a must for any software deployment.

This research aims at evaluating such visualization methods which play an important role to convince the stakeholders and customers with normalizing the usage of a single visualization method. The evaluation is performed at ABB Jokab Safety in the development of a Software Demonstrator for (Group) Safety Radar System which requires to demonstrate the improved reliability in object detection.

Radar Sensor Technologies are being deployed across different platforms and applications to track the movement of objects and humans. Data collected through these sensors are of huge volume, which requires high computational requirements and good visualization methods/techniques. This paper is concerned with one particular radar technology, namely IWR6843 Milli-Meter Wave Radar Technology and the different data visualization methods used during the development of a software demonstrator with multiple radar sensors for improved reliability in detection at ABB Jokab Safety.

An embedded software system, in this case, the safety radar system is a combination of software intelligence and hardware components which are programmed to perform a set of operations or tasks. During the developmental stages of such systems, data visualization plays an important role, which being said is to display all the requirements, features and functionality of the system. Many developers, software practitioners and team members tend to use a convenient or goal-oriented visualization method to depict the developed component. Now, it is a challenge if there are multiple visualization methods for different features of the system developed by different developers. This directly impacts on the understandability, performance and reliability of the system when the users and stakeholders try to make sense of what information is being illustrated. Post Development collaborating and synthesizing the data from various perspectives into a single data visualization method is again a time consuming and cost-effective process.

From the existing literature and company artefacts, we can learn which data visualization methods are utilized to illustrate the radar data. We would like to evaluate

these visualization methods and design an evaluation study to determine an effective Data visualization method for similar data or applications.

Following contributions are stated as potential outcomes from this research,

- Identification of the visualization methods used in depicting radar sensor data.
- An evaluation criteria with user satisfaction, effectiveness, and understandability of data.
- An analytic evaluation study with real time sensor data and the industrial requirement to determine an effective visualization method for similar problems.

Results from this study will be a significant contribution to the Information Visualization industry and radar technology. Generalization of one data visualization method which can convert all the requirements from the Company will impact the performance and increase efficiency in the development team, creating a common platform for visualization.

Secti

1.1 Research Problem

Software embedded systems often require visualization methods to help understand the system functionality and provide insight into the performance, efficiency and accuracy. An effective visualization method which can illustrate all the intended outcomes requires evaluation. The Research and Development department at ABB Jokab Safety is working towards developing demonstrator software for the Safety Radar System which is a software embedded hardware system with more than one mmW radar module to improve the detection reliability. The data generated from these radar modules are complex and huge in size, which is combined using a Coordinate transformation engine developed as a part of this research study. In order to convince the stakeholders in the organization and aide in the decision making process regarding the use of this proposed system, an effective data visualization method is required for all the features. It also helps developers, team members working in this project reduce challenges in illustrating and collaborating their work using one visualization method for all features like encircling the target, analyzing surveillance volume for clutter and object detection which are developed for the newly designed safety radar system. Further, the evaluation for an effective visualization plays a crucial role in the demonstration software to convince the stakeholders in the company and demonstrate the system intended outcome, which is improved detection. An in-depth investigation into which data visualization methods are utilized and what are their outcomes with relevance to the system should be made. An evaluation determining the effective data visualization method for the demonstration software is required, which has significant effect helping in understanding the improved detection reliability and performance of the system. The result of this study would generalize an information visualization evaluation method for the radar data set for similar Safety Radar systems or other software-intensive systems in the future.

1.2 Aims and Objectives

The overall aim of this research is to identify relevant data visualization methods and evaluate different data visualization methods used in the safety radar system to determine an effective visualization method.

A brief breakdown of the objectives are below:

- Identify the methods and techniques used to visualize the data in safety radar system detection.
- Perform an analytic evaluation of the data visualization methods in an industrial problem (Software Demonstrator for Safety Radar System).

1.3 Thesis Outline

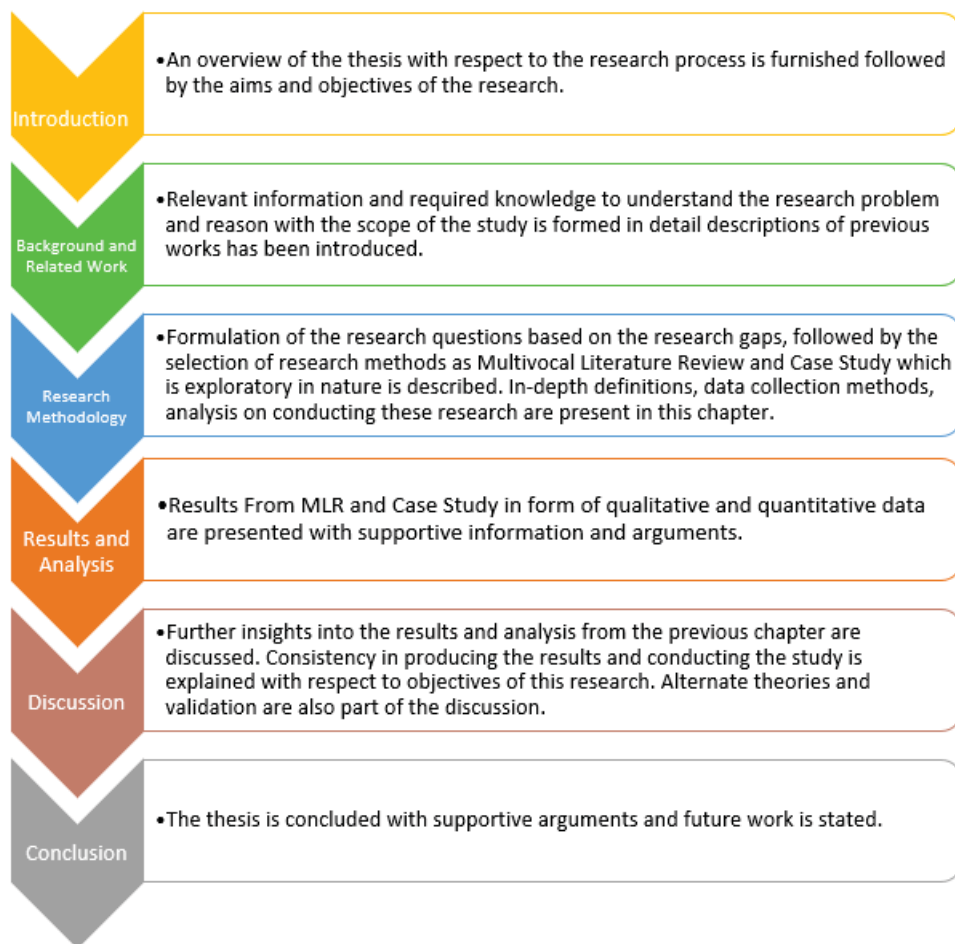


Figure 1.1: Thesis Document Structure

In this chapter, We make ourselves familiar with some relevant concepts and primary motivations to this research.

Data Visualization is a process of understanding a considerable amount of data which is crucial in making decisions for real-time problems[5][6]. Having Data in a verbal or written format is useful but comes with many challenges like when there is a large chunk of data and extracting meaningful information is difficult but can be easily navigated through simple illustrations like in graphs, charts or plots. This helps the user make decisions effectively and easily with a proper understanding of the data. Data Visualization methods have a wider scope of applications and use such as identifying patterns, trends and performing analysis.

Recent years have paved way to many growing industries like autonomous vehicles, heavy machine manufacturers, automated factories and much more. Most of these activities often include machines which work with sharp blades, hot tools, heavy payload carriers which might harm a human. Placing certain safeguards and preventive measures on these machines so that they do not harm any human is known as machine safety. Machine safety has a set of defined guidelines and rules which are to follow by specific industries and products which provide trust and safety to the users. Failing to follow the machine safety norms or policies would increase the risk of getting injured or damaged. For example, automobile manufacturing industries are equipped with heavy payload machines which assemble most of the cars; these workplaces are intervened by humans on a daily basis. If the machine is not able to sense the presence of a human nearby and come to a PAUSE state or HALT state, it could harm the human.

Radar, a critical technology was first used during world war II when there was a need to locate and know the position of enemy ships. Over the years, radar has been growing and developed at a rapid rate. Radar is defined as a detection system which uses radioactive waves to determine the range, velocity and angle of certain objects. Millimetre-wave radar technology is operated at a wide bandwidth of the radioactive spectrum with short-range detecting capabilities and material penetrating capabilities. Modern-day radar detection systems use a significant amount of software computing and utilize huge lines of code to improve the detection capability of the system. Millimetre-wave Radar(mmW) is being used in the current real-world industrial problems and experimented in autonomous vehicles due to its low cost and

high availability in harsh environmental conditions. Reliability in the safety radar systems indicates that there are more safety and availability, which means the probability of the false detection is very less compared to other detection technologies like infrared cameras, LIDAR and RGB imaging[7]. In this study, we observe the usage of more than one mmW radar module to improve the reliability by combining homogeneous data sets, instead of sensor fusion using heterogeneous data sets with two different sensor types.

Like many software embedded in hardware systems, the safety radar system also needs illustrative charts, graphs and plots to show the position of the objects initially when developing machine optimized algorithms and functions for further automation. More importantly, visualization is a core component in the demonstrating software which is presented to the stakeholders or customers to illustrate the capability and use of this system. There is no object detection when we cannot display the position of an object or human in-front of the radar modules.

In general, a radar system is a combination of hardware components and software intelligence which compose an image of the surroundings with relative positions, range, angle and velocity of the object. Producing these results consumes a lot of data processing which is binary in nature. The users need a simple and understandable form of information from the radar modules, which is illustrated using the visualization methods. It is safe to assume that, one cannot imagine what is it like the Radar modules are sensing without these visualizations and analyze the data leading to low reliability and safety. Similarly, there are different visualization methods used by other developers and practitioners while developing this new technology with multiple radar modules. Determining a single visualization method which can also be used to convince the stakeholders will impact the business intelligence and growth of the project.

2.1 Safety Radar System

The Safety Radar System proposed by ABB Jokab Safety comprises of an Evaluation Unit where the demonstration software is executed and a hardware setup with more than one radar module to improve the detection reliability using the mmW radar technology. The data is sent to the external Evaluation unit where it is transformed, processed and visualizations are made to show the positioning of the objects present in the field of contact[7][8]. This radar technology occupies the frequencies ranging from 30GHz - 300Ghz. The size of mmWave sensors is small and compact in nature as antenna size is inversely proportional to the frequency. These sensors have fair range resolution, can offer wider bandwidth of frequency and has a significant role in calculating or depiction the range, velocity and angle of the objects in the area of contact within reach of these sensors. Due to their robust build quality and low-cost feature, these have been easily accessible and offer consistent detection and operation in different environmental conditions. The imaging results illustration aspects deal with scanning the electromagnetic radiation emitted from the objects which are used to outline and detect the position of an object[9]. Bandwidth and Range Resolution

is two important factors considered in the detection of humans/objects present in the area of contact for the radar sensors. Bandwidth is only the measure of the frequency in which the sensor is operating, but the ability to differentiate between one or more objects present in short distances is known as range resolution.

$$d_{res} = \frac{c}{2B}$$

Range resolution is evaluated using a mathematical formula where d_{res} is the range resolution in m, B is the bandwidth of the radar and c the speed of light[9].

2.1.1 Dataset

Data collected from three radar modules in this system. A brief overview of the system is given in Figure 2.1. Real-time datasets are used in the visualization methods where the data is integrated using coordinate transformations. As part of this thesis, a newly developed code for the co-ordinate transformation engine is used in the software demonstrator, which combines data from multiple radar sensors, as shown in Figure 2.2. This sensor data consists of information in the form of a range, velocity, and the angle of objects with unique accuracy. Data samples are generated in a designed space with a radar reflector, object and human. The dataset has numerous variables which are used by a mathematical function with Fast Fourier transforms (FFT) that calculates the precise position of the objects by converting time-domain signals to a frequency domain.

System overview

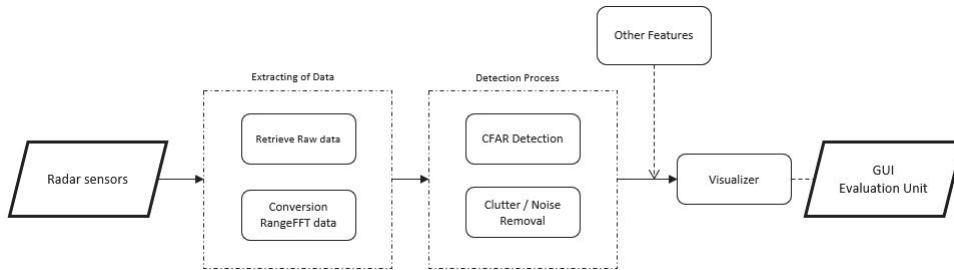


Figure 2.1: Safety Radar System Structure

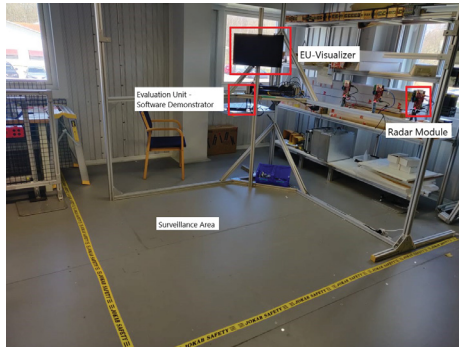


Figure 2.2: System Setup at the Company

In this architecture, we can observe that all the data from other services are sent to the coordinate transformation engine where they are integrated and plotted using a visualizer. Below is the Figure 3.2 is a code snippet which has the logic behind combining homogeneous data points of the RangeFFT from different sensors. The RangeFFT data is a collection of velocity, speed and angular information of the radio waves emitted from the sensors. The raw data recorded is converted using a Fast Fourier transform formula which represents the position of the objects being detected. After which the various features are implemented and passed through the coordinate transformation engine.

```

%Selecting only n peak indices from the data which pass the Cfar threshold.
ind_1 = ind_1(1:n);
ind_2 = ind_2(1:n);
ind_3 = ind_3(1:n);

%the detected n peaks from the data are mapped into x-y plane below
%r(1,2,3) & q(1,2,3) = indices in x-y plane for the peaks
%p_o(1,2,3) = Saving the position of object wrt sensors using peaks
%mapped from above.
for i = 1:n
    r1 = xem(ind_1(i),41)+1;
    q1 = floor(ind_1(i)/41);
    p_o1(i,:) = [cx(r1,q1),cy(r1,q1),0.2*rand-0.1];
    r2 = xem(ind_2(i),41)+1;
    q2 = floor(ind_2(i)/41);
    p_o2(i,:) = [cx(r2,q2),cy(r2,q2),0.2*rand-0.1];
    r3 = xem(ind_3(i),41)+1;
    q3 = floor(ind_3(i)/41);
    p_o3(i,:) = [cx(r3,q3),cy(r3,q3),0.2*rand-0.1];
end

%Positions of the sensors wrt global frame (Fixed for this case)
p3=[0,1,1.31]';
p2=[0,1.5,1.31]';
p1=[0,2,1.31]';
%po=[1.48,1.58,1.31]';

%4th row of the homogeneous transform matrix
T4 = [zeros(1,3),1];

%T_{i}g = transformation matrix of sensor wrt global frame
T_1g = [eye(3),p1;T4];
T_2g = [eye(3),p2;T4];
T_3g = [eye(3),p3;T4];
%T_{og} = [eye(3),po;T4];%Fixed object

%T_{g}{} = transformation matrix of global frame wrt sensor{}
T_g1 = [eye(3),-p1;T4];
T_g2 = [eye(3),-p2;T4];
T_g3 = [eye(3),-p3;T4];
%T_{go} = [eye(3),-po;T4];%Fixed object

%p_o{}g = Mapping position of object wrt sensor {} to position of object wrt global frame.
p_o1g = T_1g*[p_o1, ones(10,1)]';
p_o2g = T_2g*[p_o2, ones(10,1)]';
p_o3g = T_3g*[p_o3, ones(10,1)]';

%plotting position of object wrt global frame using 3d scatter plot
figure;
for i=1:10
    plot3(p_o1g(1,i),p_o1g(2,i),p_o1g(3,i),'+r'); hold on;
    plot3(p_o2g(1,i),p_o2g(2,i),p_o2g(3,i),'+g'); hold on;
    plot3(p_o3g(1,i),p_o3g(2,i),p_o3g(3,i),'+m'); hold on;
end

```

Figure 3.2: Code Logic for Coordinate Transformation of the data points

This thesis aims at finding relevant data visualization methods for the features in the system and then perform an evaluation regarding which visualization method offer a more effective and accurate illustration of the data for detecting objects using multiple radar sensors which is a new concept for improved reliability and safety. The different resultant visualization methods after passing through this engine are furnished in Appendix A.2.

In this chapter, we furnish with some previous works and contributions with regard to the objectives of this paper.

Evaluating Information Visualization methods is an indispensable task in the real world and with research problems. The usefulness and usability of a certain visualization method are required to be known with the increasing needs and problems in the software industry and other organization sectors. The users reasoning and sense-making abilities are at test during an Information Visualization[10].

4.1 On Information Visualization Evaluation

Promoting a better understanding and providing an insight into a data set is the primary goal of any information (data) visualization. Challenges like lack of established procedures and frameworks for the evaluations lead to inefficient results and demand new solutions specific to the problem, which are expensive and time-consuming. Such challenges and relevant solutions are discussed in the paper by Forsell and Johansson[11]. They also propose a heuristic set for evaluation of Information Visualization. They also contradict that this evaluation set is not a final product, and further validation is required for dynamic requirements. Additionally, Camilla Forsell[12] proposes a scientific evaluation study design for such real-world data visualization problems through experimentation and other empirical methods. One of which is implemented in this study. Prior studies like in Isenberg et al. [13] they utilize the traditional systematic literature review to evaluate visualization papers by introducing scenarios and goals common to the evaluation criteria.

Understanding what kind of data is useful for our analysis and which method to use for illustrative purposes is essential in any industry. For that, a detailed and concise mapping of the datatype, formats and behaviour should be clear, which is possible by analyzing various attributes, conditions and parameters relative to the problem and the application of use. Koua et al[14] discussed the usefulness and usability of data visualization in such problems outlining few general goals and tasks to determine their usefulness. Plaisant [15], in her paper, highlights few metrics and accounts for different user studies in the evaluation of visualization tools and their utility towards adopting such methods.

4.2 On Radar Data

It is observed that when data is expressed as some sort of presentation charts or figures, it's more effective in understanding the characteristics and behaviour of the system from where the information is originated. This is achieved by using suitable tools and techniques present in data visualization. There are many data visualization methods present in the software industry. Having an idea of how the data reaches the interface and what goes into processing the data to visualize it is studied in the paper by Pai[16]. Understanding the visualization of the radar data is difficult to most of the users and if this process is not streamlined or compacted it will pose challenges related to the efficiency and reliability of detecting the objects by the radar system. Effectiveness and the reliability of the data visualization can be evaluated using various methods which are discussed in the literature present[17][18]. Data Visualization helps in order to examine the discontinuities and gaps which are generally caused due to system performance or unreliable programs which are discussed by Bharadwaj et al[19]. They also discuss how to determine complete accuracy in visualizing the data in order to furnish a good illustration of data patterns and behaviour.

Like discussed in previous chapters, our overall goal in this thesis is to determine how effectively can the data be visualized in order to detect the objects more efficiently and which methods could help in improving it. A better data visualization method offers high confidence in the detection process. In order to showcase this theory or argument, data visualization is considered as an essential element in detecting objects by the radar sensors. Schneider and Wenger[20], in their paper, discuss how increasing the resolution of these radar systems could have a possible effect in how these systems generate the output image and illustrate an improved visualization of data using images which could further be processed in order to generalize the process and automate considering other aspects such as complexity, time and cost. They conclude by providing a detailed report on how the data processing plays a driving role in generating these images which basically represent what the sensors are detecting and how the hardware components exactly fit into the domain for processing the data. The environment they have chosen was a single mm-wave sensor setup with an evaluation module. The use of images is to validate the actual data visualization, which cannot be conceived to a novice user.

Grover et al[21], in their paper talk about how they have implemented a method which is a combination of two different kinds of sensors and fused the data produced by the sensors to improve the detection process in a cluttered environment. They have used a night vision camera which is kind of an image sensor fused with the mm-wave radar sensor (76Ghz). Using a combination of both sensors is a good application, but one cannot understand up to what scale is this useful in real-time problems. The data output was in the form of images, radar scans (azimuth maps) and graph plots. Again they have the data processing, visualization methods and techniques to combine the data generated from two different sensors and evaluated the effectiveness of this implementation. This clearly states the motive mentioned in our thesis which is to take the data and utilize the results to draw meaningful conclusions with the scope of potential tool development, i.e., a demonstrator to

automate this whole process of using multiple sensors of the same radar technology instead of combining two different technologies which are more feasible and provide a sustainable approach

Singh et al[9], have illustrated a clear and apt application of the mm-wave radar sensors which fits into the scope of the industry usage. Most of the robotics and automobile companies are companies that deal with the safety and reliability of using radar technology for detection. In their paper, The authors devise a new framework (HAR) for detecting the human activity recognition by generating a point clouds from the data collected and processed using various machine learning classifiers to detect various activities performed in the mm-wave radar sensors area of contact, which is also constrained to using a single sensor. Meis and Schneider[8], discuss future safety and how enhancing the sensing capabilities of the sensors would be advantageous for the environment in the automotive industry and illustrate the same using histograms, graph plots of the chirp rates emitted in the radar. They typically perform image acquisition and interpretation study to illustrate and determine the high spatial and resolution features offered by the mmWave sensors by visualizing the results and evaluating the aspects relative to their capabilities. These results show the ability to handle data and transforming it to derive at various arguments.

Hence, it can be seen that many authors surround themselves by the image data visualization methods to refer the results from the data produced. Measuring the effectiveness of data visualization methods is crucial for this thesis which could generalize the approach for other practitioners who intend to use them for similar purposes would be a good step towards the contribution in software industry[2].

With all that said, data visualization is becoming more and more established in every application. However, technologies for validation are scarce and immature, particularly in object detection and radar technology. Only a few studies focus on evaluating visualizations of radar data or sensor data. Hence, this research (Evaluating an effective Visualization method for Safety Radar System) is a small step towards high-quality results and machine Safety.

In this chapter, we will be presenting the research questions and their respective research methodologies. Each section in this chapter is concerned with addressing the process (Figure 5.1), and the flow of the method implemented to collect data and arrive at meaningful decisions.

Systematic procedures are critical for addressing research questions. These procedures should be defined to the finest level of granularity with the structured protocol for uniform adaption. A research methodology is an encapsulation of the systematic procedures exercised with rigours effort to dwell meaningful results from research[22]. Majority of contributions in Software Engineering relies on the following Research Methodologies[23][24][25][26][27]. Their descriptions are as follows:

- **Case Study:** Case studies are an empirical investigation of the contemporary phenomenon in real-life context from multiple sources of evidence[22]. They have a flexible design with many variables and primarily generate qualitative data to serve its exploratory rationale. A generalisation of results from case studies is restricted and can be extended only with triangulation. They are ideal for projects with long term commitments.
- **Experiment:** Experiment has controlled factors and generates quantitative data to practically prove theory and hypothesis[22]. It has a fixed design and less human influence. They are replicable and generalisable.
- **Survey:** Surveys collect standardized information from a sample of a population [22]. They have a fixed design with few variables and provide statistical generalisability. They rely on quantitative data generated from different collection methods to derive findings. They are highly affected by human influences like behaviour.

5.1 Research Questions

RQ I: Which Data Visualization methods are used in demonstrating the Safety Radar System?

Motivation: To understand and identify which data visualization methods are used in the safety radar system. Provides insight into how the data can be combined and what is the improved reliability. Most importantly, an overview of the detection

process involved in these systems and how it can be illustrated in a meaningful way to convince the developers, stakeholders or users to is understood.

RQ II: What is the effective Data Visualization method suitable for demonstrating the improved reliability of detection with multiple radar modules in the safety radar system?

Motivation: The motivation towards this research question is to evaluate the understandability, effectiveness and user satisfaction of the visualization methods in detecting the objects. The evaluation is key to determine an effective visualization which could illustrate all the intended outcomes of the system through the demonstration software and its relative performance. This will evidently aide in the stakeholder's decision making process regarding the business value of the proposed system. Further, this question will be a major contribution to the thesis as it will generalize similar data visualization problems with similar applications.

5.1.1 Method Selection

The following methods[28][24][29] are followed in a hierarchical order to answer the above research questions. Initially, we consider a Multi-vocal Literature review (MLR) for RQ I due to the availability of useful information from different platforms and an exploratory case study for RQ II using semi-structured interviews and task-based evaluation conducted on 14 participants which include seven working team members with couple experts from ABB jokab Safety and seven Masters students from BTH University, Karlskrona to perform the evaluation. Further, the qualitative analysis from the interviews helps to support and validate the arguments and results obtained from the case study and collect any new ideas, implementations towards this particular case. Figure 5.1 illustrates the process of conduct in this thesis.

Experiment and Survey were not chosen as there holds a strong motive in this research which is the unknown nature of the data and usage of multiple radar modules as an input. The factors for evaluation are also not known in the initial stages but is an iterative process hence eliminating the scope of an experiment. In order to understand various facts and evaluative assessments from the participants, direct interaction and discussions are required for an in-depth investigation. This helps understand the industrial perspective closely for further development in the project. Although Surveys are a good medium for data collection. It would not provide reliable solution to the problem[23] [28].

Questions like What? Why? How? are better answered through the exploratory nature of a Case study enabling us to investigate with a comprehensive nature over the problem domain. Adding to it, Case studies allow us to interact, evaluate new methods, tools frameworks with real-time data which have high standard results and provide fair reciprocity. Further detailed justification and motivations with respect to the implemented research methods are furnished in sections 4.3 & 4.4.

5.2 Research Design

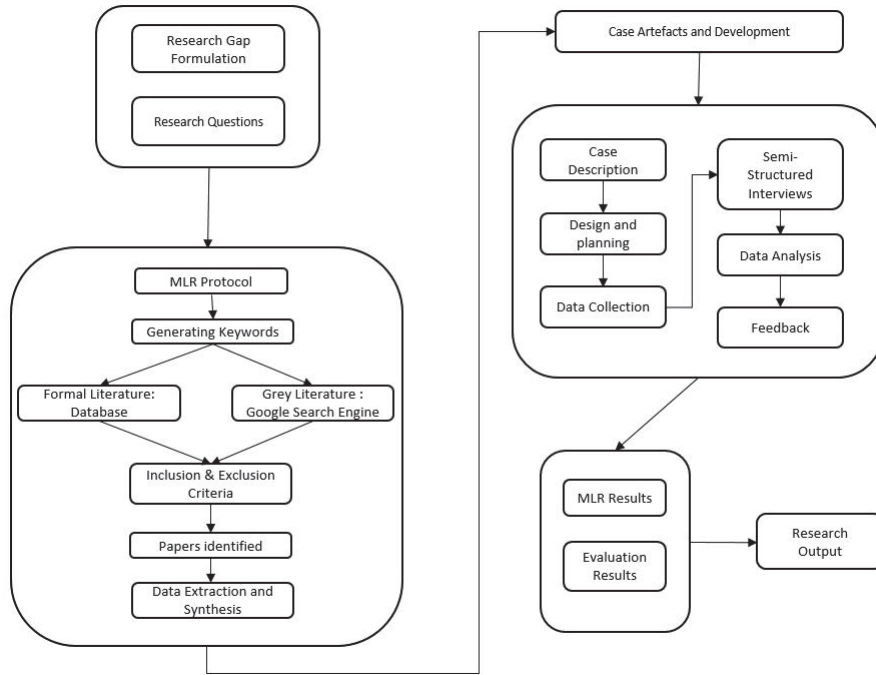


Figure 5.1: Research Process

Figure 5.1 illustrates the flow of process in this thesis. To answer both the research questions we make use of a Case study with task based evaluation using semi structured interview and a Multivocal literature review which are discussed in the following sections.

5.3 Multi Vocal - Grey Literature

Multi-Vocal is performed in order to understand the given problem and gain perspective into this area of study. As most of the related work and important information articles with regard to the working of mmW radar technology is a new area of study to authors. Knowledge kits, documentation in the form of developers guide, user manuals and other relative information are furnished at the company serve more importance. Additionally, extensive knowledge of the existing literature on what visualizations methods are used for such systems should be tabulated. This provides a base to the research and development team at the industry for selecting the suitable methods for evaluation with respect to the features developed in the demonstration software.

Multi-Vocal Literature review is a mixture of both systematic literature review and grey literature which means exploration of published literary works and other sources like blogs, articles, company artefacts, company manuals, videos or white papers is performed. As the literature on data visualizations used in radar

systems are scattered across various streams of sources and very limited work on visualizing mmW radar sensors data, this method was an apt selection to gather the required knowledge. Following are the two sections which state a well-structured protocol which is followed in collecting the relevant literature.

5.3.1 Data Collection

The MLR is categorized into two divisions wherein the 1st division we read the company artefacts, knowledge kits and search in Google for articles, websites or related documents. In the 2nd division, we perform a search in the scientific databases for related articles, journals or publications. Keyword and search strings are used for the electronic search, which is based on the scope of the problem and research questions. These are formulated in Table 5.1. All potential articles or useful information are filtered using the inclusion and exclusion criteria from source selection in Section 4.3.1.2.

5.3.1.1 Literature search string Formulation

The effort put into searching relevant studies and papers is crucial in this thesis. Accumulating the required information in the literature is only possibly by effectively choosing the search strings and formulating them to perform the search across multiple scientific database platforms. Initially, the search strings are formulated based on the scope and objective of the research questions and related keywords to the research problem. Below is the Table 5.1 are the set of search strings used in both scientific databases and in google for grey literature.

Keywords	Search Strings
Data Visualization, Radar, Sensors, Data, mmWave, mmWave Radars, Evaluation Metrics.	"Data Visualization" AND "mmWave Radars", "Data Visualization" AND "mmWave Sensors", "Evaluation of Data Visualization methods", "Data Visualization evaluation" OR "Evaluation of Data Visualization" OR "Data Visualization method evaluation" OR "Data Visualization approach evaluation"

Table 5.1: Multivocal Literature review : Search String Formulation

5.3.1.2 Literature Source Section

After the relevant literature is search initially, rules from the inclusion and exclusion criteria are used to filter out unwanted articles and repeated content. Due to the presence of grey literature in this knowledge-gathering process, a re-screening of the filtered out data is performed, which is considered as the literate insight into this research. Below Table 5.3 and Table 5.2 are the Inclusion and Exclusion criteria which are formulated to select relevant and credible studies from the gathered sources.

Exclusion Criteria

Exclusion description	Study rationale
Duplicate Studies	Replicated or duplicated studies are eliminated to maintain consistency in results.
Irrelevant and Incomplete papers	Studies with non-comprehensive protocol and results are excluded to avoid unwanted direction to the Research.
Grey literature in non-textual formats and from public opinion, discussion forums and networking	Discarded given these include general expression of thought with biased by personal choice and improper background evidence.

Table 5.2: Exclusion Criteria for the Literature

Inclusion Criteria

Inclusion description	Study rationale
Visualization techniques for radar data	Our research study essentially grounds on different visualization techniques available for radar data.
Understandability of visualizations by different stakeholders	Our research aims to draw a comparison between different visualization techniques available for radar data.
Full text access	Clear evidence and critical background is inevitable for sourcing the knowledge. Summarised variants cannot satisfy the research.
Include formal and grey literature	Given the research area is young, it is difficult to find formal literature answering the research objectives. It is important to include grey literature and organisation artefacts to answer the RQs.
Paper written in English Language	Translated sources often involve text discrepancies and expression bias inflicting and undesired direction to research.

Table 5.3: Inclusion Criteria for the Literature

5.3.1.3 Literature Quality Assessment

This step in the MLR is important as all the content relevant to the scope of this study should be free of bias and should be retrieved from credible sources. This ensures that the data and knowledge gathered is usable with good standards. Quality Assessment in this study is performed using a checklist (Table 5.4) based on the

criteria written by Garousi et al[28].

Criteria	Questions
Date	a) Does all the listed papers have the published year?
Relation	a) Are the papers listed from both formal literature and grey literature have the desired outcome ? b) Are the stated conclusions supported by the presented data?
Methods	a) Does the following papers have clearly state Research question and objective? b) Are relevant papers and studies clearly described and identified using search methods?

Table 5.4: Literature Quality Assessment Criteria

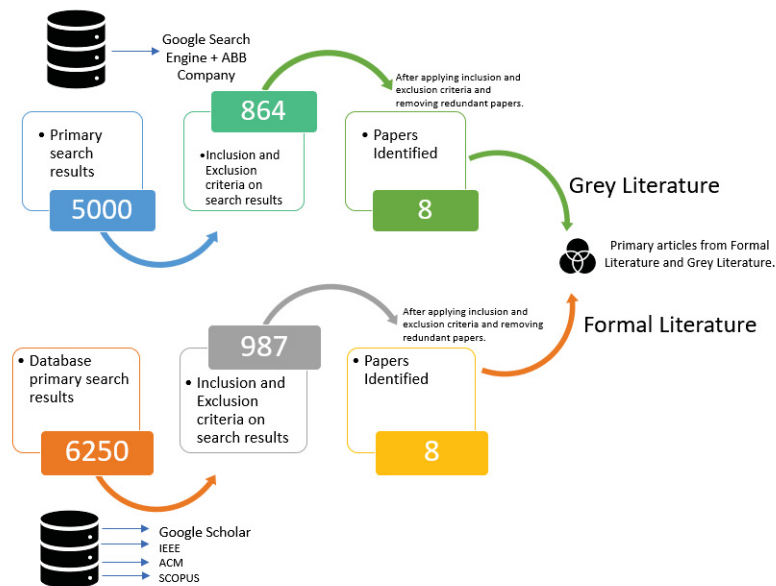


Figure 5.2: Literature Search Process results

5.3.2 Data Analysis

The literature data was extracted following the guidelines of Tore et al.[30]. The two primary channels for sourcing were:

- Scientific Databases for Formal Literature: Peer-reviewed publication from ACM, IEEE, Springer, SCOPUS and Google Scholar.
- Google search engine and ABB Jokab Safety for Grey Literature: Articles available in textual formats with appropriate detail and background granularity across the Internet and company artefacts.

1. **Manual inspection of results:** The data extracted was carefully examined manually to identify obvious similarities and set a base start for further synthesis. In our case, all the primary studies revolve around data visualizations and radar technologies. Therefore, to identify relevant visualizations, we first try to investigate the use of different visualization techniques in the context of radar data.
2. **Data encoding:** Post manual exploration of the data, initial codes are framed based on the identified similarities. The structure of these codes is intended to answer the research questions. The encoding was done based on the following steps:
 - (a) Identify different visualization techniques.
 - (b) Examine their application and usability.
 - (c) Validate its extensibility on radar data.
 - (d) Extract adaptation and mapping procedure of the visualization technique for radar data.
 - (e) Extract user-experience and understandability of the visualization on radar data.
3. **Code to theme transition:** Repetition and redundancies were eliminated based on the categorisation of codes. With the narrowed categories, themes encompassing the majority of categories were generated.
4. **Higher-order theme selection:** Combination of themes was used in relation to synthesize better results and coverage of research objectives. They were structured with the emphasis to answer the research question.

5.4 Case Study

Due to the nature of this research problem into the understanding the Data Visualization process and development of the coordinate transformation engine to combine multiple homogeneous data sets into a single frame and perform an evaluation over different visualization methods, case study seemed to do justice.

This exploratory case study aims at closely working with the industry to visualize the transformed data obtained from the hardware equipment in order to detect objects and determine the effectiveness and reliability of the radar detection using data visualizations. With context to software engineering, a case study could be defined as the empirical investigation into a contemporary problem which has real-life application and the factors, characteristics couldn't be distinguished. Both qualitative and quantitative data are collected and investigated in order to determine which data visualization method helps in an effective perception of the newly developed Safety Radar System at ABB Jokab Safety. A case study offers more flexibility and is modular in design. A well-defined plan will ensure satisfactory results.

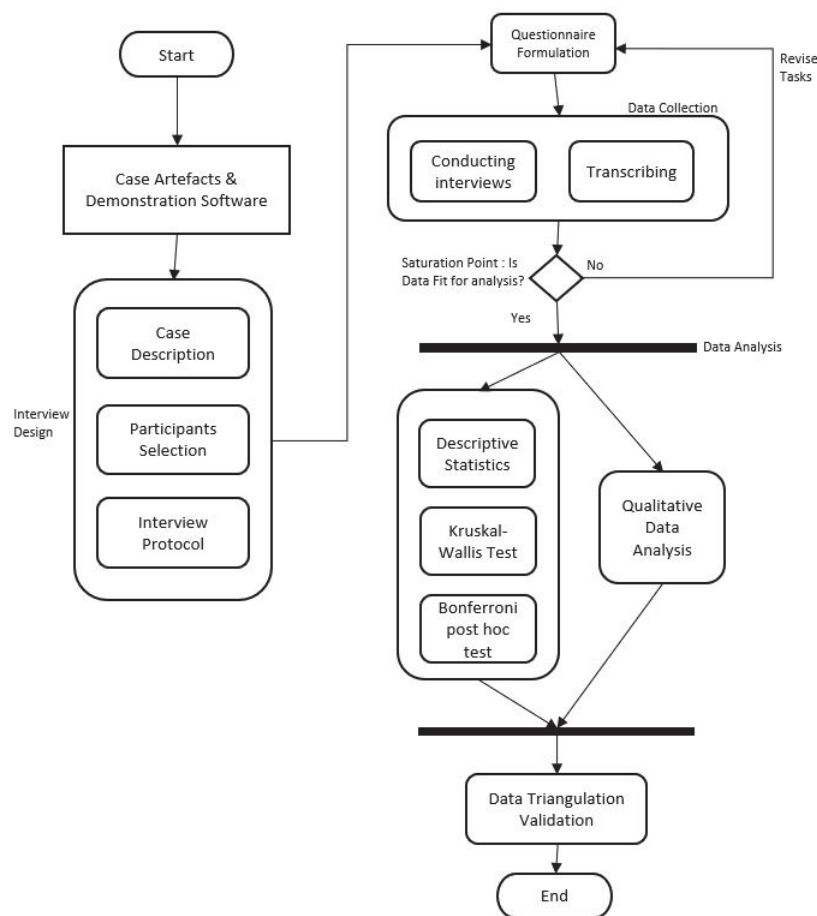


Figure 5.4: Case Study Protocol

5.4.1 Case Description

ABB Jokab Safety is a division of a Swiss-Swedish made multinational corporation ABB. It is a global pioneer in robotics, equipment for heavy electrical and famous for automation technology and has selected a case of determining the safety measure for humans and objects using Millimetre-wave(mmW) radar technology for industrial purposes. mmW is a growing radar technology which uses short-range radio frequency spectrums for object detection. An ongoing research and development project known as Safety Radar System with more than one radar modules to improve detection reliability is being deployed. With few data complexity, collaboration and size management challenges, developers and practitioners in the team need relevant data visualization methods to illustrate individual features like encircling target, analyzing surveillance volume, clutter/noise removal and combined object detection developed for the demonstrator software. Further, In order to convince the stakeholders an effective Data Visualization method which can portray all the features is to be evaluated. An in-depth investigation is required to evaluate an effective data visualization method which impacts the improved detection reliability is required. An overview of the case study process is displayed in Figure 5.4.

5.4.1.1 Unit of Analysis

From the case stated above, the data visualizations are considered as the unit of analysis (Figure 5.5) due to the fact that the results produced from the system are crucial for the company and users who might not have full knowledge in this field. Stating the functionality and results using illustrative graphs which are visualizations methods is the goal of this study, making visualizations as the unit of analysis.

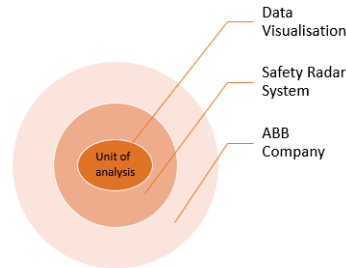


Figure 5.5: Unit of Analysis

5.4.2 User Evaluation Criteria and Metrics

Various studies propose distinguished evaluation criteria and usability heuristics to evaluate information visualization or data visualization methods. After an extensive background study and daily meetings with the developers, practitioners and experts at the industry we have come to a conclusion that effectiveness, efficiency and satisfaction of a user, when interacted with the target object in evaluation, is a good way to measure the usability of an object. ISO 9241-11:2018 defines the above mentioned metrics as[33][34]:

- **Effectiveness:** It is defined as the ability of a user to achieve the specified goals with *Accuracy and Completeness*, where accuracy is checked with respect to how good the provided outcome matches the intended outcome. Completeness is calculated on how well do the users achieve the desired outcomes.
- **Efficiency:** It is defined as the number of resources utilized to complete or meet the outcomes. Here the resources could be either time, money, effort or similar entities.
- **Satisfaction:** It is defined as the level of expectations, wish or desire fulfilled, which can be through the user's physical, emotional or cognitive abilities.

Section 4.4.3.1 utilizes these metrics as a baseline to perform the task-based evaluation study on the data visualization methods.

1. **Response Time:** Each task in the interviews is clocked for response time where the amount of time consumed by each participant to complete each task is recorded.
2. **Ease of Understanding:** Towards the end of each task, A Likert scale of 1-10 is presented to the participants where they provide their feedback of how easy what the task to understand and how easy was it to perform.

3. **Confidence:** A Likert scale of 1-5 is presented to the participants where they rate the level of confidence for their answers. This indicates the extent to which they are sure of their response.
4. **Accuracy:** Accuracy is measured by calculating the number of correctly answered tasks by a total number of tasks with the help of pre-determined results of the data sets under the guidance of the industrial supervisor.

User Feedback is collected toward the end of the interview where each participant is asked which Data Visualization method is more effective, informative and conveys the objective of the system. Adding to their selection, a Likert scale of 1-5 is presented to rate the accuracy of the selected method.

5.4.3 Data Collection

This case study aims to collect both qualitative data and quantitative data in a real-time industrial environment from concluding which data visualization method is effective and offers more understandability and accuracy over the data. Initially, raw data sets in the form of ".bin" files are delivered by the team members with the chosen data visualization methods with reference to the existing literature. The data is then combined using the developed coordinate transformation code, and the resultant visualization methods are evaluated by conducting semi-structured interviews. These interviews are conducted on 14 participants with 7 ABB employees and 7 Master thesis students at BTH University, Karlskrona with relevant experience and academic background in using data visualization methods. Refer Figure 5.4 and 5.6 for an overview of the process.

Interviews are considered to be the best way of collecting both qualitative and quantitative data in software engineering due to its nature of capturing detailed information regarding a problem and its characteristics. In total, there are three types of interviews known as fully structured interviews, semi-structured interviews and the unstructured interviews which are formulated based on the need and scope of the goal. Due to the exploratory and descriptive nature of the semi-structured interviews it is the best fit for this study. The need to question about various different visualization methods and their importance in transforming raw data to something understandable is crucial, hence a questionnaire with open-ended questions and few close-ended questions with appropriate scales are formulated[27].

5.4.3.1 Semi-Structured Interviews

Semi-structured interviews are considered to be the best data collection method to understand which data visualization is better and effective in illustrating the raw data from the safety radar system. Its modular design and flexible characteristics enable the easy collection of both qualitative and quantitative data in a real-time industrial environment. Figure 5.6 illustrates the interview protocol followed in this study.

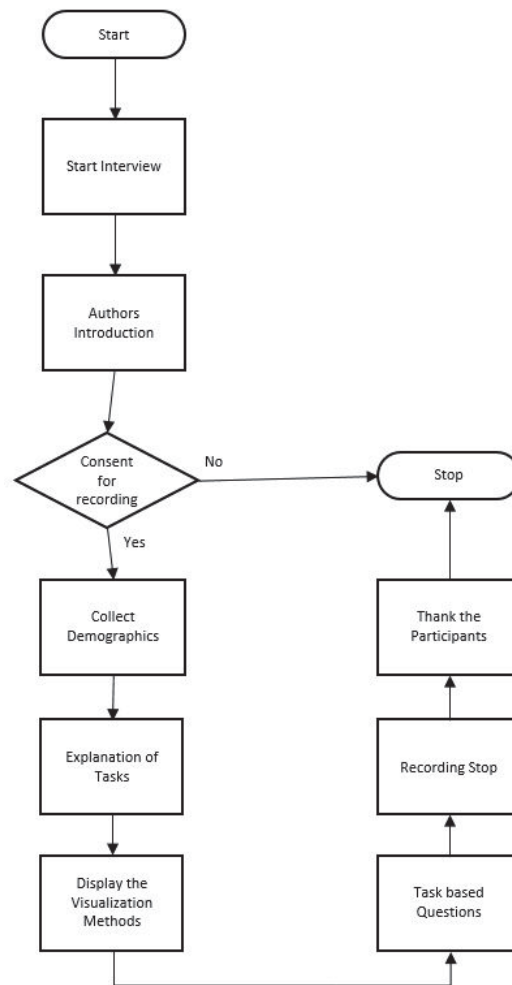


Figure 5.6: Interview Protocol

1. **Interview Planning:** The Interviews are concentrated to extract the experience and feedback given by the participants on selecting an effective visualization method. Convenience Sampling is considered for selecting the participants as we need two types of participants, One with people who are related to this study and the other people who do not have any exposure to the problem. Invitations for the interviews are sent to the participants in the company through the mailing system provided in the company and the same was done for the participants selected outside the company. The Invitation email sent to the participants consisted of the goal of the interview, date, time, place and introductory information regarding the objectives of this study and basic background with related terminology. A feature to account for the consent and rescheduling option was also attached in the invitation. The interviews concerning company employees were conducted in a conference room at the company and for those selected outside the company were conducted online through a video-conferencing tool. All these interviews were recorded only after the participant's acceptance and transcribed the next day for analysis.
2. **Participant Selection** A total of 14 participants were selected for this study

where seven are from the case company and seven Master’s Students from BTH University with relevant experience and academic background. Relevant experience accounts if they were part of any evaluation project or study, background in data analysis, pursuing research or career in the field of data visualizations, familiarity to this research objective, object detection or machine safety. Majority of the participants from ABB are considered to be skilled and experienced in this field of study. The other participants do not have any exposure on the Safety Radar System but have knowledge in visualization methods with few insights on its evaluation.

Interviewee	Background	Duration	Experience	ABB
1	SW Developer	30 min	1+ Years	No
2	Data Analyst	35 min	1+ Years	No
3	SW Developer	30 min	2+ Years	No
4	SW Developer	31 min	2+ Years	Yes
5	SW Developer	34 min	1+ Years	Yes
6	Data Analyst	28 min	1+ Years	No
7	Data Analyst	22 min	1+ Years	No
8	SW Architect	30 min	2+ Years	No
9	SW Developer	31 min	2+ Years	No
10	Industry Expert	33 min	4+ Years	Yes
11	SW Developer	30 min	2+ Years	Yes
12	ML/AI Developer	28 min	2+ Years	Yes
13	Industry Expert	29 min	5+ Years	Yes
14	Data Analyst	30 min	2+ Years	Yes

Table 5.5: Participants in the Interviews

- Interview Questionnaire** The interview is designed for a **task based evaluation** where each participant is given four tasks with some open-ended questions towards the end for each visualization method and the structure is showed below in table 5.6. Before the actual tasks, demographic information was collected, which can be found in the appendix. The detailed interview questionnaire and demographics can be found in Appendix A.1
- Transcription:** After the interviews are conducted, we store the recordings and following on the next day they are transcribed from audio format to text format using software known as the InqScribe. All the recordings are stored in a hard drive owned by the authors, and a backup is created in the company server for future reference. The converted textual recordings were accessed later for data analysis where relevant relations and feedback was to be found.
- Post Interview:** After each interview we thank the interviewees for their participation and additionally send an email containing Thank you note and the appreciation for their effort.

Task	Questions	Form of Answer
<i>(Task1)</i> Identification of Object	1Q	Two-answer (Yes/No)
	2Q	Brief Answer / Reasoning
	3Q	Likert - 10
	4Q	Likert - 5
<i>(Task2)</i> Identification of the Radar Modules with respect to the objects	1Q	Two-answer (Yes/No)
	2Q	Brief Answer / Reasoning
	3Q	Likert - 10
	4Q	Likert - 5
<i>(Task3)</i> Estimating the size of the object detected in a visualization method	1Q	Short Answer
	2Q	Brief Answer / Reasoning
	3Q	Two-answer (Yes/No)
	4Q	Likert - 10
	5Q	Likert - 5
<i>(Task4)</i> Feedback	1Q	Multiple Answer (Yes/No/Maybe)
	2Q	Brief Answer / Reasoning
	3Q	Brief Answer / Reasoning

Table 5.6: Task Evaluation Table

5.4.4 Task Based Evaluation

The semi-structured interview is designed to deploy a task-based evaluation. An in-depth insight into the design is presented in this section.

1. **Participants and Design:** Like discussed in the previous section 4.4.3.1 under participants selection, a total of 14 interviewees, are selected who will be presented with four tasks. Most of them were in the age group of 18-25 years. Initially, a demographic is established for each interviewee to understand their familiarity with the subject and relevant background information. All of the interviewees were fluent in English. Each interviewee was individually interviewed to avoid any influence from others and prior to the actual interview, an introductory research proposal was sent to the participants in the invitation. The interview was started once consent for recording was collected. All the interviewees were given four tasks for each data visualization method. To avoid any familiarity with the results from the data, each visualization method illustrated a different dataset but with the same characteristics of the data. The independent variables for this evaluation were considered to be the visualization methods, and dependent variables are the evaluation metrics response time, ease of understanding, confidence and accuracy. A design representing the same can be observed in Figure 5.7.
2. **Tasks and Materials:** The interviewees were presented with 4 tasks (Table 5.6 for each visualization method. The authors recorded the interview

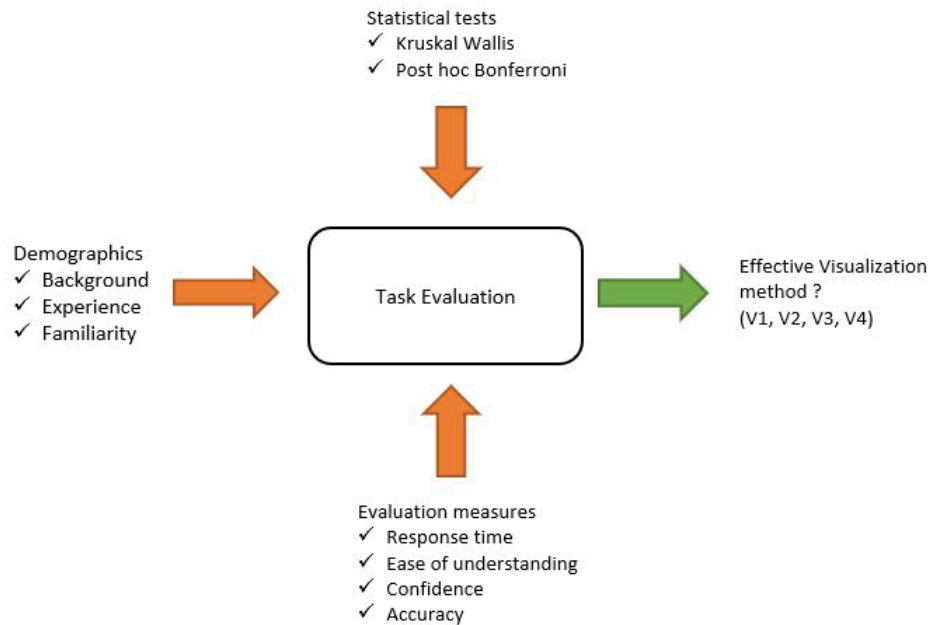


Figure 5.7: Task evaluation design

and clocked the response time for each task. The participants consumed 3-5 minutes, answering all the tasks for each visualization method. To keep the interview more engaging and avoid tiredness, the last task consists of open-ended questions to check the interviewee's presence and note any feedback. All the visualization methods were shown on a screen, and each task was questioned by the authors to the interviewees. The tasks in this evaluation were based on the system's objective to detect objects using multiple radar modules with all the features to offer increased detection reliability. The interviewees were asked to identify the object which shows the ability of a visualization method to convey the position of the objects. Then identification of the radar modules was asked, which aims at locating the radar module positions with respect to the object being detected. Size of the object which is being detected, and the relevant description was asked, which show the ability of the interviewees to estimate the cross-section of an object through the visualization. Ease of understanding in the task and confidence of their ability to answer were also collected. Finally, a feedback task to collect important suggestions and experience of the interviewee is taken for each visualization method. Accuracy of the tasks was calculated by the authors where they had access to the correct results of each visualization under the guidance of the industrial supervisor.

3. **Order of viewing:** The order of displaying the visualization methods was same for all the participants as we used different data sets for each visualization which ensured there is no significant impact on the interviewee's perception and cognitive ability. Additionally, prior to each interview, all of the interviewees had received a research proposal and background information regarding this project in order to eliminate or reduce knowledge bias over the visualization methods and previous experiences of the interviewees.

5.4.5 Data Analysis

Initially, informative demographics, descriptive statistics for the quantitative data is established in the study with the general overview of the participants. With the data collected through interviews, data analysis is divided into two parts:

1. **Quantitative Data Analysis:** Based on the evaluation criteria and selected metrics (time, confidence, ease of understanding and accuracy) the data is analyzed using descriptive statistics and statistical tests. Based on the tabulated descriptive statistics and by performing a Normal Q-Q plot using **R tool**, it is observed that all the 16 data samples (4 metrics X 4 Visualization methods) are Non-normally distributed. Hence we decided to conduct a Non-parametric test. As there are more than two individual groups (Visualization methods), the Kruskal-Wallis test is selected for the analysis followed by Bonferroni post hoc test to filter out a single group.
2. **Qualitative Data Analysis:** *Narrative Analysis* is selected to analyze the qualitative data obtained from the interviews. This method is selected in order to summarize the data from the open-ended questions in the interviews. This method aims at converting human perceptions, experiences and interpretations into summarized points or a story

5.4.5.1 Quantitative Data Analysis

Introduced in the year 1952 Kruskal-Wallis test is a non-parametric statistical test designed to check if there are differences between 3 or more independent groups with a continuous variable that has non-normal distribution [35]. This test is conducted to analyze if there are any groups with a significant difference. If so, followed by Bonferroni Post-Hoc Statistical test to filter out a single or paired group with significant differences by comparing the means of the independent groups (i.e., Data visualization methods). Further, the Kruskal-Wallis test is based on the following conditions with regard to any exploratory or inferential study:

1. **Condition 1:** There are more than two independent sampled groups of data in the study.
2. **Condition 2:** The dependent variable is continuous.
3. **Condition 3:** The data should be non-normally distributed for Kruskal-Wallis tests.

Kruskal-Wallis Test

1. **Determination of Test Hypothesis:** In order to analyze and evaluate the data from the interviews, we need a clear goal of the analysis. Hence, we consider a null hypothesis which states :

H_0 : There is no significant difference between the medians of groups over a respective metric

Which is contradicted by an alternate Hypothesis stating:

H_1 : There is at least one group of medians which have a significant difference.

2. **Calculation of Statistical tests:** Once, the hypothesis is determined with all the conditions met, we conduct a statistical test for the collected data using **R tool** which is a frequently used programming language for statistical computing, data analytics and presentation. In this R tool, the Kruskal-Wallis test is executed like shown in the below Figure to determine which Independent group(s) are significantly different with respect to the chosen metric from the evaluation criteria.

Now, to conclude which hypothesis is correct, Kruskal-Wallis makes use of a difference threshold H which can be search in the *Chi-square* : χ^2 table i.e., based on the value of α (level of significance).

3. **Decision making:** This is a crucial step in this process and evidently two kinds of assessments can be made from the analysis namely:
 - (a) **Potential Assessment 1** : H_0 is not rejected. i.e., *Calculated* $\chi^2 < \text{Critical } \chi^2$. This indicates that there is "No" significant difference between the independent groups (Data Visualization methods) and we conclude that all the medians of groups are equal. Similarly, this can be verified via a computer software if *P-value* $> \alpha\text{-value}$. Here α -value is 0.05 (Confidence level=95%)
 - (b) **Potential Assessment 2** : H_0 is rejected. i.e *Calculated* $\chi^2 > \text{Critical } \chi^2$. This indicates that there is at least one significant difference between the independent groups (Data Visualization methods). Similarly, this can be verified via a computer software if *P-value* $< \alpha\text{-value}$. Here α -value is 0.05 (Confidence level=95%)

Bonferroni Post Hoc Test

In general, the Post Hoc Test is a comparison which is made when the results of the Kruskal-Wallis test indicate that the Null Hypothesis(H_0) is rejected. In this case, it is evident that there is a significant difference between the groups. Bonferroni Post hoc test is one of the frequently used and most conservative methods which is best suited to this study as we require a multiple comparison correction[36].

By definition, The Bonferroni Post hoc Test is used to compare distinctive groups and analyze the relationships among variables[37]. This test makes use of threshold values on the basis of t-distribution and is more rigorous over other post hoc tests like Tukey, Scheffe's and more. Further, in this test, the α value should be lowered to remove unwanted positives in the data. Then a New significance value is generated, i.e.,

$$\text{New significance Value } (\alpha_n) = \alpha / \mathbf{n}$$

Where "n" is the number of comparisons between the groups. In this case, there are six comparisons (V1-V2, V1-V3, V1-V4, V2-V3, V2-V4, V3-V4).

1. Determination of Test Hypothesis:

H_0 (**Null Hypothesis**): There is no significant difference between the compared paired groups, i.e., Comparisons.

H_1 (**Alternative Hypothesis**): There is at least one paired group with a significant difference between the compared paired groups, i.e., Comparisons.

2. Decision Making: Once again two kinds of assessments can be made from the analysis namely:

- (a) **Potential Assessment 1** : H_0 is not rejected. i.e., $P_{unadjusted\ value} > New\ Significance\ value$. This indicates that there is "No" significant difference between the compared paired groups (Data Visualization methods).
- (b) **Potential Assessment 2** : H_0 is rejected. i.e., $P_{unadjusted\ value} < New\ Significance\ value$. This indicates that there is a significant difference between the compared paired groups (Data Visualization methods).

If Assessment 2 results from the test, Median values of the individual groups from box plots are considered as a way of decision making.

5.4.5.2 Qualitative Data Analysis

Narrative analysis is generally used in both the quantitative and qualitative analysis of data in empirical studies. These interviews are focused on extracting the feedback and perceptions of the participants in this study into a summarized story or points for determining an effective data visualization method and record all the reasons, views and justifications which state their decisions. This analytic method helps for an in-depth understanding of how the participants have given their answers and formulate a story based on the existing information from the questions. Hence, reading the transcribed interviews and formulating a conclusive story directly contributes as an answer to the research problem. Adding to it, the narrative analysis also helps in structuring the data in an organized form to convey better insights. Open-ended questions are better analysed using this method as the participant's data is valuable and any insights help rule out insignificant Visualization methods for this research problem. For each question in the interview questionnaire, the following are established:

1. **Scope of the Question:** Description of the question and how it is relevant to this thesis is established. Potential assumptions on how the participants might react are mentioned with what kind of results are to be expected.
2. **Participants Responses:** All the important responses from the interviews are presented which describe how? What? and why? of the feedback.

3. **Narrative analysis results:** A summary of how all the participants have responded, and their choice of answers are given with supportive arguments from the interview.

In this chapter, results and supported analysis for the research questions are furnished.

6.1 RQ I

The goal of any literature study is to gather and equip ourselves with knowledge which is required to solve or devise a path towards the outcome to a given problem. Due to the rigid planning and robust design formulated for this literature study, we have arrived at collecting the following papers and information on different data visualization methods which can be used for object detection the safety radar system using multiple mm-wave radar modules using the newly developed coordinate transformation engine.

6.1.1 Multi-vocal Literature Review Results

A total of 16 articles were filtered and presented after implementing the inclusion and exclusion criteria on the data retrieved from formal literature and grey literature. Online database systems like Google Scholar, IEEE Xplore, ACM Digital Library and Scopus was used for electronically searching the articles related to this study which contribute to the formal literature. And the company artefacts, user manuals, documentation and websites, papers from Google search engine contribute to the grey literature. Below is the table; the selected sources are furnished, which are further analysed using thematic analysis and potential data visualization methods are selected, which could be used in the case study.

During the initial search using the search strings tabulated in Table 5.1 a total of 6250 formal literature and 5000 grey literature were found which later were put through a series of the inclusion and exclusion criteria to filter out the duplicate, irrelevant studies followed by screening abstracts, the scope of the paper and related outcomes 987 of formal literature and 864 grey literature are collected. A careful and in-depth reading by the authors resulted in eight articles from the formal and eight grey literature which were identified as potential results for the MLR, as shown in Table 6.1.

Literature Paper ID	Selected articles for study	Type of Literature
LP1	Barrett, Dennis, and Adrian Alvarez. mmWave radar sensors in robotics applications. Technical report, Texas Instruments, 2017[38].	Grey Literature
LP2	Antenna alignment and positional validation of a mmWave antenna system using 6D coordinate metrology.[39].	Formal Literature
LP3	Data statistics and image properties of a large multistatic mm-wave imaging system.[40].	Formal Literature
LP4	TI, Tracking Using. Traffic Monitoring Object Detection and Tracking Reference Design Using Single-Chip mmWave Radar Sensor.[41]	Formal Literature
LP5	Introduction to automotive FMCW Radar Technologies: Using Texas Instruments mmWave AWR sensor series.[42].	Formal Literature
LP6	RAO, Sandeep. Introduction to mmWave sensing: FMCW radars. Texas Instruments (TI) mmWave Training Series, 2017[43]	Grey Literature
LP7	The Data Visualisation Catalogue - Radar Chart[44]	Grey Literature
LP8	Imaging Radar Using Cascaded mmWave Sensor Reference Design[45]	Grey Literature
LP9	Obstacle Detection Using Millimeter-Wave Radar and Its Visualization on Image Sequence. [46]	Formal Literature
LP10	mmWave Demo Visualiser [47]	Grey Literature
LP11	mmWave Studio [48]	Grey Literature
LP12	Texas Instruments AWR1642BOOST Radar Sensor Evaluation Module (EVM)[49]	Grey Literature

Literature Paper ID	Selected articles for study	Type of Literature
LP13	Single-chip 60-GHz to 64-GHz intelligent mmWave sensor integrating processing capability[50]	Grey Literature
LP14	Extending Reliability of mmWave Radar Tracking and Detection via Fusion With Camera[51]	Formal Literature
LP15	ThuMouse: A Micro-gesture Cursor Input through mmWave Radar-based Interaction, [52]	Formal Literature
LP16	Millimeterwave bistatic scattering from ground and vegetation targets.[53].	Formal Literature

Table 6.1: Multi Vocal Literature Review Results

Due to lack of contributions and significant work in this field, extraction of usable information from the literature had to be thorough and verified at every step. It is observed that minimal studies discuss this research problem and most of the data extracted was found by similar mapping results and considering the need to find data visualizations methods which support and could be used to illustrate the raw data results generated from the safety radar system with multiple radar modules at ABB.

After carefully reading the literature, a data analysis method which was responsible for identifying relevant patterns, themes and codes which could determine potential data visualization methods was adopted.

6.1.2 Thematic Analysis Results

Thematic analysis (Figure 6.1) is carried out to categorize the codes into respective themes for better understanding of the interpreted results. It is a known and common approach to recognise, analyse and report the related codes from the papers in the form of themes. Here, In Multivocal Literature Review first, the data is collected by reading the Relevant Literature Paper for our study then they are broken down into codes. Later, themes are identified for organising the listed codes. This is done to show what are the different data visualization methods, how and why are they important in the radars and to our study of research. We have drawn the conclusions on different visualization methods through a thematic analysis where initial eight codes were generated as shown in table 6.2 where Cx (1,2,3,4,. . .) denotes codes. Then later, these codes are organised into one theme.

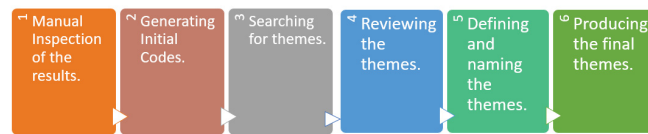


Figure 6.1: Process for thematic analysis

6.1.2.1 Familiarizing with the Data

This is the initial phase, where one intends to read the collected literature paper and includes the necessary data for our research study. It involves continuous reading and complete understanding of papers for an effective expected result. To make the collected data meaningful and identify similar ideas, they are broken into codes and organised into themes. Subsequently, after getting familiar with the data, we iterated the process couple of times to remove duplication of data and to make sure the relevant and important information is not missed.

6.1.2.2 Generating Initial Codes

Relating to the situation, the data or information can be interpreted in a meaningful way by breaking down the necessary data into codes. Where codes in the thematic analysis are the collection of unit elements. There were two ways to code the data one was automated computer coding, which was using tools, and other was manual. We preferred to take manual coding so that we can observe the data closely and take only the necessary and required information from the papers through our research knowledge. In our study, we read 16 Literature papers consisting of 8 Grey Literature and 8 Formal Literature. By these 16 papers we have generated eight codes by organising the data into meaningful form during this process as $C_x(x=1,2,3, \dots)$ from the data which is represented in below table 6.2.

Code Number	Codes
C1	Contour plot
C2	Scatter plot
C3	Azimuth heatmap
C4	Frequency plots
C5	Radar Charts
C6	Range Doppler
C7	Time domain Plot
C8	Heap Map Temperature

Table 6.2: Codes from thematic analysis

6.1.2.3 Identifying Themes

Here, the generated codes are grouped into different clusters known as themes so that all relevant data will be gathered together. In our case, we have generated 1

Potential theme based on the codes formed for our study. Namely, what are the different visualization methods in radar systems? To group the codes into theme they are listed, compared and structured manually, which was done in reference to our formed research questions. After assigning the respective and relatable codes to the themes, they were refined and reverified to check if there were any similar data found in the data extracts which helped in making the data more readable and reliable.

Theme	Explanation
Different types of data visualizations	All the codes related to different types of visualization are categorized into this theme.

Table 6.3: Theme Identification

6.1.2.4 Constructing Thematic Network

In this phase, the listed codes and themes are formulated into a network as shown in Figure 6.2 below.

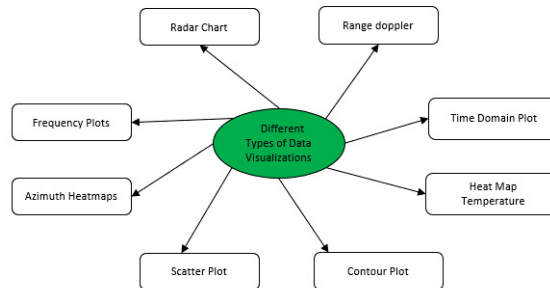


Figure 6.2: Thematic Network

6.1.2.5 Results from the literature

A detailed report of the thematic analysis is finally produced (Table 6.4) after the findings are related to our research objectives and system requirements. The selection of the relevant and utilizable data visualizations methods for evaluation is based on the features and intended outcomes of the system, as discussed in Chapter 2 & 3.

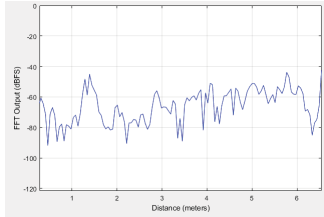
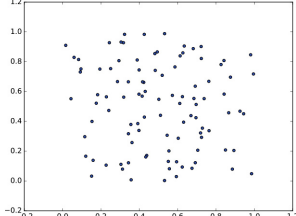
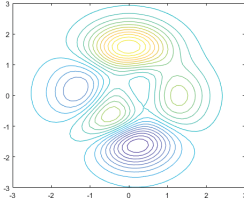
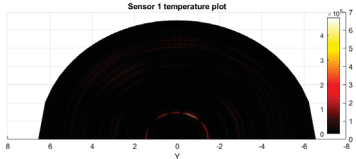
Data Visualization Method	Description	Template
Time-domain	<p>It is a commonly applied graph in radar and sonar systems where a chirp denotes a signal with varying frequency against time. It is a plot between RangeFFT vs Distance in mmWave radar Sensors. It translates the chirp index and properties in Sine Wave form. Every peak in the waves represent a possible presence of an object.</p>	
Scatter	<p>It maps Cartesian values of a combination of attributes in a data. In mmWave radar Sensors, it plots the possible arch of the object in space using continuous coordinate values. Each point on the map denotes a possible location of object in the space.</p>	
Contour	<p>Contour plots visualize three dimensional surfaces on 2 dimensional planes. They are also called Level Plots and are useful for mapping multi-variable functions. Contour plots in mmWaves visualize the object location along with other possible noises in the experimental space. The potential objects in the graph are color coded with encircling shapes.</p>	
Heat Temperature	<p>The heat temperature plot is a minimalistic visualization for object detection in space with mmWave radars Sensors. The object arc is visualized with bright red arc and remaining plane is color coded black to nullify any noise or disturbance offering a clear view to the exact position of the surface in space.</p>	

Table 6.4: Thematic Analysis: Results from the Literature

6.2 RQ II

In this section, we would be presenting the analysis and results obtained from the interview study regarding which data visualization method was more effective in illustrating the data. The metrics considered to evaluate these results are response time, ease of understanding, confidence and accuracy.

Firstly, the background of the participants was required where we collected some demographic information. Present day data analysis consider great deal with regard to the sociological variables like gender, age, qualification and other analytical variables related to the subject of this study to establish a account of suitability and quality into the study.

1. **Age of the participants:** As the figure 6.3 suggests highest frequency of the participants were in the age group of 18-25 years and lowest in the age group of 30-35.

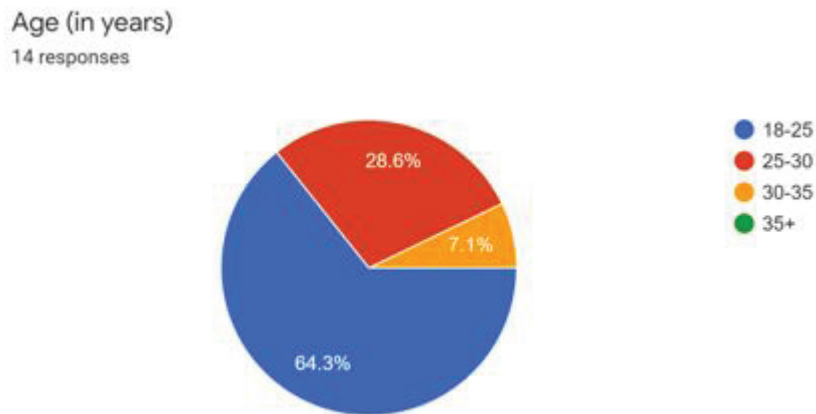


Figure 6.3: Age of the participants in the study

2. **Gender:** As the figure 6.4 suggests, approximately 85% of the participants were male and the rest were female.

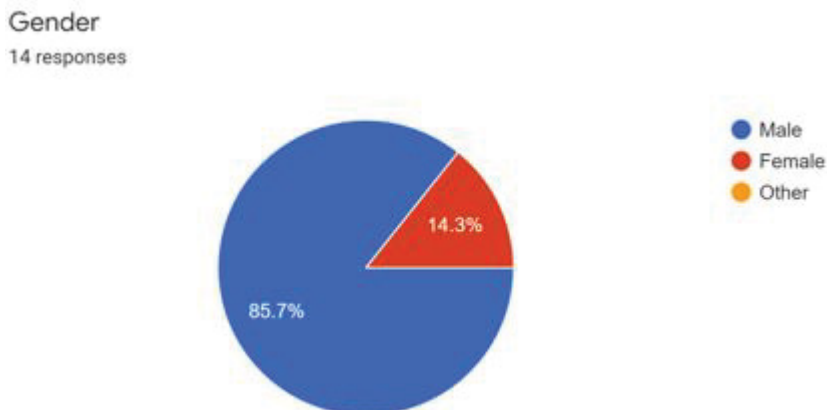


Figure 6.4: Participants Gender

3. **Highest Degree or level of school:** From the figure 6.6 we can understand that highest frequency belongs to the Master's degree qualification and the least begin Bachelor's degree.

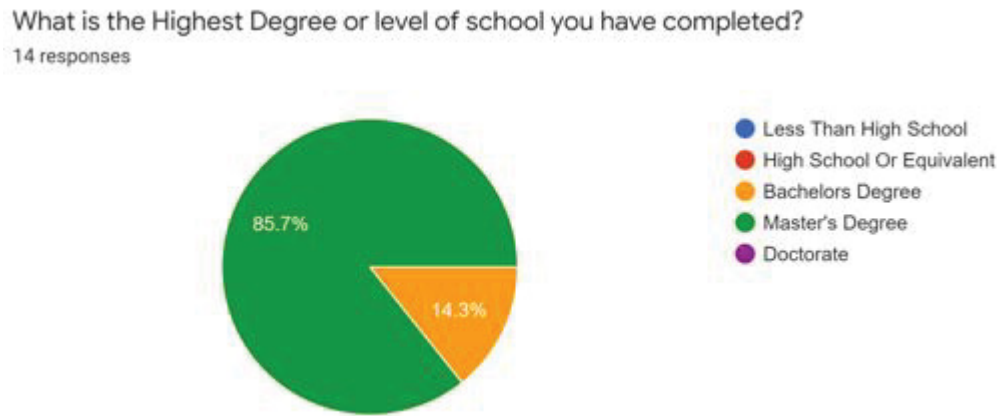


Figure 6.5: Educational Qualification level of the Participants

4. **Field of Study:** As the figure 6.6 suggests that most of the participants were from the computer science engineering and least begin electromagnetic fusion and space engineering.

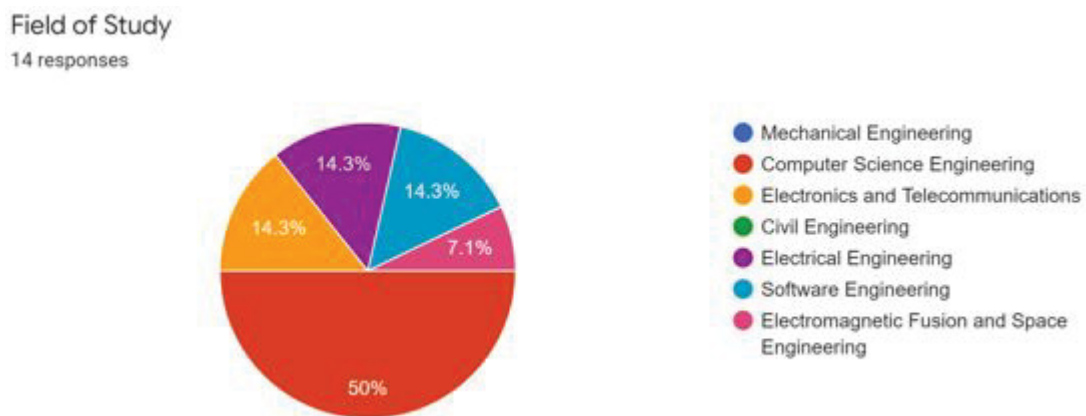


Figure 6.6: Educational Field of study of the participants

5. **Experience in research field:** A the figure 6.7 suggests, highest frequency of participants have 1 year of research experience and the less frequency with 5+ years.

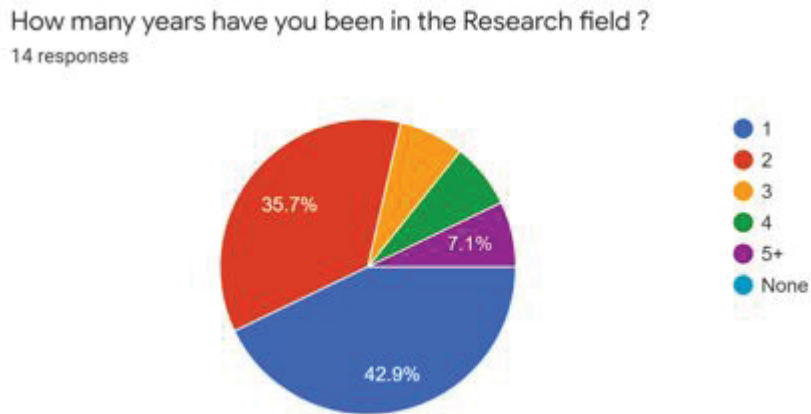


Figure 6.7: Experience in the Research field

6. **Experience in Data Visualization and their evaluation:** As the figure 6.8 suggests, majority of the participants agree that they are aware and have experience with using data visualization methods and its evaluation.

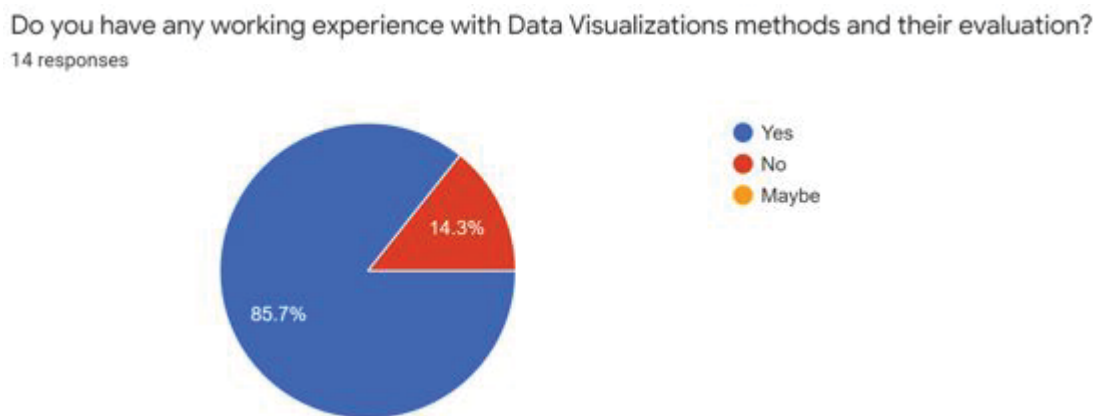


Figure 6.8: Experience with data visualization methods and their evaluation

7. **Previous Projects or Courses with data visualization as a core concept:** Looking at the figure 6.9, We can observe that majority(85.7%) of the participants have taken courses or projects with data visualization as a core concept.

Did you take any previous courses or projects with data visualization as a core concept?
14 responses

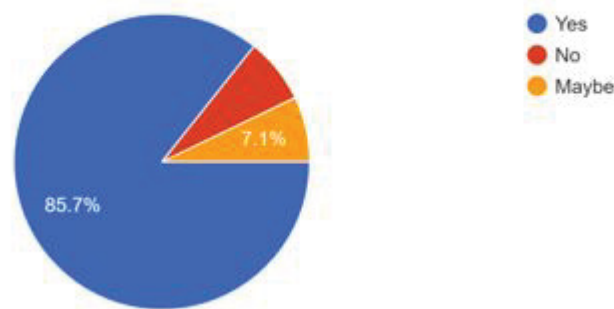


Figure 6.9: Previous encounter with data visualization in either course work or projects

8. Familiarity with the concept of Object detection (Likert scale 1 -10):

It seems that majority of the participants are Familiar with the concept of object detection in general from the figure 6.10 with 9 participants rating above and equal to 8 on the scale of 1 to 10.

How familiar are you with the Concept of Object Detection?

14 responses

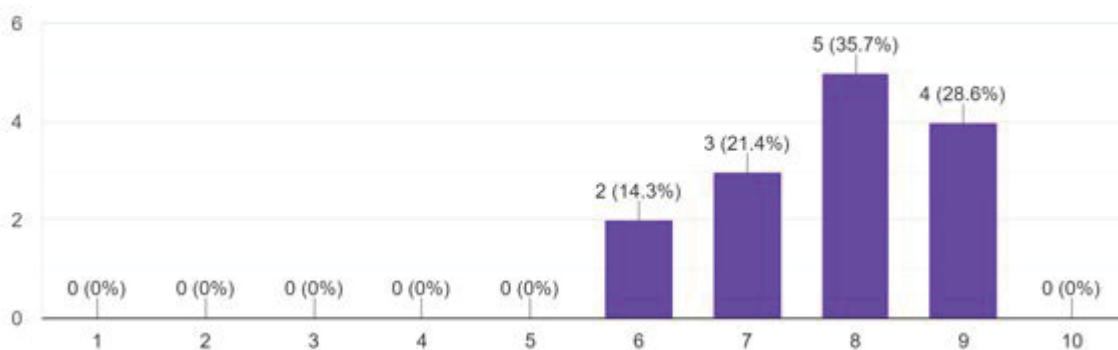


Figure 6.10: Familiarity of the participants with the concept of object detection

9. Familiarity with the concept of Machine Safety (Likert scale 1 -5):

It can be deduced that all the participants are familiar with machine safety on some level. Figure 6.11 tells that highest frequency of them have rated 3 followed by 4 on a scale of 1 to 5.

How familiar are you with the concept of Machine Safety?
14 responses

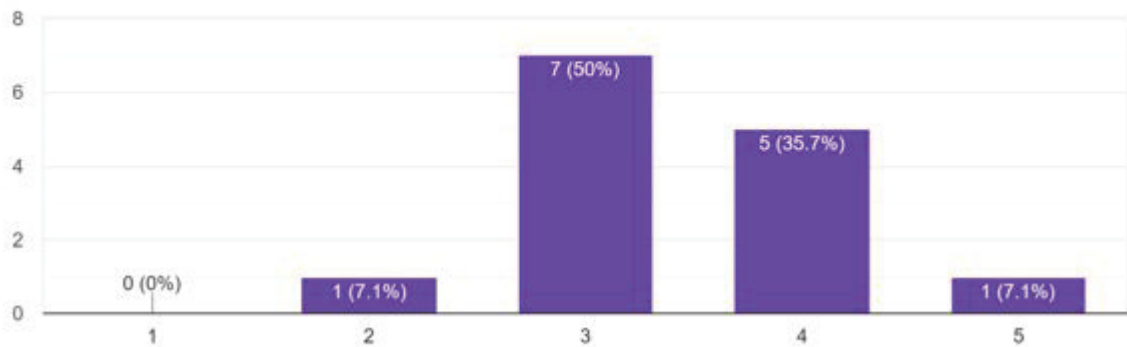


Figure 6.11: Familiarity of the participants with the concept of machine safety

10. **Familiarity with radar technology in object detection (Likert Scale 1-10):** From the Figure 6.12, it is observed that only 1 participants is not very familiar with the usage or influence of radar technology in object detection. Other participants have fairly better familiarity as they have rated above 5 on a scale 1 to 10.

How Familiar are you with the influence of Radar Technology in Object Detection
14 responses

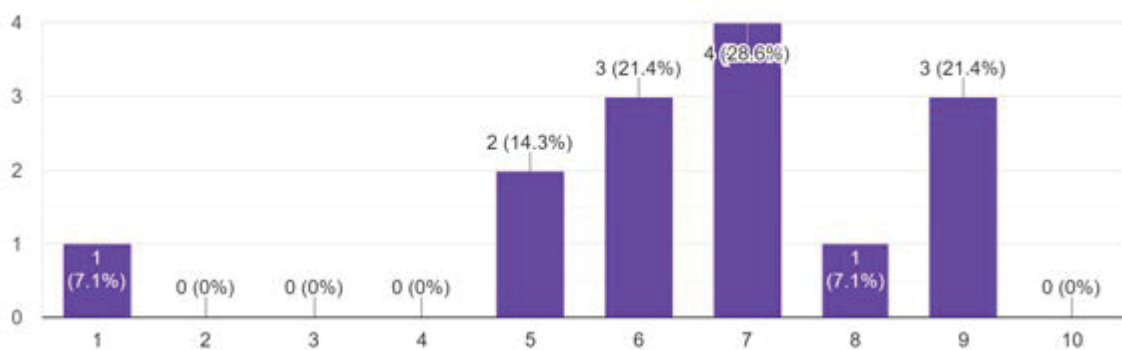


Figure 6.12: Familiarity with the usage or influence of radar technology with object detection

During an introductory conversation, the participants were asked to state their agreement or disagreement with a question in the below figure 6.13. And almost 92% of them have agreed that data visualization is key to breakdown complex data structures and sets.

Do you agree that Data Visualization is key to understanding complex data sets and structures?
14 responses

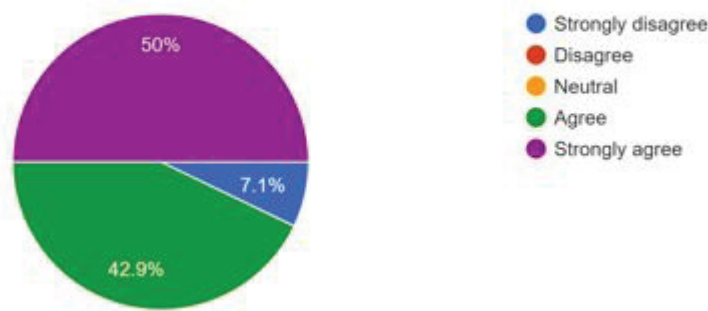


Figure 6.13: Feedback regarding Data visualization and its impact on understanding complex data

A task-based evaluation was conducted during the interviews on the participants where each participant was given three tasks, and some open-ended questions under each data visualization method and the response time was recorded. This accounts for the understandable nature and perceptibly of a user. Followed by each task, the participants were asked to rate on a Likert scale of 1 - 10 for ease of understanding the questions in the task. Another scale of 1 - 5 was used to establish the confidence levels of the participants as well for each task per visualization method. The Data visualizations used for the study are present in Appendix A.2 as V1, V2, V3 and V4, respectively.

6.2.1 Evaluation over Response Time

Below is the table 6.5 which accounts for the descriptive statistics of response time recorded for overall participants during each visualization method, namely V1, V2, V3, V4. The average time taken by 14 participants in a number of seconds is tabulated in the table 6.5 for each visualization method during the interview tasks.

Groups	Count	Sum (Sec)	Average	Variance	Median
V1	14	5257	375.5	7343.654	379.5
V2	14	4268	304.8571	5905.978	333.5
V3	14	4300	307.1429	3085.209	311.5
V4	14	3181	227.2143	2142.797	255.5

Table 6.5: Summary - Descriptive Statistics for evaluation over response time

With reference to the results in 6.5, V4 has the least average and median response time and V1 has the most average and median response time.

Statistical Test : Kruskal-Wallis Test

Null Hypothesis: There is no significant difference in Response time between the

Confidence Level	95%
Dependent Variables	Response Time
Independent Variables	Metric : Data Visualization methods V1, V2, V3, V4

Table 6.6: Quantitative Variables : Response Time

Visualization methods

Alternate Hypothesis: There is at least one visualization method which has significant difference in response time

Now, we check if all the conditions are true for this data over response time.

- **Condition 1:** There are more than two Independent sampled groups of data in the study.
Reasoning: This condition is true for this test data over response time with four independent groups(Visualization methods).
- **Condition 2:** The dependent variable is continuous.
Reasoning: All the groups occupy any value between the maximum and minimum values of the test data.
- **Condition 3:** The data should be non-normally distributed for Kruskal tests.
Reasoning: As the data points in Figure A.5 from the Response time data set are scattered instead of having close proximity along the reference line. Hence, we conclude that data is non-normally distributed for all the four visualization methods.

As all the conditions are satisfied, We perform the Kruskal-Wallis Test using R Tool (Figure 6.14) to assess the significant difference among the groups.

```
kruskal-wallis rank sum test
data: Data by visualisations
kruskal-wallis chi-squared = 21.185, df = 3, p-value = 9.636e-05
```

Figure 6.14: Kruskal-Wallis Test Result: Response Time

" The "e" stands for base-10 scientific notation which means it represents the number of zeros in the decimal number. For example 9.636e-05 is 9.636×10^{-5} simply 0.00009636 with 4 zeros after the decimal."

From the results above, It is observed that the P(Probability) value is less than the level of significance, i.e., α value (0.05) resulting in the rejection of the Null Hypothesis. This means that there is at least one pair of data visualization methods which have a significant difference with respect to response time. As the alternative hypothesis is accepted, further Bonferroni post hoc test is performed to check which pair of groups are significantly different. Additionally, we also check which group has

the highest and lowest response time.

Post Hoc Test: Bonferroni

To conduct this test, we have considered a new significance value of α_n . Where,

$$\alpha_n = \alpha / n$$

$\alpha = 0.05$: Level of significance

$n = 6$: Number of Comparisons between groups

H_0 (**Null Hypothesis**) : There is no significant difference between the compared paired groups .

H_1 (**Alternative Hypothesis**) : There is at least one paired group with significant difference between the compared paired groups .

```
Dunn (1964) kruskal-wallis multiple comparison
p-values adjusted with the Bonferroni method.
  comparison      Z      P.unadj      P.adj
1    v1 - v2  1.91795115  5.511720e-02  3.307032e-01
2    v1 - v3  1.98748413  4.686877e-02  2.812126e-01
3    v2 - v3  0.06953297  9.445654e-01  1.000000e+00
4    v1 - v4  4.57758735  4.703694e-06  2.822216e-05
5    v2 - v4  2.65963619  7.822510e-03  4.693506e-02
6    v3 - v4  2.59010322  9.594716e-03  5.756829e-02
..
```

Figure 6.15: Bonferroni Results : Response time

From the benferroni test conducted , if $P_{unadjusted}$ value is less than new significance value α_n (0.0083) for the respective comparison . Then Null hypothesis is rejected for that comparison, vice-versa for alternative hypothesis.

From Figure 6.15, we observe that there is no significant difference in V1-V2, V1-V3, V2-V3, V3-V4 as H_0 is not rejected. But in V1-V4, V2-V4, H_0 is rejected meaning that there is a significant difference among these paired groups.

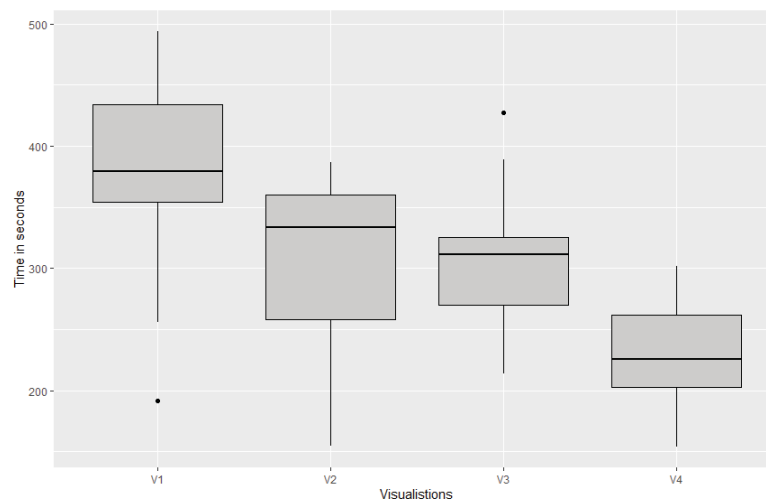


Figure 6.16: Response Time : Box plot

Figure 6.16 demonstrates the response time comparison for each of our candidate visualizations using box plots. The median values for every visualization are consistently skewed towards the upper quartile or the lower quartile, indicating that majority of the candidates took similar time to identify and understand the visualization technique. It is evident that the V4 demanded significantly lower time and for all participants against other visualization techniques. The maximum and minimum values of V4 are also reasonable without considerable dispersion. Also it is interesting to note that there are no outliers recorded with V4 thus, reflecting upon the fact that all participants were comfortable with the V4 irrespective of the backgrounds.

Now, in order to conclude which individual group among the two paired groups has the least response time, we compare the median response time between each independent group using the box-plot in R tool. Following the analysis from Figure 6.16, it is evident that V4 has the least median over Response time. And among V1, V2 and V3 majority of the participants consumed more time for V1.

Result: V4 has the least Response time

6.2.2 Evaluation over Ease of Understanding

This metric is evaluated based on the ratings given by each participant for each task per visualization method. Below is the table 6.7 which accounts for the descriptive statistics of overall Understandability ratings on a Likert scale of 1 to 10 for V1, V2, V3, V4. Average of the ratings given by 14 participants for all the tasks collectively are tabulated.

Groups	Count	Sum(Likert)	Average	Variance	Median
V1	14	78.33333	5.595238	1.114164	5.667
V2	14	100.3333	7.166667	4.029915	7.500
V3	14	110.6667	7.904762	1.323565	8.167
V4	14	113.6667	8.119048	0.967643	7.833

Table 6.7: Summary - Descriptive Statistics for evaluation over Ease of Understanding

With reference to the results in 6.7, V3 has the Highest median understandability ratings and V1 has the Least median ratings.

Statistical Test : Kruskal-Wallis Test

Confidence Level	95%
Dependent Variables	Ease of Understanding
Independent Variables	Metric : Data Visualization methods V1, V2, V3, V4

Table 6.8: Quantitative Variables : Ease of Understanding

Null Hypothesis: There is no significant difference in ease of understanding between the visualization methods.

Alternate Hypothesis: There is at least one visualization methods which have significant difference in ease of understanding.

Now we check if all the conditions are true for this data over Ease of understanding of the participant

- **Condition 1:** There are more than two Independent sampled groups of data in the study.
Reasoning: This condition is true for this test data over Ease of Understanding with four independent groups(Visualization methods).
- **Condition 2:** The dependent variable is continuous.
Reasoning: All the groups occupy any value between the maximum and minimum values of the test data.
- **Condition 3:** The data should be non-normally distributed for Kruskal tests.
Reasoning:As the data points in Figure A.6 from the ease of understanding data set are scattered instead of having close proximity along the reference line. Hence, we conclude that data is non-normally distributed for all the four visualization methods.

Now that all the conditions are satisfied, We perform the Krukal-Wallis Test using R Tool (Figure 6.17) to asses the significant difference among the groups.

```

kruskal-wallis rank sum test
data: Data by visualisations
kruskal-wallis chi-squared = 20.498, df = 3, p-value = 0.0001339

```

Figure 6.17: Kruskal-Wallis Test Result : Ease of Understanding

From the results above, It is observed that the P(Probability) value is less than the level of significance, i.e., α value (0.05) resulting in the rejection of the Null Hypothesis. This means that there is at least one pair of data visualization methods which have a significant difference with respect to ease of understanding. As the alternative hypothesis is accepted, further Bonferroni post hoc test is performed to check which pair of groups are significantly different. Additionally, we also check which group has the highest and lowest response time.

Post Hoc Test : Bonferroni

To conduct this test, we have considered a new significance value α_n .Where,

$$\alpha_n = \alpha / n$$

$\alpha = 0.05$: Level of significance
 $n = 6$: Number of Comparisons between groups

H_0 (**Null Hypothesis**): There is no significant difference between the compared paired groups in ease of understanding.

H_1 (**Alternative Hypothesis**) : There is at least one paired group with significant difference between the compared paired groups in ease of understanding.

```
Dunn (1964) kruskal-wallis multiple comparison
p-values adjusted with the Bonferroni method.
```

Comparison	Z	P.unadj	P.adj
1 V1 - V2	-2.6271256	8.610954e-03	0.051665721
2 V1 - V3	-3.7081995	2.087382e-04	0.001252429
3 V2 - V3	-1.0810738	2.796643e-01	1.000000000
4 V1 - V4	-4.1034307	4.070683e-05	0.000244241
5 V2 - V4	-1.4763051	1.398620e-01	0.839172014
6 V3 - V4	-0.3952313	6.926722e-01	1.000000000

Figure 6.18: Bonferroni Results : Ease of Understanding

From the benferroni test conducted , if $P_{unadjusted}$ value is less than new significance value α_n (0.0083). Then Null hypothesis is rejected, vice-versa for alternative hypothesis.

From Figure 6.18 we observe that there is no significant difference in V1-V2, V2-V3, V2-V4, V3-V4 so H_0 is not rejected. But in V1-V3, V1-V4 H_0 is rejected meaning that there is a significant difference among these paired groups.

Now, in order to conclude which individual group among the 2 paired groups has the least response time, we compare the median of response time between each independent group using a box-plot in R tool.

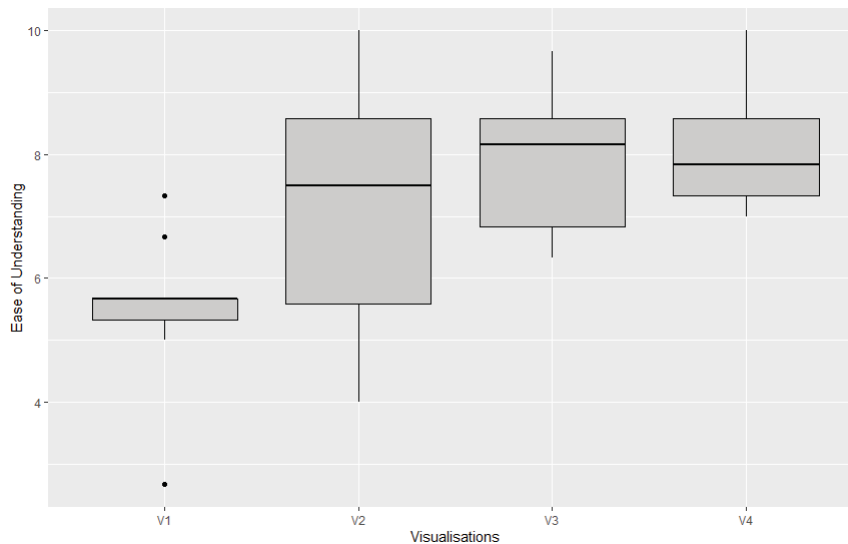


Figure 6.19: Ease of Understanding : Box plot

Figure 6.19 comments on the complexity of each visualization by comparing their ease of understanding. V3 and V4 demonstrate a smaller dispersion of data therefore, addressing that all the participants had optimal or no learning curve required to derive the message from each of the visualizations. Although V4 shows more

maximum value than V3, taking medians into consideration, V3 was clearly more favourable to understand.

It is evident from the graph in Figure 6.19 that V3 has the Highest median over Ease of understanding. And among V1, V2 and V4 majority of the participants have a better understanding of V4 as well.

Result: V3 has the Highest Ease of Understanding

6.2.3 Evaluation over Confidence

This metric is evaluated based on the confidence ratings given by each participant for each task per visualization method. Below is the table 6.9 which accounts for the Descriptive statistics of overall average Confidence level on a Likert scale of 1 to 10 for V1, V2, V3, V4 given by 14 participants.

Groups	Count	Sum(Likert)	Average	Variance	Median
V1	14	45.33333	3.238095	1.340659	3.333
V2	14	56.66667	4.047619	0.544567	4.333
V3	14	60.66667	4.333333	0.273504	4.500
V4	14	62	4.428571	0.332112	4.667

Table 6.9: Summary - Descriptive Statistics for evaluation over Confidence of the participants

With reference to the results in 6.9, V4 has the Highest median Confidence ratings, and V1 has the Least median ratings.

Statistical Test: Kruskal-Wallis Test

Confidence Level	95%
Dependent Variables	Confidence
Independent Variables	Metric : Data Visualization methods V1, V2, V3, V4

Table 6.10: Quantitative Variables : Ease of Understanding

Null Hypothesis: There is no significant difference in confidence between the Visualization methods.

Alternate Hypothesis: There is at least one visualization method which has significant difference in confidence.

Now we check if all the conditions are true for this data over Confidence of the participant

- **Condition 1:** There are more than two independent sampled groups of data in the study.

Reasoning: This condition is true for this test data over Ease of Understanding with four independent groups(Visualization methods).

- **Condition 2:** The dependent variable is continuous.

Reasoning: All the groups occupy any value between the maximum and minimum values of the test data.

- **Condition 3:** The data should be non-normally distributed for Kruskal tests.

Reasoning: As the data points in Figure A.7 from the Confidence data set are scattered instead of having close proximity along the reference line. We conclude that data is non-normally distributed for all the four visualization methods.

Now that all the conditions are satisfied, We perform the Krukal-Wallis Test using R Tool (Figure 6.20) to asses the significant difference among the groups.

```

kruskal-wallis rank sum test
data: Data by visualisations
kruskal-wallis chi-squared = 10.856, df = 3, p-value = 0.01253

```

Figure 6.20: Kruskal-Wallis Test Result : Confidence

From the results above, It is observed that the P(Probability) value is less than the level of significance, i.e., α value (0.05) resulting in the rejection of the Null Hypothesis. This means that there is at least one pair of data visualization methods which have a significant difference with respect to ease of understanding. As the alternative hypothesis is accepted, further Bonferroni post hoc test is performed to check which pair of groups are significantly different. Additionally, we also check which group has the highest and lowest response time.

Post Hoc Test: Bonferroni

To conduct this test, we have considered a new significance value α_n .Where,

$$\alpha_n = \alpha / n$$

$\alpha = 0.05$: Level of significance
 $n = 6$: Number of Comparisons between groups

H_0 (**Null Hypothesis**): There is no significant difference between the compared paired groups with confidence.

H_1 (**Alternative Hypothesis**) : There is at least one paired group with significant difference between the compared paired groups in confidence.

```

Dunn (1964) kruskal-wallis multiple comparison
p-values adjusted with the Bonferroni method.
Comparison      Z      P.unadj      P.adj
1  v1 - v2 -1.7120745  0.086882937  0.52129762
2  v1 - v3 -2.5622485  0.010399688  0.06239813
3  v2 - v3 -0.8501740  0.395228360  1.00000000
4  v1 - v4 -3.0664896  0.002165883  0.01299530
5  v2 - v4 -1.3544151  0.175603978  1.00000000
6  v3 - v4 -0.5042411  0.614091944  1.00000000

```

Figure 6.21: Bonferroni Results : Confidence

From the benferroni test conducted , if $P_{unadjusted}$ value is less than new significance value α_n (0.0083). Then Null hypothesis is rejected, vice-versa for alternative hypothesis. From the Figure 6.21 we observe that there is no significant difference in V1-V2, V2-V3, V2-V4, V3-V4, V1-V3 so H_0 is not rejected. But V1-V4 H_0 is rejected meaning that there is significant difference among these paired groups.

Now, in order to conclude which individual group among the 2 paired groups has the least response time, we compare the median of response time between each independent group using a box-plot in R tool.

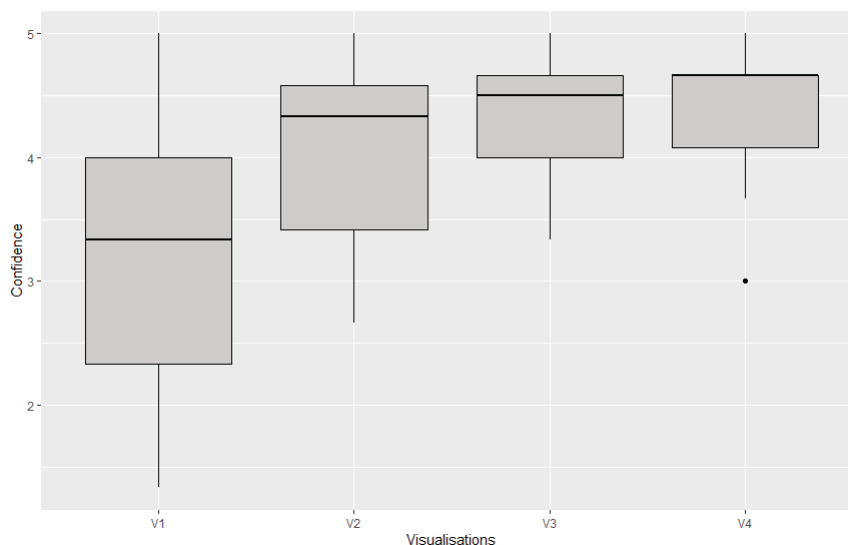


Figure 6.22: Confidence : Box plot

Figure 6.22 draws a comparison between the confidence of participants while dealing with the visualization. The median of the V4 completely overlaps with the upper quartile indicating that all the participants were highly confident when working with it expect one outlier.

It is evident from the graph in Figure 6.22 that V4 has the Highest median over Confidence in participants. And among V1, V2 and V4 majority of the participants have better confidence in answering tasks for V3 as well.

Result: V4 has the highest confidence rate

6.2.4 Evaluation over Accuracy

Accuracy is calculated after each of the participants has given their interview. Based on the interview responses, the authors assess the answers and the rate of accuracy in terms of percentage for each task they have answered. These answers are cross-checked with the known solutions of tasks with the respective visualization method under the supervision of industrial supervisor. The accuracy percentage ratings are 0%, 25%, 50%, 75% and 100% as there are four tasks in this evaluation study and each increment of 25% represents one correctly answered task for a visualization method shown that the participants have given.

Below is the table 6.11 which accounts for the descriptive statistics of overall accuracy for each visualization method of 14 participants.

Groups	Count	Sum (%)	Average	Variance	Median
V1	14	1000	71,42857143	370,8791209	75
V2	14	875	62,5	264,4230769	50
V3	14	1200	85,71428571	164,8351648	75
V4	14	1175	83,92857143	442,9945055	100

Table 6.11: Summary - Descriptive Statistics for evaluation over Accuracy of the participants

With reference to the results in the above table, V4 has the Highest median Confidence ratings, and V1 has the Least median ratings.

Statistical Test : Kruskal-Wallis Test

Confidence Level	95%
Dependent Variables	Accuracy
Independent Variables	Metric : Data Visualization methods V1, V2, V3, V4

Table 6.12: Quantitative Variables : Accuracy

Null Hypothesis: There is no significant difference in accuracy between the visualization methods.

Alternate Hypothesis: There is at least one visualization method which has significant difference in accuracy.

Now we check if all the conditions are true for this data over Confidence of the participant

- **Condition 1:** There are more than two independent sampled groups of data in the study.

Reasoning: This condition is true for this test data over Accuracy with four independent groups(Visualization methods).

- **Condition 2:** The dependent variable is continuous.
Reasoning: All the groups occupy any value between the maximum and minimum values of the test data.
- **Condition 3:** The data should be non-normally distributed for Kruskal tests.
Reasoning: As the data points in Figure A.8 from the calculated Accuracy data set are scattered instead of having close proximity along the reference line. We conclude that data is non-normally distributed for all the four visualization methods.

Now that all the Condition are satisfied, We perform the Krukal-Wallis Test using R Tool (Figure 6.23) to asses the significant difference among the groups.

```

kruskal-wallis rank sum test

data: Data by visualisations
Kruskal-wallis chi-squared = 13.037, df = 3, p-value =
0.004558

```

Figure 6.23: Kruskal-Wallis Test Result : Accuracy

From the results above, It is observed that the P(Probability) value is less than the level of significance, i.e., α value (0.05) resulting in the rejection of the Null Hypothesis. This means that there is at least one pair of data visualization methods which have a significant difference with respect to ease of understanding. As the alternative hypothesis is accepted, further Bonferroni post hoc test is performed to check which pair of groups are significantly different. Additionally, we also check which group has the highest and lowest response time.

Post Hoc Test: Bonferroni

To conduct this test, we have considered a new significance value α_n . Where,

$$\alpha_n = \alpha / n$$

$\alpha = 0.05$: Level of significance
 $n = 6$: Number of Comparisons between groups

H_0 (**Null Hypothesis**): There is no significant difference between the compared paired groups with in accuracy.

H_1 (**Alternative Hypothesis**) : There is at least one paired group with significant difference between the compared paired groups in accuracy.

```

Dunn (1964) kruskal-wallis multiple comparison
p-values adjusted with the Bonferroni method.

  comparison      Z      P.unadj      P.adj
1   v1 - v2  1.1960320  0.231684064  1.00000000
2   v1 - v3 -1.9111852  0.055980787  0.33588472
3   v2 - v3 -3.1072172  0.001888576  0.01133145
4   v1 - v4 -1.7015713  0.088835756  0.53301454
5   v2 - v4 -2.8976033  0.003760258  0.02256155
6   v3 - v4  0.2096139  0.833969062  1.00000000

```

Figure 6.24: Bonferroni Results : Accuracy

From the benferroni test conducted , if $P_{unadjusted}$ value is less than new significance value α_n (0.0083). Then Null hypothesis is rejected, vice-versa for alternative hypothesis. From the Figure 6.24 we observe that there is no significant difference in V1-V2, V3-V4, V1-V4, V1-V3 so H_0 is not rejected. But in V2-V3, V2-V4 H_0 is rejected meaning that there is significant difference among these paired groups.

Now, in order to conclude which individual group among the 2 paired groups has the least response time, we compare the median of response time between each independent group using a box-plot in R tool.

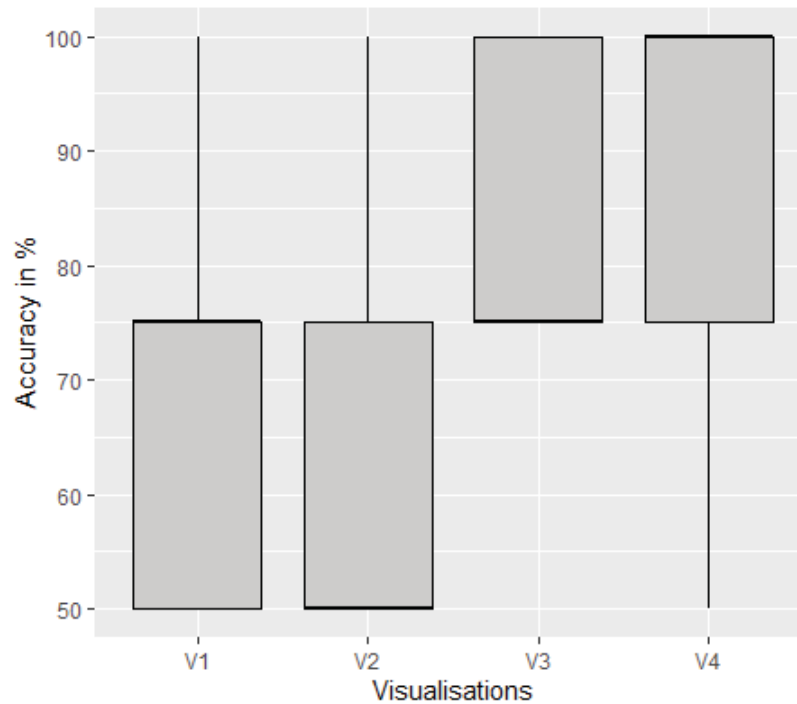


Figure 6.25: Accuracy : Box plot

Figure 6.25 compares the correctness of responses provided by the participants. Given the complexity of the subject, the correctness of the response relies on the visualization. Therefore with V1 and V2 the accuracy was poor unanimously among all the participants as represented by their respective medians. V3 despite having the complete population inside the box dispersion, majority of the candidates showed accuracy equivalent to V1. In contrast, V4 has the majority of its participants showing cent percent accuracy.

It is evident that V4 has the Highest median over accuracy in participants. Due to the shorter sample size of the participants and the accuracy being calculated in the intervals of 25% for each correctly answered task, the data is skewed towards the extreme median values.

Result: V4 has the Highest Accuracy rate

Measure	Kruskal Wallis Test	Bonferroni Test	Result
Response Time	$H_0(Rejected)$	$H_0(Rejected)$	V4 (Temperature HeatMap)
Ease of Understanding	$H_0(Rejected)$	$H_0(Rejected)$	V3 (Contour Plot)
Confidence	$H_0(Rejected)$	$H_0(Rejected)$	V4 (Temperature HeatMap)
Accuracy	$H_0(Rejected)$	$H_0(Rejected)$	V4 (Temperature HeatMap)

Table 6.13: Analysis Summary

6.2.5 Interview Summary & Data Triangulation

Towards the end of each task and interview, qualitative questions regarding detailed feedback and experience are collected with a final selection (in Figure 6.26) of what the interviewees feel is an effective visualization method in this paper. In this section, we would be presenting the results from the narrative analysis conducted on the overall semi-structured interviews with respect to the evaluation results over metrics.

6.2.5.1 Understandability of a Data Visualization method

Questions regarding the understandability of each visualization method was posed to all the interviewees. The question is "What problems did you face in understanding the data using this visualization method?"

- **Scope of the question:** This question aims to know the extent to which the cognitive perceptibly and understanding skills of the participants is present. After all the tasks for each visualization method are attempted by the participants, their reactions and inputs regarding what challenges they have faced in each visualization method were recorded.
- **Response of the participants:** The Interviewees had different opinions about different visualization methods. Some of the interesting and unique feedback given by them are presented below:
 - For V1 :
 1. "This method purely lacks the ability to provide a location in 2D or some kind of location in space."
 2. "I do not understand what this method is trying to convey exactly as The detection is not clear."
 - For V2:
 1. "No, the location is clearly visible, but there is denotion of the size or characteristics of the objects."
 2. "The scattered points are distant and, I think that the object might be a big board or similar. If this is not the case, then the detection is not accurate."
 3. "Task 3 was fairly difficult to answer, as there is no indication of any noise or presence of the idle objects."
 - For V3:

1. "Identifying the Location of sensors was not easy."
 2. "There was a little issue with distinguishing noise which needs closer observations."
- For V4:
1. "I feel that the image is more refined and provides clear description"
 2. " There is no problem with this method, object was clearly detected by all 3 sensors"
- **Narrative analysis result:** Analyzing the interview responses; it was observed that most of them had a good understanding of object detection and its relative properties but were unable to estimate the approximate size(cross-section) of the objects. V1, V2 were not able to convey the clutter/noise and objects characteristics as expected but had shown good detection, on the other side the V3, V4 had overall positive responses over-representation of all the features.

6.2.5.2 Suggestions and feedback

This is an open-ended question asked towards the end of each visualization to gather additional information in order to record what the interviews expected from that method more.

- **Scope of the question:** In order to understand what the interviewee thinks and what they feel was required in each visualization.
- **Response of the participants:** Being an open-ended question with a broad perspective; the responses were not much different for this question against few visualization methods across all the participants. Few of the responses are:
 - For V1
 1. "I think that this method lacks the core objective of object detection and its motive to show an exact location."
 2. "I think this plot can be used to check if there is any detection or not"
 - For V2
 1. "In my view, Visualizing the data using scatter plot would be effective as scatter plot is good at showing how distant are the data points located from each other, like how scattered there are. The makes detecting anomalous points easier."
 - For V3
 1. "When the data is visualized using this plot, the density of the arcs at one place shows the position of the object, but it is confusing to analyse due to the presence of many similar arcs which could be improved ."
 - For V4

1. "This plot provides good visualization over others; More feature additions could be made."
2. " I am more inclined towards understanding the data in this method."

- **Narrative Analysis Result:** A deep narrative analysis of the responses overall, has given a good understanding of what the interviewees expect from a visualization method. In particular, for object detection, the users expect the following:

- The users require a Visualization method to be able to locate the object clearly with different colour contrast or related combinations.
- Distinguishing the noise, stationery objects and surroundings from the moving objects or the target should be clear.
- For object detection to have good reliability, It is important that the visualization methods account for the external factors as well, such as environmental conditions.
- There should be a good estimation of the size of the object (in this case cross-section) so that the machine can predict the designated operation when an object is detected.

6.2.5.3 Participants selection

Towards the end of each interview, the participants were asked to provide their choice of visualization method, which they felt was more effective.

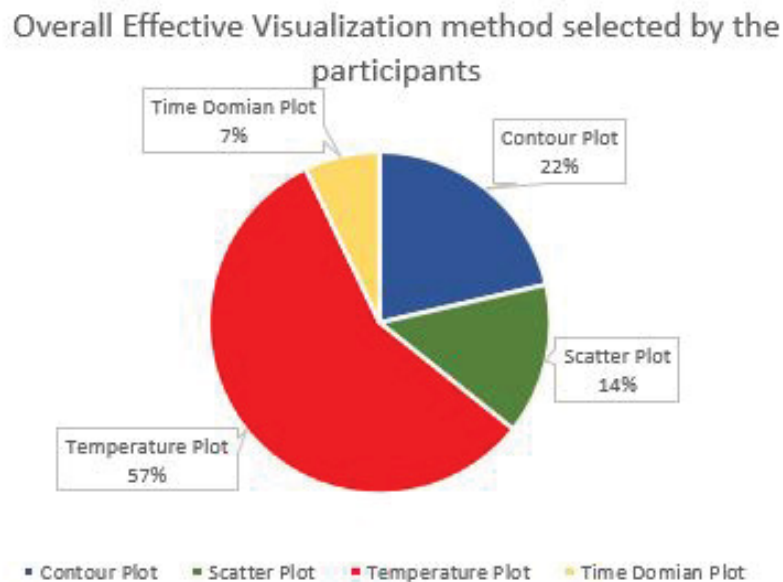


Figure 6.26: Overall Selection of an Effective Data Visualization by the participants

As the Figure 6.26 suggests, more than 50 % of the participants have selected the V4, i.e., Heat-Temperature plot to be more effective which can also be corroborated with the quantitative results from section 6.2.1, 6.2.3 and 6.2.4. With reference

to the results, It can be seen that two distinct metrics namely the response time and accuracy of V4 indicate that both the task evaluation results and individual participants feedback agree with each other and provide validation to this study. It is clear that analysis results help us determine an effective visualization method to be a Heat Temperature Plot and the evaluation feedback is positive, feasible and reliable. Even though the research results could be more detailed and improvised, the current results provide a way to evaluate and arrive at significant outcomes.

The threats to validity concerning this research study are discussed in this chapter. Based on the papers written by Ihantola et al.[54] , Sven Modell [55] , Wohlin et al.[23], we categorize the threats in this study into reliability, internal validity, external validity and construct validity.

7.1 Reliability

In general, this threat of validity concerns with the repeatability and replicability of this study with similar research objectives. A well-defined research structure and method will help tackle this threat. Although, this study mainly concentrates on the qualitative and quantitative data from the interviews and is key to make important decisions regarding the evaluation. A change of the participants may change how the outcomes are derived. Another important source of knowledge lies within how the multi-vocal literature review results. This also needs consistency in the planning and execution of the literature review. Careful selection of the sources and considering all the parameters during a search will surely reduce the risk[55][23].

Feedback and inputs from the interviews are based on the participant's selection which could mean that different kind of participants would result in different responses. Lack of communication skills or misunderstandings are a problem, all things considered, Fairly due to the size of the chosen participants and their work experience, the outcomes from the interviews should be credible and acceptable.

7.2 Internal Validity

The internal validity is responsible for ensuring that the conclusions obtained from this study conducted are valid with supportive facts and information. It overlooks the parameter which is generally hard to control and supervise. Potential threats to this study are the collection of quantitative data or lack of quality in the collected data, which directly have an impact on the outcomes of the research. Another threat would be arriving at a conclusion without proper review and consideration for wrongful execution. These are resolved by collecting the quantitative data under the supervision of industry supervisor and verifying the results by having daily review meetings in the company. All the literature material and carefully analyzed and

mapped to the conclusions by the authors towards the end of the paper.

Another threat to our study would be if the search strings formulation went bad or if the search method itself was at fault. This was avoided by following the selection bias approach to formulating strings under the supervision of both authors. In order to avoid search method threat, only reputed and recognized databases like Google Scholar, Scopus, IEEE and ACM Digital library followed by grey literature from Google Search engine are considered. The selected articles were cross verified by the authors and mapped with each other's results to make sure there was no culture bias and publication bias as well.

7.3 External Validity

External validity threats are concerned with how the results of this study are relevant to other problems or studies. The overall scope of this study aims at providing a better and effective data visualization method which will help developers, stakeholders and customers understand the outcomes of a product of which they have no knowledge. The literature method used can be replicated to any similar problem or similar problem with a different system. Due to limited work in this research problem, we have searched papers from 2002-2020 for eliminating the bottlenecks in literature time filter[54].

It is important to the study that the results are derived through proper channels to provide solid relevance to other similar problems. Another threat to this study was directed towards the data analysis from the interviews and the personal bias of the authors. For this, all the interviews were transcribed and the narrative analysis performed in the presence of both the authors which were later verified by a re-screening of the interview results.

7.4 Construct Validity

Construct validity concerns with how well do the methods implemented in the study are efficient to answer the research questions. This case study was divided into two parts for an effective and efficient collection and analysis of crucial data which would contribute to making decisions based on the results. The interviews which would evidently contribute to this thesis outcomes was designed based on the guidelines and support provided from the university supervisor and all the participants were selected based on the study objectives. Seven participants from the company involved in the same project with different backgrounds and contributions to the project. The other 7 participants who were former colleagues to the authors with respect to their knowledge in the data visualizations and software development but no in-depth knowledge of radar technologies. Adding to this, a well structured and the designed interview was conducted by describing all the elements of this study to the participants[23].

7.5 Limitations

Familiarity of the visualization methods: This is a potential limitation which could affect the participant's perception and reasoning towards the tasks with regard to their previous experiences in using any familiar data visualization method. Generally, most of the practitioners and developers show a little familiarity towards the time domain or frequency plots and scatter plots. To avoid this, prior to conducting the interviews, all the participants were sent a document with the research proposal and relevant background information in the invitation about the system and data visualizations methods used in the evaluation. Although one cannot be certain if all the participants have read the material before the interview, it is an effective counter measure to reduce the familiarity which can also be said to how the results are derived.

Lack of resources due to Covid-19: Due to the pandemic situation in the world, more number of the industry officials, practitioners were not available. Although a fair amount of the interviewees were selected and interviewed, additional data points would be good merit to the study.

In this chapter, we will discuss the results and other findings related to the thesis. Adding to it, we will look upon some unwanted encounters and experiences faced during the study. Some important opinions and feedback would also be furnished in light to other related work which drives the necessity of conducting such research. The visualizations methods referred here are present in Appendix A.2

As we all know data visualization is an important concept, tool or medium through which we graphically represent data to understand the raw form of data has its importance in many applications, especially in the software industry. The need to determine an effective and efficient visualization method is also important based on need and usage. Similarly, the results in this research indicate that a task-based evaluation will evidently provide developers, stakeholders and other practitioners the expected outcome with supportive background analysis which will help them to make critical business decisions. When compared to the work by Forsell and Johansson[11] like using heuristics for evaluation which are time-consuming, this paper concentrates on the design and formulation of the tasks which can effectively collect the metric related data. Due to the real-time implementation of this research work in an industrial environment, extensive background work was done to make sure the results were reliable, and the study was replicable.

Effectiveness, Efficiency, Understandability and Satisfaction play an important role in convincing people to use a product or service. The design, procedure and analysis behind the evaluation study is a major contribution in this research work due to industrial exposure. Working in a real-time project and conducting an evaluation to determine an effective visualization method opens many challenges and unknown opportunities to gather knowledge on different aspects and usage of a visualization method which was discussed by Koua et al[14] to similarly determine their usefulness.

Unlike what Grover et al[21] have done in their paper for object detection. Detection reliability, despite many efforts over the past years, has something new to discover and improve. In this case, using multiple millimeter wave radar modules and its data to visualize the detection has been an important aspect given the different data features and functionalities being developed. And using a single effective visualization method for all the requirements in the newly developed software demonstrator is required, which demands the current knowledge of existing methods and their evaluation to convince stakeholders regarding the business value this system

has to offer.

Initially, due to the nature and unfamiliarity of the radar data and its respective visualizations in this research, It was unclear on how to gather knowledge and facts which will help design an evaluation study. It was evident over conducting rigorous research on existing literature; There aren't many studies with similar research objectives in this field of object detection. Brainstorming for few days and discussing with both Industry and University supervisors, we arrived at implementing a multivocal literature review which gave equal importance to the articles such as company artefacts, User manuals and other papers within other online sources. Upon collecting relevant visualization methods, a critical part of the analysis included mapping the relative visualization methods with the research objectives and daily development in the industry where the raw data is involved. Interacting with the industrial developers in the safety radar project provided more feedback into what Visualization methods are suitable to different individual functionalities of the detection process. Adding to this, the design and development of coordinate transformation engine were also based on how the data will be visualized.

Designing a responsive task evaluation interview questionnaire is a milestone which was executed by following the guidelines and work from existing literature. The evaluation criteria, being a core part of the study was actually derived from the needs of some developers at industry and through Christina et al. [33] The concept for evaluating any targeted subject was clearly stated, and its related metrics have provided direction to why evaluation was needed. There isn't much existing Literature on Radar detection and its visualizations which motivated the use of evaluation criteria rather than a heuristic approach due to time constraints and Industrial requirement.

The evaluation metrics, namely Response time, Ease of Understanding, Confidence and Accuracy are crucial to determine the effectiveness and efficiency of a visualization method. In these only Response time and accuracy are not dependent or biased towards the participants, i.e., the data collected for these metrics should evidently provide answers to which visualization is effective and is perceptible. This research is a huge contribution to radar data applications and the formulation of study design to the software industry. Statistical analysis and validation to determine an effective visualization method is a research-oriented implication followed by the resultant visualization being used in the developed Software Demonstrator an industry-oriented implication. The narrative analysis and statistical analysis are self-validating with their findings.

Although for this study, a sample of 14 participants was considered, which is an ideal size, increased number of participants with different roles will have a significant impact on the rate of effectiveness.

One can argue that why other statistical methods like Anova analysis or simple a t-test was performed on the interview data. The justification is simple, The type of statistical test selected is dependent on the nature of data distribution. We

check for Normal and Non-Normal distribution characteristics using the R tool and ensure that the dependent variable (metrics) is continuous. We observe the Normal Q-Q plots for each data set and make a decision. Many researchers have found that Kruskal Wallis analysis is a better alternative to Anova analysis when the data is non-normally distributed. Moving further, one of the major conflicting work in this thesis was the formulation of hypotheses which was tackled carefully by dividing the objective of evaluation into steps namely; firstly we needed to know if there was any significant difference between any visualization methods with respect to the metric then we needed to check the same which pair combinations leading to the utilization of Bonferroni Post hoc test. When a combination is ruled out, a decision was made with reference to the median boxplots of the interview data.

The interview results indicate that more than 50% of the participants rate the heat temperature plot to be more effective over other methods. While the qualitative results focus on participants feedback, The quantitative data analysis states that heat temperature plot (V4) is more effective under response time, Confidence and accuracy. On the other hand, the Contour plot(V3) has the highest ratings for Ease of understanding, although the heat temperature plot was close and has similar feedback in the interviews as well. It can be also be argued that why did V3 have better ease of understanding and shouldn't the easily understandable method be more effective?. In the study, two metrics namely ease of understanding and confidence, are not perceived directly by the participants and are prone to human error which is resolved considering the metrics response time and accuracy. Most of the participants also share their interest in scatter plot due to its ability to pinpoint the detection.

On a positive note, this research has many implications with regard to the methods, procedure, evaluation, analysis and implementation. To name a few, we now have a well-executed evaluation study to determine effective data visualization methods given the external requirements. The results from this study directly impact on the detection reliability of the radars and help develop new algorithms, features and convince stakeholders. Due to repeated verification and validation from existing literature and real-time experience in an industrial scenario, this study is reliable and can be replicated for other data visualization evaluation problems which is a recognizable contribution in the software industry.

In this section, we narrow down the results and valuable conclusions with supportive feedback. Future works with potential limitations are also discussed.

This thesis significantly presents a recognizable contribution to the information visualization, otherwise known as the data visualization community, where evaluation plays a crucial role. An exploratory case study with the intention of determining an effective data visualization method has been carried out with applications in radar technology and object detection. This thesis presents a refinement over evaluation studies and statistical analysis using empirical methods in an industrial scenario. The aim of this thesis was to address the which existing data visualization methods are being utilized in safety radar system with millimetre wave radar data and perform an evaluation to determine which data visualization method is more effective.

Initially, for RQ I, a multi-vocal literature review was conducted to explore and identify relevant data visualization methods which provided the baseline for the evaluation. Thematic analysis was implemented to filter out 4 data visualization methods with reference to the safety radar system. Further, evaluation criteria and study design are defined in order to perform a task-based evaluation.

RQ II is answered by performing task evaluation based on the responses from the participants over the defined metrics from the evaluation criteria, namely Response time, Ease of Understanding, Confidence and Accuracy. Statistical analysis using Kruskal Wallis and Bonferroni post hoc tests were executed to narrow the comparison into combinations of 2 visualization methods followed by conclusions made from the median boxplots over the respective metrics. Static validations using participants feedback and narrative analysis was performed on the qualitative data. According to the qualitative responses, it was observed that around 57% of the participants consider the heat temperature plot (V4) to be more effective when compared to other data visualization methods. 22% think that contour plot (V3) is effective and 14% think the scatter plot (V2) with remaining 7% for the time domain plot (V1) to be effective. While the quantitative analysis over Response time, confidence and accuracy also state that V4 is more effective leaving V3 with higher ease of understanding . Upon the discussions from the previous chapter and observing the results closely in the box plots with acceptable difference , All the results broadly state the data visualization method V4, namely Heat temperature plot to be more effective and also satisfied the industrial stakeholders to make important business decisions using

this method.

With the results being furnished in this research, there is always scope to improvise the methods and relevant areas of the study. Time and industrial requirements have always played an important role in this study.

As part of the future work, we can consider implementing similar evaluation study on a similar data visualization problem with more participants and more number of data samples. There is also the matter of addressing critical descriptions and tasks for the evaluation if need be to a new problem. One could attempt at finding more metrics and provide additional evaluation criteria which could provide a quantitative approach for better evaluation. Different treatments and evaluation design like grouping the participants based on experience, familiarity and usability characteristics could be perceived. On the other hand, another study with similar research objectives could be made with more insight into the features of the data. Additionally, As this thesis was part of an Industry oriented research development with real-time implications, only a few employees with expertise in this domain were selected. More employees globally could be included as an extension over this research.

It is also important that investigation into the replicability of this study and its implementation should be made for additional contributions regarding new problems which depend on data visualizations and their evaluation.

Bibliography

- [1] Lidong Wang, Guanghui Wang, and Cheryl Ann Alexander. Big data and visualization: methods, challenges and technology progress. *Digital Technologies*, 1(1):33–38, 2015.
- [2] Ying Zhu. Measuring effective data visualization. In *International Symposium on Visual Computing*, pages 652–661. Springer, 2007.
- [3] Prachi Mishra, Naina Bharadwaj, and Prajyoti Dsilva. Interactive data visualization for census data. 2016.
- [4] Peter Rodgers, Gem Stapleton, Bilal Alsallakh, Luana Michalief, Rob Baker, and Simon Thompson. A task-based evaluation of combined set and network visualization. *Information Sciences*, 367:58–79, 2016.
- [5] Muzammil Khan and Sarwar Shah Khan. Data and information visualization methods, and interactive mechanisms: A survey. *International Journal of Computer Applications*, 34(1):1–14, 2011.
- [6] Alessandro Comai. Decision-making support: the role of data visualization in analyzing complex systems. *World Future Review*, 6(4):477–484, 2014.
- [7] FH Krantz and WD Murray. A survey of digital methods for radar data processing. In *Papers presented at the December 13-15, 1960, eastern joint IRE-AIEE-ACM computer conference*, pages 67–82, 1960.
- [8] Urban Meis and Robert Schneider. Radar image acquisition and interpretation for automotive applications. In *IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No. 03TH8683)*, pages 328–332. IEEE, 2003.
- [9] Akash Deep Singh, Sandeep Singh Sandha, Luis Garcia, and Mani Srivastava. Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar. In *Proceedings of the 3rd ACM Workshop on Millimeter-Wave Networks and Sensing Systems*, mmNets’19, page 51–56, New York, NY, USA, 2019. Association for Computing Machinery.
- [10] Xu Zhuge, Hao Wang, and Girts Strazdins. Evaluating the data visualization for demanding marine operations. In *2016 Techno-Ocean (Techno-Ocean)*, pages 471–477. IEEE, 2016.
- [11] Camilla Forsell and Jimmy Johansson. An heuristic set for evaluation in information visualization. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 199–206, 2010.
- [12] Camilla Forsell. A guide to scientific evaluation in information visualization. In *2010 14th International Conference Information Visualisation*, pages 162–169. IEEE, 2010.

- [13] Tobias Isenberg, Petra Isenberg, Jian Chen, Michael Sedlmair, and Torsten Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013.
- [14] Etien L Koua, Alan MacEachren, and M-J Kraak. Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science*, 20(4):425–448, 2006.
- [15] Catherine Plaisant. The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces*, pages 109–116, 2004.
- [16] T Pai et al. Big data new challenges, tools and techniques. *International Journal of Engineering Research and Modern Education (IJERME), ISSN (Online)*, pages 2455–4200, 2017.
- [17] Leonel Merino, Mohammad Ghafari, Craig Anslow, and Oscar Nierstrasz. A systematic literature review of software visualization evaluation. *Journal of systems and software*, 144:165–180, 2018.
- [18] Melanie Tory and Torsten Moller. Evaluating visualizations: do expert reviews work? *IEEE computer graphics and applications*, 25(5):8–11, 2005.
- [19] Naina Bharadwaj, Prachi Mishra, and Prajyoti Dsilva. Interactive data visualization for census data. *International Journal of Computer Applications*, 149(4), 2016.
- [20] R Schneider and J Wenger. High resolution radar for automobile applications. *Advances in Radio Science*, 1(B. 3):105–111, 2003.
- [21] Richard Grover, Graham Brooker, and Hugh F Durrant-Whyte. A low level fusion of millimeter wave radar and night-vision imaging for enhanced characterization of a cluttered environment. In *Proceedings 2001 Australian Conference on Robotics and Automation*, 2001.
- [22] Colin Robson. *Real world research: A resource for social scientists and practitioner-researchers*, volume 2. Blackwell Oxford, 2002.
- [23] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. *Experimentation in software engineering*. Springer Science & Business Media, 2012.
- [24] Per Runeson and Martin Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14(2):131, 2009.
- [25] Dewayne E. Perry, Susan Elliott Sim, and Steve Easterbrook. Case studies for software engineers. In *Proceedings of the 28th International Conference on Software Engineering, ICSE '06*, page 1045–1046, New York, NY, USA, 2006. Association for Computing Machinery.
- [26] Teade Punter, Marcus Ciolkowski, Bernd Freimut, and Isabel John. Conducting on-line surveys in software engineering. In *2003 International Symposium on Empirical Software Engineering, 2003. ISESE 2003. Proceedings.*, pages 80–88. IEEE, 2003.
- [27] Rory O’Brien. An overview of the methodological approach of action research.

- 1998.
- [28] Vahid Garousi, Michael Felderer, and Mika V Mäntylä. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*, 106:101–121, 2019.
 - [29] Daniel Méndez Fernández and Stefan Wagner. Case studies in industry: What we have learnt. In *Proceedings of the 4th International Workshop on Conducting Empirical Studies in Industry*, CESI '16, page 25–31, New York, NY, USA, 2016. Association for Computing Machinery.
 - [30] Daniela S Cruzes and Tore Dyba. Recommended steps for thematic synthesis in software engineering. In *2011 international symposium on empirical software engineering and measurement*, pages 275–284. IEEE, 2011.
 - [31] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.
 - [32] Sohaib Shahid Bajwa, Xiaofeng Wang, Anh Nguyen Duc, and Pekka Abrahamsson. “failures” to be celebrated: an analysis of major pivots of software startups. *Empirical Software Engineering*, 22(5):2373–2408, 2017.
 - [33] Christina Sonebo and Joel Ekelöf. Creating and evaluating an interactive visualization tool for crowd trajectory data, 2018.
 - [34] Stephen Few. Data visualization effectiveness profile. *Perceptual Edge*, 10:12, 2017.
 - [35] Barry M Katz and Maryellen McSweeney. A multivariate kruskal-wallis test with post hoc procedures. *Multivariate Behavioral Research*, 15(3):281–297, 1980.
 - [36] Christoffer Cambronero. Multiplex protein data and how to handle it, 2016.
 - [37] Juliet Shaffer. Multiple hypothesis testing. *Annual Review of Psychology*, 46:561–84, 01 1995.
 - [38] Dennis Barrett and Adrian Alvarez. mmwave radar sensors in robotics applications. Technical report, Technical report, Texas Instruments, 2017.
 - [39] D Novotny, J Gordon, and J Guerrieri. Antenna alignment and positional validation of a mmwave antenna system using 6d coordinate metrology. *Proceedings of the Antenna Measurement Techniques Association*, 2014.
 - [40] Andreas Schiessl, Sherif Sayed Ahmed, and Lorenz-Peter Schmidt. Data statistics and image properties of a large multistatic mm-wave imaging system. In *2013 14th International Radar Symposium (IRS)*, volume 1, pages 202–206. IEEE, 2013.
 - [41] Tracking Using TI. Traffic monitoring object detection and tracking reference design using single-chip mmwave radar sensor. 2017.
 - [42] Jan Luca Uphoff. Introduction to automotive fmcw radar technologies: Using texas instruments mmwave awr sensor series, 2018.
 - [43] Sandeep Rao. Introduction to mmwave Sensing: FMCW Radars. page 70.
 - [44] The Data Visualisation Catalogue.
 - [45] url to bibtex - google search. https://www.google.com/search?q=url+to+bibtex&rlz=1C1CHBF_enSE885SE885&oq=url+to+bib&aqs=chrome.

- 0.0j69i57j016.3019j0j1&sourceid=chrome&ie=UTF-8. (Accessed on 10/11/2020).
- [46] Shigeki Sugimoto, Hayato Tateda, Hidekazu Takahashi, and Masatoshi Okutomi. Obstacle detection using millimeter-wave radar and its visualization on image sequence. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 3, pages 342–345. IEEE, 2004.
- [47] MMWAVE-STUDIO mmWave studio | TI.com.
- [48] MMWAVE_studio 02_01_01_00 - TI.com.
- [49] IWR6843 data sheet, product information and support | TI.com.
- [50] AWR1642BOOST AWR1642 single-chip 76-GHz to 81-GHz automotive radar sensor evaluation module | TI.com.
- [51] Renyuan Zhang and Siyang Cao. Extending reliability of mmwave radar tracking and detection via fusion with camera. *IEEE Access*, 7:137065–137079, 2019.
- [52] Z. Li, Z. Lei, A. Yan, E. Solovey, and K. Pahlavan. Thumouse: A micro-gesture cursor input through mmwave radar-based interaction. In *2020 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–9, 2020.
- [53] Fawwaz T Ulaby, TE Van Deventer, JR East, TF Haddock, and ME Coluzzi. Millimeter-wave bistatic scattering from ground and vegetation targets. *IEEE Transactions on Geoscience and Remote Sensing*, 26(3):229–243, 1988.
- [54] Jennifer Grafton, Anna M Lillis, Eeva-Mari Ihantola, and Lili-Anne Kihn. Threats to validity and reliability in mixed methods accounting research. *Qualitative Research in Accounting & Management*, 2011.
- [55] Sven Modell. Triangulation between case study and survey methods in management accounting research: An assessment of validity implications. *Management accounting research*, 16(2):231–254, 2005.

Appendix A

Supplemental Information

A.1 Interview demographics and questionnaire

Task	Questions
1	1Q - Identify the object being detected in this visualization method? 2Q - Please Elaborate your Answer for either of the Response 3Q - How easy was it to identify the object in this method ? 4Q - How Confident are you with your answer ?
2	1Q - Identify the position of the Radar modules with respect to the Object detected? 2Q - Please elaborate 3Q - How easy was it to identify the positions of the radar modules 4Q - How confident are you with your answer
3	1Q - Estimate the size of the object or human? 2Q - What is the surveillance area in the proposed environment ? 3Q - Do you observe any clutter or noise in this visualization method? 4Q - How easy was it to answer these questions? 5Q - How confident are you with your answers?
4	1Q - Were you able to distinguish the object, radar modules and surveillance ? 2Q - What problems did you face in understanding the data using this visualization method ? 3Q - Are there any suggestions and feedback to this visualization method?

Table A.1: Interview Questionnaire

QNO.	Demographic Questions	Form of Answer
1	Age (in years)	<i>Multiple Answer</i> a) 18-25 b) 25-30 c) 30-35 d) 35+
2	Gender	<i>Multiple Answer</i> a) Male b) Female c) Others
3	What is the Highest Degree or level of school you have completed?	<i>Multiple Answer</i> a) Less Than High School b) High School Or Equivalent c) Bachelors Degree d) Master's Degree e) Doctorate
4	Field of Study	<i>Multiple Answer</i> a) Mechanical Engineering b) Computer Science Engineering c) Electronics and Telecommunications d) Civil Engineering e) Electrical Engineering f) Software Engineering g) Others
5	How many years have you been in the Research field ?	<i>Multiple Answer</i> a) 1 b) 2 c) 3 d) 4 e) 5+ f) None
6	Do you have any working experience with Data Visualizations methods and their evaluation?	<i>Multiple Answer</i> a) Yes b) No c) Maybe
7	Did you take any previous courses or projects with data visualization as a core concept?	<i>Multiple Answer</i> a) Yes b) No c) Maybe
8	How familiar are you with the Concept of Object Detection?	<i>Likert - 10</i>

QNO.	Demographic Questions	Form of Answer
9	How familiar are you with the concept of Machine Safety?	<i>Likert - 5</i>
10	How Familiar are you with the influence of Radar Technology in Object Detection?	<i>Likert - 10</i>
11	In your Opinion, What makes a good visualization?	<i>Short Answer</i>
12	Do you agree that Data Visualization is key to understanding complex data sets and structures?	<i>Multiple Answer</i> <i>a)Strongly disagree</i> <i>b)disagree</i> <i>c)Neutral</i> <i>d)agree</i> <i>e)Strongly agree</i>

Table A.2: Demographics Questionnaire

A.2 Data visualization methods

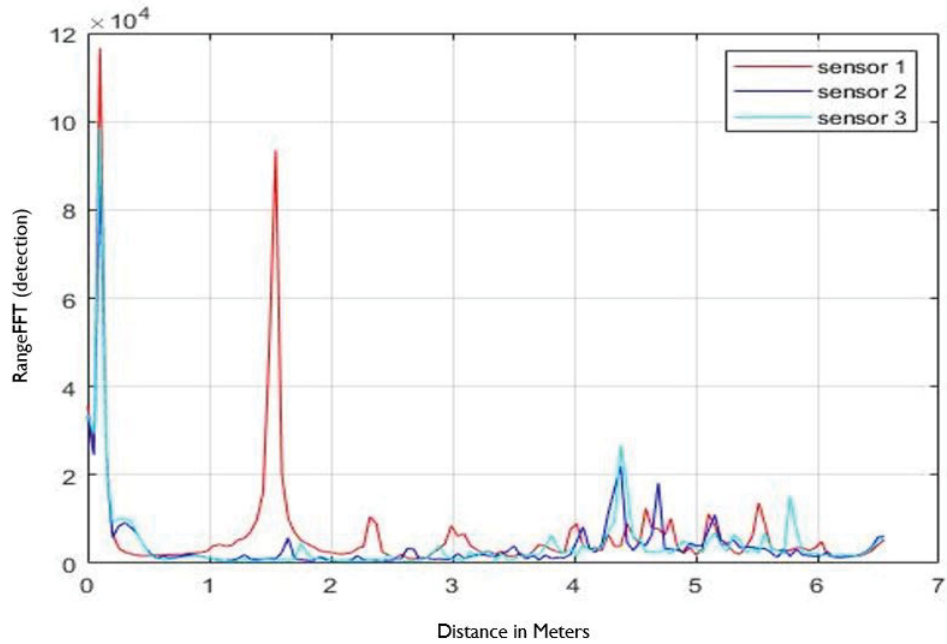


Figure A.1: V1 : Time Domain plot

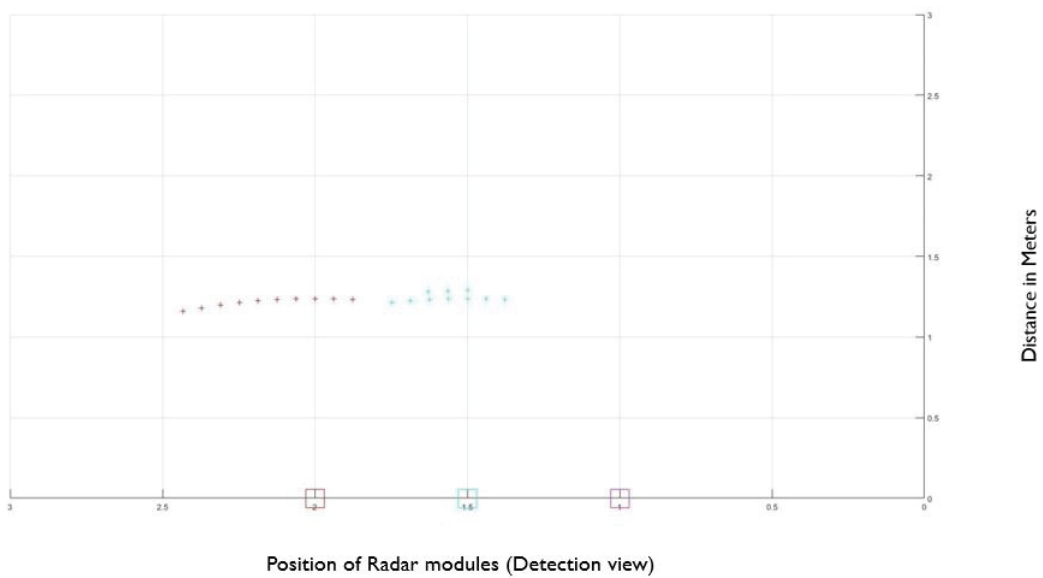


Figure A.2: V2 : Scatter Plot

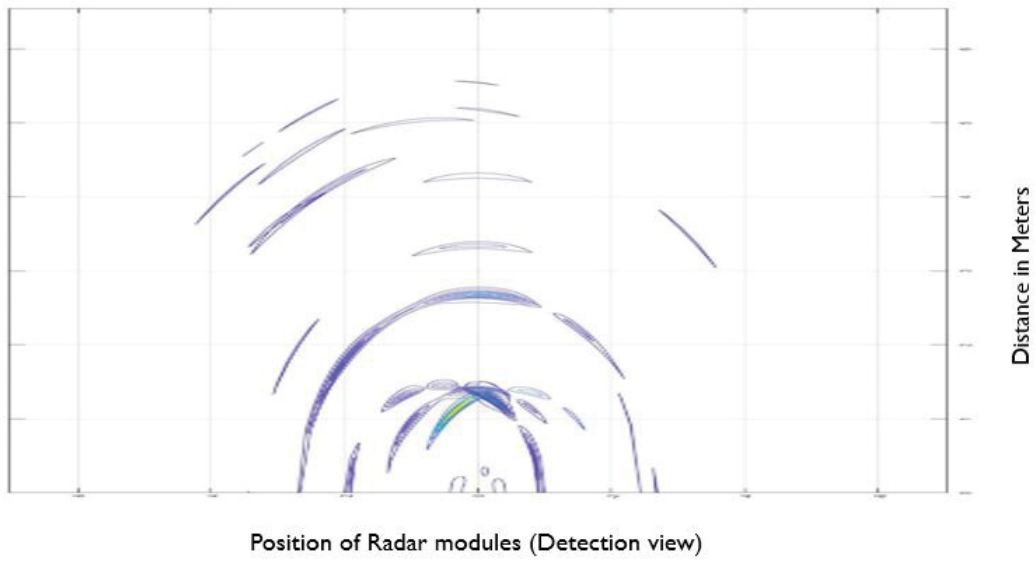


Figure A.3: V3 : Contour Plot

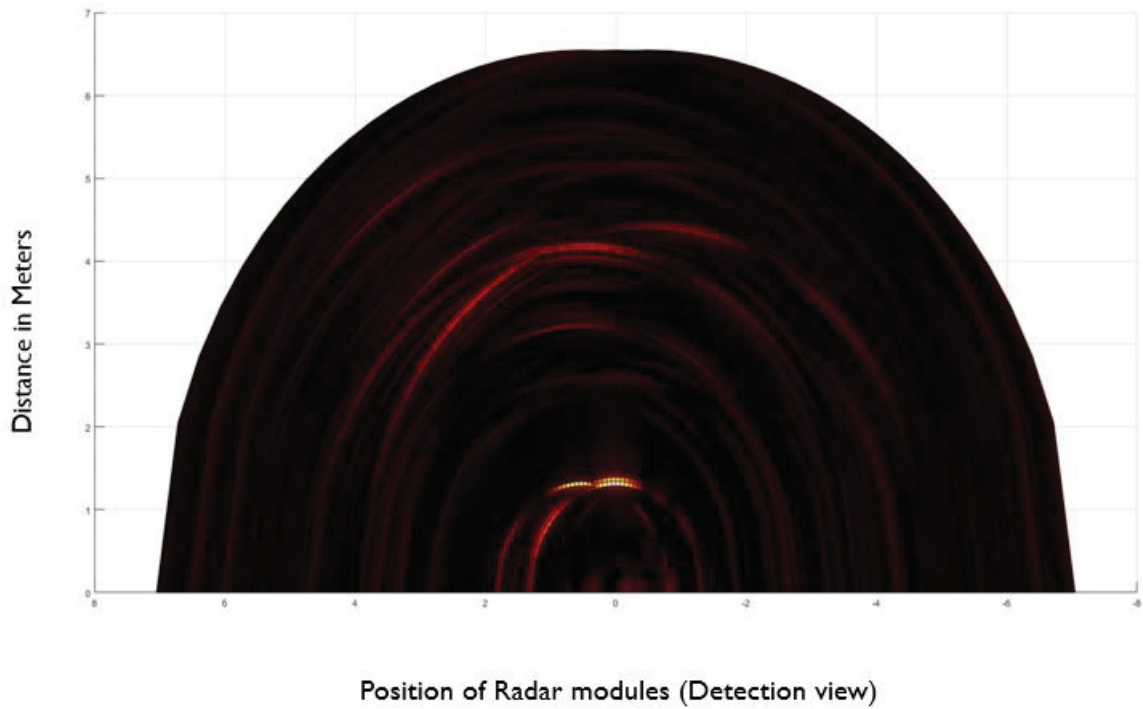
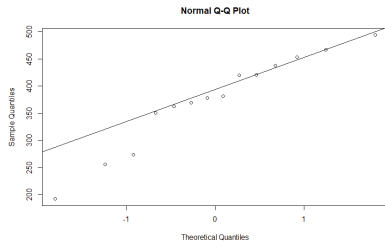
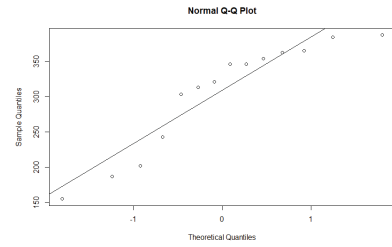


Figure A.4: V4 : Heat Temperature Plot

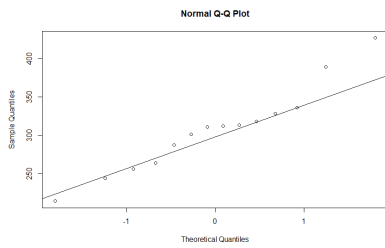
A.3 Data Distribution : Quantitative Data Analysis



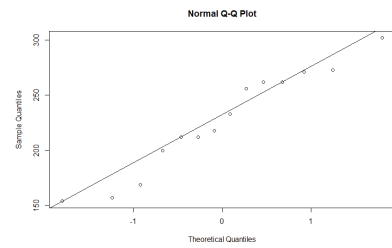
(a) Visualization 1



(b) Visualization 2

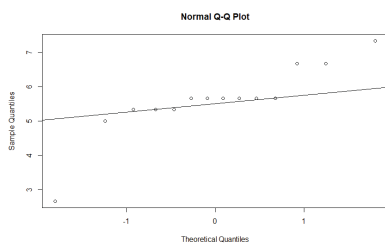


(c) Visualization 3

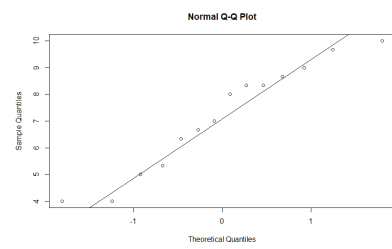


(d) Visualization 4

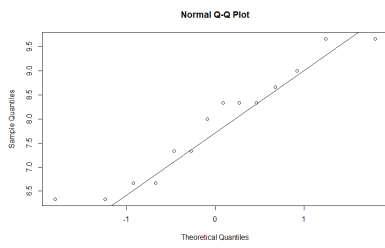
Figure A.5: Normal Q-Q Plot : Response time



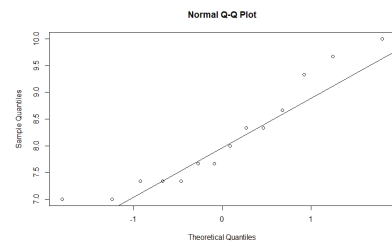
(a) Visualization 1



(b) Visualization 2

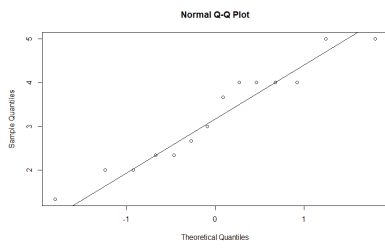


(c) Visualization 3

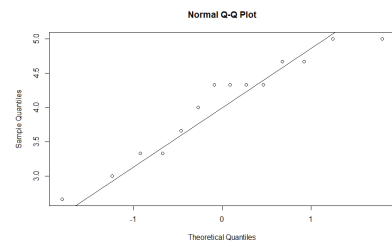


(d) Visualization 4

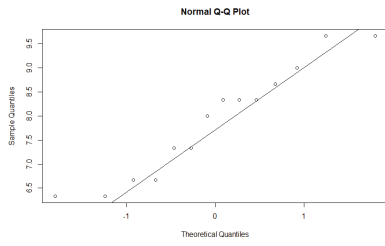
Figure A.6: Normal Q-Q Plot : Ease of Understanding



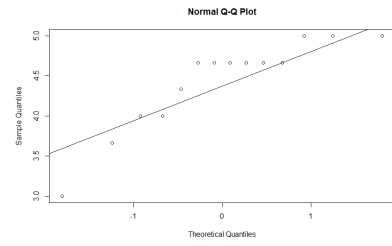
(a) Visualization 1



(b) Visualization 2

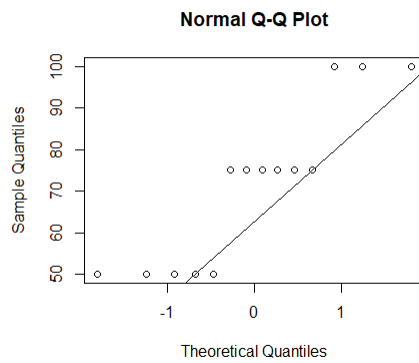


(c) Visualization 3

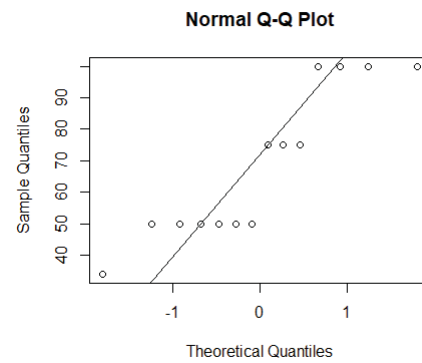


(d) Visualization 4

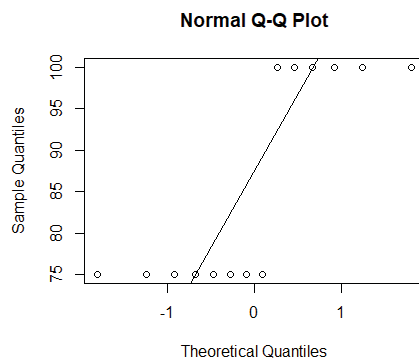
Figure A.7: Normal Q-Q Plot : Confidence



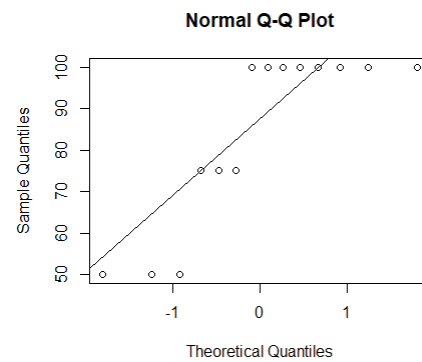
(a) Visualization 1



(b) Visualization 2



(c) Visualization 3



(d) Visualization 4

Figure A.8: Normal Q-Q Plot : Accuracy

