ejpmr, 2023,10(7), 540-556

SJIF Impact Factor 6.222

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Review Article

ISSN 2394-3211

EJPMR

DATA VISUALIZATION IN CLINICAL DATA MANAGEMENT (CDM)

Kombathula Jayasheela*, Shaikh Sameer and Sruthi

[1]B PHARMACY.
[2]BSC, MSW, MBA.
[3]B PHARM, M PHARM.

***Corresponding Author: Kombathula Jayasheela**

B. PHARMACY.

Article Received on 21/05/2023 **Article Revised on 11/06/2023** **Article Accepted on 02/07/2023**

ABSTARCT

Data visualization techniques can assist Healthcare providers in understanding large sets of clinical data, Data visualization in the context of clinical data management refers to the graphical representation of clinical data to facilitate analysis, interpretation, and communication of information related to healthcare and patient outcomes. data visualization in clinical data management transforms complex clinical data into visual representations that simplify analysis, support decision-making, and enhance communication. It enhances data understanding, aids in pattern identification and insights, supports efficient exploration, facilitates communication and collaboration, assists in decision-making, ensures data quality, and improves presentation and reporting. The purpose of data visualization in clinical data management is to effectively communicate and understand complex clinical data through visual representations. Enhancing data understanding and gaining valuable insights through data visualization in clinical data management is crucial. Facilitating the identification of patterns, trends, and outliers through data visualization is of great importance in clinical data management. The right visual aid is the key to preventing user confusion and making sure our analysis should be accurate. To design visualizations for solving real-world problems, we typically conduct design studies. In a design study, we engage closely with domain experts and their data analysis problems. The very first step of a design study is the domain characterization. Domain characterization describes domain-specific problems, such as what data is needed. A graphical depiction of data and information is called data visualisation. Data visualisation tools can offer a simple way to anticipate and comprehend trends and patterns in data by employing visual elements like graphs, charts, and maps. They assist us in categorising and organising data based on categories and themes, which makes it simpler to digest and comprehend. Using visualisation tools like CDMS, EDC, Epro, etc. Addressing data complexity and visualization, ensuring interpretability and accuracy considerations in data visualization, clinical data management systems can ensure that visualizations are reliable, trustworthy, and facilitate accurate insights.

I. INTRODUCTION

A. Explanation of data visualization in the context of clinical data management

Data visualization techniques can assist Healthcare providers in understanding large sets of clinical data, Data visualization in the context of clinical data management refers to the graphical representation of clinical data to facilitate analysis, interpretation, and communication of information related to healthcare and patient outcomes. It involves transforming raw data into visual elements such as charts, graphs, and diagrams, which allow healthcare professionals, researchers, and stakeholders to gain understandings and make informed decisions based on the data. The primary goal is to extract meaningful insights, identify patterns, and communicate information effectively to healthcare professionals, researchers, and decision-makers. Here's an explanation of data visualization in the context of clinical data management:

- Data Analysis: Visualizations provide a way to analyse and summarize complex clinical data effectively. By using various chart types, researchers can compare different variables, identify correlations, and detect patterns. This analysis helps in understanding the relationships between patient demographics, medical conditions, treatment outcomes, and other relevant factors.
- Simplifying Complex Data: Clinical data often consists of numerous variables and large datasets. Data visualization simplifies this complexity by condensing the information into visual forms that are easier to comprehend. By presenting data visually, patterns and relationships become more apparent, enabling users to extract meaningful insights.
- Decision Support: Visualizing clinical data supports evidence-based decision-making in healthcare. By presenting data visually, decision-makers can quickly grasp the implications and trends, helping

them make informed choices regarding patient care, resource allocation, and policy development. Visualizations facilitate the identification of areas that require improvement, enabling data-driven decision-making.

- Communication and collaboration: Data visualization plays a crucial role in communicating complex clinical information to diverse audiences, including healthcare professionals, researchers, policymakers, and patients. Well-designed visualizations enable clear and effective communication of findings, research outcomes, and insights. It promotes collaboration and facilitates interdisciplinary discussions by providing a common visual language that transcends disciplinary boundaries.
- Monitoring and tracking: Data visualization enables the monitoring and tracking of patient outcomes, treatment effectiveness, and disease progression over time. By representing data in visual formats, clinicians can identify changes, trends, or potential issues in patient health. Visualizations also help in evaluating the impact of interventions, tracking adherence to protocols, and identifying opportunities for intervention or improvement.
- Quality Assurance: Visualizations play a role in quality assurance by enabling data validation and error detection. By visually examining the data, researchers and data managers can identify outliers, inconsistencies, or missing values that may indicate data quality issues. This ensures the accuracy and reliability of the clinical data being analysed.
- Reporting and Presentations: Data visualization is essential for creating informative and engaging reports and presentations. Visual representations enhance the clarity and impact of the information being conveyed. By using charts, graphs, and diagrams, researchers can effectively communicate complex findings to diverse audiences, including clinicians, administrators, policymakers, and patients.

Common visualization techniques used in clinical data management include:

- Line charts: Used to display trends and changes over time, such as disease progression or treatment response.
- Bar charts: Suitable for comparing different categories or groups, such as comparing the effectiveness of different treatments.
- Scatter plots: Effective for exploring relationships between variables, such as the correlation between risk factors and disease incidence.
- Heat maps: Useful for representing large datasets and identifying patterns or clusters of data, such as identifying hotspots of disease prevalence.
- Geographic maps: Employed to visualize data spatially, such as the distribution of diseases across different regions or the density of healthcare facilities.

Thus, data visualization in clinical data management transforms complex clinical data into visual representations that simplify analysis, support decision-making, and enhance communication. It enables users to explore data, monitor patient outcomes, identify patterns, and share insights effectively, leading to improved healthcare delivery and patient outcomes.

B. Significance of data visualization for effective data analysis and interpretation

Data visualization is significant for effective data analysis and interpretation in Clinical Data Management. It enhances data understanding, aids in pattern identification and insights, supports efficient exploration, facilitates communication and collaboration, assists in decision-making, ensures data quality, and improves presentation and reporting. By harnessing the power of data visualization, healthcare professionals can derive valuable insights from clinical data, leading to improved patient care, research outcomes, and decision-making processes. Here are some key reasons why data visualization is essential in this context:

- Enhanced Data Understanding: Data visualization transforms complex clinical data into visual representations such as charts, graphs, and diagrams. These visual representations make it easier to understand and interpret the data. By presenting data in a visual format, patterns, trends, and relationships become more apparent, enabling users to gain a deeper understanding of the data.
- Pattern and Anomaly Detection: Visual representations help identify patterns, trends, and anomalies in clinical data. By visualizing the data, clinicians and researchers can quickly spot irregularities, outliers, or unexpected relationships. Visual cues provide immediate visual feedback, enabling the detection of important patterns and outliers that may require further investigation or analysis.
- Efficient Data Exploration: Visualization techniques enable efficient exploration of large and complex clinical datasets. Interactive visualizations allow users to interact with the data, filter and manipulate variables, and drill down into specific subsets of information. This interactivity facilitates a deeper understanding of the data, supports hypothesis testing, and promotes data-driven discoveries.
- Communication and Collaboration: Data visualization facilitates effective communication and collaboration among healthcare professionals, researchers, and stakeholders. Visual representations of data make it easier to convey complex information, research findings, and insights to diverse audiences. Visualizations promote better understanding, encourage interdisciplinary discussions, and support collaborative decision-making.
- Error Detection and Data Quality Assurance: Data visualization can assist in identifying data errors, outliers, and inconsistencies. Visual inspection of

the data reveals data quality issues, missing values, or anomalies that may require further investigation or cleaning. By ensuring data accuracy and integrity, visualization contributes to robust and reliable clinical data management.

- Presentation and Reporting: Data visualization enhances the presentation and reporting of clinical data. Visual representations make it easier to convey complex information in a clear and engaging manner. By using appropriate charts, graphs, and visual elements, researchers can effectively communicate their findings to diverse audiences, including healthcare professionals, administrators, and patients.

C. Purpose and objectives of the review article

The purpose of data visualization in clinical data management is to effectively communicate and understand complex clinical data through visual representations. It aims to facilitate data exploration, analysis, and interpretation, ultimately supporting evidence-based decision-making and improving patient care. Thus, it involves transforming raw data into visual representations, such as charts, graphs, and dashboards, to facilitate understanding, decision-making, and collaboration. The objectives of data visualization in clinical data management include enhancing data insights, facilitating exploratory analysis, monitoring and tracking patient progress, identifying trends and Patterns, supporting decision-making, enhancing communication and collaboration, ensuring data quality and accuracy, and enhancing reporting and presentation. Lam and researchers conducted an extensive literature review on evaluations of data visualizations (2012).

II. Importance of Data Visualization in CDM

A. Enhancing data understanding and insights

Enhancing data understanding and gaining valuable insights through data visualization in clinical data management is crucial. It enables healthcare professionals to simplify complex data, make informed decisions, identify trends and patterns, generate hypotheses, facilitate effective communication and collaboration, detect anomalies, and derive data-driven insights. By harnessing the power of data visualization, clinical data management can contribute to improved patient care, research outcomes, and healthcare practices. Here's why it is important:

Complex Data Simplification, Improved Decision-Making, Identification of Trends and Patterns, Hypothesis Generation and Research Direction, Effective, Communication and Collaboration, Early Detection of Anomalies or Outliers, Enhanced Data-Driven Insights.

B. Facilitating identification of patterns, trends, and outliers

Facilitating the identification of patterns, trends, and outliers through data visualization is crucial in clinical data management. It enables early detection of anomalies, identifies trends and temporal patterns, analyses correlations and relationships, identifies subpopulations, supports monitoring and predictive analytics, aids in quality assurance, and enhances insights for evidence-based practice. By leveraging data visualization techniques, healthcare professionals can uncover valuable insights and make informed decisions that improve patient care and outcomes. Here are the key reasons why it is significant:

- Early Detection of Anomalies and Outliers: Data visualization enables the quick identification of anomalies and outliers in clinical data. These could include unusual patient outcomes, unexpected responses to treatments, or outliers in laboratory measurements.
- Identification of Trends and Temporal Patterns: Visual representations of clinical data allow for the identification of trends and temporal patterns over time.
- Correlation and Relationship Analysis: Data visualization aids in identifying correlations and relationships between variables within clinical data. By plotting data points on a scatter plot or using other visualization techniques, healthcare professionals can visually assess the strength and direction of relationships between different variables. This information can inform treatment plans, risk assessments, and predictive models.
- Identification of Subpopulations and Subgroups: Visualizing clinical data helps in identifying subpopulations or subgroups within the patient population. By segmenting the data and representing it visually, healthcare professionals can discern patterns and characteristics unique to specific groups. This understanding can support personalized medicine, targeted interventions, and the development of tailored treatment approaches.
- Monitoring and Predictive Analytics: Data visualization supports the monitoring of patient health and the identification of potential predictive indicators. By visualizing patient data over time, healthcare professionals can track changes in vital signs, disease progression, or treatment response. This information can guide timely interventions, facilitate early diagnosis, and support predictive analytics for risk assessment or prognostic modelling.
- Quality Assurance and Error Detection: Data visualization aids in quality assurance by helping healthcare professionals identify data errors, inconsistencies, or missing values. Visual inspection allows for the identification of data quality issues that may affect the integrity and reliability of the data. By detecting such errors, steps can be taken to rectify or validate the data, ensuring accurate analysis and interpretation.
- Enhanced Insights and Evidence-Based Practice: Visualizing patterns, trends, and outliers in clinical data leads to enhanced insights and evidence-based practice. By visually examining the data, healthcare

professionals can gain a deeper understanding of patient characteristics, treatment outcomes, and disease progression. These insights drive evidence-based decision-making, enabling healthcare professionals to make informed choices regarding patient care, treatment strategies, and resource allocation (NCBI, 2022).

C. Supporting data quality assessment and anomaly detection

Supporting data quality assessment and anomaly detection are essential aspects of data visualization in clinical data management. Supporting data quality assessment and anomaly detection through data visualization is critical in clinical data management. It enables early detection of data anomalies, ensures data accuracy and integrity, visualizes data distribution, supports comparisons and benchmarks, facilitates interactive data exploration, aids in compliance with regulatory standards, and enhances communication and documentation of data quality issues. By leveraging the power of data visualization, healthcare professionals can improve data quality, enhance data reliability, and foster trust in the clinical data being used for decision-making and research.

1. Early Detection of Data Anomalies
2. Ensuring Data Accuracy and Integrity
3. Visualization of Data Distribution
4. Comparisons and Benchmarks
5. Interactive Data Exploration
6. Supporting Data Auditing and Compliance:
7. Facilitating Communication and Documentation

D. Enabling effective communication of findings to stakeholders

Enabling effective communication of findings to stakeholders is a crucial aspect of data visualization in clinical data management. Enabling effective communication of findings to stakeholders through data visualization is vital in clinical data management. It simplifies complex information, enhances accessibility and engagement, supports data-driven decision-making, facilitates knowledge exchange and collaboration, improves presentation and reporting, aids in storytelling and narrative construction, and promotes stakeholder engagement and buy-in. By leveraging the power of data visualization, clinical data management can effectively communicate insights, drive informed decision-making, and promote positive change in healthcare practices. Here's why it is important:

- Simplifying Complex Information: Data visualization simplifies complex information by presenting it visually, making it easier for stakeholders to grasp the key findings and insights. Visual representations such as charts, graphs, or diagrams provide a clear and concise overview of the data, allowing stakeholders to quickly understand.
- Enhancing Clarity and Interpretation: Data visualization enhances the clarity and interpretation of findings in clinical data management. Visual representations provide a visual context that aids in understanding the relationships, trends, and patterns within the data. By presenting data visually, stakeholders can easily interpret the information.
- Facilitating Decision-Making: Effective data visualization supports evidence-based decision-making by providing stakeholders with a visual representation of the data insights.
- Engaging and Persuasive Presentations: Data visualization makes presentations more engaging and persuasive.
- Promoting Knowledge Sharing and Collaboration: Data visualization facilitates knowledge sharing and collaboration among stakeholders.
- Supporting Stakeholder Engagement: Data visualization encourages stakeholder engagement by involving them in the data analysis and interpretation process.
- Enhancing Transparency and Trust: Data visualization enhances transparency and trust in clinical data management. Visual representations make the data analysis process more transparent and understandable to stakeholders. By clearly presenting the data and the analytical methods used, stakeholders can have confidence in the validity and reliability of the findings.

III. Types of Data Visualization Techniques

Some of the most common types of data visualization chart and graph formats include:

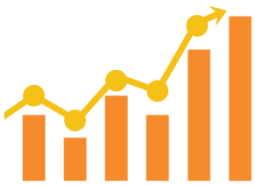
- Column Chart
- Bar Graph
- Stacked Bar Graph
- Stacked Column Chart
- Area Chart
- Dual Axis Chart
- Line Graph
- Mekko Chart
- Pie Chart
- Waterfall Chart
- Bubble Chart
- Scatter Plot Chart
- Bullet Graph
- Funnel Chart
- Heat Map

The right visual aid is the key to preventing user confusion and making sure our analysis should be accurate. Let’s dive into 10 of these 15 types of charts and graphs below. There are myriad different types of charts, graphs and other visualization techniques that can help analysts represent and relay important data. Let’s take a look at 10 of the most common ones:

1. Column Chart

It is one of the most common type of data visualization tools. It is a simple, time-honoured way to show a comparison among different sets of data. And it can also use a column chart to track data sets over time. Column

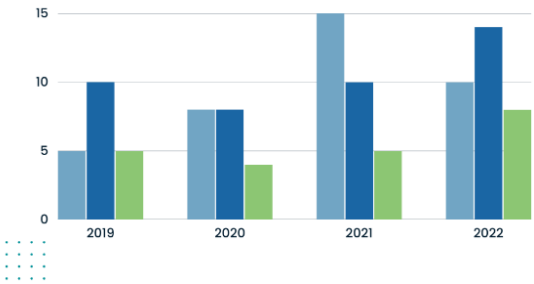
charts are used to track monthly sales figures, revenue per landing page, or similar measurements. Consistent colours help keep the focus on the data itself, though can introduce accent colours to emphasize important data points or to track changes over time.



2. BAR GRAPH

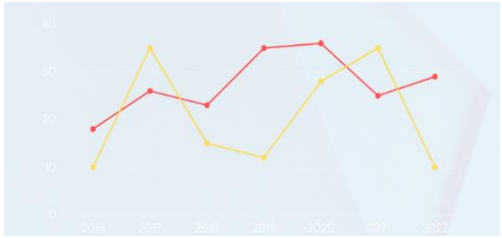
Bar graph and column chart will be the same way, though column charts limit label and compare space. It's best to stick with a bar graph if:

- Working with lengthier labels
- Displaying negative numbers
- Comparing 10 or more items



3. Stacked Bar Graph

For comparing many different items, we want to track the individual growth of each data set itself, along with the group's growth as a collective whole. To reveal this part-to-whole relationship, should create a stacked bar graph. It would look similar to a standard bar chart. The "stacked" layout represents this chart's contrasting colour scheme. These colours map back to a legend that accompanies a map.



4. LINE GRAPH

It is another one of those standard chart types that's instantly recognizable. A line graph is designed to reveal trends, progress, or changes that occur over time. As such, it works best when data set is continuous rather than full of starts and stops.

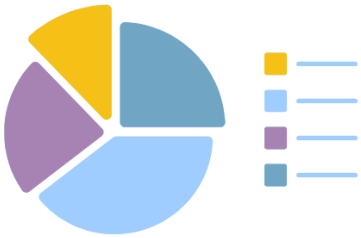
5. DUAL AXIS CHART

This layout allows to show a relationship between different variables, and it works best when working with three data sets as follows:

- One set of continuous data
- Two data sets grouped by category

6. MEKKO CHART

This chart might be less familiar with unless in the data analysation space. Standing for Marimekko chart, a Mekko chart has a similar layout to a stacked bar graph, with one major exception: Instead of tracking time progression, the X-axis measures another dimension of data sets. With this layout, can compare values, measure the composition of each value, and analyse data distribution all at the same time.



7. Pie Chart

A pie chart represents one static number, divided into categories that constitute its individual portions. It represents numerical amounts in percentages.

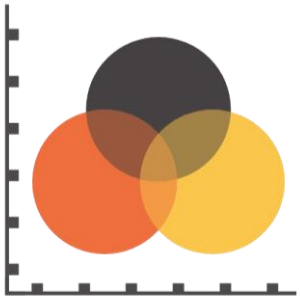
These are especially helpful in digital marketing, as it can be used to show a breakdown of:

- Market shares
- Marketing expenditures
- Customer demographics
- Customer device usage
- Online traffic sources

Pie chart have plenty of differentiation between slices. It's best to limit the number of categories to illustrate.

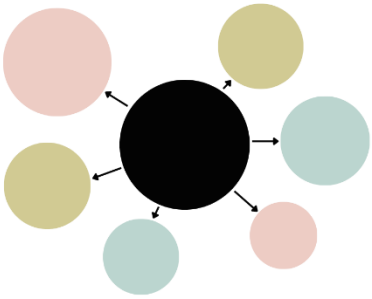
8. Scatter Plot

This type of visualization is also called a scattergram, and These types of data visualization work best when analysing multiple data points and looking for any similarities within the data set. And can notice any outliers and also gain a clearer understanding of overall data distribution.



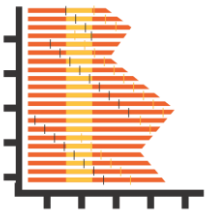
9. Bubble Chart

Like a scatter chart, a bubble chart can also show relationships or distribution. In this variation, however, replace the data points with bubbles. And also vary the sizes of the bubble to represent a third data set. As with a scatter chart, a bubble chart does not use a category axis. The data sets as X-values, Y-values and now, Z-values (bubble size).



10. Bullet Graph

A bullet graph can help them visually track the progress. Similar in layout to a bar graph, these also incorporate other visual elements. When using a bullet graph, it begins with a one, main measure, and then compare that measure to another (or multiple) measure to find a deeper meaning and connection.



WHAT ARE VISUAL ANALYTICS TOOLS

Visual analytics tools extract data from multiple data sources by using modern technologies which enable to explore information in a visual, interactive, and graphic manner. Such software helps to easily identify key trends and extract actionable business insights.

Businesses today must manage and process large volumes of datasets and information. Techniques and procedures to collect, monitor, evaluate, and analyse their data have become a necessity for both large enterprises and small businesses.

Professional visual analytics tools offer these critical features and benefits

An interactive view of data: Revealing patterns behind voluminous, ambiguous, and often contradictory data with interactive and visual presentations that offer additional analytical results and insights empower business users to become analysts on their own. Effective understanding, reasoning, and decision-making are more precise since these tools enable people to interact with visualizations and discover the unexpected quickly and without profound IT knowledge.

Unlock and access big data sets: The essence of visual analytics software lies within the visualization of complex and large datasets. Display of data in clear, well-defined, and graphical business dashboards ensures quick interpretation and ease of understanding. Facts and figures are communicated concisely, offering visual context that the human eye is processing on a much faster level, enabling quick consumption of information.

Data exploration: By using a professional tool it is more than efficient to analyse real-time data. Blending various data connectors into a centralized point of access ensures generating insights and ease of access to all business information, Quarterly updates don't give instant feedback on a company's performance, therefore, cannot provide the number of insights needed at any given point in time. The continuous exploration of data, on the other hand, enables users to instantly identify variables that need further analysis.

Accessibility and collaboration: Making data accessible is one of the biggest challenges businesses faces when dealing with their information. With the help of powerful visual data analysis tools that are user-friendly. Everyone in organization will be empowered to work with data. This will not only improve the way, perform the individual tasks but will also considerably enhance interdepartmental collaboration, leading to the implementation of a data-driven culture that will your business apart from the competition.

More data in less time: Thousands of rows in a spreadsheet cannot cope with the responsiveness of a visual analytics tool. The speed and accuracy on which these tools perform the analysis are incomparable with the traditional forms of analysis. More data can be processed in a significantly less amount of time, which makes these tools an unparalleled force in the business world.

Upscaled reporting: Creating visual reports no longer takes hours or days
to generate. Since visual analytics software offers automation of the reporting practice, managers and employees can focus on what really matters: the data at hand. With powerful dashboards and automated reports, each business user can look beyond the data and generate actionable, real-time insights.

Multiple visualization options: Traditional means of data analysis are static and full of numbers that are hard to understand most of the recipients. Modern online data visualization tools offer a wide range of interactive graphs and charts so that can visualize the most important key performance indicators and extract actionable insights in the most efficient way. Being able

to quickly understand what's behind your data is the basis for business success and having multiple charting options such as speedometer, area, maps, line charts, column, or bar charts, in hand is the way to accomplish it.

Customized dashboards: The visual nature of a modern business dashboard is to make it so successful for companies of all sizes. To take the power of these visual data analytics tools even one step further, dashboards can be customized according to the colours and logo of brand. This way can give the reports a more professional look and use them as visual support for meetings with clients.

Interactive dashboard is an business data management tool that allows users to interact with data by tracking, analysing, monitoring, and displaying key business metrics. By using an interactive dashboard, users can dig deeper into an organization's operational information and filter it in several ways. Financial and operational data can be merged onto this central platform to handle all functions and related processes at a corporate level. Interactive dashboards also allow organizations to create and track key performance indicators (KPIs). This allows for data to be inspected from various perspectives, right down to every tiny detail. Data-driven decisions are only possible when all relevant data is coordinated to create a clear, understandable display of business-related figures. The analytics software takes the bulk of the work and feeds information to display on interactive dashboards. Because dashboards are shared with relevant stakeholders from business colleagues to partners and customers everyone who needs to access the information. Using interactive dashboards, organizations can effectively communicate relevant business figures to those within and outside the organization. These dashboards can be tailored so employees can have limited access and only see what is appropriate for their role. Interactive dashboards add a more visual element to presentations.

HEAT MAP

Heatmap visualization or heatmap data visualization is a method of graphically representing numerical data where the value of each data point is indicated using colours. The most commonly used colour scheme used in heatmap visualization is the warm-to-cool colour scheme, with the warm colours representing high-value data points and the cool colours representing low-value data points.

The reason that visualization of data through methods like heatmap has become so popular is that humans, in essence, are visual beings. Many studies on human psychology and perception suggest that view and process visuals way more efficiently than written text or written data.

TREE MAPS

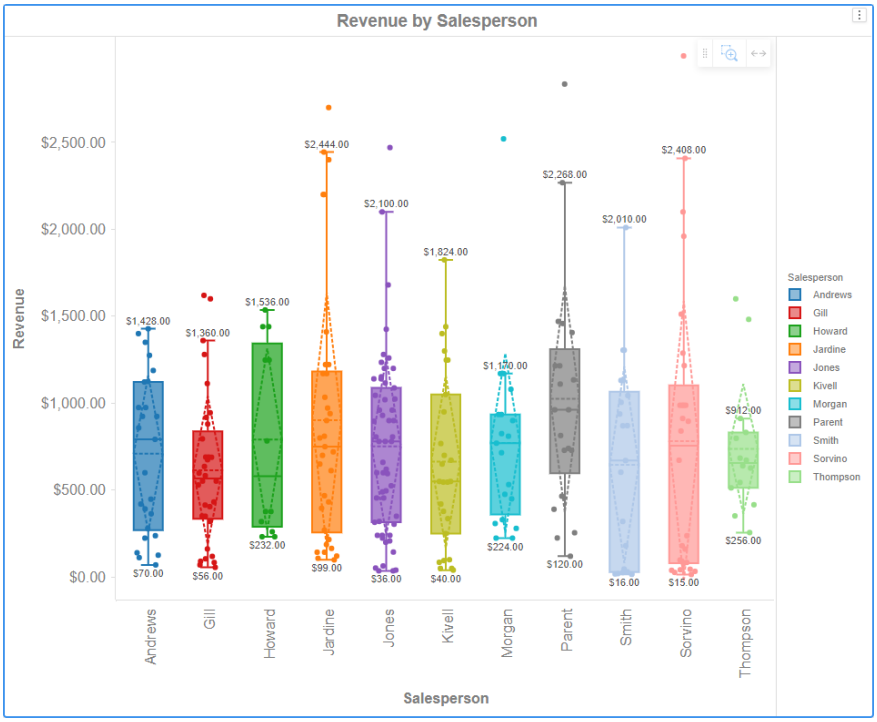
Tree maps provide an accessible way for viewers to interpret their data at a glance. Colour can represent dimensions (such as categories) or measures (such as KPIs). If used to represent a KPI, a darker colour may highlight extremes, high or low. For dimensions, a user might use a categorical palette, assigning a different colour for each available shipping mode. For measures, a continuous colour palette would show a company's sales numbers or profit.

GEOGRAPHIC INFORMATION SYSTEM

A geographic information system (GIS) is a system that creates, manages, analyses, and maps all types of data. GIS connects data to a map, integrating location data with all types of descriptive information. This provides a foundation for mapping and analysis that is used in science and almost every industry. GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication and efficiency as well as better management and decision making

BOX PLOT

A box plot visualization allows to examine the distribution of data. One box plot appears for each attribute element. Each box plot displays the minimum, first quartile, median, third quartile, and maximum values. In addition, it can choose to display the mean and standard deviation as dashed lines. Outliers appear as points in the visualization. It can adjust the spacing between points (that is, jitter) to avoid overlap. A box plot must include at least one metric and at least one attribute.



HISTOGRAM

A histogram is a chart that displays numeric data in ranges, where each bar represents how frequently numbers fall into a particular range. Histograms consist of a series of vertical bars along the x-axis. Histograms are most commonly used to depict what a set of data looks like in aggregate. At a quick glance, histograms tell whether a dataset has values that are clustered around a small number of ranges or are more spread out.

DISTRIBUTION PLOT

Distribution plots plot a univariate distribution of data. They are used to visually assess the distribution by comparing the observed data and expected outcomes. A distribution plot displays two different plots for the same variable

SANKEY DIAGRAM

A Sankey Diagram is a visualisation technique that allows to display flows. Several entities (nodes) are represented by rectangles or text. Their links are represented with arrow or arcs that have a width proportional to the importance of the flow.

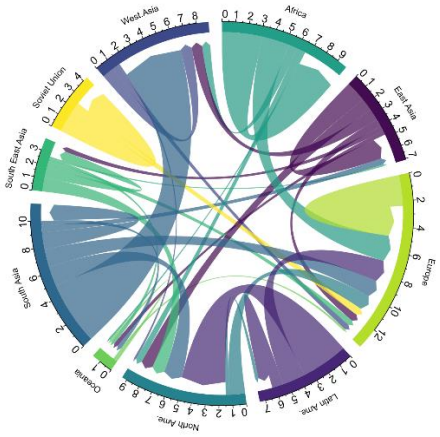
NETWORK GRAPHS

A network graph is a chart that displays relations between elements (nodes) using simple links. Network graph allows us to visualize clusters and relationships between the nodes quickly; the chart is often used in industries such as life science, cybersecurity, intelligence, etc.

CHORD DIAGRAM

A chord diagram represents flows or connections between several entities (called nodes). Each entity is represented by a fragment on the outer part of the circular layout. Then, arcs are drawn between each entity. The size of the arc is proportional to the importance of the flow.

Chord diagrams are eye catching and quite popular in data visualization. They allow to visualize weighted relationships between several entities.



IV. Visualization Considerations in CDM

To design visualizations for solving real-world problems, we typically conduct design studies. In a design study, we engage closely with domain experts and their data analysis problems. The very first step of a design study is the domain characterization. Domain characterization describes domain-specific problems, such as what data is needed. Humans are the only species that can learn tacit knowledge through cognitive processes. While cognitive science knowledge elicitation techniques might be useful in this situation, additional research is required to

determine how applicable they are for domain characterisation in visualisation design studies.

A Data preparation and preprocessing for effective visualization

A five-stage data gathering process is used to gradually focus on important cognitive points.

1. Incident Choice: The purpose of the event's selection stage is to choose a task requiring competencies outside of everyday knowledge. The situation's main decision-maker should be the participant. one must anticipate a difference between the decisions of a professional and those of a beginner in order to extract actual expertise.

2. Incident Recall: The Unstructured Incident's objective the purpose of the recall stage is to gain a first impression of the scenario and to get the participants' memories going. The person who participated is asked to provide a detailed account of the entire incident. This could involve storing a data set until an intriguing association is discovered, for instance. Interviewers should concentrate on comprehending the narrative. Interrupting for anything other than a quick clarification is dangerous.

3. Establishing a timeline: The Timeline's objective Establishing a shared understanding between interviewers and participants is the purpose of the construction stage. The interviewers create a timeline of the situation based on what they hear. It includes the order and length of events. A single occurrence (like a data point) can be an event.

4. Identification: The decision point identification stage's objective is to locate crucial turning moments in the timeline so that a thorough study may be conducted. The interviewers elicit those instances where the situation may be understood in a variety of ways or where several possible actions could be taken. Some are clear from verbal cues (such as "I had to decide whether to include this predictor in the selection"), while others are less obvious. Others require selecting one of several options (such as focusing on one aspect of the data first) or making a decision it has an impact on the action.

5. Decision Point Probing: The purpose of the Decision Point Probing stage is to help participants better appreciate the significance of data for their situational evaluation. Interviewers are employed by the decision areas and request clarification. Coding is typically used to lay the foundation for transforming interview data into various representations that depict domain knowledge, reasoning, and task activity. provide visuals that were effective for their applications. We specifically point out two of these artefacts because we believe they best serve the goal of domain characterization. The situation assessment record (SAR) captures the expert's perception of the incident's dynamic evolution. The situation's turning points are described, together with their underlying cues, experience, knowledge, objectives, and actions.

DATA REPROCESSING

Any kind of processing of raw data to prepare it for another data processing operation is called data preprocessing and is part of data preparation. Historically, this was an important first step in the data mining process. Data preparation methods have recently changed to train AI and machine learning models to draw conclusions. Data preprocessing transforms data into a format that can be processed faster and more efficiently in data mining, machine learning, and other data science tasks. To ensure reliable results, these techniques are typically applied in the first stages of machine learning and AI development pipelines.

Preprocessing data can be done using a variety of tools and techniques, such as the following

- Sampling, which chooses a representative subset from a large population of data;
- Transformation, which modifies raw data to create a single input; denoising, which eliminates noise from data;
- Imputation, which creates statistically relevant data for missing values;
- Normalisation, which arranges data for easier access; and
- Feature extraction, which isolates a relevant feature subset that is significant in a particular context.

These techniques and tools can be applied to a range of data sources, including streaming data and data that has been stored in files or databases.

The steps used in data preprocessing include

1. Data Profiling. Data profiling is the process of examining, analysing, and reviewing data to gather statistics about its quality. It starts with collecting existing data and its properties. A data scientist identifies a dataset relevant to the problem at hand, takes inventory of its important attributes, and hypothesizes features that may be relevant to a proposed analysis or machine learning task. Also, tie your data sources to related business concepts and consider which preprocessing libraries are available.
2. Sanitize data. The goal is to find the easiest way to solve quality problems. B. Make sure the raw data is suitable for feature engineering by removing erroneous data or filling in missing data.
3. Data reduction. Raw data sets often contain redundant data resulting from different characterizations of phenomena, or data that is not relevant to any particular ML, AI, or analytics task. Data reduction uses techniques such as principal component analysis to transform raw data into a simpler form suitable for specific use cases.
4. Data conversion. Here, the data scientist considers how different aspects of the data should be organized to make the goal most meaningful. This includes structuring unstructured data, combining

key variables where meaningful, and identifying key areas of focus.

5. Data Enrichment. In this step, the data scientist applies various feature engineering libraries to the data to perform the desired transformations. The result is a dataset organized to give the best balance between new model training time and required computational power.
6. Data Verification. This phase splits the data into her two sets. The first set is used to train a machine learning or deep learning model. The second set consists of test data used to measure the accuracy and robustness of the resulting model. This second step of his helps identify hypothetical problems used in data cleaning and feature engineering. Once the data scientist is satisfied with the results, delegate the preprocessing task to the data engineer, who then finds a way to scale it into production. If not, the data scientist can go back and modify the way they implemented the data cleansing and feature engineering steps.

B Selection of appropriate visualization techniques based on data types and objectives

C Design principles for creating clear and informative visualizations

The purpose of data visualization is to give viewers faster insights from data and make trends clearer. Visualizations are often poor at conveying insight because they were not created with a clear purpose or audience in mind. Whether your visualization is for yourself, an internal or external audience, narrow the purpose of your visualization and understand which audience it will affect. This is useful for both data exploration and data presentation.

1. A clear objective is where to start when choosing different datasets to explore, when deciding how to compare different categories of data, or when creating a final visualization to decide the finishing touches. It helps to know what They don't passively dig into data when goals are set. Based on the purpose and clarity of the main "question" you are trying to answer, you can decide which secondary questions need to be addressed. After answering the secondary questions, the journey of discovery begins. This helps us draw conclusions based on facts. A clear purpose helps you identify the key characteristics and datasets needed to create visualizations that best address key concerns.

2. Depending on the type of chart you use; your data history can be profitable or costly. Various useful charts exist. Therefore, it is important to clarify the purpose and audience of data visualization. Both can affect which chart is most appropriate. Adjust the data literacy level of the target group. The types of data visualizations you need to create are heavily influenced by your audience's data literacy. If you want to make an impression and really pave the way for action, you need to target people who really need to take action.

3. Colour has meaning. Without using words, you can communicate a lot. However, there must be a careful balance when using colour. Be as straightforward as possible. Information can be highlighted using colour. Too many single colours or shades of colour can muddle data, and too many colours can cause dissonance. Also crucial are colour connections. Make the content easier for your readers to absorb and process by using intuitive colours. You should utilise red for heat and blue for cold when dealing with temperatures. Colour selection also has a significant effect. useful for highlighting contrasts in your data or demonstrating consistency between numbers.

4. Size helps highlight important information and provide more context. The shapes of the endangered animals in the previous figure mimic the size of the animals compared to other animals. However, size can also mean scaled values. Shape sizes can be adjusted based on data values instead of colours. For maps, using size to indicate values also works well. When you have multiple data points of the same size in your visualization and intermingled with each other, it's hard to distinguish between values. Visualization is easier when size and value are associated and colour is included as an additional cue.

5. keep simple less is more French author Antoine de Saint-Exupéry never studied data science, but he understood the power of simplicity and clarity. "Perfection is achieved not when there is nothing left to add, but when there is nothing left to remove," he once said. In this age of information overload, we have to choose carefully what to draw. Remove anything from the visualization that doesn't support the data. Feel free to remove it if it does not further your message or is not relevant to the purpose of your visualization.

D Addressing privacy and data protection considerations in visual representations:

Privacy protection has developed into an exact opposite of this idea. data-driven, digital era. even the modern smart devices Information about the websites we visit and the online services we use We are recording our behaviours, identities, routines, and preferences at an unprecedented rate. Often, the right to privacy is an essential one. Collateral harm in the drive for customization and commercialization is this. offering people commercial services. Researchers have recently asserted that the data environment has privacy problems. A multi-stakeholder effort within the data ecosystem is required to resolve the crisis, which calls for quick cooperation.

Re-identification by link: For data exchange, it is required but insufficient to merely suppress personally identifying information (PII), such as names, social security numbers, and email addresses. Attackers may use quasi-identifiers like age, gender, and postal codes as well as correlating characteristics of publicly accessible data sources (like voter registration) with personally accessible information (like hospital data) to undermine

privacy. The data link problem and numerous data anonymization methods, such as generalisation, suppression, destruction, grouping, etc., are widely used to describe this. used to deal with this problem. These techniques typically use customised data mining algorithms, anonymized static data tables, or anonymous visualisation. The majority of these techniques use non-interactive privacy settings, giving data owners no control over their own data or, once given, their data's results from mining and other uses.

Anonymization method: One of the most popular methods the anonymization method is the K-anonymity model. If a written record is k-anonymous Record is indistinguishable dataset form at least k − 1 other records. For example, if k = 3, k is anonymized. The dataset contains at least 3 similar combinations in each dataset May identify variables. However, K-anonymity does not provide any guarantee that the attacker has background knowledge or knowledge uniform attack. Another anonymization technique, the l-diversity model, which ensures adequate diversity in the value of characteristics, addresses this issue.

Evaluation: A dataset is completely private if all sensitive attributes and quasi-identifiers have been removed from it. But since it cannot be used for the bulk of real-world data analysis activities, that dataset will lose its analytical value. Datasets comprising attributes describing people's behaviour and qualities can be used to understand the origins of diseases, economic trends in different states, cuisine trends that are popular in a place, etc. As a result, before being made public, the personal data is changed using anonymization techniques.

V. Benefits of Data Visualization in CDM

A graphical depiction of data and information is called data visualisation. Data visualisation tools can offer a simple way to anticipate and comprehend trends and patterns in data by employing visual elements like graphs, charts, and maps. They assist us in categorising and organising data based on categories and themes, which makes it simpler to digest and comprehend. Data visualisation tools and technologies are essential in the world of big data to analyse massive volumes of information and make data-driven decisions.

1. Unlocking key values

First and foremost, data visualisation makes sure that critical values can be extracted from enormous data volumes. Particularly large amounts of data can be daunting and challenging to comprehend. This is made possible through data visualisation, which makes the important data values obvious and simple to see. Everyone in the firm can easily comprehend and interpret it as a result.

2. Identify patterns

Second, data visualisation reveals additional patterns that were previously hidden. These additional emergent characteristics in the data can create novel, useful insights that were previously unknown. Business users may see links and trends in the data through visualisation, which also provides the data more significance. Users can assess the importance of these areas to advance their business by focusing on the specific regions in the data that require attention by analysing these patterns.

3. **Easy to understand:** By transforming the data into a format that is simple for everyone to grasp, the goal is to effectively tell a story. The noise in the data would eventually be reduced as relevant information was highlighted.

4. More attractive and user-engaged

The use of visually appealing methods of data presentation and adherence to design best practises help to attain this attractiveness. By utilising the most suitable chart and formatting options, next-level visualisations convey data in a very logical manner. Additionally, graceful transitions make it possible to move between various points in a visualization's narrative in an appealing and fluid manner. As a result, a user will be more engaged with the visualisation, making interpretation and understanding simple and rapid. As a result, the audience deeply connects with the message. The most impressive visualisation may detract from the data or speak volumes, depending on whether it is intended to draw attention to it or make a statement. It's crucial to understand that in order to effectively communicate a message, data and graphics must collaborate.

5. Display complex relationships

When displaying complex interactions, conventional graphics like bar charts and line graphs frequently fall short. For instance, it is almost never possible to show a dataset with over a million unique data points in a uniform way. In that situation, an interactive graphic with hierarchies and exploration is a far better choice. An interactive data visualisation can make it easier to explore data at a higher level and tailor it to a user's needs.

VI. Visualization Tools and Technologies in CDM

There are various software tools available for CDM, such as:

- **Clinical Data Management System (CDMS)** People are the primary focus of CDMSs. To clarify, people utilise them to store information on other people. The connection between the nurse and the subject is the foundation of information elicitation, a crucial procedure and a separate art form. The contact between research personnel and subjects should be supported by a well-designed CDMS, which is why the developers must be aware of this interaction. The only way to fully grasp this is to actively engage with the research nurses.

- **Agile software** development is a method for creating software that is primarily concerned with the needs of its users. This is due to the rapid cycles that agile software development is built upon, during which a new version is first generated before being tested by users. After that, the user comments are examined, and the procedure moves forward. A developer can gain a better understanding of how a CDMS should be constructed to promote communication between the research subject and the research nurse by regularly interacting with the users. The Manifesto for Agile Software Development states that developers should put people and interactions ahead of procedures and equipment. This claim illustrates how agile development is people-oriented. In multicentric trials, a CDMS has become essential to handle the huge amount of data. Most of the CDMS used in pharmaceutical companies are commercial, but a few open-source tools are available as well. Commonly used CDM tools are ORACLE CLINICAL, CLINTRIAL, MACRO, RAVE, and eClinical Suite.
- **Electronic Data Capture (EDC):** EDC forms often include skip logic, data limits, and preselected values. Skip logic is the electronic version of the instruction, "It speeds data entry by allowing users to automatically move to the correct field instead of having to tab through irrelevant data fields or click on "Next Page." However, if the skip logic takes the user to the wrong field, then the right data can end up in the wrong place. Skip logic should therefore always be tested. Testing is simple and can be documented using a flowchart.
- **Clinical Trial Management System (CTMS):** The crucial set of tools for efficiently planning, managing, and tracking your clinical research portfolio is the clinical trial management system. The study team is guided by this specialised, all-inclusive project management tool from study launch through enrolment and monitoring to study close.

The CTMS is a collection of eClinical technologies that perform best when they all operate in unison:

1. Managing Contacts - Sites & Teams
2. Monitoring & Calendar
3. eTMF and document management
4. Study Milestones & Tasks Project Plan
5. System of Contract & Payment
6. Integration of subject tracking and EDC
7. Authoring Visit Reports and Letters
8. Business reporting and analytics

- **Electronic Patient Reported Outcomes (ePRO):** A patient-reported result that is gathered electronically is known as an "electronic patient-reported outcome" (ePRO). Clinical trials are where ePRO approaches are most frequently utilised, although they are also applied in other areas of health care. Most ePRO questions go through the language validation process as part of the regulatory process. Data that is collected for a clinical study is seen as an example of electronic source data.
- **Randomization and Trial Supply Management (RTSM):** An RTSM system manages drug dosing and dispensing, automates Investigational Product (IP) supply and resupply to depots, sites, and direct-to-patient, and regulates patient randomization with efficiency and accuracy. Trials now have access to a larger and more varied patient population because to the automation of supply and logistics management. Modern systems may be swiftly set up and quickly modified when necessary. This automation can shorten the trial period by a few weeks, saving money and ultimately accelerating the release of new treatments.

These software tools & clinical trial data management system is designed to support various CDM activities such as data entry, data cleaning, data validation, data analysis, and data reporting and to facilitate the overall management of the data collected during a clinical trial.

Discussing real -world use cases and their outcome

1. **Clinical Trail Data Visualization:** A pharmaceutical company used data visualization to analyse clinical trial data and identify potential safety issues. The company created a dash board that displayed adverse event data, as well as data on vital signs and laboratory values. This allowed the company to quickly identify potential safety and take appropriate action to protect the safety of trail participants. The outcome was improved safety monitoring and faster identification of potential safety issues.
2. **Patient Data Visualization:** A hospital used data visualization to analyse patient data visualization to analyse patient data and identify trends and patterns. The hospital created a dashboard that displayed key metrics, such as length of stay, readmission rates, and patient satisfaction scores. This allowed the hospital to identify areas for improvement and make data-driven decisions to improve patient outcomes. The outcome and increased efficiency in hospital operations.
3. **Drug Utilization Visualization:** A health insurance company used data visualization to track drug utilization and identify potential cost savings opportunities' company created a dashboard that displayed drug utilization patterns, including the number of prescriptions, and the total cost of each drug. This allow the company to identify drugs that were being over prescribed or were being overprescribed or were too expensive and take action to reduce costs. The outcome reduces costs and improve efficiency in drug utilization.
4. **Clinical Trial progress visualization:** A clinical research organization used data visualization to monitor the progress of multiple clinical trials. The

organization created a dashboard that displayed key performance indicators, such as enrolment rates, protocol deviations, and data quality metrics. This allows the organization to quickly identify issues and take corrections action to keep the trails on track. The outcome improved efficiency in clinical trial management and faster completion of tails.

VIII. Challenges and Future Directions

Addressing data complexity and visualization scalability

Addressing data complexity and visualization scalability through these techniques, clinical data management systems can effectively handle and visualize complex clinical datasets. These challenges can be addressed:

- Data Simplification and Aggregation: Dealing with complex clinical data requires techniques to simplify and aggregate the data effectively. This can involve data pre-processing steps such as data cleaning, data transformation, and feature extraction. By simplifying the data and reducing its complexity, visualization techniques can handle large and complex datasets more efficiently.
- Hierarchical and Multilevel Visualization: Clinical data often exhibits hierarchical structures, with multiple levels of granularity and nested relationships. Utilizing visualization techniques that support hierarchical data representation can help manage data complexity. Techniques such as tree maps, sunburst charts, or nested graphs enable the visualization of hierarchical structures, allowing stakeholders to explore data at different levels of detail.
- Advanced Visual Encoding: To handle complex clinical data, advanced visual encoding techniques can be employed. These techniques involve mapping multiple dimensions of data to visual properties such as colour, size, shape, and texture. By utilizing visual encodings effectively, complex data relationships and patterns can be represented visually, reducing the cognitive load on users and enhancing data understanding.
- Interactive and On-Demand Visualization: Interactive visualization techniques allow users to interact with the data and explore it in real-time. By incorporating interactive features such as zooming, panning, filtering, and brushing, users can navigate through large datasets and focus on specific subsets of interest. On-demand visualization techniques load and render data dynamically, enabling efficient handling of large and complex datasets without overwhelming system resources.
- Scalable Visualization Platforms: To address scalability challenges, leveraging scalable visualization platforms and technologies is essential. Cloud-based solutions, distributed computing frameworks, and parallel processing techniques can be employed to handle large volumes of clinical data efficiently. These technologies enable the distribution of data processing and visualization tasks across multiple computing resources, resulting in faster and more scalable visualization capabilities.
- High-Performance Computing and Graphics Processing Units (GPUs): High-performance computing and GPUs offer substantial computational power that can handle complex data visualization tasks. GPUs excel at parallel processing and can significantly accelerate the rendering and visualization of large and complex datasets. By leveraging these technologies, clinical data management systems can achieve faster and more responsive visualization performance.
- Data Reduction and Sampling Techniques: When dealing with massive datasets, data reduction and sampling techniques can be employed to reduce data size while maintaining important patterns and characteristics. Techniques such as dimensionality reduction, clustering, or stratified sampling can be applied to obtain a representative subset of the data for visualization. This approach reduces the computational burden and enables faster visualization rendering.
- Progressive Data Loading and Visualization: Progressive data loading and visualization techniques allow for the initial display of partial data, with additional data being loaded and rendered progressively as the user interacts with the visualization. This approach enhances scalability by prioritizing the visualization of relevant and immediate data, while gradually incorporating additional data based on user interactions or data exploration needs.

These strategies enable efficient exploration, interpretation, and analysis of clinical data, supporting evidence-based decision-making and enhancing insights for healthcare professionals and researcher (Gotz and Stavropoulos, 2014).

B. Ensuring interpretability and accuracy of visualizations

Ensuring interpretability and accuracy considerations in data visualization, clinical data management systems can ensure that visualizations are reliable, trustworthy, and facilitate accurate insights. This enhances the quality of decision-making, research, and patient care based on the visualized clinical data. These aspects can be addressed:

- Clear and Intuitive Visual Design: Visualizations should be designed in a clear and intuitive manner, prioritizing the understanding of the information being presented. Avoiding clutter, using appropriate labelling, and providing clear legends and annotations help users interpret the visualizations accurately. Aesthetics and visual appeal should not compromise the clarity and interpretability of the data.
- Proper Context and Explanation: Visualizations should be accompanied by proper context and explanation to provide a comprehensive understanding of the data. Including relevant metadata, definitions of variables, units of

measurement, and clear captions or titles help users interpret the visualizations accurately. Additionally, explanations of data processing steps, statistical methods used, and any assumptions made ensure transparency and promote accurate interpretation.

- Data Validation and Quality Assurance: Before visualizing the data, thorough data validation and quality assurance processes should be implemented. This includes data cleaning, removing outliers, handling missing values, and ensuring data consistency. By addressing data quality issues, visualizations can accurately represent the underlying data and minimize potential misinterpretations due to erroneous or misleading information.
- Incorporation of Statistical Measures: Visualizations should incorporate appropriate statistical measures to support accurate interpretation. This can include measures such as means, medians, standard deviations, confidence intervals, or p-values. Adding these measures helps users understand the statistical significance and variability of the data, enabling them to make informed interpretations and draw reliable conclusions.
- User Interaction and Exploration: Interactive features in visualizations allow users to explore the data and gain a deeper understanding. Enabling users to interactively manipulate variables, apply filters, or drill down into subsets of the data promotes accurate interpretation. Users can explore different scenarios, test hypotheses, and validate findings, ensuring that visualizations are not treated as static representations but as dynamic tools for exploration and analysis.
- User Feedback and Iterative Design: Seeking feedback from users and stakeholders is essential for improving the interpretability and accuracy of visualizations. Regularly gathering user feedback on the clarity, understand ability, and accuracy of the visualizations helps identify areas for improvement. Iterative design processes that involve multiple rounds of user testing and refinement ensure that visualizations effectively convey the intended information and align with users' needs and expectations.
- Documentation and Reproducibility: Proper documentation of the visualization process, including data sources, pre-processing steps, visualization techniques used, and any transformations applied, is crucial for ensuring accuracy and reproducibility. Detailed documentation allows others to validate and replicate the visualizations, promoting transparency and facilitating trust in the findings.
- User Training and Education: Providing user training and education on data visualization principles, best practices, and common pitfalls enhances the accurate interpretation of visualizations. By educating users on how to read and interpret visualizations correctly, organizations can promote a culture of data literacy and enable stakeholders to make informed decisions based on the visualized data (AHRQ).

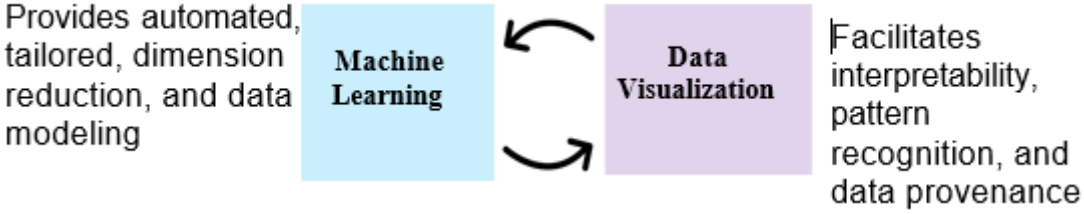
C. Incorporating advanced visualization techniques (e.g., machine learning-based visualization)

It's important to note that the adoption of machine learning-based visualization techniques in clinical data management should be accompanied by rigorous validation, transparency, and interpretability. It is crucial to ensure that the results and visualizations derived from these techniques are reliable, explainable, and align with domain expertise. Collaboration between data scientists, healthcare professionals, and researchers is vital to leverage the power of these advanced visualization techniques effectively and responsibly in the clinical setting. It can bring numerous benefits. Here's how these techniques can be applied:

- Enhanced Data Exploration: Machine learning-based visualization techniques can help healthcare professionals and researchers explore large and complex clinical datasets more effectively. By leveraging machine learning algorithms, these techniques can automatically identify patterns, clusters, and relationships within the data, enabling users to gain deeper insights and discover hidden information that may not be immediately apparent through traditional visualization approaches.
- Anomaly Detection: Machine learning-based visualization techniques can aid in the detection of anomalies or outliers in clinical data. By training machine learning models on normal patterns and behaviours, deviations from these patterns can be identified and visualized. This allows healthcare professionals to quickly identify abnormal patient conditions, potential errors in data entry, or other outliers that may require further investigation.
- Personalized Medicine and Treatment Planning: Machine learning-based visualization techniques can contribute to personalized medicine by analysing patient-specific data and visualizing personalized treatment plans. By integrating patient characteristics, genomic data, clinical records, and other relevant variables, these techniques can generate personalized visualizations that assist healthcare professionals in tailoring treatment strategies to individual patients' needs.
- Predictive Modelling and Forecasting: Machine learning-based visualization techniques can facilitate predictive modelling and forecasting in clinical data management. By training machine learning models on historical clinical data, these techniques can generate predictions and visualize future trends, outcomes, or risk factors. This enables healthcare professionals to make data-driven decisions and interventions based on anticipated scenarios.
- Decision Support Systems: Machine learning-based visualization techniques can be integrated into decision support systems in clinical data management. By combining machine learning

algorithms with visualizations, these systems can provide real-time insights, suggestions, and recommendations to healthcare professionals. This enhances clinical decision-making by presenting relevant information in an interpretable and actionable format.

- Image Analysis and Segmentation: Machine learning-based visualization techniques can be utilized for image analysis and segmentation in clinical data management. By training deep learning models on large image datasets, these techniques can automatically identify and segment specific features or abnormalities in medical images, such as tumours or lesions. The visualizations generated can assist healthcare professionals in accurate diagnosis, treatment planning, and monitoring of patients.
- Clinical Trials and Research: Machine learning-based visualization techniques can aid in the analysis and interpretation of clinical trial data. By visualizing patterns, treatment effects, and patient outcomes, researchers can gain insights into the effectiveness and safety of interventions. These techniques can also help identify patient subgroups that may benefit from specific treatments, enabling more targeted and efficient clinical trial design (MedTech Intelligence Staff, 2023).



D. Future trends and developments in data visualization for CDM

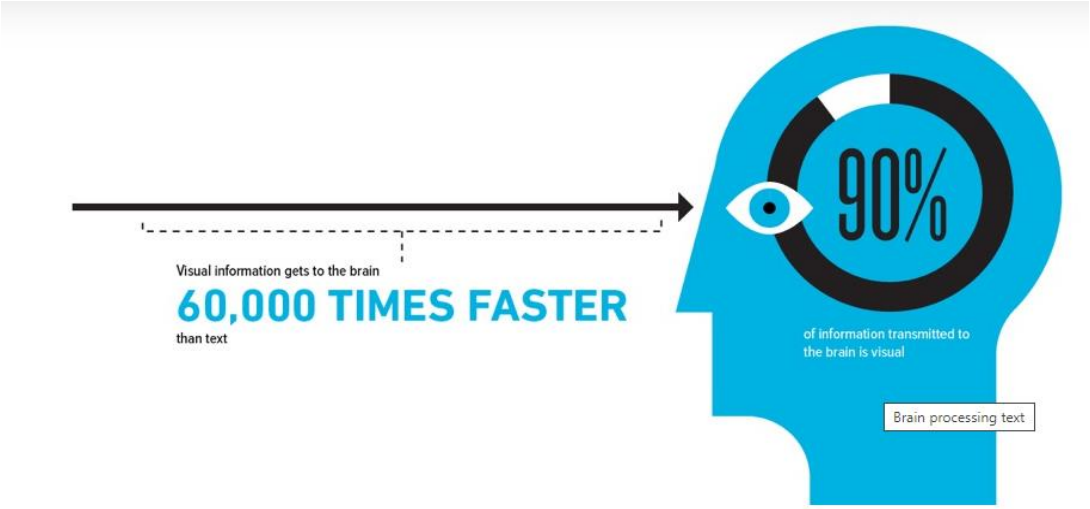
As technology advances and the volume and complexity of clinical data continue to grow, data visualization in clinical data management will evolve to address these challenges. These future trends and developments have the potential to enhance healthcare delivery, improve patient outcomes, and facilitate evidence-based decision-making in the clinical setting. Future trends and developments in data visualization for CDM are likely to focus on the following areas:

- Interactive and Immersive Visualization: The development of interactive and immersive visualization techniques will enable users to explore and interact with clinical data in more engaging and intuitive ways. Virtual reality (VR), augmented reality (AR), and mixed reality (MR) technologies may be leveraged to provide immersive visual experiences, allowing healthcare professionals to manipulate and analyse data in a three-dimensional and interactive environment.
- Real-Time and Streaming Visualization: As healthcare systems generate vast amounts of real-time data, the demand for real-time and streaming visualization techniques will increase. These techniques will enable healthcare professionals to monitor and analyse data as it is collected, allowing for timely interventions, and proactive decision-making, and immediate insights into patient conditions.
- Integration of Multiple Data Sources: The integration of diverse data sources, such as electronic health records, wearables, genetic data, and social determinants of health, will become more prevalent. Visualization techniques will need to handle the integration and visualization of these multi-modal data sources, enabling a comprehensive view of patient health and supporting personalized medicine.
- Explainable AI and Visual Analytics: There will be a growing emphasis on developing visualization techniques that provide explanations and insights into AI and machine learning models used in clinical data management. Explainable AI techniques will be integrated with visual analytics to provide interpretable and transparent results, enabling healthcare professionals to understand and trust the decisions made by AI models.
- Big Data Visualization: With the increasing availability of big data in healthcare, visualization techniques will need to handle and visualize large-scale and complex datasets. Scalable visualization platforms, distributed computing frameworks, and parallel processing techniques will be utilized to handle the volume, velocity, and variety of big data in clinical data management.
- Collaborative Visual Analytics: Collaboration will play a crucial role in clinical data management, and visual analytics techniques will support collaborative data exploration and decision-making. Collaborative visualization tools will enable multiple stakeholders, including healthcare professionals, researchers, and patients, to interactively explore data, share insights, and collaborate on data analysis and interpretation.
- Data Privacy and Security Visualization: As data privacy and security concerns continue to rise, visualization techniques will need to address these issues. Visualizations that convey the privacy and security status of data, highlight potential risks, and provide transparency regarding data usage and access will become increasingly important.
- Mobile and Portable Visualization: With the widespread use of mobile devices in healthcare

settings, there will be a need for visualization techniques that are optimized for mobile platforms. Mobile-friendly visualizations will enable healthcare professionals to access and analyse data on the go, supporting point-of-care decision-making and enhancing mobility in clinical data management.

- Natural Language Processing and Text Visualization: Clinical data includes a significant amount of unstructured text, such as clinical notes, research papers, and patient narratives. Visualization techniques that integrate natural language processing (NLP) and text mining capabilities will enable the visualization of text-based data, extracting key insights, and supporting text-based exploration and analysis.
- Ethical and Social Impact Visualization: Visualization techniques will address the ethical and social impact of clinical data management. They will provide visualizations that help stakeholders understand the potential biases, fairness considerations, and ethical implications associated with the collection, analysis, and use of healthcare data.

As technology advances and the volume and complexity of clinical data continue to grow, data visualization in clinical data management will evolve to address these challenges. These future trends and developments have the potential to enhance healthcare delivery, improve patient outcomes, and facilitate evidence-based decision-making in the clinical setting.



(Image source: killervisualstratergies.com)

REFERENCES

1. Expeed (2022), Storytelling Wit Data Using Visualization to Help Stakeholders Understand Data Findings https://expeed.com/blog-posts/storytelling-with-data-using-visualizations-to-help-stakeholders-understand-data-findings/
2. Gotz, D., & Stavropoulos, H. Decisionflow: Visual analytics for high-dimensional temporal event sequence data. IEEE transactions on visualization and computer graphics, 2014; 20(12): 1783-1792.
3. Kaisler, S. Armour, F. Espinosa, J. A. and Money, W. Big data: Issues and challenges moving forward In *Hawaii International Conference on System Sciences*, HICSS 2013, pages 995-1004,IEEE, 2013.
4. Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. Empirical studies in information visualization: Seven scenarios. IEEE transactions on visualization and computer graphics, 2011; 18(9): 1520-1536.
5. Learning Health Systems. Agency Healthc. Res. Qual. https://www.ahrq.gov/professionals/systems/learning-health-systems/index.html.
6. Lesselroth, B. J., & Pieczkiewicz, D. S. *Data visualization strategies for the electronic health record*. Nova Science Publishers, Inc., 2011.
7. Lin, Y. L., Guerguerian, A. M., Tomasi, J., Laussen, P., & Trbovich, P. Usability of data integration and visualization software for multidisciplinary paediatric intensive care: human factors approach to assessing technology. BMC medical informatics and decision making, 2017; 17(1): 122.
8. [GB13] GAO B., BERENDT B.: Circles, posts and privacy in egocentric social networks: An exploratory visualization approach. In 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013), 2013; IEEE, 792–796.
9. [GB16] GOTZ D., BORLAND D.: Data-driven healthcare: Challenges and opportunities for interactive visualization. IEEE computer graphics and applications, 2016; 36(3): 90–96. 9.

10. [GDLS14] GKOULALAS-DIVANIS A., LOUKIDES G., SUN J.: Publishing data from electronic health records while preserving privacy: A survey of algorithms. Journal of biomedical informatics, 2014; 50: 4–19. 2, 4, 9 [GH20] GEORGE HEILNEIJER D.: The Heilmeijer Catechism, 1977 (accessed January 23, 2020).
11. URL: https://www.darpa.mil/ work-with-us/heilmeier-catechism. 13 [GHK∗16] GABOARDI M., HONAKER J., KING G., MURTAGH J., NISSIM K., ULLMAN J., VADHAN S.: Psi ({\Psi}): a private data sharing interface. arXiv preprint arXiv:1609.04340 (2016).
12. MedTech Intelligence Staff (2023), FDA Releases Two Discussion Papers on Artificial Intelligence and Machine Learning in Drug Development and Manufacturing. https://www.medtechintelligence.com/news_article/fda-releases-two-discussion-papers-on-artificial-intelligence-and-machine-learning-in-drug-development-and-manufacturing/?hsamp_network=linkedin&hsamp=bb0RgS2Zi2aE
13. NCBI (2022), Visualization Techniques in Healthcare Applications: A Narrative Review https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9741729/
14. Rind, A., Wang, T. D., Aigner, W., Miksch, S., Wongsuphasawat, K., Plaisant, C., & Shneiderman, B. Interactive information visualization to explore and query electronic health records. *Foundations and Trends® in Human–Computer Interaction*, 2013; *5*(3): 207-298. Retrieved July 4, 2019, http://www.cs.umd.edu/hcil/trs/2010-19/2010-19.pdf
15. Springer Open (2022), Big data in healthcare: management, analysis and future prospects https://journalofbigdata.springeropen.com/articles/10.11