SOCIAL MEDIA AND WEB ANALYTICS

Dr. M. Soundarya | Dr. P. Hameem Khan
Mrs. R. Geetha | Dr. Kalidhass

SOCIAL MEDIA AND WEB ANALYTICS

Authors

Dr. M. Soundarya
Dr. P. Hameem Khan
Dr. Geetha
Dr. Kalidhass

FIRST BOOK
PUBLISHING

Title: Social Media and Web Analytics

Author's Name: Dr. M. Soundarya
Dr. P. Hameem Khan
Dr. Geetha
Dr. Kalidhass

Published by: First Book Publishing
Madurai – 625 011.

Publisher's Address: First Book Publishing
82, Anna Main Street,
Jaihindupuram,
Madurai – 625 011.

Edition Details (I,II,III): I

ISBN: 978-93-92151-74-3

Month & Year: April, 2025

Pages: 125

SYLLABUS

UNIT 1 Web and Social Media

Introduction - Web and social media - Web sites, web apps, mobile apps and social media - Usability, user experience, customer experience, customer sentiments, web marketing, conversion rates, ROI, brand reputation, competitive advantages - Web analytics and a Web analytics 2.0 framework - clickstream, multiple outcomes analysis, experimentation and testing, voice of customer, competitive intelligence, Insights.

UNIT 2 DATA

Data - Structured data, unstructured data, metadata, Big Data and Linked Data. - Lab testing and experiment design: selecting participants, within-subjects or between subjects study, counterbalancing, independent and dependent variable; A/B testing, multivariate testing, controlled experiments. - Data analysis basics: types of data, metrics and data, descriptive statistics, comparing means, correlations, nonparametric tests, presenting data graphically.

UNIT 3 Usability Metrics

Usability metrics: performance metrics, issues-based metrics, self-reported metrics - Planning and performing a usability study: study goals, user goals, metrics and evaluation methods, participants, data collection, data analysis. - Typical types of usability studies and their corresponding metrics: comparing alternative designs, comparing with competition, completing a task or transaction, evaluating the impact of subtle changes.

UNIT 4 Web Metrics and Web Analytics

Web metrics and web analytics: PULSE metrics - active users on business and technical issues; - HEART metrics - success on user behaviour issues; - On-site web analytics: off-site web analytics, the goal-signal-metric process. - Social media analytics: social media KPIs- Performing social media analytics: business goal, data gathering, analysis, measure and feedback.

UNIT 5 Data Analysis Language and Tools

Data analysis language and tools - Ready-made tools for Web and social media analytics - Key Google Analytics metrics, dashboard, social reports - Statistical programming language (R) - its graphical development environment for data exploration and analysis, and its social media analysis packages.

CONTENTS

Unit No.	Title	Page No.
1	Social Media Web Analytics	1
2	Data	40
3	Usability Metrics	69
4	Web Metrics and Web Analytics	99
5	Data Analysis Language and Tools	112

UNIT – I

SOCIAL MEDIA WEB ANALYTICS

1.1 Evolution of Online Communities

Online communities are groups of individuals who interact and engage with each other over the internet, typically sharing common interests, goals, or activities. These communities can take various forms and serve a wide range of purposes. The evolution of online communities is a fascinating journey that has unfolded over several decades, shaped by technological advancements, social dynamics, and the changing needs and interests of internet users. Here is a brief overview of the key stages in the evolution of online communities:

[1] Bulletin Board Systems (BBS)

In the late 1970s and 1980s, Bulletin Board Systems (BBS) were one of the earliest forms of online communities. Users could connect to BBS using their modems and engage in text-based discussions, share files, and play text-based games. BBSs were typically run by individuals or small groups and catered to niche interests.

[2] Usenet Newsgroups

Usenet, established in the late 1970s, consisted of a vast collection of discussion forums organized into newsgroups. It allowed users to post and read text-based messages on various topics. Usenet was decentralized and distributed across multiple servers. Internet Forums: Internet forums, also known as message boards or discussion boards, gained popularity in the 1990s. They offered a user-friendly interface for discussions on various topics, often organized into categories and sub-forums. Popular platforms like phpBB and Bulletin emerged during this period.

[3] Social Networking Sites

The early 2000s witnessed the rise of social networking sites like Friendster, MySpace, and eventually Facebook. These platforms allowed users to connect with friends, share personal information, and engage in discussions. They represented a shift toward more personal and interconnected online communities. Online Gaming Communities: Online gaming communities began to grow with the popularity of multiplayer online games like World of War craft and Second Life. Gamers could form groups, guilds, and clans to collaborate in virtual worlds.

[4] Web 2.0 and Social Media

The mid-2000s marked the advent of Web 2.0, characterized by user-generated content and interactivity. Platforms like YouTube, Twitter, and Reddit allowed users to create, share, and discuss content on a massive scale, fostering large and diverse online communities.

[5] Specialized Niche Communities

As the internet matured, there was a surge in specialized niche communities. Websites and forums dedicated to specific interests, hobbies, and professions became more prominent. Examples include Stack Overflow for programming and GitHub for software development.

[6] Mobile and Messaging Apps

The rise of smartphones and mobile apps led to the growth of online communities within apps like WhatsApp, Snapchat, and Slack. These platforms facilitated real-time communication and group collaboration.

[7] Online Marketplaces and Review Sites

Online marketplaces like Airbnb and Etsy and review sites like Yelp and TripAdvisor integrated community aspects by allowing users to leave reviews, rate products, and connect with sellers and other customers.

Decentralized and Blockchain-Based Communities: More recently, blockchain technology has enabled the creation of decentralized communities and social platforms, often with a focus on user privacy and ownership of data.

[8] Virtual Reality (VR) Communities

With advancements in VR technology, virtual reality communities have emerged, allowing users to interact in immersive digital environments. The evolution of online communities continues, adapting to changing technology and user preferences. Emerging trends include decentralized social networks, AI-driven communities, and an increasing emphasis on user privacy and data ownership. Online communities have become an integral part of our digital lives, providing spaces for connection, collaboration, and the exchange of knowledge and experience.

1.2 Introduction of Social Media

"Social media refers to a set of online platforms and technologies that enable individuals and communities to create, share, and interact with content, information, and each other". These platforms are designed for social interaction and often involve the exchange of user-generated content.

Social media has become an integral part of modern communication and has transformed the way people connect, communicate, and share in both personal and professional contexts. The history and evolution of social media is a captivating journey that spans several decades, characterized by technological advancements and changes in how people connect, communicate, and share information. Here is a brief overview of the key milestones in the history of social media:

1.2.1 Arpanet and Bulletin Board Systems (BBS) (1960s-1970s)

The precursor to the internet, Arpanet, was used to exchange messages and share information among academics and researchers. Bulletin Board Systems (BBS) allowed users to post messages, share files, and communicate via dial-up modems.

1.2.2 Usenet (1980s)

Usenet newsgroups facilitated discussion and file sharing among early internet users. It was one of the first instances of online community interaction.

1.2.3 Six Degrees (1997)

Six Degrees, often considered the first social media platform, allowed users to create profiles and connect with friends. It laid the foundation for the concept of an online social network.

1.2.4 Friendster (2002)

Friendster was one of the earliest social networking sites that enabled users to connect with friends, post photos, and write testimonials. It had millions of users but eventually declined due to technical issues.

1.2.5 MySpace (2003)

MySpace gained tremendous popularity, especially among young people, and allowed users to create customizable profiles, connect with others, and share music and media. It was eventually overtaken by Facebook.

1.2.6 Facebook (2004)

Founded by Mark Zuckerberg and his colleagues, Facebook rapidly became one of the world's largest social networks. It initially targeted college students and gradually opened to a wider audience. Facebook introduced the News Feed, which transformed how users consumed content.

1.2.7 YouTube (2005)

YouTube revolutionized the way people shared and consumed video content, enabling users to upload, view, and share videos on a massive scale.

1.2.8 Twitter (2006)

Twitter introduced the concept of micro blogging, allowing users to post short messages (tweets) to a global audience. It played a significant role in real-time communication and news sharing.

1.2.9 IPhone and Mobile Social Media (2007)

The release of the iPhone and other smartphones led to a surge in mobile app development, including social media apps. This made social media more accessible and immediate.

1.2.10 Instagram (2010)

Instagram, a photo-sharing platform, gained rapid popularity for its simplicity and visual focus, and it was eventually acquired by Facebook.

1.2.11 Pinterest (2010)

Pinterest allowed users to discover and save visual content, primarily in the form of images and ideas.

1.2.12 Snapchat (2011)

Snapchat introduced ephemeral messaging, where messages and media disappear after being viewed. It appealed to a younger demographic.

1.2.13 WhatsApp (2011)

WhatsApp provided a secure and convenient platform for instant messaging and phone calls over the internet. It became particularly popular for international communication.

1.2.14 TikTok (2016)

TikTok (formerly Musical.ly) became a sensation for short-form video content, with users creating and sharing entertaining videos set to music. Evolving Trends: Recent trends in social media include live streaming, influencer marketing, stories (ephemeral content), augmented reality (AR) filters, and increased emphasis on privacy and data protection.

The history of social media reflects the dynamic nature of technology and society, with platforms constantly evolving to meet changing user needs and preferences. Social media has not only transformed how people interact but also had significant impacts on politics, marketing, journalism, and many other aspects of modern life.

1.3 Websites

A website can be identified by a common domain name and it is also considered as the collection of the webpages. The published pages belong to one server. The websites can provide us with all the necessary information needed for people across the globe within seconds. The major websites are google.com, wikipedia.org, Entri.com, etc. All the websites which are publicly accessible together comprise the World Wide Web. The private networks can access private websites.

In 1989, World Wide Web (WWW) was developed by Tim Berners-Lee, the British CERN physicist. The world Wide Web was later announced to be used free by anyone across the globe turned out to be a decisive action in the history of the website. the announcement was commenced on 30 April 1993. This led to the growth of the web. To retrieve individual files from the server, initially, File Transfer Protocol and the gopher protocol were used and is later replaced with Hypertext Transfer Protocol (HTTP).

1.3.1 Types of Websites

Type	Purpose	Examples
Informational Websites	To provide information or educate visitors	Government websites, blogs, news portals.
E-Commerce Websites	To facilitate online buying and selling of goods and services.	Amazon, Flip kart, Etsy
Business or Corporate Websites	To represent a company, showcase services/products, and establish credibility.	Coca-Cola, TCS, IBM.
Portfolio Websites	To showcase creative work or personal achievements	Artist portfolios, designer showcases, personal brands
Blogs	To share ideas, opinions, expertise, or personal experiences	Travel blogs, food blogs, tech blogs
Educational Websites	To provide learning resources and educational tools	Khan Academy, Coursera, Duolingo.
Entertainment Websites	To entertain users through media content like videos, music, and games	YouTube, Netflix, Spotify.
Social Media Websites	To connect people and foster communication and content sharing.	Facebook, Instagram, Twitter.
Non-profit or Charity Websites	To promote causes, raise awareness, and collect donations.	UNICEF, Red Cross, WWF
Personal Websites	To create an online presence for an individual, often as a digital resume or personal blog.	John Doe's resume site, personal travel diaries.

Community and Forum Websites	To foster discussions among users on specific topics	Reddit, Quora, Stack Overflow.
Media or Magazine Websites	To publish news, articles, or multimedia content.	Forbes, National Geographic.
Search Engine Websites	To help users find information online.	Google, Bing, DuckDuckGo.

1.3.2 WEB APP

A web app is a software application that is accessed through a web browser and operates over the internet. Unlike traditional desktop software, a web app does not require installation on a device, making it easily accessible from anywhere with an internet connection. It is typically built using web technologies such as HTML, CSS, and JavaScript, often coupled with modern frameworks like React, Angular, or Vue.js for a dynamic and interactive user experience.

Web apps are designed to be platform-independent, meaning they can run on any operating system or device with a web browser, including desktops, tablets, and smartphones. They are hosted on web servers, allowing developers to update them centrally; ensuring users always access the latest version without the need for manual updates.

Examples of popular web apps include **Google Docs, online banking portals, and social media platforms like Twitter**.

While most web apps require an internet connection, some can function offline with advanced features like service workers. Their versatility, ease of deployment, and cross-platform compatibility make them a popular choice for businesses and developers.

1.3.3 Mobile App

A mobile app is a software application specifically designed to run on mobile devices such as smartphones and tablets.

These apps are developed to provide a tailored user experience, leveraging the unique features of mobile devices, such as touch screens, GPS, cameras, and sensors. Mobile apps are typically platform-specific, with separate versions for Android and IOS, built using native programming languages like Kotlin or Swift.

However, cross-platform development frameworks like Flutter and React Native enable apps to function across multiple platforms using a single codebase. Mobile apps are downloaded and installed from app stores, such as the Google Play Store or Apple App Store, and often require periodic updates for new features or bug fixes. Many mobile apps can function offline, offering a seamless experience without constant internet access. They are widely used for a variety of purposes, including social media **(e.g., Instagram), communication (e.g., WhatsApp),** gaming, e-commerce, and productivity tools. Their ability to integrate deeply with device hardware and provide high-performance, responsive experiences makes mobile apps indispensable in today's digital ecosystem.

Comparison Table of Mobile App and Web APP

Feature	Web App	Mobile App
Installation	No installation required	Requires installation
Platform	Browser-based, universal	Platform-specific
Access	Internet-dependent	May work offline
Updates	Automatic (server-side)	User-triggered via updates
Performance	Slower, dependent on browser	Faster, uses device features
Development	Web technologies	Native or hybrid tools

1.3.4 Mobile App and Social Media

A Mobile app is a software application designed specifically for mobile devices like smartphones and tablets. Social media platforms, a subset of mobile apps, are among the most widely used applications globally. Social media mobile apps are created to facilitate communication, content sharing, and community building, allowing users to connect, interact, and share media in real-time.

Apps like Instagram, Facebook, TikTok, and Twitter (X) have revolutionized how people engage with social networks, offering features like instant messaging, multimedia uploads, live streaming, and push notifications to keep users updated. These apps often leverage the capabilities of mobile devices, such as cameras for capturing photos and videos, GPS for location tagging, and sensors for interactive experiences like augmented reality filters.

Social media mobile apps are optimized for ease of use, offering responsive designs and personalized feeds powered by algorithms to enhance user engagement. They often include integration with other apps and platforms, making it easier to share content and manage accounts. The portability and accessibility of these apps on mobile devices have significantly contributed to the global rise of social networking, enabling users to stay connected anytime, anywhere.

Usability in Social Media Web Analytics

1. **Ease of Navigation**: Users should be able to intuitively find key metrics (e.g., engagement rates, impressions, or audience demographics) without extensive searching or training.
   Example: A dashboard with clear menus for filtering data by platform, time frame, or campaign type.
2. **Clarity of Information**: Data visualizations like charts, graphs, and tables should be easy to interpret, avoiding unnecessary complexity or jargon.

Example: Bar charts showing weekly follower growth or pie charts depicting audience distribution by age group.

3. **Efficient Reporting**: Generating and exporting reports should require minimal steps and support common formats like PDF, CSV, or Excel.
   Example: A "one-click export" feature to download analytics summaries.
4. **Error Recovery**: Users should be able to correct mistakes, such as choosing the wrong date range or filters, without losing progress.
   Example: "Reset filter" and "Undo" options on the interface.
5. **Responsiveness**: The platform should load quickly and adapt seamlessly to different screen sizes and devices.

1.3.6 User Experience in Social Media Web Analytics

1.3.6.1 Personalization

Providing users with customized dashboards based on their goals, such as monitoring campaign performance or audience engagement.

Example: Widgets that allow users to drag-and-drop their most relevant metrics into a personal view.

1.3.6.2 Actionable Insights

UX goes beyond presenting raw data by offering meaningful insights and recommendations.

Example: Highlighting posts with the highest engagement and suggesting optimal posting times.

1.3.6.3 Aesthetic Design

A visually appealing interface with modern typography, color-coded data, and meaningful icons enhances the overall experience.

Example: Using green and red to indicate positive or negative trends in performance metrics.

1.3.6.4 Feedback and Guidance

Providing interactive elements that guide users through the analytics platform.

Example: Tooltips explaining metrics like "CTR" (click-through rate) or a walkthrough for new users.

1.3.6.5 Emotional Connection

The platform should evoke trust and reliability by presenting accurate data and maintaining a polished, professional appearance.

Example: Clear branding and intuitive error messages that reassure users about data accuracy.

1.3.7 Customer Experience (CX) in Social Media Analytics

Customer experience refers to the overall perception customers have of a brand based on their interactions across various touch points, including social media. In social media web analytics, CX focuses on understanding how users interact with a brand and how those interactions impact satisfaction and loyalty.

1.3.7 Key Metrics for CX in Social Media Analytics

1. **Engagement Rates**: Likes, comments, shares, and clicks indicate how customers interact with your content.
   Higher engagement suggests positive CX, while low engagement may indicate disinterest or poor content alignment.
2. **Response Times**: How quickly a brand replies to customer queries or concerns on social media.
   Faster response times lead to better CX and increased customer satisfaction.
3. **Content Performance**: Monitoring which types of posts (videos, polls, images) resonate most with the audience.
   Insights can guide content strategies to enhance CX.

4. **Audience Growth**: Tracking follower growth or decline helps assess how your brand is perceived over time.
5. **Customer Journey Mapping**: Understanding how customers move through touch points on social media from discovery to interaction and purchase.

1.3.8 Customer Sentiments in Social Media Analytics

Customer sentiment refers to the emotions or attitudes that customers express about a brand on social media, ranging from positive to negative. Sentiment analysis, often powered by AI and natural language processing (NLP), helps businesses measure these emotions at scale.

1.3.8.1 How Sentiments are Analysed

1. **Sentiment Scores**: Text analysis assigns scores to posts or comments, categorizing them as positive, negative, or neutral.
   Example: "I love this product!" is positive; "Terrible service!" is negative.
2. **Emotion Detection**: Advanced tools identify emotions like happiness, anger, or frustration in customer posts or reviews.
   Example: Identifying trends of anger during a product recall.
3. **Keyword Tracking**: Monitoring keywords associated with the brand, competitors, or industry to gauge sentiment.
   Example: Tracking words like “excellent,” “broken,” or “refund” in customer comments.
4. **Hashtag Sentiment**: Evaluating the sentiment of hashtags associated with campaigns or events.
   Example: Sentiment trends for #NewLaunch or #BrandFail.

1.3.8.2 Integrating CX and Sentiments in Social Media Analytics

1. **Real-Time Monitoring**: Tools like Hootsuite or Sprout Social offer real-time dashboards for tracking customer interactions and sentiments, enabling businesses to respond promptly.
   Example: Addressing negative sentiment during a product launch to mitigate backlash.
2. **Trend Analysis**: Identifying shifts in sentiment over time to evaluate how campaigns, product updates, or incidents affect customer experience.
   Example: A spike in positive sentiment after announcing a discount.
3. **Feedback Loop**: Use sentiment insights to refine CX strategies. Positive feedback can inspire more of what customers love, while addressing negative feedback improves the overall experience.
   Example: Customers praise faster delivery; the brand highlights this in future promotions.
4. **Personalized Engagement**: Combining sentiment analysis with CX metrics allows brands to engage with customers in meaningful ways.
   Example: Thanking customers for positive reviews or addressing complaints with tailored solutions.

1.3.9 Web Marketing

Web Marketing**, also known as** online marketing **or** digital marketing**,** involves leveraging internet-based platforms and strategies to promote products, services, or brands to a targeted audience. It encompasses a wide range of activities aimed at driving traffic, engaging users, and converting them into customers

Key Components of Web Marketing

1. **Search Engine Optimization (SEO)**
   - The process of optimizing a website to rank higher on search engine results pages (SERPs) for relevant keywords.
   - Techniques include on-page optimization (content, meta tags, keywords) and off-page optimization (backlinks, social signals).
   - Example: Writing blog posts targeting specific keywords to attract organic traffic.
2. **Content Marketing**
   - Creating and distributing valuable, relevant, and consistent content to attract and retain a target audience.
   - Formats include blogs, videos, infographics, podcasts, and eBooks.
   - Example: A software company publishing tutorials and case studies to educate and engage potential clients.
3. **Pay-Per-Click (PPC) Advertising**
   - Paid advertisements displayed on search engines, social media, or other websites where advertisers pay per click.
   - Platforms include Google Ads and Bing Ads.
   - Example: Running a PPC campaign for a new product launch to appear at the top of search results.
4. **Social Media Marketing**
   - Promoting products or services on social platforms like Facebook, Instagram, LinkedIn, Twitter, and TikTok.
   - Activities include creating organic content, running ads, and engaging with followers.
   - Example: A brand using Instagram Reels to showcase its products.
5. **Email Marketing**
   - Sending targeted messages to a list of subscribers to nurture relationships, promote products, and drive conversions.

- Includes newsletters, promotional emails, and personalized campaigns.
- Example: A retailer sending discount codes to subscribers during holiday sales.

6. **Affiliate Marketing**
   - Partnering with affiliates who promote a business's products or services in exchange for a commission on sales generated through their referral links.
   - Example: Influencers sharing affiliate links to products on their blogs or social media.
7. **Influencer Marketing**
   - Collaborating with influencers to reach their followers and promote a brand authentically.
   - Example: A beauty brand partnering with makeup artists on YouTube for product reviews.
8. **Web Analytics**
   - Using tools like Google Analytics to track website traffic, user behavior, and campaign performance.
   - Helps marketers refine strategies and optimize conversions.
   - Example: Identifying high-performing pages and improving underperforming ones.
9. **Conversion Rate Optimization (CRO)**
   - Improving website elements to increase the percentage of visitors who take desired actions, such as making a purchase or signing up for a newsletter.
   - Example: Simplifying a checkout process to reduce cart abandonment.
10. **Mobile Marketing**
    - Optimizing marketing strategies for mobile users through responsive websites, SMS campaigns, and mobile apps.
    - Example: Sending push notifications through a mobile app about limited-time offers.

1.3.10 Conversion Rates

Conversion rates in **social media web analytics** measure the percentage of users who complete a desired action after interacting with a social media campaign or content. These actions could include making a purchase, signing up for a newsletter, downloading an app, or any other goal set by the brand.

Conversion rate is calculated as:

Conversion Rate (%) = (Number of Conversions/ Total Interactions or Clicks)× 100

For example, if a social media campaign receives 1,000 clicks and 50 people complete the desired action, the conversion rate is:

Conversion Rate=(50/1000)×100=5%

Types of Conversions in Social Media

1. **Website Conversions**:

   Actions like purchases, form submissions, or account registrations triggered by social media traffic.

2. **Lead Conversions**

   Collection of user information such as email addresses or phone numbers for future marketing.

3. **Engagement Conversions**

   Actions such as downloading an eBook, watching a video, or participating in a webinar.

4. **App Conversions**

   Installs, sign-ins, or in-app purchases resulting from social media campaigns.

5. **Physical Conversions**

   Offline actions like visiting a store or attending an event based on social media promotions.

1.3.11 Return on Investment

ROI (Return on Investment) **in** social media analytics measures the effectiveness of social media marketing efforts in generating financial returns relative to the costs involved. It helps businesses determine whether their social media activities are worth the investment, such as advertising spend, content creation costs, and team resources.

The basic formula for ROI is:

ROI=Net Profit /Cost of Investment×100

Where:

- **Net Profit** = Revenue generated from social media efforts – Total costs
- **Cost of Investment** = The total cost associated with social media activities (e.g., ad spend, content creation, tools, staff salaries).

ROI (Return on Investment) in **social media analytics** measures the effectiveness of social media marketing efforts in generating financial returns relative to the costs involved. It helps businesses determine whether their social media activities are worth the investment, such as advertising spend, content creation costs, and team resources.

Steps to Calculate Social Media ROI

1. **Set Clear Objectives**: Identify the goals you aim to achieve through social media (e.g., lead generation, sales, brand awareness).
   Example: Increase sales by 20% through Instagram ads.
2. **Track Revenue or Conversion Data**: Use tools like Google Analytics, Facebook Ads Manager, or CRM systems to track revenue or conversions generated directly from social media campaigns.
   Example: Track how many people purchase a product after clicking an Instagram ad.

3. **Determine Social Media Costs**: Calculate the total investment in the social media campaign, including ad spend, content creation, management tools, and personnel costs.
   Example: $5,000 spent on Facebook ads and $2,000 for influencer collaborations.
4. **Calculate Net Profit**: Subtract the costs of investment from the revenue generated through social media activities.
   Example: If the revenue generated from the campaign is $15,000, then the net profit is $15,000 - $7,000 = $8,000.

1.3.12 Brand Reputation

Brand Reputation refers to the perception and overall image that customers, stakeholders, and the public have about a company or its products and services. It is built over time through consistent actions, communication, and experiences, and can significantly influence consumer behaviour, loyalty, and business success.

Factors Influencing Brand Reputation

1. **Quality of Products or Services**
   - Offering high-quality products or services is the cornerstone of a positive brand reputation. Poor product quality can quickly damage consumer trust.
   - Example: A smartphone company that consistently releases well-designed, reliable products builds a strong reputation over time.
2. **Customer Service**
   - The way a company interacts with customers and addresses complaints plays a key role in shaping its reputation.
   - Example: A company with responsive, helpful customer service will be perceived positively, even when issues arise.

3. **Corporate Social Responsibility (CSR)**
   - A brand's involvement in social, environmental, and community initiatives can enhance its reputation.
   - Example: Brands like Patagonia and Ben & Jerry's have built strong reputations through environmental and social activism.
4. **Public Relations and Media Coverage**
   - How a brand is portrayed in the media and its ability to manage public relations crises can impact its reputation.
   - Example: A company handling a public crisis with transparency and responsibility is often viewed more favourably.
5. **Online Reviews and Social Media Presence**
   - Customer reviews, ratings, and online discussions significantly influence brand reputation in the digital age.
   - Example: Positive reviews on platforms like Yelp or Trustpilot can improve a brand's image, while negative reviews can harm it.
6. **Consistency in Communication**
   - A brand's messaging and values should be consistent across all channels, including advertising, social media, and customer interactions.
   - Example: A luxury brand that consistently delivers a premium customer experience across all touch points strengthens its reputation.
7. **Brand Transparency and Authenticity**
   - Brands that are open about their business practices, product sourcing, and mistakes tend to build stronger reputations.
   - Example: A company that admits product flaws and works to improve them can earn consumer respect for its honesty.

1.3.13 Competitive Advantages

1. **Understanding Consumer Behaviour**

Insight into Preferences and Needs: Social media analytics allows businesses to track and understand customer preferences, behaviours, and pain points by monitoring likes, comments, shares, mentions, and hash tags. These insights help brands tailor their offerings to meet customer needs.

Example: A clothing brand can analyse the social media posts of its followers to identify popular fashion trends and adjust its inventory or marketing campaigns accordingly.

2. **Targeted and Personalized Marketing**

Precise Audience Segmentation: By analysing data from social media platforms, brands can identify different customer segments based on demographics, interests, online behaviour, and more. This allows for highly targeted campaigns that are more likely to resonate with specific audience groups.

Example: A company selling athletic shoes can use social media analytics to segment its audience by age, location, and activity level, and then run personalized ads that appeal directly to each group, increasing the likelihood of conversions.

3. **Real-Time Performance Tracking**

Agility in Strategy Adjustments: Social media analytics provides real-time insights into campaign performance, allowing businesses to track metrics such as engagement rates, click-through rates, and conversions. This enables brands to quickly adjust strategies, optimize content, and maximize the effectiveness of their campaigns.

Example: If a social media ad is underperforming, brands can quickly tweak the messaging, target audience, or visuals based on performance data, ensuring that they don't waste budget on ineffective campaigns.

4. **Identifying Trends and Opportunities**

Trend Analysis: By analysing social media conversations, hashtags, and influencer activity, businesses can stay ahead of trends, identify emerging consumer preferences, and capitalize on market shifts faster than competitors.

Example: A brand can use social listening tools to detect a rise in interest for sustainable fashion and introduce an eco-friendly product line ahead of competitors, positioning itself as an industry leader.

5. **Competitive Benchmarking**

Competitor Analysis: Social media web analytics allows businesses to track and monitor competitor activity on social platforms. By analyzing competitors' engagement, content strategies, and sentiment, businesses can identify areas where they are outperforming or areas requiring improvement.

Example: A business can track its competitor's social media campaigns to learn which type of content resonates most with their audience, and then apply similar strategies to boost their own social media engagement.

6. **Enhanced Customer Engagement**

Interactive and Real-Time Engagement: Social media platforms offer brands the opportunity to engage directly with customers in real time. Analytics can track sentiment, engagement rates, and customer feedback to ensure that a brand is responding appropriately to customer concerns, questions, and comments, creating better brand loyalty.

Example: If customers frequently mention a particular feature of a product in their posts, brands can use this feedback to improve products or respond to customer concerns, strengthening relationships with their audience.

7. **Improving Brand Reputation**

Sentiment Analysis: Social media analytics tools can track mentions of a brand and analyze customer sentiment (positive, negative, or neutral).

By monitoring sentiment in real time, businesses can quickly address negative feedback or capitalize on positive sentiment.

Example: If a brand receives negative feedback about a product on social media, the company can respond quickly and mitigate any damage to its reputation, potentially turning a dissatisfied customer into a loyal one.

1.3.14 Web Analytics

Definition: Web analytics is the measurement, collection, analysis, and reporting of Internet data for the purposes of understanding and optimizing Web usage. (Source: Web Analytics Association)

Importance of Web Analytics

Today, Web interactions between commercial businesses and their customers take place via e-commerce stores, customer service sites, interactive real-time chat, e-mail, and social media streams. These interactions are as important, if not more so, for a business's growth as customer touches through traditional voice and bricks-and-mortar channels

- Web data integrated with other channels provides a better picture of the customer–business relationship and helps in identifying customer trends.
- It is also useful in assessing the effectiveness of marketing campaigns and optimizing marketing spend.
- And it improves the customer experience through faster service, thereby driving business growth and enhancing reputation

1.3.15 Web Analytics 2.0

Web Analytics 2.0 is an advanced, modern approach to web analytics that goes beyond the traditional tracking of website traffic and basic metrics. It focuses on the integration of data from multiple sources, deeper insights into customer behaviour, and advanced strategies for improving user engagement and

optimizing digital marketing. Web Analytics 2.0 aims to offer a more comprehensive understanding of how users interact with websites, apps, and digital content, enabling businesses to make data-driven decisions that improve overall performance.

Key Characteristics of Web Analytics 2.0

1. Data Integration

Web Analytics 2.0 integrates data not just from website traffic but from multiple sources, including social media platforms, mobile apps, email campaigns, and offline touch points. This holistic approach provides a more complete view of customer behaviour.

Example: A business might integrate data from its website, social media platforms (Facebook, Twitter, etc.), email marketing campaigns, and sales databases to understand how different channels contribute to conversions.

2. Customer-Centric Metrics

Web Analytics 2.0 shifts focus from just tracking basic traffic metrics (such as page views) to understanding user behaviour, engagement, and conversion paths. This involves tracking metrics such as customer lifetime value (CLV), user interactions, and the effectiveness of marketing campaigns across different channels.

Example: Instead of merely measuring how many people visited a product page, Web Analytics 2.0 might track how many visitors actually added the product to their cart, initiated checkout, and completed a purchase, providing a more complete understanding of user behaviour.

3. Advanced Segmentation

Web Analytics 2.0 enables businesses to segment users more effectively based on behaviour, demographics, location, referral sources, or even predictive models. These segments can then be targeted with personalized marketing efforts or product recommendations.

Example: A retailer may segment visitors into groups like first-time visitors, returning customers, and cart abandoners, and then deliver personalized content or ads to each group based on their behaviour.

4. Real-Time Analytics

Unlike traditional web analytics, which often focuses on historical data, Web Analytics 2.0 provides real-time insights into customer actions. This helps businesses to quickly adjust their strategies, optimize content, and respond to customer behaviour immediately.

Example: If a brand sees an unexpected spike in traffic to a specific product page, they can immediately trigger promotional offers or adjust their advertising to capitalize on the momentum.

5. Conversion and Goal Tracking

Web Analytics 2.0 emphasizes setting up specific goals (such as form submissions, sign-ups, downloads, or purchases) and tracking the conversion funnel. By understanding where users drop off, businesses can optimize the user experience and improve conversions.

Example: A SaaS company could track user sign-ups, usage, and subscription upgrades, allowing them to identify friction points in their user journey and make data-driven decisions to improve conversions.

6. Multi-Device Tracking

With the increase in mobile usage, Web Analytics 2.0 incorporates multi-device tracking, which allows businesses to see how users interact with content across different platforms (e.g., desktop, tablet, smartphone). This helps businesses understand cross-device behaviour and optimize experiences across devices.

Example: A customer may browse products on a mobile device and later complete the purchase on a desktop. Web Analytics 2.0 enables businesses to track this multi-device journey and provide a consistent experience.

7. Enhanced User Experience (UX) Insights

Web Analytics 2.0 focuses on understanding and optimizing the user experience by tracking elements like click heatmaps, scroll depth, and on-page behaviour. This enables businesses to make design and content decisions that improve usability and engagement.

Example: By tracking where users are clicking most on a webpage, a business can adjust its design and call-to-action buttons to improve conversion rates.

8. Predictive Analytics

Advanced predictive analytics is a hallmark of Web Analytics 2.0, allowing businesses to forecast future trends, user behaviour, and potential outcomes based on historical data and machine learning models.

Example: By analysing past customer behaviour, a retailer may predict which products are likely to become popular and adjust their inventory or marketing campaigns accordingly.

9. Attribution Modelling

Web Analytics 2.0 introduces advanced attribution models to better understand the customer journey and how various touch points (such as email campaigns, social media ads, and organic search) contribute to conversions.

Example: A multi-channel attribution model can help a company determine the exact role of a Facebook ad in the conversion process, compared to an email campaign or search engine optimization (SEO).

10. Customizable Dashboards and Reporting

With Web Analytics 2.0, businesses can create tailored dashboards and reports that highlight the most relevant metrics for different teams and decision-makers. Custom reports help in gaining actionable insights quickly.

Example: A marketing team may have a dashboard focused on conversion rates, while a product team may want to track customer feedback and engagement metrics.

1.3.16 Web Analytics Framework with Clickstream Analysis

A **clickstream** is the sequence of clicks that a user makes while navigating a website or app. It's a detailed record of how users interact with a site, from the first page they visit to the final page before exiting. **Clickstream analysis** in web analytics is crucial for understanding user behaviour, identifying patterns, and optimizing website or app performance.

Web Analytics Framework with Clickstream Analysis

The clickstream analysis framework in web analytics involves tracking the path a user takes through a site or app. This framework focuses on identifying the interactions and behaviours of users as they move from page to page, providing insights into user engagement, content preferences, and potential friction points in the user journey.

Here's how the clickstream framework fits within a web analytics strategy:

1. Data Collection: Tracking User Clicks

What is Clickstream Data?

Clickstream data refers to the detailed records of users' interactions with a website or app. This data is collected through tracking code implemented on web pages or within mobile apps, capturing every click, scroll, and interaction.

Data points can include:

- Page views
- Clicks on specific elements (buttons, links, etc.)
- Time spent on each page
- Scroll depth
- Referrals (where users came from)
- Exit pages (where users left the site)

Tools for Clickstream Data Collection

Google Analytics: Captures page views, click events, and user flows.

Hotjar: Tracks clicks, scroll behaviour, and user engagement with on-site elements.

Mixpanel: Tracks detailed user actions and events across both websites and mobile apps.

2. Clickstream Data Analysis: Understanding User Behaviour

Path Analysis: The sequence of pages that a user visits on a website, also known as the "user journey," can be analysed to identify common pathways or unexpected detours. This can help businesses optimize navigation or layout.

Example: If users consistently follow a specific path (e.g., homepage → product page → checkout), this suggests an effective user flow that can be enhanced further.

Example: If a significant number of users exit the site after visiting a particular page, it may indicate a need for optimization on that page.

Funnel Analysis: Funnel analysis tracks the steps users take toward completing a specific goal, such as making a purchase or signing up for a service. Clickstream data can help businesses visualize where users drop off in the conversion process, enabling them to focus on improving conversion rates.

Example: If users frequently click through the checkout process but abandon their carts before completing the purchase, the business can test changes (like reducing form fields or offering a discount) to improve conversion rates.

Click-Through Rate (CTR): Analysing the CTR on links, buttons, or ads can show which parts of the site engage users most effectively.

Example: A product page with a high click-through rate on the "Add to Cart" button indicates a well-designed page that motivates users to take the next step.

A **clickstream** is the sequence of clicks that a user makes while navigating a website or app. It's a detailed record of how users interact with a site, from the first page they visit to the final page before exiting.

Clickstream analysis in web analytics is crucial for understanding user behaviour, identifying patterns, and optimizing website or app performance.

3. User Segmentation: Identifying Different Behaviours

Segmentation allows businesses to group users based on different behaviours and characteristics. By segmenting clickstream data, companies can tailor their strategies to different types of users (e.g., new visitors, repeat visitors, users from different geographies, or those who abandon carts).

Example: A website might segment users who click through to product pages and add items to the cart but don't complete the purchase. This segment can be retargeted with specific ads or reminders about abandoned items.

Behavioural Segmentation: This helps identify user intent and categorize them into segments like browsers, buyers, frequent users, or one-time visitors.

Example: Frequent users may be shown personalized content, while first-time visitors may be shown introductory or educational content.

1.3.17 Multiple Outcome Analysis

Multiple Outcome Analysis in Social Media Web Analytics refers to the process of evaluating and understanding the impact of social media activities on multiple business objectives or outcomes simultaneously. Instead of focusing on a single metric or conversion, this approach tracks how different social media campaigns or activities influence various aspects of a business, such as brand awareness, customer engagement, lead generation, sales, customer retention, and overall ROI.

By employing multiple outcome analysis, businesses can gain a more holistic view of the effectiveness of their social media efforts, allowing for better decision-making and optimization of campaigns. Here's how multiple outcome analysis works in the context of social media web analytics:

1. Defining Multiple Outcomes

Before diving into data analysis, businesses need to define which outcomes they want to measure. These can include:

- **Brand Awareness**: Metrics like reach, impressions, and social mentions that track how many people are exposed to the brand.
- **Customer Engagement**: Metrics such as likes, shares, comments, and interactions that reflect how users are engaging with content.
- **Lead Generation**: Tracking form submissions, click-through rates (CTR) on links, and sign-ups that indicate the level of interest and conversion from social media content.
- **Sales**: Measuring direct conversions or sales from social media platforms, including purchase rates, order values, and the effectiveness of social media ads.
- **Customer Retention**: Analysing returning customer behaviours, engagement with loyalty programs, or customer satisfaction feedback from social channels.
- **Sentiment Analysis**: Analysing the tone of user comments or feedback to gauge how positive or negative the public perception is regarding a brand or campaign.

2. Collecting Data Across Multiple Channels

To perform multiple outcome analysis, data must be collected from multiple social media platforms and integrated into a centralized system. Social media platforms provide rich insights and metrics through their own analytics tools:

- **Facebook Insights**: Provides data on reach, engagement, clicks, conversions, and audience demographics.
- **Twitter Analytics**: Measures impressions, engagement rate, and link clicks, as well as follower growth.
- **Instagram Insights**: Offers engagement metrics, reach, impressions, and actions on posts and stories.

- **LinkedIn Analytics**: Focuses on engagement, follower demographics, post views, and shares.
- **Social Media Listening Tools**: Tools like Brand watch, Hoot suite, or Sprout Social help track brand mentions, sentiment, and user-generated content across different platforms.

3. Linking Outcomes to Business Goals

For the analysis to be valuable, the outcomes must be tied directly to the business's overall objectives. This can involve aligning social media outcomes with key performance indicators (KPIs) that are relevant to the business:

- **Sales and Revenue Impact**: Assess the direct contribution of social media posts or campaigns to sales or revenue.
- **Customer Lifetime Value (CLV)**: Measure how social media engagement leads to higher retention or repeat purchases, and how social media can influence long-term customer value.
- **Cost-Effectiveness**: Analyse how efficiently social media campaigns are driving business results (e.g., ROI) across various outcomes like brand awareness or lead generation.

4. Advanced Analytics Techniques for Multiple Outcomes

Multiple outcome analysis typically requires the use of advanced data analytics techniques to handle the complexity of the data:

- **Attribution Modelling**: This technique assigns value to different touch points in the customer journey, helping determine how various social media interactions contribute to multiple outcomes. For instance, if a user first sees an ad on Instagram and later converts via a Facebook post, attribution modelling helps understand the combined effect of both channels.

- **Example**: A user interacts with a brand on Instagram by liking a post (brand awareness), later clicks a Facebook ad (lead generation), and then makes a purchase (sales). Attribution modelling allows businesses to attribute value to all these touch points.
- **Predictive Analytics**: By using historical social media data and machine learning algorithms, businesses can predict how changes in social media strategies might affect multiple outcomes. For instance, predictive models can forecast how increased engagement on Instagram will impact future sales or customer retention.
- **Multivariate Analysis**: This statistical method helps businesses understand the relationships between multiple variables, such as how different types of posts (image, video, poll) on different platforms impact both engagement and sales simultaneously.
- **Correlation Analysis**: Helps in understanding how social media metrics (such as engagement, reach, or sentiment) correlate with multiple business outcomes (sales, customer retention). By analysing these relationships, businesses can prioritize actions that drive the most significant results.

5. Analysing and Reporting on Multiple Outcomes

Once the data is collected and analysed, businesses can visualize and report on the outcomes across different dimensions:

- **Dashboards**: Interactive dashboards like Google Data Studio or Tableau can be set up to monitor multiple outcomes simultaneously. These dashboards provide a real-time, visual representation of key metrics related to brand awareness, engagement, lead generation, and more.

Example: A dashboard might display a breakdown of how social media posts contribute to brand awareness (reach), customer engagement (likes, shares), and sales (conversions) all on the same page.

- **Heat maps & Visualizations**: Heat maps of click-through rates, engagement maps, or sentiment charts help visualize where interactions and positive/negative sentiments occur, offering insights into what type of content performs best across multiple outcomes.
- **Segmentation Reports**: Segment reports are helpful in understanding how different audience groups (age, location, interests) interact with various social media platforms and the associated outcomes. For example, the analysis may show that younger demographics engage more on Instagram, while older demographics convert better via Facebook.

1.3.18 Experimentation and Testing in Web Analytics

Experimentation and Testing in Web Analytics involves conducting systematic experiments (like A/B testing, multivariate testing, and other forms of controlled trials) to determine the effectiveness of different website or app features, designs, content, or marketing strategies. These methods are crucial for improving user experience, conversion rates, and overall performance by gathering data-driven insights to make informed decisions.

1. A/B Testing (Split Testing)

A/B testing is one of the most widely used experimentation methods in web analytics. It involves comparing two versions (A and B) of a webpage or app to see which one performs better on a specific goal, such as increasing conversions or engagement.

How A/B Testing Works

- **Hypothesis**: Before the test, you form a hypothesis. For example, "Changing the colour of the call-to-action button will increase the click-through rate."
- **Variants**: You create two versions of the page—Version A (the original) and Version B (the variation with the proposed change).
- **Random Split**: Traffic is randomly split between the two versions, ensuring that each user sees either A or B.
- **Measure and Compare**: Key metrics (such as click-through rates, bounce rates, or conversion rates) are measured to see which version performs better.

Best Practices for A/B Testing

- Test only one variable at a time (e.g., changing the button color, headline, or layout) to isolate the effect of the change.
- Use a sufficient sample size to ensure the results are statistically significant.
- Define clear goals or KPIs (e.g., form submissions, purchases) to measure success.
- Conduct the test for a reasonable period to account for various traffic conditions (e.g., weekdays, weekends).

Tools for A/B Testing

- **Google Optimize**: A free tool for A/B testing and personalization.
- **Optimizely**: A robust experimentation platform offering A/B testing and multivariate testing.
- **VWO (Visual Website Optimizer)**: Another popular tool for A/B testing and conversion rate optimization.

2. Multivariate Testing

Multivariate testing is a more advanced version of A/B testing. Instead of testing two variants, it tests multiple elements (e.g., headlines, images, CTAs) at once to determine which combination produces the best result.

How Multivariate Testing Works

- **Multiple Variants**: You create different combinations of variables (e.g., different versions of a headline, image, and button placement).
- **Random Split**: Like A/B testing, traffic is randomly split among all the variants, with each user seeing a different combination.
- **Measure and Compare**: You track how each combination performs against the predetermined goals.

Best Practices for Multivariate Testing

- Use when you have enough traffic to test multiple variations at once (since multivariate testing requires more data to reach statistical significance).
- Keep track of the performance of each element (e.g., headlines, images) separately, as well as their combinations.
- Ensure the test duration is long enough to account for different user behaviours across time.

Tools for Multivariate Testing

- **Google Optimize**: Supports multivariate testing alongside A/B testing.
- **Optimizely**: Offers multivariate testing capabilities for more complex experiments.
- **Adobe Target**: Another tool that supports multivariate testing with personalization features.

1.3.19 Voice of the Customer (VOC)

Voice of the Customer (VOC) in Social Media Web Analytics refers to the practice of capturing and analysing customer feedback, opinions, preferences, and sentiments expressed through social media platforms.

It provides businesses with real-time insights into customer needs, perceptions, and behaviours, enabling them to refine strategies, improve customer experiences, and build stronger relationships with their audience.

Key Aspects of Voice of the Customer in Social Media Analytics

1. What is Voice of the Customer (VOC)?

Voice of the Customer encompasses all customer feedback and sentiment shared across various channels, including social media, blogs, reviews, forums, and direct communications. It helps businesses:

- Understand customer expectations and pain points.
- Identify trends and emerging customer preferences.
- Prioritize product or service improvements based on customer insights.

Social media, in particular, is a goldmine for VOC data due to its interactive nature and the volume of conversations happening in real time.

2. Why is VOC Important in Social Media Analytics?

Social media platforms allow customers to share their thoughts freely, making them a critical source for understanding the customer's voice. Key benefits include:

- **Real-Time Insights**: Gain immediate feedback on campaigns, products, or services.
- **Sentiment Understanding**: Analyse customer emotions (positive, negative, or neutral) about a brand or product.
- **Trend Spotting**: Detect new customer demands or issues before they become widespread.
- **Enhanced Engagement**: Improve customer relations by responding to feedback effectively.

3. Key Metrics for Capturing VOC in Social Media

To extract meaningful insights, businesses focus on specific metrics:

- **Mentions**: Track the volume and frequency of brand mentions.
- **Sentiment Analysis**: Analyse whether mentions and comments reflect positive, negative, or neutral sentiment.
- **Engagement**: Measure likes, shares, comments, and replies to understand the impact of social media posts.
- **Customer Queries and Complaints**: Monitor customer service-related interactions to address issues promptly.
- **Feedback Volume**: Assess the amount of direct or indirect feedback about campaigns or product launches.

1.3.20 Competitive Intelligence

Competitive Intelligence in Social Media Web Analytics involves gathering and analysing data from social media platforms to understand competitors' strategies, performance, audience engagement, and market positioning. This information enables businesses to benchmark their performance, identify market trends, and uncover opportunities for gaining a competitive edge.

Key Aspects of Competitive Intelligence in Social Media Analytics

1. What is Competitive Intelligence?

Competitive intelligence (CI) refers to the process of systematically collecting and analyzing information about competitors to make informed business decisions. In the context of social media analytics, this means monitoring competitors' social media activities, audience interactions, content strategies, and campaign effectiveness.

2. Why is Competitive Intelligence Important in Social Media Analytics?

- **Benchmarking**: Understand how your brand compares to competitors in terms of engagement, reach, and follower growth.
- **Trend Identification**: Spot industry trends and best practices by analyzing competitors' successful campaigns.
- **Audience Insights**: Learn about shared or niche audience interests by examining competitor followers and interactions.
- **Strategic Planning**: Identify gaps in competitors' strategies and position your brand to fill those gaps effectively.

3. Metrics for Competitive Intelligence in Social Media Analytics

a) Audience Engagement

- **Likes, Shares, and Comments**: Track the level of engagement competitors receive on their posts to identify what resonates with audiences.
- **Engagement Rate**: Compare engagement relative to the total number of followers to assess content effectiveness.

b) Follower Growth

- **Rate of Growth**: Monitor how competitors' followers grow over time to identify successful campaigns or strategies.
- **Demographics**: Understand the demographics of their audience to see if you're targeting similar or complementary groups.

c) Content Performance

- **Post Frequency and Timing**: Analyze how often and when competitors post to optimize your own schedule.

- **Content Types**: Assess the effectiveness of different content formats (videos, images, polls, etc.).
- **Top-Performing Posts**: Identify posts with the highest engagement to understand what works best in your industry.

d) Sentiment Analysis

Monitor the tone of customer interactions with competitors (positive, negative, or neutral) to assess brand reputation and customer satisfaction.

e) Hashtag and Keyword Analysis

Analyse competitors' use of hashtags and keywords to optimize your own content for better reach and visibility.

f) Campaign Effectiveness

Track the success of competitors' campaigns (e.g., discounts, contests, influencer partnerships) and their impact on audience engagement and growth.

UNIT - 2

DATA

2.1 Introduction of Data

Data is a collection of information gathered by observations, measurements, research or analysis. They may consist of facts, numbers, names, figures or even description of things. Data is organized in the form of graphs, charts or tables. There exist data scientists who does data mining and with the help of that data analyse our world.

In general, data is a distinct piece of information that is gathered and translated for some purpose. If data is not formatted in a specific way, it does not valuable to computers or humans. Data can be available in terms of different forms, such as bits and bytes stored in electronic memory, numbers or text on pieces of paper, or facts stored in a person's mind. Since the invention of computers, people have used the word data to mean computer information, and this information is transmitted or stored. There are different kinds of data; such are as follows:

- Sound
- Video
- Single character
- Number (integer or floating-point)
- Picture
- Boolean (true or false)
- Text (string)

In a computer's storage, data is stored in the form of a series of binary digits (bits) that contain the value 1 or 0. The information can be in terms of pictures, text documents, software programs, audio or video clips, or other kinds of data.

The computer data may be stored in files and folders on the computer's storage, and processed by the computer's CPU, which utilizes logical operations to generate output (new data) form input data.

As the data is stored on the computer in binary form (zero or one), which can be processed, with the help of various media devices or a network connection. Furthermore, if you use data multiple times, it does not deteriorate over time or lose quality.

2.2 Structure Data

Structured data is data that has a standardized format for efficient access by software and humans alike. It is typically tabular with rows and columns that clearly define data attributes. Computers can effectively process structured data for insights due to its quantitative nature. For example, a structured customer data table containing columns - name, address, and phone number- can provide insights like the total number of customers and the locality with the maximum number of customers. In contrast, unstructured data, like a list of social media posts, is more challenging to analyse.

2.2.1 Benefits of Structured Data

There are several benefits of using structured data.

Ease of use

Anyone can quickly comprehend and access structured data. Operations such as updating and amending structured data are straightforward. Storage is efficient, as fixed-length storage units can be allocated for data values.

Scalability

Structured data scales algorithmically. You can add storage and processing power as your data volume increases. Modern systems that process structured data can scale to several thousand TB of data.

Analytics

Machine learning algorithms can analyse structured data and identify common patterns for business intelligence. You can use structured query language (SQL) to generate reports as well as modify and maintain data. Structured data is also useful for big data analytics.

2.3 Unstructured Data

Unstructured data is all that information that isn't predefined and searchable on a table, such as text messages, images, videos, audio files, and emails. Unlike structured data, which is easily placed into tables like those found in Microsoft Excel, unstructured data can't be quickly analysed and searched without further processing.

Unstructured data can take various forms, including text documents, emails, social media posts, images, videos, audio recordings, presentations, and more. It often contains free-form text, natural language, and multimedia content. In other words, then, unstructured data encompasses all the different kinds of qualitative data produced by individuals every day that lack clear-cut quantitive data points.

In turn, data professionals can find unstructured data from a wide variety of different sources. Some particularly rich sources of unstructured data include:

- Customer reviews
- Social media conversations
- News articles
- Sensor data from Internet of Things (IOT) devices

2.4 Metadata

Metadata is data about data. It describes attributes of your data that can be important to know but may not be immediately relevant to your data's primary function. This could include information about how data was collected, where it's stored, and how it's used.

For example, metadata is often attached to digital files, such as photos, videos, and audio, to describe aspects of the file, such as its name, size, when it was captured, and where it was captured.

Metadata can be helpful for verification and tracking purposes. You can refer to metadata to confirm data is accurate and reliable (data integrity) or to create an organisational system for your data. Certain software and programs also use metadata to interpret your data. For example, search engines use metadata to categorise web pages.

2.4.1 Types of Metadata

Several types of metadata exist, but three common types are descriptive, structural, and administrative:

- **Descriptive metadata** describes identification attributes, such as the file name and author.
- **Structural metadata** describes how data is organised, such as versions and relationships to other pieces of data.
- **Administrative metadata** describes technical attributes, such as the file size and creation date.

2.5 Big Data

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

2.5.1 Characteristics of Big Data

Big data can be described by the following characteristics:

- Volume
- Variety
- Velocity
- Variability

(i) Volume – The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data. Hence, **'Volume'** is one characteristic which needs to be considered while dealing with Big Data solutions.

(ii) Variety – The next aspect of Big Data is its **variety**.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spread sheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analysing data.

(iii) Velocity – The term **'velocity'** refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.

(iv) Variability – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

2.6 Advantages of Big Data Processing

Ability to process Big Data in DBMS brings in multiple benefits, such as-

- Businesses can utilize outside intelligence while taking decisions

Access to social data from search engines and sites like Facebook, Twitter are enabling organizations to fine tune their business strategies.

- Improved customer service

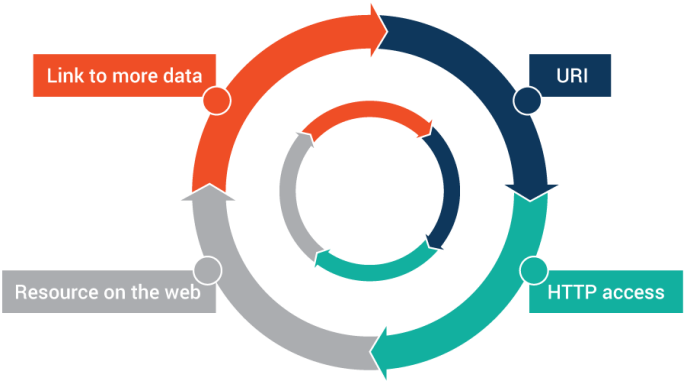
Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

- Early identification of risk to the product/services, if any
- Better operational efficiency

Big Data technologies can be used for creating a staging area or landing zone for new data before identifying what data should be moved to the data warehouse. In addition, such integration of Big Data technologies and data warehouse helps an organization to offload infrequently accessed data.

2.7 Linked Data

Linked Data is a set of design principles for sharing machine-readable interlinked data on the Web. When combined with Open Data (data that can be freely used and distributed), it is called Linked Open Data (LOD). An RDF database such as Ontotext's Graph DB is an example of LOD. It is able to handle huge datasets coming from disparate sources and link them to Open Data, which boosts knowledge discovery and efficient data-driven analytics.



Linked Data is one of the core pillars of the Semantic Web, also known as the Web of Data. The Semantic Web is about making links between datasets that are understandable not only to humans, but also to machines, and Linked Data provides the best practices for making these links possible. In other words, Linked Data is a set of design principles for sharing machine-readable interlinked data on the Web.

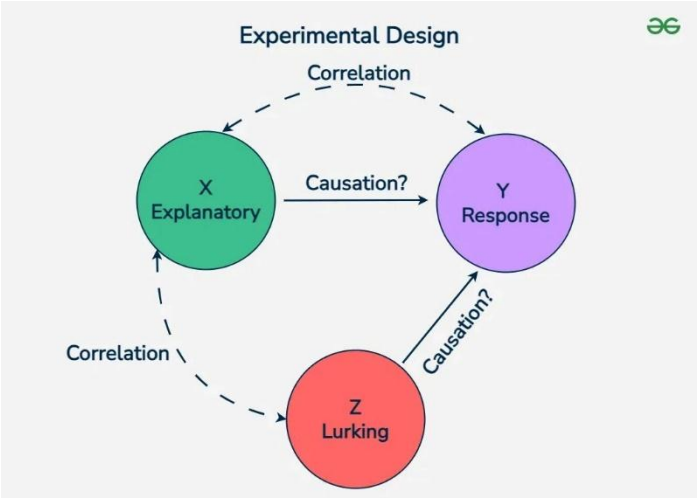
2.8 Lab Testing

Lab tests are conducted by a User Centric Design / Human Factors expert, who is typically supported by a note taker. Key stake holders connected to the website (or product) participate as Observers and their job is to get a close understanding of the customer experience. Stake holders can be business owners, engineers and developers, web analysts), product managers etc, anyone who has something to do with the website or customer experience. Tests can be conducted with live version of the website, beta versions, on-screen HTML or PowerPoint prototypes, or even with paper printouts. These paper prototypes, sometimes called wire-frames, approximate what a user might otherwise see on a computer screen, but save the development team from having to produce an on-screen product.

2.9 What is Experimental Design?

Experimental design simply refers to the strategy that is employed in conducting experiments to test hypotheses and arrive at valid conclusions. The process comprises firstly, the formulation of research questions, variable selection, specifications of the conditions for the experiment, and a protocol for data collection and analysis. The importance of experimental design can be seen through its potential to prevent bias, reduce variability, and increase the precision of results in an attempt to achieve high internal validity of studies.

By using experimental design, the researchers can generate valid results which can be generalized in other settings which helps the advancement of knowledge in various fields.



2.9.1 Types of Experimental Design

Experimental design encompasses various approaches to conducting research studies, each tailored to address specific research questions and objectives. The primary types of experimental design include:

- Pre-experimental Research Design
- True Experimental Research Design
- Quasi-Experimental Research Design
- Statistical Experimental Design

Pre-experimental Research Design

A preliminary approach where groups are observed after implementing cause and effect factors to determine the need for further investigation. It is often employed when limited information is available or when researchers seek to gain initial insights into a topic.

Pre-experimental designs lack random assignment and control groups, making it difficult to establish causal relationships.

Classifications

- One-Shot Case Study
- One-Group Pretest-Posttest Design
- Static-Group Comparison

True-experimental Research Design

The true-experimental research design involves the random assignment of participants to experimental and control groups to establish cause-and-effect relationships between variables. It is used to determine the impact of an intervention or treatment on the outcome of interest. True-experimental designs satisfy the following factors:

Factors to Satisfy

- Random Assignment
- Control Group
- Experimental Group
- Pretest-Posttest Measures

Quasi-Experimental Design

A quasi-experimental design is an alternative to the true-experimental design when the random assignment of participants to the groups is not possible or desirable. It allows for comparisons between groups without random assignment, providing valuable insights into causal relationships in real-world settings. Quasi-experimental designs are used typically in conditions wherein the random assignment of the participants cannot be done or it may not be ethical, for example, an educational or community-based intervention.

Statistical Experimental Design

Statistical experimental design, also known as design of experiments (DOE), is a branch of statistics that focuses on planning, conducting, analysing, and interpreting controlled tests to evaluate the factors that may influence a particular outcome or process. The primary goal is to determine cause-and-effect relationships and to identify the optimal conditions for achieving desired results.

2.10 Selecting Participants for Lab Testing and Experimental Design Requires Careful Consideration to Ensure Reliability, Validity, and Ethical Integrity. Here's a Step-By-Step Guide

2.10.1 Define Research Objectives

- **Purpose:** Clarify what you aim to achieve through the experiment (e.g., testing a hypothesis, evaluating a new process).
- **Variables:** Identify dependent, independent, and control variables.

2.10.2 Identify Target Population

- **Demographics:** Determine the specific characteristics needed (e.g., age, gender, geographic region, occupation).
- **Eligibility Criteria:** Set inclusion (who qualifies) and exclusion (who does not qualify) criteria based on research needs.

2.10.3 Decide on Sampling Method

- **Probability Sampling:** Every individual in the population has an equal chance of selection (e.g., simple random sampling, stratified sampling).
- **Non-Probability Sampling:** Selection is based on specific criteria or convenience (e.g., convenience sampling, purposive sampling).

2.10.4 Calculate Sample Size

Use statistical methods or tools to determine an adequate sample size based on:

- Population size.
- Confidence level (e.g., 95%).
- Margin of error (e.g., ±5%).
- Effect size (magnitude of the expected effect).

2.10.5 Recruitment Strategies

- **Methods:** Utilize advertisements, flyers, emails, or recruitment platforms to reach potential participants.
- **Incentives:** Offer compensation or benefits (monetary, gift cards, certificates) if appropriate and ethical.

2.10.6 Screen Participants

- **Pre-Screening Surveys:** Use questionnaires to ensure participants meet the criteria.
- **Consent Forms:** Explain the study's purpose, risks, benefits, and rights to participants.
- **Medical Screening:** Conduct tests if health conditions might affect results.

2.10.7 Randomization (if applicable)

- Assign participants to different experimental groups randomly to reduce bias.
- Methods: Random number generators, software tools, or stratified randomization.

2.10.8 Ethical Considerations

- **IRB Approval:** Obtain clearance from an Institutional Review Board (IRB) or Ethics Committee.
- **Informed Consent:** Ensure participants voluntarily agree after being informed of all relevant details.

- **Privacy:** Protect participant data and maintain confidentiality.

2.10.9 Plan the Testing Environment

- **Standardization:** Create consistent conditions across all tests.
- **Training:** Train researchers and staff to ensure uniform protocols.

2.10.10 Conduct Pilot Testing

- Test the experimental design with a small sample to identify and fix issues before the main study.

2.11 When deciding between a within-subject or between-subject

1. Understand the Two Study Designs

Within-Subject Design

Definition

Each participant is exposed to all experimental conditions or treatments.

Key Advantages

- Requires fewer participants.
- Eliminates individual differences as a confounding variable.
- Provides more statistical power.

Key Challenges

- Risk of **carryover effects** (e.g., learning, fatigue).
- May require **counterbalancing** to mitigate order effects.

Between-Subject Design

Definition

Different groups of participants are exposed to different conditions or treatments.

Key Advantages

- No carryover or order effects.

- Easier to implement when testing interventions that cannot be undone or reversed.

Key Challenges

- Requires more participants to achieve statistical power.
- Potential for group differences to introduce variability.

2. Key Considerations for Selection

1. Nature of the Experiment

Use a **within-subject design** when studying relative effects (e.g., comparing reactions to different stimuli in the same individual).

Opt for a **between-subject design** when the intervention is irreversible (e.g., long-term behaviour change).

2. Practical Constraints

Within-Subject: Suitable when participant recruitment is difficult.

Between-Subject: Preferred when there's concern about fatigue or learning effects in repeated testing.

3. Ethical Implications

Ensure that participants in a **between-subject design** are not disadvantaged by being in a control group.

Minimize discomfort or confusion in **within-subject designs** due to repeated exposures.

4. Statistical Power

Within-subject designs generally require fewer participants and have higher sensitivity to detecting differences.

Between-subject designs rely on more participants to average out individual differences.

3. Steps for Experimental Design

1. Define the Research Question

Clearly state the objective and determine if the study requires comparing conditions within the same individuals or across different groups.

2. **Recruit Participants**

Ensure a diverse and representative sample.

Calculate the required sample size based on the study design and statistical goals.

3. **Randomization and Counterbalancing**

For **within-subject designs**, use counterbalancing to address order effects (e.g., Latin Square design).

For **between-subject designs**, randomize participants to groups to reduce bias.

4. **Control Variables**

In **between-subject designs**, ensure groups are equivalent by controlling for confounding variables.

In **within-subject designs**, minimize environmental factors that could influence results between conditions.

5. **Pilot Testing**

Conduct a small-scale trial to identify potential issues such as carryover effects, task complexity, or participant fatigue.

2.12 Counterbalancing

Counterbalancing is a technique used in within-subject experimental designs to reduce the impact of order effects, such as learning, fatigue, or carryover effects, that arise when participants experience multiple conditions.

2.12.1 Why Use Counterbalancing?

1. **Order Effects**

Participants' responses may be influenced by the sequence in which conditions are presented.

Examples: Improved performance due to practice (learning effect) or reduced performance due to exhaustion (fatigue effect).

2. **Carryover Effects**

Experiencing one condition might influence responses to subsequent conditions.

Example: A strong emotional stimulus in the first condition could alter responses in later conditions.

3. Fair Representation

Ensures that no single condition benefits or suffers disproportionately due to its position in the sequence.

2.12.2 Types of Counterbalancing

1. Complete Counterbalancing:

All possible orders of conditions are presented.

Best suited for studies with a small number of conditions.

Example:

For three conditions (A, B, C), all six orders are tested:

ABC, ACB, BAC, BCA, CAB, CBA.

Challenges

- Number of orders grows factorially (*n!*, where *n* is the number of conditions).
- Difficult to implement with many conditions.

2. Partial Counterbalancing

A subset of all possible orders is used.

Often employs techniques like a **Latin Square design** to ensure each condition appears in every position at least once.

Example:

For three conditions (A, B, C):

- Order 1: ABC
- Order 2: BCA
- Order 3: CAB.

Benefits

Reduces complexity while still controlling order effects.

3. Randomized Counterbalancing

The order of conditions is randomized for each participant.

Works well for large sample sizes where randomization averages out biases.

4. Block Randomization

Conditions are divided into blocks, and the order within each block is randomized.

Useful for longer studies to prevent prolonged exposure to a single condition.

2.13 Independent and Dependent Variables

In experimental research, the independent and dependent variables are crucial for establishing cause-and-effect relationships. Here's an overview:

1. Independent Variable (IV)

Definition

The variable that the researcher manipulates or controls to observe its impact on the dependent variable.

Purpose

To determine how changes in the independent variable influence the outcome.

Examples:

- In a study on teaching methods: **Teaching method** (e.g., traditional vs. online) is the IV.
- In a drug efficacy test: **Dosage of the drug** is the IV.

Key Characteristics

- It is the presumed cause.
- Typically has different levels or conditions (e.g., high vs. low dosage).

2. Dependent Variable (DV)

Definition

The variable that is measured or observed to assess the effect of changes in the independent variable.

Purpose

To evaluate the outcome or result influenced by the IV.

Examples:

- In a study on teaching methods: **Students' test scores** are the DV.
- In a drug efficacy test: **Improvement in symptoms** is the DV.

Key Characteristics

- It is the presumed effect.
- Depends on the manipulation of the IV.

2.14 A/B Testing: An Overview

Definition

A/B testing is a method of comparing two versions of a product, webpage, or experience (Version A and Version B) to determine which performs better based on a defined metric, such as conversion rate, click-through rate, or user engagement.

2.14.1 Key Components of A/B Testing

1. **Version A (Control)**

   The original version that serves as the baseline.

2. **Version B (Variant)**

   A modified version of the control, incorporating a change designed to improve performance.

3. **Hypothesis**

   A testable prediction about how the change in Version B will affect the outcome.

   Example: "Changing the button color from blue to green will increase click-through rates."

4. **Metrics**

   The measurable outcome used to evaluate success (e.g., conversions, time spent on a page, purchases).

5. **Target Audience**

   The segment of users randomly divided into groups to receive either Version A or Version B.

2.14.2 How A/B Testing Works: Step-by-Step

1. **Define the Goal**

   Identify the key performance indicator (KPI) you want to optimize (e.g., newsletter sign-ups, purchases).

2. **Develop Hypotheses**

   - Predict how a specific change will impact the KPI.

- Example: "Adding customer testimonials to the homepage will increase trust and lead to higher sales."

3. **Create Variants**
   - Develop the control (Version A) and the variant (Version B).
   - Ensure that the change is isolated to one element to accurately measure its impact.
4. **Split the Audience**
   Randomly divide the audience into two groups:
   - Group 1 sees Version A (Control).
   - Group 2 sees Version B (Variant).
5. **Run the Test**
   Expose the versions to users simultaneously under identical conditions.
6. **Measure Results**
   Collect data on the predefined metric (e.g., clicks, purchases).
7. **Analyze and Interpret Data**
   Use statistical analysis to determine if the difference between Version A and Version B is significant.
8. **Implement the Winning Version**
   If the variant performs better, implement it for all users.

2.15 Multivariate Analysis (MVA)

Multivariate analysis refers to a set of statistical techniques used to analyze data that involves multiple variables simultaneously. It explores relationships, patterns, and influences between variables, often with the goal of understanding complex phenomena or making predictions.

2.15.1 Why Use Multivariate Analysis?

- To identify relationships between variables.
- To determine the influence of one or more independent variables on one or more dependent variables.
- To reduce the dimensionality of data while preserving key information.

- To test hypotheses involving multiple variables.

2.15.2 Key Types of Multivariate Analysis

1. Descriptive Multivariate Techniques

Used to explore and summarize relationships between variables without focusing on causation.

Principal Component Analysis (PCA)

Reduces data dimensionality by identifying key components that explain the most variance.

Example: Identifying main factors influencing customer satisfaction.

Factor Analysis

Groups related variables into factors to uncover latent constructs.

Example: Grouping survey items into broader dimensions like "customer loyalty" or "brand trust."

Cluster Analysis

Groups similar observations into clusters.

Example: Segmenting customers into groups based on purchasing behaviour.

2. Inferential Multivariate Techniques

Used to test hypotheses and examine cause-and-effect relationships.

Multiple Regression Analysis

Explores the relationship between one dependent variable and multiple independent variables.

Example: Predicting house prices based on size, location, and number of bedrooms.

Multivariate Analysis of Variance (MANOVA)

Extends ANOVA to multiple dependent variables.

Example: Studying the effect of teaching methods on test scores and student motivation.

Discriminant Analysis

Determines which variables best differentiate between predefined groups.

Example: Identifying factors that distinguish between high- and low-performing employees.

3. Predictive Multivariate Techniques

Used for forecasting or predicting outcomes based on input variables.

Logistic Regression

Predicts a categorical dependent variable using one or more independent variables.

Example: Predicting customer churn (yes/no) based on usage data.

Time Series Multivariate Analysis

Examines multiple time-dependent variables to predict future trends.

Example: Forecasting sales based on advertising spend and seasonality.

Canonical Correlation Analysis (CCA)

Assesses relationships between two sets of variables.

Example: Linking financial performance indicators with market conditions.

2.16 Controlled Experiments

A controlled experiment is a scientific study where one or more independent variables are manipulated while all other variables are kept constant to observe their effect on the dependent variable. This method is fundamental in establishing cause-and-effect relationships.

2.16.1 Key Elements of Controlled Experiments

1. **Independent Variable**: The factor that is intentionally changed or manipulated to observe its effect.

2. **Dependent Variable**: The outcome or response that is measured to assess the effect of the independent variable.
3. **Control Group**: A group that does not receive the experimental treatment or is subjected to standard conditions. This group serves as a baseline for comparison.
4. **Experimental Group**: A group that receives the treatment or condition being tested.
5. **Controlled Variables**: All other factors are kept constant to prevent them from influencing the results.

2.16.2 Steps in Conducting Controlled Experiments

1. **Define the Hypothesis**: Clearly state what you aim to test.
2. **Identify Variables**: Determine the independent, dependent, and controlled variables.
3. **Set up Groups**: Divide subjects into control and experimental groups randomly.
4. **Apply Treatment**: Administer the experimental condition only to the experimental group.
5. **Measure Outcomes**: Collect data on the dependent variable for both groups.
6. **Analyze Results**: Compare the outcomes between the control and experimental groups to draw conclusions.

2.17 Basics of Data Analysis

Data analysis involves examining, organizing, and interpreting data to uncover patterns, draw conclusions, and inform decisions. Whether qualitative or quantitative, it is a structured approach to making sense of raw data.

Steps in Data Analysis

1. **Define Objectives**
   - Identify what you want to analyse and why.
   - Formulate clear questions or hypotheses.

2. **Collect Data**
   - Gather data through surveys, experiments, observations, or secondary sources.
   - Ensure the data is relevant, accurate, and reliable.
3. **Clean Data**
   - Handle missing, inconsistent, or irrelevant data.
   - Standardize formats and remove outliers if necessary.
4. **Organize Data**
   - Use tools like spread sheets or databases to sort and structure data.
   - Categorize or segment data into meaningful groups.
5. **Analyze Data**
   - Apply statistical methods, tools, or qualitative techniques.
   - Use exploratory data analysis (EDA) to identify patterns or trends.
6. **Interpret Results**
   - Relate findings back to the objectives or hypotheses.
   - Identify insights and implications of the results.
7. **Present Findings**
   - Use visualizations (charts, graphs, tables) to convey insights clearly.
   - Write reports or summaries to communicate results effectively.

2.18 Types of Data

Data can be broadly categorized into two main types: **qualitative** and **quantitative**. Each has subtypes that describe its characteristics and applications.

1. Qualitative Data

Qualitative data is descriptive and non-numerical. It captures qualities, characteristics, or attributes.

Key Features

- Cannot be measured numerically but can be categorized.
- Often collected through interviews, surveys, or observations.
- Helps in understanding underlying opinions, motivations, or trends.

Subtypes of Qualitative Data

- **Nominal Data**
  - Categories without a logical order.
  - Example: Gender (male, female), Colors (red, blue, green).
- **Ordinal Data**
  - Categories with a logical order but no consistent difference between ranks.
  - Example: Customer satisfaction (poor, average, good, excellent)

2. Quantitative Data

Quantitative data is numerical and can be measured or counted.

Key Features

- Used for statistical analysis.
- Can show trends, correlations, and cause-effect relationships.

Subtypes of Quantitative Data

- **Discrete Data**
  - Countable, finite values.
  - Example: Number of employees, number of cars in a parking lot.
- **Continuous Data**:
  - Measurable and can take any value within a range.
  - Example: Height, weight, temperature.

2.19 Metrics Data

Metrics data refers to quantifiable measurements used to assess, track, and analyze the performance, progress, or quality of specific processes, systems, or objectives. It is essential in fields like business, research, IT, healthcare, and education to make informed decisions based on evidence.

2.19.1 Types of Metrics Data

1. **Key Performance Indicators (KPIs)**
   - High-level metrics that reflect critical success factors.
   - Example:
     - Business: Revenue growth rate, customer retention rate.
     - Education: Graduation rates, student satisfaction scores.
2. **Operational Metrics**
   - Day-to-day performance measures for operational processes.
   - Example: Average response time, production efficiency.
3. **Financial Metrics**
   - Metrics focusing on the financial performance of an organization.
   - Example: Profit margins, return on investment (ROI), cash flow.
4. **Productivity Metrics**
   - Measures of efficiency and output.
   - Example: Employee output per hour, sales per employee.
5. **Quality Metrics**
   - Used to measure the quality of outputs or processes.
   - Example: Customer complaint rate, defect rate.
6. **Customer Metrics**
   - Metrics to understand customer behaviour and satisfaction.
   - Example: Net Promoter Score (NPS), customer lifetime value (CLV).

7. Website/Software Metrics

- Data about user engagement, system performance, or functionality.
- Example: Page views, bounce rate, server uptime.

2.20 Comparing Means

Comparing means is a statistical process used to evaluate whether the average values of two or more groups differ significantly. It is commonly applied in fields like research, business, and education to test hypotheses and make data-driven decisions.

Key Methods for Comparing Means

1. Two-Sample Comparison

- **Independent Samples t-Test**
  - Compares the means of two independent groups.
  - Example: Examining whether two different teaching methods result in different average test scores.
  - Assumptions:
    - ❖ Data in each group is normally distributed.
    - ❖ Equal variance between groups (use Welch's t-test if this assumption is violated).
    - ❖ Observations are independent.
- **Mann-Whitney U Test**
  - A non-parametric alternative to the t-test, used when data is not normally distributed.
  - Example: Comparing customer satisfaction scores between two products.

2. Paired Comparison

- **Paired Samples t-Test**
  - Compares means from the same group at two different time points or under two conditions.
  - Example: Testing the effect of a training program by comparing employee performance scores before and after the program.

- Assumptions:
  - ❖ Data is normally distributed.
  - ❖ Paired observations (e.g., same subjects measured twice).

- **Wilcoxon Signed-Rank Test**

3. Multiple Group Comparison

- **One-Way ANOVA (Analysis of Variance)**
  - Compares the means of three or more independent groups.
  - Example: Evaluating whether average customer ratings differ across three store locations.
  - Assumptions:
    - ❖ Data in each group is normally distributed.
    - ❖ Homogeneity of variances (use Welch's ANOVA if violated).
    - ❖ Groups are independent.
- **Kruskal-Wallis Test**
  - A non-parametric alternative to one-way ANOVA for non-normally distributed data.
  - A non-parametric alternative to the paired t-test for non-normally distributed data.

2.21 Correlations

Correlation measures the strength and direction of a relationship between two variables. It does not imply causation but helps understand how variables move together.

Types of Correlation

1. Positive Correlation
- Both variables increase or decrease together.
- Example: As income increases, spending increases.

2. Negative Correlation
- One variable increases while the other decreases.
- Example: As exercise time increases, weight decreases.

3. **No Correlation**
    - No discernible relationship between the variables.
    - Example: Hair color and exam scores.

2.22 Non-Parametric Tests

Non-parametric tests are statistical tests that do not assume a specific distribution for the data (e.g., normal distribution). These tests are ideal for analysing ordinal data, non-normal distributions, or when sample sizes are small.

2.22.1 Common Non-Parametric Tests

1. **Comparing Two Groups**

- **Mann-Whitney U Test**
    - Compares medians of two independent groups.
    - Example: Comparing customer satisfaction scores between two companies.
    - Parametric Equivalent: Independent t-test.
- **Wilcoxon Signed-Rank Test**:
    - Compares medians of two related groups.
    - Example: Comparing pre- and post-test scores of students.
    - Parametric Equivalent: Paired t-test.

2. **Comparing Multiple Groups**

- **Kruskal-Wallis Test**
    - Compares medians of three or more independent groups.
    - Example: Comparing average ratings for three products.
    - Parametric Equivalent: One-way ANOVA.
- **Friedman Test**
    - Compares medians of three or more related groups.
    - Example: Comparing performance under three different teaching methods.
    - Parametric Equivalent: Repeated Measures ANOVA.

3. Correlation

- **Spearman's Rank Correlation**
  - Measures the monotonic relationship between two variables.
  - Example: Relationship between ranks in sales and advertising spend.
  - Parametric Equivalent: Pearson's correlation.
- **Kendall's Tau**
  - Measures association between two ordinal variables.
  - Example: Examining the association between ranks given by two judges.

4. Goodness of Fit

- **Chi-Square Goodness of Fit Test**
  - Tests whether observed categorical data matches an expected distribution.
  - Example: Testing if dice rolls are fair.
  - Parametric Equivalent: None (used for categorical data).

5. Test of Association

- **Chi-Square Test of Independence**
  - Tests the association between two categorical variables.
  - Example: Analysing if gender is related to product preference.
- **Fisher's Exact Test**
  - Similar to the Chi-Square test but used for small sample sizes.
  - Example: Testing association between two binary variables.

2.22 Non –Parametric Test

Non-Parametric Tests: Overview

Non-parametric tests are statistical tests that do not rely on specific assumptions about the underlying data distribution. These tests are versatile and often used when:

1. The data is ordinal, nominal, or non-normal.
2. The sample size is small.

3. The assumptions of parametric tests are violated.

Categories of Non-Parametric Tests

1. Tests for One Sample

- **Chi-Square Goodness-of-Fit Test**
  - Purpose: Tests if the observed data fits an expected distribution.
  - Example: Checking if the distribution of votes matches predicted proportions.
- **Kolmogorov-Smirnov Test**
  - Purpose: Compares a sample to a reference probability distribution.
  - Example: Testing if data follows a uniform distribution.
- **Runs Test**
  - Purpose: Tests for randomness in a sequence.
  - Example: Analyzing if coin toss outcomes are random.

2. Tests for Two Independent Samples

- **Mann-Whitney U Test**:
  - Purpose: Compares the medians of two independent groups.
  - Example: Comparing customer satisfaction scores between two brands.
  - Parametric Equivalent: Independent t-test.
- **Kolmogorov-Smirnov Test for Two Samples**
  - Purpose: Compares the distribution of two independent samples.
  - Example: Testing if two groups have the same income distribution.

UNIT- 3

USABILITY METRICS

3.1 Meaning -Usability Metrics

Usability metrics are quantitative measures used to evaluate the user-friendliness and effectiveness of a product, system, or service. These metrics assess how efficiently users can complete tasks, how satisfied they feel with the experience, and how prone they are to errors while interacting with the system. Key usability metrics include task success rate, which measures the percentage of tasks successfully completed by users, and error rate, which tracks the frequency of mistakes made during use. Time on task evaluates how long it takes users to complete specific actions, while user satisfaction is typically gauged through surveys or questionnaires. Learnability examines how easily new users can accomplish basic tasks, and efficiency reflects the speed and ease with which experienced users perform actions. Additionally, retention rate measures how well users remember how to use the system over time. Together, these metrics provide valuable insights into user behaviour and experience, guiding improvements to make systems more intuitive, accessible, and enjoyable.

Some Key Usability Metrics include

1. **Effectiveness**: How accurately users can complete a task. This is often measured by the success rate or task completion rate.
2. **Efficiency**: How quickly users can complete a task. This metric is often measured by the time it takes to perform a task or the number of steps needed to complete it.
3. **Satisfaction**: How users feel about their experience with the system. It can be measured using surveys, feedback, or tools like the System Usability Scale (SUS).

4. **Error Rate**: The number of errors users make while performing a task. A higher error rate typically indicates a usability issue.
5. **Learnability**: How easy it is for users to learn to use the system. This is often measured by how quickly users can become proficient with the system.
6. **Retention**: How well users remember how to use the system after a period of time.

3.2 Performance Metrics

Performance metrics are quantitative measures used to assess the effectiveness and efficiency of a system, process, or organization in achieving its goals. These metrics provide insights into how well resources are utilized, how objectives are met, and where improvements are needed. Common performance metrics include productivity, which measures the output generated relative to the input used, and efficiency, which assesses the ability to achieve results with minimal waste of time or resources. Accuracy evaluates the correctness of outputs, while speed or response time measures how quickly a task or process is completed. Other important metrics include throughput, which tracks the volume of work processed within a given timeframe, and reliability, which gauges the consistency of performance over time. In organizational contexts, performance metrics may also include financial indicators like revenue, profit margins, and return on investment (ROI). These metrics are essential for monitoring progress, identifying areas for improvement, and ensuring that objectives are achieved effectively and sustainably.

Performance metrics have several key features that make them valuable tools for evaluating and improving the efficiency and effectiveness of systems, processes, or organizations. These features include:

- **Measurability**: Performance metrics are quantifiable, allowing for objective assessment and comparison. They provide numerical data that can be tracked over time.
- **Relevance**: Metrics are aligned with specific goals and objectives, ensuring they measure what truly matters for the success of a system or organization.
- **Specificity**: Good metrics focus on a particular aspect of performance, providing clear and actionable insights rather than vague or generalized information.
- **Timeliness**: Metrics can be measured at appropriate intervals (real-time, daily, weekly, etc.) to provide on going insights and facilitate timely decision-making.
- **Accuracy**: Reliable data collection and analysis ensure that performance metrics are precise and reflect actual performance levels.
- **Comparability**: Metrics allow for comparisons across time periods, teams, or industry standards, enabling benchmarking and identifying areas for improvement.
- **Actionability**: Effective performance metrics provide insights that lead to informed decisions, helping to address inefficiencies, capitalize on strengths, or correct course where needed.
- **Scalability**: Metrics can be applied at various levels (individual, team, organization) and scales (small projects to large systems).
- **Consistency**: Metrics are consistently defined and measured to ensure reliability and build trust in the data.
- **Transparency**: Metrics are easily understandable and clearly communicated to stakeholders, fostering accountability and collaboration.

3.3 Issue- Based Metrics

Issue-based metrics focus on identifying, tracking, and analyzing problems or challenges encountered during a process, project, or system operation.

These metrics aim to provide insights into the nature, frequency, and impact of issues, helping organizations prioritize and address them effectively. They are particularly useful in quality assurance, risk management, and continuous improvement efforts. Common features of issue-based metrics include:

- **Issue Frequency**: Tracks how often specific issues occur over a given period, providing a sense of recurring problems.
- **Issue Severity**: Measures the impact or criticality of each issue, helping prioritize those with the most significant effect on performance or outcomes.
- **Resolution Time**: Captures the time it takes to resolve an issue, reflecting the efficiency of the problem-solving process.
- **Root Cause Analysis Metrics**: Evaluates the underlying causes of issues, focusing on patterns or systemic problems that need addressing.
- **Reopen Rate**: Measures the percentage of issues reopened after being marked as resolved, indicating the effectiveness of initial solutions.
- **Impact Metrics**: Assesses how issues affect overall performance, customer satisfaction, or other critical outcomes.
- **Detection Rate**: Monitors how quickly issues are identified, highlighting the efficiency of monitoring and detection processes.
- **Escalation Rate**: Tracks the proportion of issues escalated to higher levels of management or technical support, reflecting the complexity of problems.
- **Preventive Action Metrics**: Evaluates the effectiveness of measures taken to prevent similar issues in the future.

- **Trend Analysis**: Identifies patterns or trends in issue occurrence, offering insights into potential long-term risks or opportunities for improvement.

3.4 Self-Reported Metrics

Self-reported metrics are data points collected directly from individuals, typically through surveys, interviews, or feedback forms, where users or stakeholders report their experiences, opinions, or perceptions. These metrics are subjective but provide valuable insights into user satisfaction, engagement, and overall experience, often complementing objective performance data. Key aspects of self-reported metrics include:

- **User Satisfaction**: Measures how satisfied individuals feel about a product, service, or process, often using tools like Likert scales or Net Promoter Scores (NPS).
- **Perceived Ease of Use**: Assesses how easy users believe a system or process is to use, which can highlight areas for usability improvement.
- **Reported Time on Task**: Captures how long users think they spent on a task, which can differ from actual time and provide insights into user perceptions.
- **User Confidence**: Reflects how confident individuals feel about completing tasks or using a system effectively.
- **Feedback on Features**: Gathers opinions about specific features or functions, helping to identify what's most useful or needs improvement.
- **Emotional Responses**: Tracks users' feelings, such as frustration, delight, or satisfaction, during their interaction with a product or service.
- **Intention Metrics**: Measures the likelihood of future behaviour, such as intentions to reuse, recommend, or purchase a product.

- **Perceived Value**: Evaluates how users perceive the worth or benefits of a product or service relative to its cost or effort.
- **Open-Ended Feedback**: Collects qualitative insights that provide context or explain user ratings and opinions.
- **Demographic Information**: Gathers self-reported details like age, location, or preferences, which can help segment and analyse user groups.
- While self-reported metrics offer rich, qualitative insights into user experiences and perceptions, they can be influenced by biases like memory errors or emotional states. For this reason, they are often used in conjunction with objective metrics to provide a more comprehensive understanding of performance and user satisfaction.

3.5 Planning and Performing Usability Study

A usability study is a structured process for evaluating how effectively and efficiently users interact with a product, system, or service. Proper planning and execution are crucial to ensure meaningful and actionable results. Below are the key steps:

Planning Phase

1. **Define Objectives**
   - Identify the purpose of the study, such as improving user experience, identifying pain points, or validating design changes.
   - Specify the metrics to be measured (e.g., task success rate, error rate, satisfaction).
2. **Understand the Target Audience**
   - Define user personas based on demographics, behavior, and needs.
   - Choose representative participants who align with the intended user base.

3. **Scope the Study**
   - Determine the tasks or scenarios to be tested (e.g., logging in, completing a purchase).
   - Decide the scope: focus on a specific feature, the overall system, or comparative analysis.
4. **Select the Study Type**
   - **Moderated Testing**: In-person or remote, guided by a facilitator.
   - **Unmoderated Testing**: Participants perform tasks independently.
   - **Exploratory Testing**: For early-stage feedback.
   - **Comparative Testing**: To evaluate multiple designs or systems.
5. **Develop Test Materials**
   - Prepare task scenarios with clear instructions.
   - Design pre- and post-test questionnaires to gather background and feedback.
   - Create a script for consistency in moderation.
6. **Set Up the Environment**
   - Choose the location (lab, remote, or field).
   - Ensure the necessary tools, such as screen-recording software, cameras, or usability platforms, are available.
7. **Recruit Participants**
   - Screen participants to ensure they match the target user profile.
   - Offer incentives, if appropriate, to encourage participation.

II. Performing Phase

1. **Pilot Testing**
   - Conduct a trial run to refine tasks, instructions, and tools.
2. **Facilitate the Study**
   - Guide participants through the test scenarios without leading or influencing their behavior.

- Observe and record interactions, noting any difficulties, errors, or delays.

3. **Capture Data**
   - Use tools to record metrics like task completion time, error rates, and screen interactions.
   - Collect qualitative data through think-aloud protocols, where participants verbalize their thoughts.
4. **Administer Questionnaires**
   - Pre-test surveys gather demographic and background information.
   - Post-test surveys measure satisfaction, ease of use, and overall impressions.
5. **Maintain Neutrality**
   - Avoid prompting or assisting participants unless absolutely necessary to proceed.
   - Ensure a non-judgmental environment where participants feel comfortable.

3.6 Planning and Performing a Usability - Study Goals

1. **Identify Usability Issues**
   - Detect barriers that prevent users from achieving their goals effectively.
   - Uncover interface design problems, such as confusing navigation or unclear instructions.
2. **Measure User Performance**
   - Assess how efficiently and accurately users complete tasks.
   - Quantify key usability metrics like task success rate, time on task, and error rates.
3. **Evaluate User Satisfaction**
   - Understand how users feel about the product through surveys, interviews, or ratings.
   - Measure perceptions of ease of use, aesthetics, and overall experience.

4. **Validate Design Decisions**
   - Confirm whether the current design aligns with user needs and expectations.
   - Compare design alternatives to choose the most effective solution.
5. **Enhance Learnability and Efficiency**
   - Assess how quickly new users can learn to use the product.
   - Ensure experienced users can perform tasks efficiently and consistently.
6. **Prioritize Improvements**
   - Identify and prioritize critical issues that require immediate attention.
   - Provide actionable recommendations to improve usability.
7. **Ensure Accessibility**
   - Verify that the system is usable for a diverse audience, including individuals with disabilities.
   - Test compliance with accessibility standards (e.g., WCAG).
8. **Optimize User Engagement**
   - Evaluate how well the system holds users' attention and encourages continued use.
   - Identify factors contributing to user retention or abandonment.
9. **Support Business Goals**
   - Align the usability of the product with broader business objectives, such as increasing customer satisfaction, reducing support costs, or driving conversions.
10. **Inform Future Development**
    - Gather insights to guide future design and development efforts.
    - Establish a benchmark for comparing usability improvements over time.

3.7 Planning and Performing a Usability - User Goals

When planning and performing a usability study, the first key step is to define and understand the **user goals**. These goals are the specific objectives that users aim to achieve when interacting with a product or system. By identifying user goals, you can ensure that the product meets the users' needs, offering a better overall experience. Here's how you can approach it:

1. **Identify Target Users**
   - **Demographics**: Understand who your users are—age, gender, education, profession, and experience level.
   - **User Segments**: Divide users into segments based on behaviors, needs, or tasks they are trying to accomplish.
2. **Define User Goals**
   - **Functional Goals**: What tasks do users want to accomplish? For instance, users may want to complete a purchase, sign up for an account, or retrieve specific information.
   - **Experiential Goals**: How do users want the interaction to feel? This could include ease of use, speed, or pleasure.
   - **Contextual Goals**: Consider the environment and situations in which users will interact with the system (e.g., using a mobile app while commuting).
3. **Create User Personas**
   - Based on the target audience, create personas to represent different user types.
   - These personas help clarify user goals and preferences by creating a fictional, yet representative, user for each group.
4. **Map out the User Journey**
   - **Tasks**: Identify the tasks users need to complete to achieve their goals.
   - **Pain Points**: Highlight potential obstacles or challenges users might face during the process.

- **Touch points**: Identify points of interaction with the system, such as buttons, forms, or notifications.

5. **Set Priorities for User Goals**
   - Not all goals are of equal importance. Prioritize them based on frequency, importance, and how they impact the user's success in using the system.
   - Focus on the most critical tasks that allow users to achieve their main objectives.
6. **Collect Data**
   - **Interviews**: Speak directly with potential or existing users to gather insights into their goals.
   - **Surveys and Questionnaires**: Use structured questions to understand user preferences and needs.
   - **Observation**: Observe users as they interact with a similar product or prototype.
7. **Usability Testing**
   - Test the system with real users, keeping their goals in mind.
   - Gather feedback on whether the system helps them achieve their objectives efficiently and effectively.
   - Evaluate task success rates, time taken to complete tasks, and user satisfaction.

By focusing on user goals during your usability study, you ensure that the product is designed with the end user in mind, ultimately leading to a more intuitive and effective experience.

3.8 Metrics and Evaluation Method

Metrics and Evaluation Methods are critical in assessing the success and usability of a system. These help in quantifying the user experience, identifying areas for improvement, and ensuring the product meets user goals effectively. Here's an overview:

Key Metrics for Usability

1. Effectiveness

Measures the accuracy and completeness with which users achieve their goals.

- **Task Success Rate**: Percentage of tasks completed correctly.
- **Error Rate**: Number and type of errors users make during tasks.

2. Efficiency

Measures the resources expended in relation to the accuracy and completeness of goals achieved.

- **Time on Task**: Average time taken to complete a specific task.
- **Time to First Action**: Time taken for users to make their first interaction after encountering a task.
- **Navigation Steps**: Number of clicks or steps required to complete a task.

3. Satisfaction

Measures user perceptions and satisfaction with the system.

- **System Usability Scale (SUS)**: A standardized questionnaire providing a usability score.
- **Net Promoter Score (NPS)**: Measures how likely users are to recommend the product.
- **User Satisfaction Rating**: Collected via post-task or post-test surveys, often on a Likert scale.

4. Learnability

Assesses how easy it is for new users to learn the system.

- **Time to Learn**: Time taken to become proficient in completing key tasks.
- **Number of Help Requests**: Frequency of consulting help materials or support.

5. Error Frequency and Severity

Evaluates how often users encounter issues and the impact of those errors.

- **Number of Errors**: Total errors made by users during tasks.
- **Recovery Time**: Time taken to recover from an error.

6. Engagement

Evaluates how users interact with the system over time.

- **Retention Rate**: Percentage of users returning to the system after initial use.
- **Task Abandonment Rate**: Percentage of tasks started but not completed.

3.9 Evaluation Methods

1. **Quantitative Evaluation**

Focuses on collecting numerical data to measure performance.

- **A/B Testing**: Comparing two versions of a system to determine which performs better.
- **Analytics Tools**: Use tools like Google Analytics to track user interactions and behavior.
- **Benchmarking**: Comparing metrics to industry standards or competitor products.

2. **Qualitative Evaluation**

Gathers subjective feedback and insights about user experience.

- **Think-Aloud Protocol**: Users verbalize thoughts while interacting with the system.
- **Contextual Inquiry**: Observing users in their natural environment.
- **Interviews and Focus Groups**: In-depth discussions about user experiences and challenges.

3. **Usability Testing**

- **Moderated Testing**: A facilitator guides users through tasks and collects feedback.
- **Unmoderated Testing**: Users complete tasks on their own, often using online platforms like us ability Hub or User Testing.

- **Remote Testing**: Testing conducted via online tools to observe and record user interactions.

4. **Surveys and Questionnaires**
   - Tools like SUS, NPS, or custom surveys help gather user satisfaction and qualitative feedback.
   - Ideal for post-test evaluations to understand overall impressions.
5. **Heuristic Evaluation**
   - Expert evaluators assess the system against established usability principles (heuristics).
   - Identifies usability issues without direct user involvement.
6. **Eye Tracking and Heatmaps**
   - Measures where users look on the screen and how they navigate the interface.
   - Useful for evaluating visual design and information hierarchy.
7. **Cognitive Walkthrough**
   - Evaluators simulate the user's thought process to identify potential usability issues.

3.10 Usability –Participants

Characteristics of Participants

- **Representative of Target Users**: Choose participants whose demographics, behaviors, and needs align with the intended users of the system.
- **Experience Level**:
  - **Novice Users**: To test learnability and identify barriers for new users.
  - **Experienced Users**: To assess efficiency and advanced functionality.
- **Diversity**:
  - Include users with different levels of familiarity with similar products.

- Ensure inclusivity in terms of age, gender, cultural background, and accessibility needs.

3.11 Data Collection

Data Collection in usability studies involves gathering qualitative and quantitative information from participants to evaluate and improve a product's usability. A well-structured data collection process ensures you obtain reliable, actionable insights. Here's how to approach it:

1. Types of Data Collected

a. Quantitative Data

- Provides measurable, numeric insights about usability.
- Examples:
  - Task completion rates.
  - Time taken to complete tasks.
  - Number of errors or attempts.
  - System Usability Scale (SUS) scores.
  - Click or navigation paths.

b. Qualitative Data

- Captures subjective experiences, opinions, and feedback.
- Examples:
  - Verbal feedback during think-aloud protocols.
  - Observations of user behavior (hesitations, confusion).
  - Written responses in post-test surveys.

2. Data Collection Methods

a. Direct Observation

- Watch participants as they interact with the product.
- Record behaviors, challenges, and strategies used.
- Use tools like notebooks or pre-defined observation sheets.

b. Think-Aloud Protocol

- Ask participants to verbalize their thoughts while performing tasks.
- Helps uncover user reasoning, expectations, and frustrations.

c. Usability Testing

- Participants complete a series of tasks, either:
  - **Moderated**: With a facilitator present to guide and observe.
  - **Unmoderated**: Conducted independently by participants using tools like UsabilityHub or UserTesting.
- Collect task performance metrics and feedback.

d. Surveys and Questionnaires

- Administer before, during, or after the usability test.
- Examples:
  - Pre-test surveys: Gather background information about users.
  - Post-test surveys: Assess satisfaction (e.g., SUS, NPS).
  - Open-ended questions: Collect detailed feedback.

e. Analytics Tools

- Use heatmaps, click tracking, or session recording tools.
- Tools like Google Analytics, Hotjar, or Crazy Egg can provide insights on user behavior patterns.

f. Interviews

- Conduct one-on-one interviews to delve deeper into user experiences.
- Can be done before or after the test for additional context.

g. Screen and Video Recording

- Record user interactions with the interface to analyze later.
- Include facial expressions and voice (if relevant) for emotional cues.

h. Logs and System Data

- Collect system-generated data such as:
  - Error logs.
  - Time stamps.
  - Interaction counts.

Tools for Data Collection

- **Screen Recording**: Lookback, Camtasia, Zoom.
- **Surveys**: Google Forms, Typeform, SurveyMonkey.
- **Analytics**: Hotjar, Crazy Egg, FullStory.
- **Task Metrics**: Morae, Optimal Workshop.
- **Eye-Tracking Tools**: Tobii, GazePoint for visual behavior.
- Data Analysis:

Data Analysis in usability studies involves interpreting the collected data to identify patterns, insights, and areas for improvement. It combines quantitative metrics with qualitative feedback to provide a holistic view of the user experience.

1. Organizing Data for Analysis

- **Quantitative Data**: Compile metrics like task completion rates, time on task, and error rates into spreadsheets or databases.
- **Qualitative Data**: Transcribe notes, audio recordings, and observations. Organize responses into themes or categories.
- **Tools**: Use software like Excel, SPSS, NVivo, or R for data management and analysis.

2. Quantitative Data Analysis

Focuses on measurable, numeric data to identify trends and performance indicators.

Key Steps:

1. **Calculate Descriptive Statistics**
   - Task success rate (percentage of completed tasks).
   - Average time on task.
   - Frequency and severity of errors.
2. **Identify Outliers**
   - Detect anomalies in task completion times or error rates.
3. **Compare Against Benchmarks**
   - Use industry standards or previous studies as a baseline.
4. **Statistical Analysis** (if needed)
   - Perform tests like t-tests or ANOVA for comparing groups (e.g., novice vs. expert users).

3. Qualitative Data Analysis

Captures user perceptions, emotions, and experiences.

Key Techniques:

1. **Thematic Analysis**:
   - Identify recurring themes or patterns in user feedback.
   - Example: "Confusion about navigation" might emerge as a theme.
2. **Affinity Mapping**:
   - Group related observations and comments to identify common issues.
3. **Sentiment Analysis**:
   - Categorize feedback as positive, negative, or neutral to gauge satisfaction.
4. **Behavioral Insights**:
   - Analyze think-aloud data to understand user reasoning and challenges.

4. Combining Quantitative and Qualitative Data

- Cross-reference findings:
  - If a task shows a low success rate (quantitative), examine user feedback or recordings (qualitative) to understand why.

- Pair time-on-task data with observations of user hesitations or errors.

5. **Identifying Usability Issues**
   - Categorize problems based on:
     - **Frequency**: How often the issue occurs.
     - **Severity**: Impact on user experience (e.g., critical, major, minor).
     - **Persistence**: Whether the issue affects all users or specific groups.
   - Use a **Priority Matrix**:
     - High Frequency + High Severity = Top Priority for resolution.

3.12 Typical Studies of Usability Studies and their Corresponding Metrics

1. **Task-Based Usability Testing**
   - **Purpose**: Assess how well users can complete specific tasks using the system.
   - **Metrics**:
     - **Task Success Rate**: Percentage of tasks completed successfully.
     - **Time on Task**: Average time taken to complete a task.
     - **Error Rate**: Number of errors made per task.
     - **Help Requests**: Frequency of participants seeking assistance.
     - **Task Satisfaction Rating**: User-reported ease of completing tasks.
2. **Exploratory Usability Testing**
   - **Purpose**: Understand user behaviors, expectations, and challenges in early-stage designs or prototypes.
   - **Metrics**:
     - **Navigation Paths**: Routes users take to accomplish a task.

- **Think-Aloud Insights**: Observations of user thought processes.
- **Success Rate for Undefined Tasks**: How intuitively users approach ambiguous scenarios.
- **Feature Discovery Rate**: How many features users identify on their own.

3. **Comparative Usability Testing**
   - **Purpose**: Compare the usability of two or more designs, products, or systems.

   - **Metrics**:
     - **Task Completion Time**: Compare average time across versions.
     - **Error Rates**: Frequency of mistakes for each version.
     - **Satisfaction Scores**: System Usability Scale (SUS) or Net Promoter Score (NPS) for each option.
     - **Preference Rankings**: User-stated preferences between systems.
4. **Accessibility Testing**
   - **Purpose**: Ensure the product is usable for people with disabilities.
   - **Metrics**:
     - **Screen Reader Compatibility**: Percentage of tasks completed using assistive tools.
     - **Keyboard Navigation Efficiency**: Time and effort using only keyboard inputs.
     - **Error Rates for Assistive Technology Users**: Errors encountered by users with disabilities.
     - **WCAG Compliance**: Percentage adherence to Web Content Accessibility Guidelines.
5. **Remote Usability Testing**
   - **Purpose**: Test usability in users' natural environments without a moderator.

- **Metrics**:
  - **Completion Rates**: Tasks completed successfully in unsupervised conditions.
  - **Time on Task**: Average time per task.
  - **Click Tracking**: Number of clicks or taps needed to complete a task.
  - **Abandonment Rate**: Percentage of users who quit tasks before completion.

6. **First-Time Use or Learnability Testing**
   - **Purpose**: Evaluate how easily new users can learn and use the system.
   - **Metrics**:
     - **Time to First Task Completion**: Time taken to complete the first assigned task.
     - **Error Rates During First Use**: Mistakes made by first-time users.
     - **Retention Rate**: Percentage of users who remember learned processes in follow-up tests.
     - **Satisfaction After First Use**: User-reported ease and comfort with initial use.

3.13 Comparative Alternative Designs

Comparative Alternative Designs refer to evaluating and comparing different design options or iterations of a product or interface to determine which one best meets user needs and achieves usability goals. This approach is commonly used during the design phase to refine and optimize solutions.

1. **Purpose of Comparative Alternative Designs**
   - To identify the most effective design for usability and user experience.
   - To determine how different designs impact user behavior, satisfaction, and performance.
   - To inform decisions about the final design based on evidence rather than assumptions

2. Types of Comparative Studies

a. A/B Testing

- **Description**: Compare two versions of a single variable (e.g., Design A vs. Design B).
- **Use Case**: Determine which version performs better in terms of task success, user preference, or other metrics.
- **Example**: Testing two layouts for a product page to see which leads to higher conversion rates.

b. Multivariate Testing

- **Description**: Test multiple variables or design elements simultaneously (e.g., layout, color, navigation).
- **Use Case**: Explore interactions between design elements and their combined impact.
- **Example**: Comparing combinations of button color, text style, and page layout.

c. Side-by-Side Comparisons

- **Description**: Present users with multiple designs at once for direct comparison.
- **Use Case**: Collect qualitative feedback on which design users find more appealing or intuitive.
- **Example**: Comparing two dashboard designs for clarity and ease of use.

d. Sequential Testing

- **Description**: Present users with designs in sequence rather than side by side.
- **Use Case**: Evaluate designs independently to reduce bias from direct comparisons.
- **Example**: Testing one version of an app interface, followed by another, and asking users to rate each.

e. Iterative Comparisons

- **Description**: Test successive iterations of a single design over time.
- **Use Case**: Optimize a design incrementally based on user feedback.

- **Example**: Refining a search feature through three iterations of testing.

3. Key Metrics for Comparison

- **Quantitative Metrics**:
  - Task completion rates.
  - Time on task.
  - Error rates or frequency.
  - Conversion rates (e.g., sign-ups, purchases).
  - System Usability Scale (SUS) scores.
- **Qualitative Metrics**:
  - User preferences and opinions.
  - Verbal feedback from think-aloud sessions.
  - Observations of user behavior and frustration points.

4. Best Practices for Comparative Design Testing

a. Define Clear Objectives

- Specify what you want to compare and the success criteria (e.g., ease of navigation, visual appeal, task efficiency).

b. Use Controlled Variables

- Focus on one or a few design differences to isolate their effects (e.g., button placement or font size).

c. Select the Right Participants

- Ensure participants represent your target audience for reliable insights.

d. Randomize Presentation

- Randomize the order in which designs are shown to prevent order bias.

e. Collect both Quantitative and Qualitative Data

- Use metrics like completion time alongside user feedback for a holistic evaluation.

f. Analyze Data Objectively

- Use statistical analysis for quantitative data and thematic analysis for qualitative insights.

3.14 Comparing with Competition

1. Purpose of Comparing with Competition

- **Benchmarking**: Establish a baseline for usability by measuring your product against industry standards.
- **Competitive Analysis**: Understand how competitors address similar user needs.
- **Improvement Identification**: Discover areas where your product can outperform competitors.
- **Market Positioning**: Highlight unique value propositions in your product's design or functionality.

2. Key Aspects to Compare

a. Usability Metrics

- Task success rate.
- Task completion time.
- Error rates.
- System Usability Scale (SUS) scores.
- Learnability for first-time users.

b. User Experience

- Visual design appeal.
- Navigation ease and clarity.
- Accessibility features.
- Responsiveness and performance.

c. Features and Functionality

- Range of features offered.
- Efficiency of workflows (e.g., checkout process, account setup).
- Customizability and personalization.

d. Customer Support

- Availability of help resources (FAQs, chatbots, support teams).
- Ease of finding and using support.
- User satisfaction with support quality.

e. Emotional Impact

- Overall user satisfaction.
- Likelihood of recommendation (Net Promoter Score - NPS).
- Emotional connection with the brand or interface.

3. Study Design for Competitive Analysis

a. Participant Recruitment

- Recruit participants who represent your target audience.
- Ensure participants are familiar with competitors' products for more informed feedback.

b. Define Comparable Tasks

- Choose similar tasks across products to ensure a fair comparison.
- Example: Searching for a product, signing up for an account, or completing a purchase.

c. Collect Metrics

- Use the same usability metrics for all products.
- Record task performance, user preferences, and satisfaction ratings.

d. Use Surveys and Interviews

- Collect qualitative feedback on user perceptions of each product.
- Ask participants to compare the ease of use, design appeal, and overall satisfaction.

e. A/B Testing

- Present participants with your product and a competitor's version to directly compare usability and features.

3.15 Completing a Task or Transactions -Usability Metrics

Usability metrics are critical in assessing how effectively users can complete tasks or transactions within a system, such as a digital banking platform or any user interface.

These metrics help gauge the ease, efficiency, and satisfaction with which users interact with a system. Key usability metrics typically include:

5. **Task Success Rate**: Measures the percentage of tasks or transactions completed successfully by users. A higher rate indicates that the system is easy to use and users can achieve their goals.
6. **Time on Task**: The amount of time it takes for users to complete a task or transaction. Shorter times usually imply that the system is efficient, although in some cases, longer times may reflect a need for thoroughness or complexity.
7. **Error Rate**: The frequency of errors made by users during the task or transaction. A lower error rate is indicative of an intuitive and user-friendly design. It could include things like incorrect inputs, navigating to the wrong page, or failing to complete a transaction.
8. **Cognitive Load**: Measures the mental effort required by users to complete a task. This can be gauged using techniques like NASA-TLX (Task Load Index) or subjective user ratings. Systems that minimize cognitive load allow users to perform tasks without requiring too much mental effort.
9. **User Satisfaction**: This is often measured through surveys or interviews after a task is completed, asking users about their experience. Satisfaction can be quantified through tools like the System Usability Scale (SUS) or Net Promoter Score (NPS).
10. **Task Abandonment Rate**: The percentage of users who begin but do not finish a task or transaction. High abandonment rates can indicate that users are frustrated or find the process too complex.

11. **Learnability**: How easily new users can learn to use the system or complete a task. This metric can be evaluated by tracking the time it takes for users to perform tasks correctly on their first attempt.
12. **Retention Rate**: In the context of transactions, this measures how many users return to the system to perform tasks or transactions after an initial use. High retention indicates users find value in the service.

These usability metrics are essential for optimizing a system and ensuring that users have a smooth, efficient, and satisfying experience when completing tasks or transactions.

Evaluating the impact of subtle changes in a system or interface (such as a digital banking platform) involves understanding how small modifications influence user experience, behavior, and overall performance. These changes can range from minor design tweaks, such as altering button placement, font size, or color schemes, to slight adjustments in the process flow or features.

Here's a structured approach to evaluate the impact of subtle changes:

1. **Define Clear Objectives**
   - **What are the goals of the changes?** This could be improving task efficiency, reducing errors, enhancing visual appeal, or increasing user satisfaction.
   - **What specific metrics will be used to measure impact?** These can include usability metrics like task success rate, time on task, error rate, or satisfaction.
2. **A/B Testing (Split Testing)**
   - **What is A/B testing?** A/B testing involves comparing two versions of a system or interface: the original (A) and a modified version (B).
   - **How does it work?** Randomly assign users to either the original or the new version, and track their behavior and performance across the selected usability metrics.

- **Analysis**: Measure the differences between the two groups and determine whether the changes had a statistically significant impact on performance or user experience.

3. **Usability Testing**
   - **User Feedback**: Conduct usability testing where users are asked to complete tasks using both the original and modified versions of the system. Observe their behaviors and ask for subjective feedback (e.g., which version they prefer and why).
   - **Behavioral Metrics**: Track the time it takes to complete tasks, error rates, and the number of successful transactions on each version.
4. **Cognitive Walkthroughs**
   - **What is a cognitive walkthrough?** This is a method where usability experts simulate the user's journey through the system to identify potential problems caused by changes.
   - **How does it work?** Evaluators step through the interface and assess whether the change has made it easier or more difficult for users to accomplish their goals. They focus on identifying cognitive load, confusion, or obstacles introduced by the subtle changes.
5. **Surveys and Post-Task Questionnaires**
   - **System Usability Scale (SUS)**: This is a widely-used survey to evaluate the perceived usability of the system after a subtle change. Users rate their experience, and the results can show if there has been an improvement or decline in overall satisfaction.
   - **Net Promoter Score (NPS)**: Measures users' likelihood to recommend the service to others, indicating overall satisfaction and how changes affect user loyalty.

6. **Behavioral Analytics**
    - **Tracking User Interactions**: Use analytics tools to track how users interact with the interface, focusing on how the subtle changes affect user behavior. For example, did users spend more time on certain steps? Did they complete tasks more efficiently? Are there any shifts in user flow?
    - **Heatmaps**: These show where users click most often or how they move their mouse over the interface, helping identify areas that are either confusing or more engaging after a subtle change.
7. **Benchmarking Pre- and Post-Change Data**
    - **Before and After Comparison**: Collect baseline data (pre-change) on key performance indicators (KPIs) such as task success rate, user satisfaction, and time on task. Then, after implementing the change, compare the new data to the baseline.
    - **Statistical Significance**: Use statistical methods (e.g., t-tests, chi-square tests) to ensure that the differences observed are not due to random variation, especially when dealing with subtle changes.
8. **Long-Term Impact Assessment**
    - **Retention and Repeat Usage**: In some cases, the impact of subtle changes may not be immediate. Assess long-term retention rates or repeated task completion to see if the changes lead to more sustained user engagement.
    - **User Behavior Trends**: Observe whether there are gradual changes in how users interact with the platform over time, particularly if the subtle changes influence their trust or familiarity with the system.
9. **Contextual Inquiry or Interviews**
    - **User Interviews**: Conduct interviews with users to gain qualitative insights into their experience with the subtle changes.

- This helps uncover deeper insights into how the changes affected their overall perception of the system, beyond just task completion metrics.
- **Contextual Inquiry**: Observe users in real-world settings as they interact with the system. This can provide insights into whether the subtle changes are working as intended or if unexpected problems arise.

Conclusion

The impact of subtle changes can often be significant, though not always immediately apparent. A combination of quantitative (A/B testing, usability metrics) and qualitative (user feedback, interviews) approaches will provide a comprehensive view of how the changes have influenced the user experience. By carefully measuring both the objective outcomes and subjective perceptions, you can assess whether the changes improve the overall usability and effectiveness of the system.

UNIT – 4
WEB METRICS AND WEB ANALYTICS

4.1 Web Metrics and Web Analytics: An Introduction

Web Metrics refer to the quantitative data that is collected from websites to measure, analyze, and assess the effectiveness of various digital strategies, user engagement, and overall performance. These metrics help businesses understand how well their website is performing, identify potential issues, and make data-driven decisions.

Web Analytics is the process of analysing and interpreting the data collected from web metrics to gain insights into user behaviour, optimize website performance, and improve user experience. It involves tracking key performance indicators (KPIs), user actions, and interactions across a website.

4.2 Key Elements of Web Metrics

1. Traffic Metrics

These include the number of visitors to a website, session duration, page views, and bounce rates. They help determine how many people are visiting the site and how engaging the content is.

- *Visitors*: Total number of users who visit the site.
- *Sessions*: Total number of times users interact with the website.
- *Page Views*: The number of times a page is viewed by visitors.
- *Bounce Rate*: The percentage of visitors who leave the website after viewing only one page.

2. Engagement Metrics

These metrics help track how users interact with the website's content.

- *Average Time on Page*: The average amount of time users spend on a specific page.

- *Pages per Session*: The number of pages a user visits during a session.
- *Click-Through Rate (CTR)*: The percentage of users who click on a specific link or ad.
- *Conversion Rate*: The percentage of visitors who complete a specific action (such as signing up, making a purchase).

3. Behavioural Metrics

These metrics focus on how users navigate through a website.

- *Exit Pages*: The last pages users visit before leaving the site.
- *Heat maps*: Visual representations showing where users click or scroll on a page.

4. Acquisition Metrics

These metrics track where the website traffic is coming from.

- *Referral Traffic*: Traffic generated from external websites that link to the site.
- *Organic Search Traffic*: Visitors who arrive through search engines.
- *Paid Search Traffic*: Visitors from paid advertisements (such as Google Ads).

5. E-commerce Metrics

For e-commerce sites, these metrics focus on sales performance.

- *Revenue per Visitor (RPV)*: The total revenue generated divided by the number of visitors.
- *Shopping Cart Abandonment Rate*: The percentage of users who add items to their cart but leave before completing the purchase.

4.3 Tools for Web Analytics

1. **Google Analytics**: One of the most widely used tools, providing insights on traffic, conversions, user behaviour, and acquisition.
2. **Adobe Analytics**: A powerful analytics platform offering advanced reporting and analysis features.
3. **Mixpanel**: A tool that focuses on tracking user behaviour and engagement through detailed event tracking.
4. **Hotjar**: A tool that provides heat maps, session recordings, and user surveys to understand user behaviour.

4.4 PULSE Metrics:

"Pulse metrics" generally refers to key performance indicators or data points that provide a real-time snapshot of the health, progress, or performance of a system, organization, or process. They can vary depending on the context. Below are some examples in different scenarios:

1. Business/Organizational Context

- Employee Engagement: Surveys, Net Promoter Scores (NPS), or feedback participation rates.
- Customer Metrics: Customer satisfaction (CSAT), retention rate, or churn rate.
- Sales Performance: Daily sales figures, conversion rates, or sales pipeline health.

2. Healthcare

- Patient Vital Signs: Heart rate, blood pressure, oxygen saturation, or respiratory rate.
- Operational Metrics: Average wait time, bed occupancy rate, or patient satisfaction.

3. Digital Platforms or Apps

- User Engagement: Daily Active Users (DAU), Monthly Active Users (MAU), or time spent on the platform.
- Performance Metrics: Latency, uptime percentage, or error rates.

4. Education

- Student Performance: Attendance rates, test scores, or assignment completion rates.
- **Faculty Metrics**: Research output, student feedback, or professional development sessions.

5. Personal Fitness

- **Health Tracking**: Heart rate, sleep patterns, calorie burn, or steps taken.
- **Workout Metrics**: Duration, intensity levels, or recovery rates.

Key PULSE Metrics for Business and Technical Issues

1. Business Issues

- **Daily Active Users (DAU)**: Number of unique users engaging with the system daily to address business-related tasks or queries.
- **Monthly Active Users (MAU)**: Number of unique users engaging monthly to resolve business issues.
- **Ticket Volume**: Total number of business-related issues or support tickets raised.
- **Issue Resolution Time**: Average time taken to resolve business issues.
- **Escalation Rate**: Percentage of unresolved business issues escalated to higher support levels.
- **Customer Satisfaction (CSAT)**: Feedback score from users regarding the resolution of business issues.
- **First Contact Resolution (FCR)**: Percentage of business issues resolved on the first attempt.

2. Technical Issues

- **Technical Active Users (TAU)**: Users actively engaging with the system or support team for technical problems.
- **Error/Failure Reports**: Number of technical issues logged per day/week.
- **Resolution Time for Technical Issues**: Time taken to address bugs or technical glitches.

- **System Uptime**: Percentage of time the system remains operational and error-free.
- **Reoccurrence Rate**: Percentage of technical issues that reoccur after resolution.
- **Developer or Support Team Response Time**: Average time to respond to technical queries.
- **Bug Fix Deployment Frequency**: Rate of implementing fixes for technical issues.

4.5 Heart Metrics

HEART metrics is a user-centric framework developed by Google to evaluate the quality of user experience in digital products. It focuses on five dimensions: **Happiness**, which measures user satisfaction and sentiment through metrics like Net Promoter Score (NPS) and Customer Satisfaction Score (CSAT); **Engagement**, which tracks the frequency and depth of user interactions, such as Daily Active Users (DAU) and session duration; **Adoption**, which assesses how many new users start using the product or specific features over time; **Retention**, which evaluates how effectively the product retains users and prevents churn; and **Task Success**, which measures how efficiently and effectively users accomplish their goals, using metrics like task completion rate, time on task, and error rate. By aligning these metrics with specific business goals and user needs, teams can identify areas for improvement and iteratively enhance user satisfaction and product performance.

4.5.1 HEART Metrics - Success on user Behaviour Issues

1. Happiness

- **User Sentiment**: Use surveys or feedback forms to understand frustrations (e.g., "What did you find challenging in the app?").
- **Metrics**:
  - Increase in Net Promoter Score (NPS) after resolving behavioural blockers.

- Improvement in satisfaction ratings on specific tasks.

2. Engagement

- **User Interaction**: Measure how frequently users engage with features where behaviour issues were identified.
- **Metrics**:
  - Increase in feature usage rates after implementing improvements.
  - Reduced drop-off rates in workflows or tasks that were previously problematic.

3. Adoption

- **Behaviour Change in New Users**: Track if new users adopt corrected features or workflows more easily than before.
- **Metrics**:
  - Faster on boarding completion rates.
  - Higher percentage of new users successfully adopting improved processes.

4. Retention

- **Long-Term Behaviour**: Monitor if users return to the product after resolving behaviour issues.
- **Metrics**:
  - Reduced churn rate linked to behaviour-related pain points.
  - Increased retention rates for features or areas where friction was reduced.

5. Task Success

- **Behavioural Efficiency**: Assess how effectively users complete tasks after changes.
- **Metrics**:
  - Higher task completion rates in areas previously impacted by behavior issues.
  - Reduced error rates or retries on critical actions.
  - Shortened time-to-task completion.

4.6 On-Site Web Analytics: Off-Site Web Analytics

On-site web analytics and **off-site web analytics** are two complementary approaches to understanding and optimizing web performance. On-site web analytics focuses on analysing visitor behaviour within a specific website, tracking metrics such as page views, time on site, bounce rate, and conversion rates. This helps businesses understand how users interact with their website, identify bottlenecks, and optimize content or navigation. In contrast, off-site web analytics examines the broader context, including a website's visibility, traffic sources, and brand reputation on the internet. This involves monitoring metrics like search engine rankings, social media mentions, competitor comparisons, and referral traffic. While on-site analytics provide detailed insights into user behaviour within the website, off-site analytics offer a macro view of external factors influencing a website's success. Together, they form a holistic approach to web analytics, enabling organizations to improve both user experience and overall online performance.

4.7 Goal-Signal-Metric (GSM)

Goal-Signal-Metric (GSM) process is a framework used for aligning business objectives with measurable outcomes, ensuring that actions taken are in line with the overarching goals. It is widely used in various fields, including business management, product development, and performance tracking. Here's how the process works:

1. **Goal**

The first step is defining **clear, actionable goals** that the organization or team seeks to achieve. These goals should be specific, measurable, achievable, relevant, and time-bound (SMART). Setting goals helps to ensure that everyone is aligned on what needs to be accomplished and why.

Example: Increase monthly active users (MAUs) by 20% over the next 6 months.

2. Signal

Once the goals are set, the next step is to identify **signals** that indicate whether progress is being made toward achieving those goals. Signals are qualitative or quantitative indicators that give early indications of success or failure. They help in understanding if the actions being taken are moving the organization in the right direction.

Example: A signal for increasing MAUs could be an increase in user registrations or engagement with the app's features in the first month.

3. Metric

Finally, **metrics** are defined to measure and track the signals over time. These metrics provide concrete data that can be analyzed to assess whether the goal is being met and to guide decision-making. Metrics should be closely linked to the signals and help evaluate the efficiency and effectiveness of the actions being taken.

Example: Specific metrics for tracking MAUs could include the number of new sign-ups per week, average session duration, and the number of repeat visits in a given period.

4.8 Social Media Analytics: Social Media KPIs

Social media analytics involves the process of collecting and analysing data from social media platforms to assess performance, optimize strategies, and improve engagement. By measuring various metrics, businesses and organizations can understand their audience, refine content, and measure the success of campaigns.

Key Social Media KPIs (Key Performance Indicators)

1. **Engagement Metrics**
   - **Likes, Shares, and Comments**: Basic engagement actions showing how users interact with posts.

- **Engagement Rate**: The percentage of people who interacted with your content compared to those who saw it (calculated as [Total Engagements/Total Reach] * 100).
- **Post Interactions**: Total interactions across different types of posts (likes, comments, retweets, etc.).
- **Mentions**: Number of times your brand is mentioned on social platforms.

2. **Reach and Impressions**
   - **Reach**: The number of unique users who saw your post or content. It helps measure the breadth of exposure.
   - **Impressions**: The total number of times your content was shown, regardless of whether it was clicked.
   - **Follower Growth Rate**: The speed at which your social media following is increasing. (calculated as [New Followers/Total Followers] * 100).
3. **Traffic Metrics**
   - **Website Clicks**: The number of clicks on links to your website from social media posts.
   - **Referral Traffic**: Amount of website traffic originating from social media platforms.
   - **Conversion Rate**: The percentage of visitors who take the desired action (sign-ups, purchases, etc.) after being referred from social media.
4. **Audience Demographics**
   - **Age, Gender, Location, and Interests**: Understanding the demographic profile of your followers helps tailor content and improve targeting.
   - **Audience Growth**: Tracking the increase or decrease in followers, segmented by different demographics.
5. **Content Performance**
   - **Top Performing Posts**: Identifying which posts generate the most engagement, reach, or shares.

- **Video Views and Watch Time**: Measuring how long users watch videos, a critical metric for platforms like YouTube, Instagram, or TikTok.

6. **Customer Satisfaction Metrics**
   - **Sentiment Analysis**: Analyzing the tone of mentions or comments (positive, negative, neutral) to gauge public sentiment about your brand.
   - **Customer Response Time**: The average time it takes to respond to customer inquiries or comments on social media.
7. **Social Share of Voice (SOV)**
   - **SOV**: A measure of the total mentions your brand receives compared to competitors. It indicates your brand's visibility in relation to others in the same industry or category.
8. **Return on Investment (ROI)**
   - **Social Media ROI**: The measurable return on money and resources spent on social media activities, including conversions, leads, and revenue generated from social media campaigns.

4.9 Performing Social Media Analytics: Business Goal, Data Gathering, Analysis, Measure and Feedback

1. Business Goal Definition

The first step in social media analytics is to define clear and measurable business objectives. These goals serve as the foundation for all subsequent steps.

Common Social Media Goals

- **Increase Brand Awareness**: Expand reach and visibility of the brand.
- **Engagement Growth**: Boost likes, shares, comments, and interactions.
- **Lead Generation**: Attract potential customers and capture their information.

- **Sales and Conversions**: Drive traffic to websites or sales platforms for purchases.
- **Customer Support**: Address customer queries and improve satisfaction via social platforms.

Example Goal: "Increase website traffic from social media by 20% in the next quarter."

2. Data Gathering

Data gathering involves collecting raw information from various social media platforms and tools. This data can include quantitative metrics as well as qualitative insights.

Tools for Data Collection

- Native Platform Analytics: Facebook Insights, Twitter Analytics, Instagram Insights, LinkedIn Analytics.
- Third-party Tools: Hootsuite, Sprout Social, Google Analytics (for tracking referral traffic), and Brandwatch.
- Sentiment Analysis Tools: Tools like Mention or Talkwalker for assessing brand sentiment.

- **Types of Data Collected**:
  - **Engagement Metrics**: Likes, shares, retweets, comments, mentions.
  - **Reach and Impressions**: Number of users exposed to the content.
  - **Traffic and Conversions**: Click-through rates, referral traffic, and conversions from social platforms.
  - **Audience Demographics**: Age, gender, location, interests.
  - **Content Performance**: Metrics for specific posts, stories, or campaigns.

3. Data Analysis

Once data is collected, it needs to be processed and analysed to derive actionable insights. This step identifies trends, patterns, and areas for improvement.

Key Analytical Approaches

Quantitative Analysis

- Calculate engagement rates, click-through rates (CTR), and ROI.
- Identify top-performing posts and campaigns.

Qualitative Analysis

- Assess user sentiment through comments, reviews, and mentions.
- Understand audience preferences based on feedback.

Comparative Analysis

- Compare performance across platforms or against competitors.
- Benchmark results against industry standards or past campaigns.

Visualization Tools

- Use tools like Tableau, Power BI, or Google Data Studio to create visual dashboards for better understanding.

4. Measure and Evaluate Performance

Measurement involves comparing the analysed data against predefined goals to determine success.

KPIs (Key Performance Indicators)

- Engagement Rate = (Total Engagements / Total Reach) * 100
- Conversion Rate = (Conversions / Total Clicks) * 100
- ROI = (Revenue - Cost of Campaign) / Cost of Campaign * 100
- Customer Response Time: Average time to reply to customer queries.

Reporting

- Create detailed reports highlighting achievements, failures, and areas of improvement.
- Segment data by platform, demographics, or campaign for a granular view.

5. Feedback and Optimization

Feedback is critical to refine strategies and ensure continuous improvement in social media performance.

Implementing Feedback

- Adjust posting schedules, content types, or platform focus based on audience behaviour.
- Invest more in high-performing campaigns or platforms while revising underperforming ones.
- Address negative feedback promptly to improve brand sentiment.

Continuous Monitoring:

- Social media is dynamic; regular monitoring ensures agility in adapting to trends or audience behaviour.
- Use A/B testing to experiment with new strategies and evaluate their success.

UNIT-5
DATA ANALYSIS LANGUAGE AND TOOLS

5.1 Data Analysis Language and Tools –Meaning

Data analysis relies on a range of programming languages and tools tailored to diverse analytical tasks. Python is a popular choice, offering libraries like Pandas, NumPy, and Scikit-learn for data cleaning, visualization, and machine learning. R is widely used for statistical analysis and visualization, featuring tools such as ggplot2 and dplyr, particularly favoured in academia and research. SQL remains essential for querying and managing structured data in databases, while Julia is gaining traction for high-performance numerical and scientific computing. MATLAB is also a prominent tool for mathematical modelling and simulation. Together, these languages and tools empower analysts to extract insights and solve complex problems efficiently.

5.1.1 Programming Languages for Data Analysis

1. **Python**
   - Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow.
   - Widely used for data cleaning, visualization, and machine learning.
   - Popular for its ease of use and extensive library support.
2. **R**
   - Libraries: ggplot2, dplyr, tidyr, Shiny.
   - Especially strong in statistical analysis and visualization.
   - Preferred by statisticians and academic researchers.
3. **SQL**
   - Primary language for querying and managing databases.
   - Tools: MySQL, PostgreSQL, SQLite, Microsoft SQL Server.
   - Essential for extracting and manipulating structured data.

4. **Julia**
   - A high-performance language for numerical and scientific computing.
   - Libraries: DataFrames.jl, Plots.jl.
   - Gaining traction for computationally intensive tasks.
5. **MATLAB**
   - Used for mathematical modelling, simulation, and data visualization.
   - Common in engineering and academic research.

5.1.2 Tools and Platforms for Data Analysis

1. **Spread sheet Software**
   - Examples: Microsoft Excel, Google Sheets.
   - Best for small datasets and quick visualizations.
   - Add-ons like Power Query enhance functionality.
2. **Business Intelligence Tools**
   - Examples: Tableau, Power BI, QlikView.
   - Focused on data visualization and dashboard creation.
   - Simplify the communication of insights to stakeholders.
3. **Big Data Tools**
   - Examples: Apache Hadoop, Apache Spark.
   - Handle massive datasets for distributed computing and analytics.
   - Used for real-time processing and analytics.
4. **Data Visualization Tools**
   - Examples: D3.js, Plotly, Tableau, Matplotlib.
   - Specialized in creating charts, graphs, and interactive visuals.
5. **Machine Learning Platforms**
   - Examples: TensorFlow, PyTorch, RapidMiner, KNIME.

5.1.3 Ready-Made Tools for Web and Social Media Analytics

1. **Google Analytics**
   - **Purpose**: Website traffic analysis.
   - **Features**: Tracks user demographics, behaviour, and real-time data; measures campaign performance and website engagement.
   - **Best For**: Website owners and marketers.
2. **Hootsuite**
   - **Purpose**: Social media management and analytics.
   - **Features**: Tracks engagement, reach, and follower growth across multiple platforms; schedules posts and monitors trends.
   - **Best For**: Social media managers.
3. **Sprout Social**
   - **Purpose**: Social media analytics and customer engagement.
   - **Features**: Provides audience insights, content performance reports, and brand monitoring.
   - **Best For**: Medium to large-scale enterprises.
4. **HubSpot**
   - **Purpose**: Marketing, sales, and customer relationship management (CRM).
   - **Features**: Analyzes web traffic, SEO, social media performance, and campaign ROI.
   - **Best For**: Comprehensive inbound marketing.
5. **Buzz Sumo**
   - **Purpose**: Content and influencer analytics.
   - **Features**: Identifies trending topics, analyzes competitor content, and finds key influencers.
   - **Best For**: Content marketers.
6. **SEM rush**
   - **Purpose**: SEO and web traffic analytics.
   - **Features**: Tracks keyword rankings, backlink profiles, and site audits; provides competitor insights.
   - **Best For**: SEO specialists and marketers.

7. **Tableau**
   - **Purpose**: Data visualization and analytics.
   - **Features**: Processes large datasets for website and social media metrics; creates dynamic dashboards.
   - **Best For**: Data analysts and decision-makers.
8. **Brand watch**
   - **Purpose**: Social listening and analytics.
   - **Features**: Monitors online mentions, tracks sentiment analysis, and identifies brand trends.
   - **Best For**: Brand managers and PR teams.
9. **Zoho Social**
   - **Purpose**: Social media management and analytics.
   - **Features**: Provides multi-channel monitoring, custom reports, and audience targeting.
   - **Best For**: Small to medium-sized businesses.
10. **Adobe Analytics**
    - **Purpose**: Advanced web analytics.
    - **Features**: Offers real-time web tracking, predictive analytics, and customer segmentation.
    - **Best For**: Enterprises requiring in-depth analysis.

5.1.4 Key Google Analytics Metrics

Google Analytics provides a wide range of metrics to measure website performance and user behavior. Here are the key ones:

Audience Metrics

1. **Users**: The total number of unique visitors to the website within a specified period.
2. **New Users**: First-time visitors during the selected timeframe.
3. **Sessions**: A group of interactions by a user within a set time frame (default is 30 minutes).
4. **Session Duration**: The average amount of time users spend on the site.

5. **Bounce Rate**: The percentage of single-page visits where users leave without further interaction.

Acquisition Metrics

1. **Traffic Sources**: Breakdown of traffic origins (e.g., organic search, direct, referral, social).
2. **Referral Traffic**: Visitors coming to the website via external links.
3. **Campaign Performance**: Insights into the effectiveness of marketing campaigns using UTM parameters.

Behavior Metrics

1. **Page views**: Total number of pages viewed, including repeat views by the same user.
2. **Pages per Session**: The average number of pages a user visits during a session.
3. **Landing Pages**: The pages where users enter the site, offering insights into which pages attract the most traffic.
4. **Exit Pages**: The last pages users view before leaving the site.

Conversion Metrics

1. **Goals Completion**: Number of user actions that meet predefined objectives (e.g., form submissions, purchases).
2. **Conversion Rate**: The percentage of sessions that result in a goal completion.
3. **E-commerce Metrics**: Metrics like revenue, average order value (AOV), and product performance for online stores.

Real-Time Metrics

1. **Active Users**: The number of users currently browsing the site.

2. **Live Traffic Sources**: Real-time insights into how users are arriving at the site.

Dashboard and Social Reports in Google Analytics

Google Analytics offers **Dashboards** and **Social Reports** to help you visualize data and monitor performance across various channels. Here's an overview:

1. Dashboards

Dashboards are customizable panels in Google Analytics that provide a summary of key metrics and reports. They allow you to monitor the most important aspects of your website or campaign in one view.

Features

- **Widgets**: Add multiple widgets (e.g., timeline, pie charts, tables) to track different metrics like traffic, conversion rates, or bounce rates.
- **Custom Dashboards**: Create dashboards tailored to specific goals (e.g., SEO performance, e-commerce tracking).
- **Prebuilt Templates**: Use existing templates for quick setup (e.g., traffic analysis or social media performance).
- **Real-Time Data**: Monitor live user activity, pageviews, and sources.

2. Social Reports

Social Reports in Google Analytics help you measure the impact of social media on your website's traffic, engagement, and conversions. They provide insights into how social platforms contribute to your marketing goals and audience interaction.

Types of Social Reports

1. **Overview Report**
   - **Purpose**: Provides a summary of all social media traffic and its role in conversions.

- **Key Metrics**: Total social sessions, conversions, and assisted conversions.
- **Use**: Understand the overall contribution of social media channels to your site performance.

2. **Network Referrals**
   - **Purpose**: Shows traffic from individual social networks.
   - **Key Metrics**: Sessions, average session duration, pages per session.
   - **Use**: Identify which platforms drive the most engaged users.
3. **Landing Pages**
   - **Purpose**: Tracks the top landing pages that attract visitors from social media.
   - **Key Metrics**: Pageviews, social sessions, and bounce rates for each landing page.
   - **Use**: Optimize high-performing pages for better social media traffic.
4. **Conversions**
   - **Purpose**: Measures how social media contributes to website goals, including purchases or sign-ups.
   - **Key Metrics**: Assisted conversions, last-click conversions, and conversion value.
   - **Use**: Assess the ROI of social media campaigns.
5. **Plugins**
   - **Purpose**: Tracks user interactions with social sharing buttons on your site.
   - **Key Metrics**: Clicks on "Share" or "Like" buttons.
   - **Use**: Evaluate the effectiveness of social sharing tools.
6. **User Flow**
   - **Purpose**: Visualizes the path users take from social media to your website and beyond.
   - **Key Metrics**: Entry points, drop-offs, and navigation patterns.
   - **Use**: Understand user behavior to improve site design and content strategy.

Statistical Programming Language (R)

R is a powerful statistical programming language widely used for data analysis, visualization, and statistical modeling. Designed specifically for statistical computing, it offers an extensive library of packages, such as ggplot2 for data visualization, dplyr for data manipulation, and caret for machine learning. R excels in handling complex data analysis tasks, including regression, time series analysis, and hypothesis testing. Its open-source nature allows users to contribute custom packages, making it highly versatile for various research and business applications. Additionally, R integrates seamlessly with other tools like Python, SQL, and Tableau, making it an indispensable tool for statisticians, data scientists, and analysts worldwide.

ABOUT THE AUTHORS

Dr. M. Soundarya, Assistant Professor, School of Management Studies, Sathyabama Institute of Science and Technology, Chennai. She has Completed the Doctorate in Full time from the Department of Bank Management, Alagappa University, Karaikudi. She has done MBA (Banking), MBA (Marketing), M.Phil, B.Com.,CA, PGDVAT & Service Tax from Alagappa University. Her area of specialization is Banking, Insurance, Credit Management, Marketing and Systems. She has been worked as Guest Lecturer in the Department of Bank Management, Alagappa University. She has a five years research experience and Two years teaching experience. She has published more than sixty Journals in various reputed journals in Scopus, UGC Care Listed, ABDC Category Journals and Peer reviewed Journals. She has also participated and presented in National and International Conference in more than 30 colleges and also got Best paper Award. She has also published many articles in Hindu Business Line. She is strong in SPPS, AMOS and Power BI. Honored with Best Faculty Award on December 2023 in Knowledge Research Academy in Coimbatore. She acted as resource person for various institutions and she has been invited as a trainer for hands on training on SPSS and she has also published the books such as Principles and practice of banking, Financial Management and Indian Financial System.

Dr. P. Hameem Khan is an academician and consultant, he has master's degrees in Management, Computer Science, Industrial & Personnel management. He is doctorate in management from Jawharlal Nehru Technological University Anantapur. He has published numerous research papers in ABDC, Scopus and SCI indexed international journals. He worked as a freelancing ERP consultant in software industry for more than a decade. He is an advisor to many startup companies and editor for international journals. His research interests include Business Analytics, Block Chain technology and artificial intelligence in Human resource management.

Mrs. R. Geetha , Assistant Professor, School of Management Studies, Sathyabama Institute of Science and Technology, Chennai. She is pursuing Doctorate from Vellore Institute of Technology Business School, Vellore. She has completed MBA from D.G. Vaishnav College, Chennai and B. Com from Ethiraj College, Chennai. Her area of specialization is Marketing and Human Resource. She has more than 7 years of work experience in both industries and academics. She had also worked as a research associate for ICSSR funded project. She has published around 4 papers in Scopus indexed journals along with impact factor and ABDC. She has also participated and presented more papers in the International Conferences.

Mr. K. Kalidhass, Assistant Professor, School of Management Studies, Sathyabama Institute of Science and Technology, Chennai. He is pursuing Doctorate in Full time from the Department of International Business, Alagappa University, Karaikudi. He has Completed the Master of Philosophy in the Department of International Business, Alagappa University, Karaikudi and he did his MBA (Marketing) from SRM University, Kattankulathur, Chennai. His area of specialization is Management, Marketing and Systems. He has published more than six Journals in various reputed journals in UGC Care Listed and Peer reviewed Journals. He has also participated and presented in National and International Conference in colleges. He did publish two articles in Hindu Business Line. He had experience as one of the organizing committee members for ICSSR sponsored Ten days workshop, happened in Alagappa University.

ISBN
9789392151743

FIRST BOOK PUBLISHING
Madurai