

Prepared by Dr. Karim Ahmed

MKT631

Digital Analytics



The background features five vertical panels of varying heights and widths. The first panel on the left is dark blue. The second panel is dark brown and shows hands typing on a laptop. The third panel is dark brown and shows a laptop screen with a play button icon. The fourth panel is light brown and shows a target icon, a book icon, and a magnifying glass icon. The fifth panel is light brown and shows a document icon. The text 'MKT631 Digital Analytics' is centered on the right side of the slide.

MKT631 Digital Analytics

Lecture 1

Prepared by
Dr. Karim Ahmed

The Evolution of Digital Analytics and the Internet

Objectives:

- Define the field of digital analytics and explain its historical emergence.
- Differentiate between structured, semi-structured, and unstructured data.
- Describe the evolution of the World Wide Web from Web 1.0 through Web 4.0.
- Distinguish first-party from third-party data .
- Outline the major milestones in the history of digital analytics.
- Articulate the “three roles” of a data analyst.
- Assess the evolving expectations for analysts within organization.
- Analyze the debate over AI’s impact on the analyst’s profession.

Lecture Content

- **How Digital Analytics Began?**
- **Structured versus Unstructured Data**
- **The Evolution of the Internet and Web**
- **Introduction to First-Party and Third-Party Data**
- **Digital Analytics History**
- **The Era of Massive Data Growth**
- **The Three Roles of the Analyst**
- **The Evolution of the Analyst in the Firm**
- **Characteristics of the Successful Analyst**
- **Will AI Replace Analysts or Empower Them?**

How Digital Analytics Began?

Digital Analytics is the study of various forms of business data to improve the online experience of a business and its customers. The evolution of Digital Analytics took place simultaneously with the development of digital marketing. For our purposes, once Internet technology was developed, it was made to be accessible and extensible to the public. The Internet allowed search engines and social media to emerge as specific online marketing disciplines that students and practitioners could study and master. Also, the development of the Google search engine and its associated Web services supplemented this change and revolutionized the manner and ease with which we access and use data.

Structured versus Unstructured Data

More than 50 years ago, the data we had collected was mostly “structured.”

Structured data is data with a high level of organization, such as balance sheets (cash, accounts payable), medical devices (heart rate, blood pressure), or census records (birth, employment). The structure is provided in the collection of the data through certain objectives so that once the data is captured, it occupies a specified field. Structured data can be processed easily in a relational database or spreadsheet.

Structured versus **Unstructured** Data

Nowadays, we capture a great deal of “unstructured” or “semistructured” data.

Unstructured data is information that does not exist in a row and column database. To be able to capture unstructured data, we must modify and create special tools that manipulate the data into clear results. Otherwise, when left unmanaged, data can become overwhelming, making it difficult to organize and work with.

Structured versus Unstructured Data

There are three types of data from this perspective:

- data that is easy for data scientists and business analysts to organize and extract meaning from (**Structured**)
- data that is not (**Unstructured**)
- Data that is nearly prepared and ready to be organized, but it lacks the cleaning and organization necessary for it to be processed; for example, resolving misspellings and spacing errors in a column of data (**semi-structured**)

STRUCTURED DATA



UNSTRUCTURED DATA



Introduction to **First-Party** and **Third-Party** Data

First-Party Data:

- **Source:** First-party data is data collected directly from your own sources. These sources can include your website, mobile app, email marketing campaigns, and customer databases.
- **Ownership:** You have full control and ownership over first-party data since it originates from interactions with your brand.
- **Examples:** First-party data may include customer demographics, purchase history, website behavior, and email engagement metrics.
- **Usage:** It's valuable for personalizing user experiences, segmenting your audience, and measuring the effectiveness of your marketing efforts.

Introduction to **First-Party** and **Third-Party** Data

Third-Party Data:

- **Source:** Third-party data is data that is collected by external sources, not directly affiliated with your organization. These sources can include data brokers, social media platforms, and other companies.
- **Ownership:** Unlike first-party data, you do not have direct control or ownership over third-party data. It is typically purchased or acquired from these external sources.
- **Examples:** It can include data like consumer interests, income levels, location data, and more.
- **Usage:** Marketers use third-party data to improve their advertising strategies and reach a wider, more diverse audience.

The Evolution of the Internet and Web

The Internet, formerly called Arpanet (Advanced Research Projects Agency Network) by the Department of Defense (DOD), is a global network of interconnected networks that was created in 1983 as a way for researchers to quickly and freely communicate.

In 1990, the now famous computer scientist Tim Berners-Lee created the World Wide Web, which opened the network for everyone, nationally and internationally.

The Evolution of the Internet and Web

Internet Timeline

- 1950s: After WWII, the Cold War spawns the need for new ways to communicate more quickly, hence Arpanet originated for military uses.
- 1960s: Arpanet became available to researchers around the world.
- 1970s: Apple and IBM developed personal computers in the late 1970s, which appeal to a new audience who can run simple word processing and spreadsheet applications on them.
- 1980s: Transmission Control Protocol/Internet Protocol (TCP/IP) created as a basic communication language for the Internet. The National Science Foundation operated the backbone and banned commercial traffic.
- 1980s: Web 1.0, or the first stage in the World Wide Web, developed, which was entirely made up of static, non-interactive Web pages connected by hyperlinks.
- 1990s: Venture capital funding poured into untried and uniquely individual business models.
- 1991: Gopher developed as a TCP/IP for searching and distributing documents over the Internet. Hyper Text Markup Language (HTML) developed for creating Web pages and Web applications.

- 1993: Mosaic was the first freely available browser, which soon spawned Netscape. Mosaic supported Gopher, file transfer protocol, and network news transfer protocol.
- 1995: Yahoo! Search Engine debuted and initially used Google algorithms up until 2004 where it began investing in its algorithms.
- 1998: Google Search founded on September 4, 1998 by Larry Page and Sergey Brin. Google Search differentiated itself from other search engines by focusing on the relevance of pages based on patterns formed in hyperlinks.
- 2000–2002: Businesses failed during the 2000s, which led to the disappearance of venture capital funding.
- 2004–2009: The popularity of the WWW equated to a rapid growth in users, content, and sales, both B2C and B2B. As a result of the popularity, there is an emergence of new players, such as Facebook, Twitter, Foursquare, YouTube, and LinkedIn—the social networks we are familiar with today.

The Evolution of the Internet and Web

Web 1.0, “read-only web,” is an early version of the Internet and is static in nature. It has a one-to-many fashion of sharing information. In Web 1.0, Web developers and designers create websites and content for users to consume. The design of the site prevents users from contributing to the content or crafting a response. Web 1.0 users are only passive recipients of the information and content, which means they would be considered the readers, not contributors, of the content. Web 1.0 is another channel of narrow information distribution like other conventional one-to-many technologies, such as radio and television. Web 1.0 websites are only used for information presentation purposes and not used for generating information or content.

The Evolution of the Internet and Web

Web 2.0 is the version of the World Wide Web that most of us have experience of. It is characterized by the shift from static webpages to user-generated content and the emergence of social media and mobile technologies. Web 2.0 alters the WWW landscape by turning the Web into a collaborative ecosystem where users could create content, share ideas, and offer actual products and services. Web 2.0 allows two-way and many-to-many information flow and user-generated content. The content generated by users over social media platforms is known as user-generated content (UGC). Some tools of this collaborative ecosystem are podcasts, blogs, tags, social bookmarks, social networks, wikis, etc.

The Evolution of the Internet and Web

One thing to note is that Web 2.0 is not a technical standard or an update to the first standard, Web 1.0. However, it reflects the changes in the way people use the Web and how programmers design websites. Due to changing societal norms, webmasters and programmers develop platforms allowing users to create, monetize, and share their content.

The Evolution of the Internet and Web

Web 3.0 marks the era of a connected Web operating system where most software components (e.g., application programs and operating systems) and data processing reside on the Internet. Web 3.0 is smarter, quicker, and more reliable at connecting data, concepts, applications, and people. A critical dimension of Web 3.0 is the Semantic Web or Linked Data. The Semantic Web is known as the “web of data” (W3 2015) and aims to make huge amounts of data and the relationship among the datasets Web-available in a machine-readable format, such as Resource Description Framework (RDF) format, for applications to query it. RDF is a general-purpose language for representing information on the web (W3 2015). The World Wide Web became more personalized with an “on demand” approach.

The Evolution of the Internet and Web

The Semantic Web refers to a web where information is structured in a way that machines can read and understand it. This means that data can be linked together in a much more meaningful way, making it easier to find and use.

The Evolution of the Internet and Web

Web 4.0 is a term that refers to the future of the internet, where data is more accessible and easier to process by machines. Web 4.0 is also known as the symbiotic web, where humans and machines interact in harmony. Some of the features of Web 4.0 are:

- Artificial intelligence: Web 4.0 will use advanced AI technologies to create intelligent agents that can understand, learn, and generate content based on user preferences and needs.

The Evolution of the Internet and Web

- Big data: Web 4.0 will handle massive amounts of data from various sources and analyze them to provide insights and solutions for users and businesses.
- Cloud computing: Web 4.0 will use cloud services to store, process, and deliver data and applications across different devices and platforms.
- Internet of things: Web 4.0 will connect all devices and objects in the real and virtual world, enabling them to communicate and exchange information in real-time.

Web 4.0 is still an emerging concept and there is no exact definition of how it will be. However, it is expected that Web 4.0 will make the internet more user-friendly, efficient, and secure for everyone.

The Evolution of the Internet and Web

The symbiotic web is a term that refers to a possible future stage of the web's evolution, where humans and machines can interact in a mutually beneficial way. The symbiotic web would enable machines to learn from human behavior and feedback, and provide insights and recommendations that can assist in decision-making. The symbiotic web would also allow humans to communicate and interact with machines through natural language, brain-computer interfaces, biometric sensors, and other technologies. Some examples of potential applications of the symbiotic web are smart glasses and holograms.

The Evolution of the Internet and Web

Table 1.1 Web 1.0 to Web 4.0.

Web 1.0	Web 2.0	Web 3.0	Web 4.0
Read-Only	Read-Write	Personal, Portable, Extensible	Converged
Company & Business focused	Community, Social focused	Focused on the Individual	Personal Assistants (Siri/ Google Now Cortana)
Home Page	Blogs, Blogging, Wikis	Streaming media, Live streaming, Data Waves	Bots/Intelligent Agents/Drones

The Evolution of the Internet and Web

Web 1.0	Web 2.0	Web 3.0	Web 4.0
Own your content	Share your content	Collect & Content	Programmable Content & Content on Demand
Web Forms	Web Applications	Smart/ Autonomous Applications	Artificial Intelligence/ Algorithms
Web Directories and Folders	Content Tagging	User Behavior / Quantified Self	Private Cloud, Internet of Things
Page Views	Cost Per Click (CPC)	User Engagement via collected metrics	Cross Channel Measurement
Banner Ads	Interactive Ads	Behavioral Ads & Targeting	Programmatic Advertising
Britannica Online	Wikipedia	Semantic Web	Transparent Web/WikiLeaks
HTML and Web Portals	XML / RSS	RDF/RDFS/OWL	Quantum Computing, AI

The Evolution of the Internet and Web

The differences between Web 3.0 and Web 4.0 are:

- Web 3.0 is the last stage of the web's development, while Web 4.0 is a future stage that is still under speculation and research.
- Web 3.0 aims to make the web more efficient and intelligent by reusing and linking data across websites, while Web 4.0 envisions a web that is more embedded and connected to our brains, bodies, and environments.
- Web 3.0 enables users to access, analyze, and manipulate information in various formats, such as text, images, videos, and audio, while Web 4.0 could enable users to communicate and interact with the web through brain-computer interfaces, biometric sensors, augmented reality, and virtual reality.

The Evolution of the Internet and Web

The differences between Web 3.0 and Web 4.0 are:

- Web 3.0 incorporates artificial intelligence, natural language processing, and blockchain technologies to create more personalized, interactive, and decentralized web experiences, while Web 4.0 could create a web that is more autonomous, self-organizing, and adaptive, using artificial intelligence, machine learning, and the Internet of Things.
- Some examples of Web 3.0 applications are voice assistants, smart contracts, recommendation systems, and social networks, while some possible examples of Web 4.0 applications are neural implants, smart glasses, holograms, and ambient intelligence.

The Evolution of the Internet and Web

The advantages of Web 4.0 in digital analytics are:

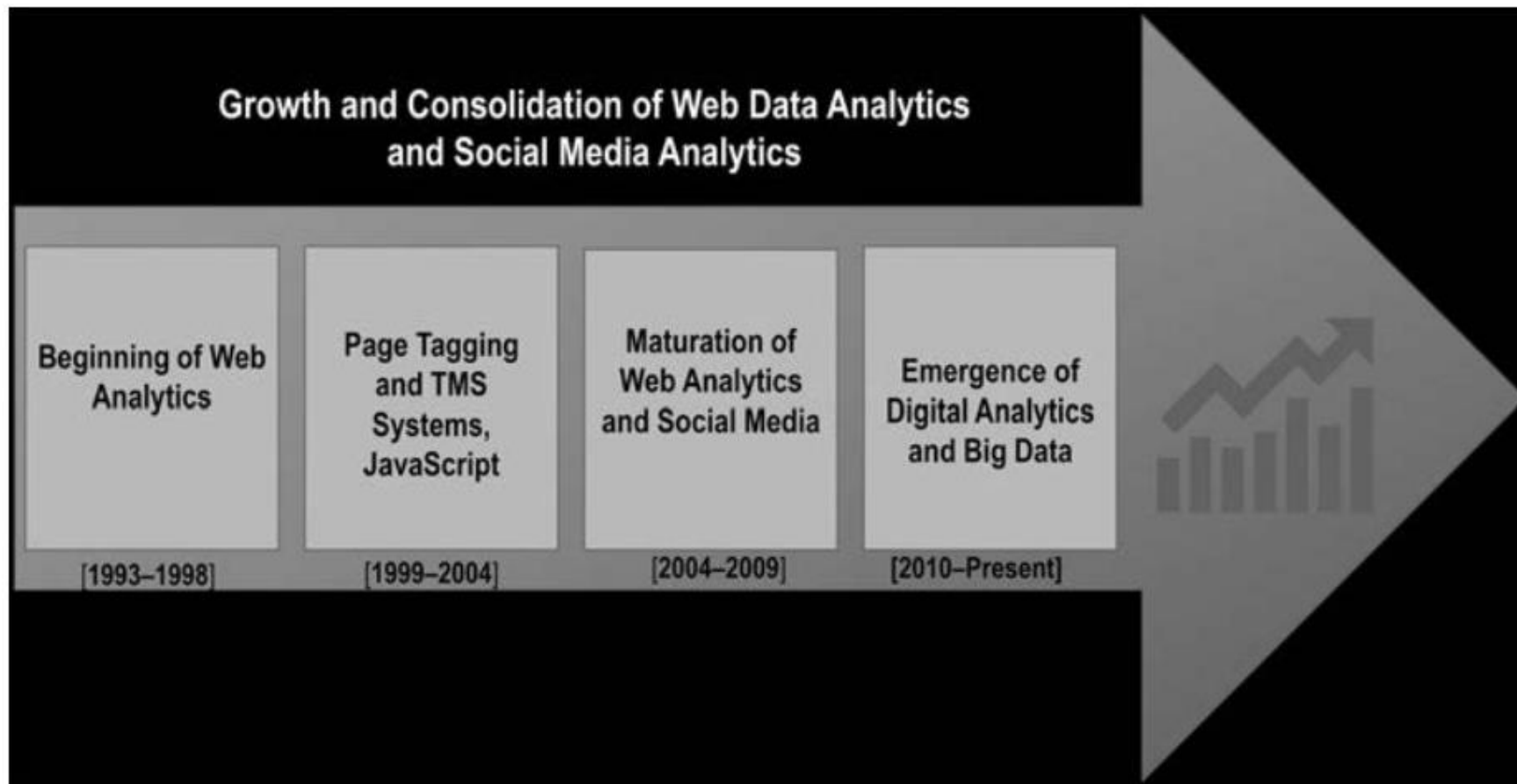
- Web 4.0 will enable a more **interactive** and **personalized** web experience, which will allow users to access and analyze data more easily and efficiently. For example, users can use voice or gesture commands to query data, or use biometric sensors to tailor data to their preferences and emotions.
- Web 4.0 will use **artificial intelligence** and **machine learning** to enhance the quality and accuracy of data analysis. For example, users can use natural language processing to extract insights from unstructured data, or use deep learning to discover patterns and trends from large and complex data sets.

The Evolution of the Internet and Web

The advantages of Web 4.0 in digital analytics are:

- Web 4.0 will support **software agents** that can work autonomously or collaboratively with other systems to solve problems. For example, users can delegate some of the data analysis tasks to intelligent agents that can learn from data and provide recommendations or solutions.
- Web 4.0 will use **blockchain** and **the internet of things** to ensure the security and reliability of data analysis. For example, users can use blockchain to verify the provenance and integrity of data, or use the internet of things to collect and transmit data from various sources and devices.

Digital Analytics History



Digital Analytics History

Digital Analytics is as old as the Internet.

Web Analytics is a solution for businesses that want to gather behavioral data from their websites. Web data helps businesses gain a better understanding of their customers and thus improve customer experience, which leads to benefits, such as more sales or conversions.

The existing platforms were developed for webmasters and IT professionals in the early to mid-1990s to keep track of downloads, file input or output, etc. The first Web Analytics platforms were developed to analyze Web server logs and provide diagnostic readouts.

Digital Analytics History

With the advent of JavaScript in the late 1990s, newer and more adaptive tracking methods using Web cookies evolved. **A Web cookie** is a small piece of data sent from a website and stored on the user's computer by the user's Web browser while the user is browsing.

Cookies are chosen by the Web server to identify users, prepare customized webpages, or to save site login information about users, so they do not need to re-enter the same login information every time they visit a website.

Web tracking shifted away from installing software to a third-party hosted solution.

Web Analytics no longer needed dedicated servers to collect or process the data. It became the basis of a new page-tagging approach using JavaScript beacon scripts.

The Era of Massive Data Growth

Every year, the amount of information being generated surpasses almost all the information humanity has created in its entire history, and most of the data being generated is unstructured. Welcome to the world of Big Data! Structured data is what we have in spreadsheets with rows and columns and the information already organized. Unstructured data is becoming just about everything else. Data is created so massively and quickly that new ways to process and store it have become necessary. **Data growth requires individuals who know how to work with the data.**

The Era of Massive Data Growth

The term “data is the new oil” reflects the current consensus that the most valuable asset an organization has is its data. Examples of turning data into dollars:

- Credit card companies detect unusual spending patterns of their customers using sophisticated algorithms that examine massive datasets. Customers are immediately alerted of the suspicious activities by phone or email. This activity saves customers and credit card companies billions of dollars a year.
- Major retailers use unstructured data collected for customer transactions to refine their online search engines and to encourage customers to buy more products.
- Food and beverage suppliers merge customers, logistics, and manufacturing data to improve their plant operations significantly.

The Era of Massive Data Growth

Rise of Online Advertising

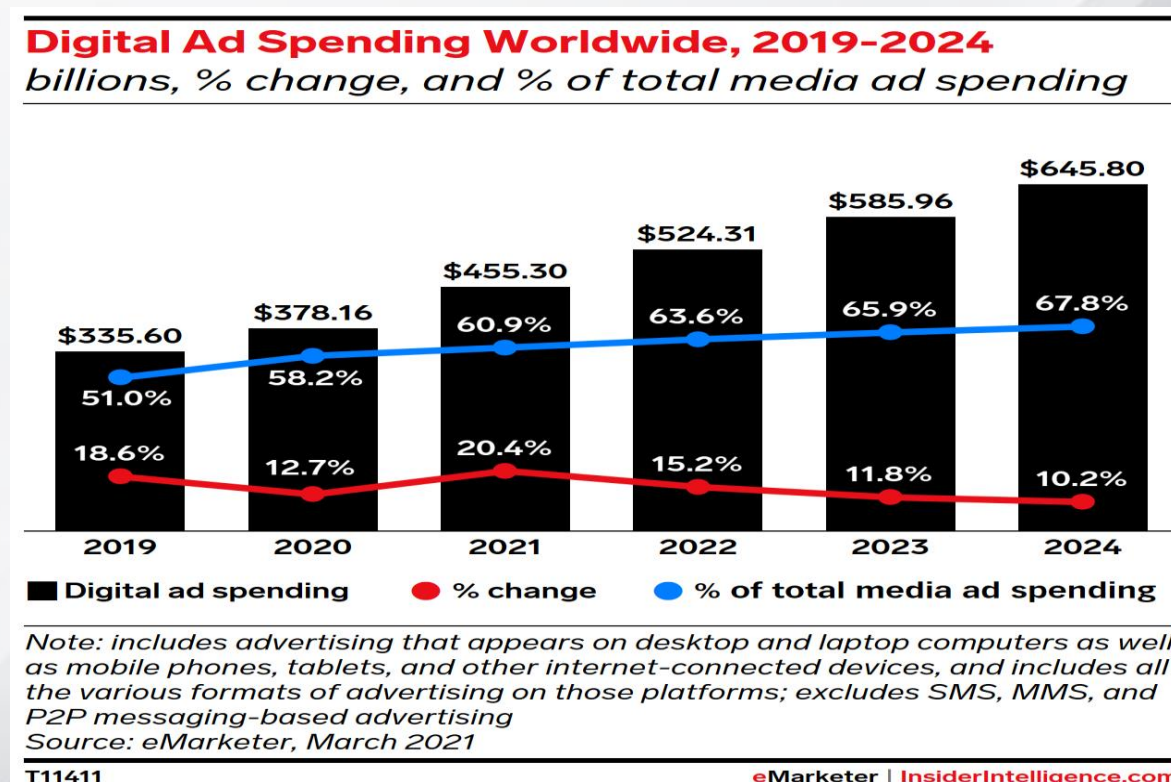
As popularity is increased with the Internet, online marketing and advertising rose to match with businesses' new strategy of “staying connected” by increasing the resources and capabilities a consumer could perform. In this context, “staying connected” refers to the emphasis we place on being reachable through online communications or various messaging applications at any given time.

Overall, the amount of digital advertising spending in 2019 is projected to double by 2024.

Digital advertising as a percentage of total advertising is continuing to grow, which highlights the need for increasing the analytics around the sector.

The Era of Massive Data Growth

Rise of Online Advertising



The Three Roles of the Analyst

Analysts should develop skills and expertise in three functional areas. Think of these functional areas as hats that successful analysts wear throughout any analysis project or endeavor. Each area of expertise has always been an important skill for successful data analysts, regardless of their industries or titles. These skills are interdependent: strength in one area will be throttled by weakness in another. As a result, analysts should seek to balance their abilities across the three roles.

- The ‘Data Strategist’
- The ‘Techie’
- The ‘Data Designer’

The Three Roles of the Analyst

- The ‘Data Strategist’

The first role the analyst must play is that of the “Data Strategist.” In this role, analysts must bridge the data and marketing worlds and understand their firms’ business operations, objectives, and environment. The successful Data Strategist understands how data can support the firm and can interpret data in the context of what they mean to the firm’s performance.

The Three Roles of the Analyst

- The ‘Techie’

The second role an analyst must play is that of the “Techie.” Here, the analyst is the data owner and must know how to collect and manage data in ways that ensure quality and promote efficiency. The successful Techie has the technical know-how to solve difficult data challenges and understands a breadth of data tools.

- The ‘Data Designer’

The last role is what we’ll call “Data Designer.” The successful Data Designer has strong creative skills and can turn insights into innovative visuals that people remember. This role requires the analyst to know how to express data stories and use tools to visualize those stories.

The Evolution of the Analyst in the Firm

As industries mature and explore new applications for data, the need for analysts is emerging in traditional businesses as well as in so-called “mission-driven roles” that have a societal or environmental benefit, such as those that focus on the application of data to advance health outcomes, energy efficiency, or nonprofit work.

Several trends have expanded analysts’ opportunities:

- Data analytics tools are becoming more accessible

The advancement of open-source and vendor-developed tools has improved accessibility for challenges that were previously reserved for complex and rigorous computer science solutions. While the market’s most sophisticated tools are far from being “point and click,” they’re evolving quickly and becoming easier for a wider audience of analysts to use.

The Evolution of the Analyst in the Firm

Several trends have expanded analysts' opportunities:

- Analytics know-how is increasingly becoming a requirement for all Roles

With many analytics teams increasing in size and scope, businesses increasingly face the challenge of connecting technical teams to other business unit stakeholders. As a result, companies expect nonanalyst business leaders to have a greater understanding of data analytics. In this way, they use their expertise to be representatives for the quantitative team and free up data scientists to do what they do best: advanced technical analysis.

The Evolution of the Analyst in the Firm

Several trends have expanded analysts' opportunities:

- The Evolution of Teaching Digital Data Analytics: Adapting to a Data-Driven World

The teaching of digital data analytics has evolved significantly in recent years. With the rapid advancement of technology and the increasing importance of data-driven decision-making in various industries, educational institutions have adapted to provide more accessible and comprehensive analytics training. This evolution includes the integration of analytics-focused curriculums and degree programs within traditional educational settings. This transformation in analytics education reflects the dynamic nature of the field and the growing demand for data analytics skills.

Characteristics of the Successful Analyst

Being a successful analyst requires more than mastering the three roles discussed above. It demands a mindset. Indeed, the successful analyst demonstrates a mindset characterized by three important traits: curiosity, optimism, and a healthy dose of boldness.

- *Curiosity*

Data analytics operates on the leading edge of marketing, advertising, and technology, and that's exactly what the analyst will love about it. A new search is always on the horizon: a new framework to develop, a new challenge to quantify, a new learning to discover. The analyst should embrace these challenges with the vigor, zeal, and inquisitive nature of a "Data Frontiersperson." Curiosity compels the successful analyst forward.

Characteristics of the Successful Analyst

- *Optimism*

The challenges data analytics pose aren't for the weak of mind or spirit. Analysts should expect to run into more brick walls than open doors. The arduous process of analyzing data means that failure is routine. I've seen analysts who abandoned their work when it became too challenging. More often than not, this was because they were limited in imagination, resilience, and zeal. Those who succeed do so because their tenacity and passion outweigh their frustrations. Perhaps, most importantly, analysts are successful because their optimism lets them believe solutions are there to be found.

Characteristics of the Successful Analyst

- *Boldness (Confidence)*

People who work with marketing data, especially data analysts, are deeply involved in using numbers for marketing purposes. The successful ones are those who go beyond what was previously believed to be the extent of what data can do. They fearlessly ask challenging questions and pursue answers with determination. They've earned the right to be confident. In my experience, this confidence leads analysts to pay close attention to the small things as they aim to deliver top-notch work. This focus on detail is what distinguishes exceptional analysis from merely decent analysis.

Will AI Replace Analysts or Empower Them? Debate pros and cons.

Argument 1: AI Will Replace Analysts

Pros of Replacement (Why It Could Happen):

1. Automation of Routine Tasks:

AI excels at repetitive tasks like data cleaning, report generation, and basic pattern recognition. Tools like AutoML (Automated Machine Learning) can build models without human intervention.

Example: A/B testing analysis that once took hours is now automated by platforms like Optimizely.

2. Cost Efficiency:

Companies may prefer AI tools over hiring junior analysts for tasks like dashboard updates or cookie-cutter insights.

3. Speed and Scale:

AI processes terabytes of data in seconds, outperforming humans in handling massive datasets (e.g., real-time social media sentiment tracking).

Will AI Replace Analysts or Empower Them? Debate pros and cons.

Argument 1: AI Will Replace Analysts

Cons of Replacement (Why It's Limited):

1.Lack of Contextual Judgment:

AI struggles with ambiguous scenarios (e.g., interpreting sarcasm in reviews or cultural nuances).

Example: An AI might flag a negative review saying, "This product is *sick!*" without understanding slang.

2.Ethical and Creative Gaps:

AI can't navigate ethical dilemmas (e.g., balancing profit goals with user privacy) or invent disruptive strategies.

Will AI Replace Analysts or Empower Them? Debate pros and cons.

Argument 2: AI Will Empower Analysts

Pros of Empowerment:

1. Augmented Decision-Making:

1. AI acts as a "co-pilot," providing analysts with deeper insights (e.g., Google Analytics 4's predictive churn metrics).
2. Analysts focus on *why* trends occur, not just *what* happened.

2. Upskilling Opportunities:

Analysts transition from data processors to strategic advisors. New roles emerge:

1. **AI Trainers:** Fine-tuning models for business-specific needs.
2. **Storytellers:** Translating AI outputs into actionable business narratives.

3. Simplifying Data Analysis for Everyone:

Low-code/no-code AI tools (e.g., Microsoft Power BI) let non-technical analysts run advanced analyses.

Will AI Replace Analysts or Empower Them? Debate pros and cons.

Argument 2: AI Will Empower Analysts

Cons of Empowerment:

1.Skill Gaps:

Analysts must learn AI tools (Python, ML basics) or risk obsolescence.

2.Over-Reliance Risks:

Blind trust in AI outputs can lead to errors (e.g., biased algorithms reinforcing flawed strategies).

Will AI Replace Analysts or Empower Them? Debate pros and cons.

The Middle Ground: AI Transforms, Not Replaces

The Future of Analysts:

1.Role Evolution:

1. **Pre-AI Era:** Analysts = data gatherers/reporters.
2. **Post-AI Era:** Analysts = strategic consultants + AI auditors.

2.Hybrid Skills in Demand:

1. Technical: Understanding AI limitations (e.g., "garbage in, garbage out").
2. Soft: Critical thinking, ethics, and communication.

3.Jobs Created by AI:

AI Ethics Officer, Predictive Analytics Strategist, Customer Journey AI Architect.

In this lecture, we discussed:

- How Digital Analytics Began?
- Structured versus Unstructured Data
- The Evolution of the Internet and Web
Web 1.0 – Web 2.0 – Web 3.0 – Web 4.0
- Introduction to First-Party and Third-Party Data
- Digital Analytics History
- The Era of Massive Data Growth
- The Three Roles of the Analyst
- Characteristics of the Successful Analyst
- Will AI Replace Analysts or Empower Them?



Thank You