Contents

[What is data analytics? 2](#_Toc160198887)

[Types of data analytics 2](#_Toc160198888)

[Predictive Data Analytics 2](#_Toc160198889)

[Prescriptive Data Analytics 2](#_Toc160198890)

[Diagnostic Data Analytics 2](#_Toc160198891)

[Descriptive Data Analytics 3](#_Toc160198892)

[Reasons for data analytics – for business 3](#_Toc160198893)

[Reasons for data analytics – general 3](#_Toc160198894)

[References 3](#_Toc160198895)

# What is data analytics?

Data analytics encompasses a multifaceted process that involves analysing raw data to uncover trends, answer questions, and generate actionable insights across various industries. At its core, data analytics involves the utilization of techniques and tools to identify patterns and trends within data sets, ultimately supporting informed decision making.

The process typically begins with descriptive analytics, where historical trends and data summaries are examined to gain an understanding of past events and behaviours. This phase sets the foundation for further analysis by providing context and insights into existing data.

Advanced analytics represents another critical component of data analytics, where sophisticated tools such as machine learning and deep learning are employed to extract insights and make predictions. These advanced techniques enable organizations to leverage vast data sets and computational power to uncover hidden patterns and trends that may not be apparent through traditional analysis methods.

The advent of machine learning tools. Along with the expansion of big data, has revolutionized the field of data analytics, allowing businesses to derive meaningful conclusions from complex and diverse data sources. This capability is particularly valuable in today’s data-driven business landscape, where organizations seek to gain a competitive edge by leveraging data to optimize processes, understand consumers, and identify new opportunities.

Overall, data analytics serves as a powerful tool for organizations looking to harness the potential of their data to drive better business outcomes. By employing a combination of techniques and tools, data analytics enables organizations to address specific challenges, make informed decisions, and ultimately achieve their strategic objectives.

## Types of data analytics

### Predictive Data Analytics

Predictive analytics involves using historical data to predict future outcomes. It utilizes techniques such as predictive modelling and statistical modelling to identify trends, correlations, and causations. For instance, businesses can use predictive analytics to forecast customer behaviour, sales trends, or market demand. This type of analytics helps organizations make informed decisions by anticipating potential future scenarios based on data patterns.

### Prescriptive Data Analytics

Prescriptive analytics goes beyond predicting outcomes; it recommends actions to achieve desired outcomes. It combines artificial intelligence and big data to analyse data and suggest the best course of action. This includes optimization techniques and random testing to determine the most effective strategies. Prescriptive analytics helps businesses answer questions like “What action should we take?” by evaluating various scenarios and suggesting the most favourable options based on data analysis.

### Diagnostic Data Analytics

Diagnostic analytics focuses on understanding why certain events occurred by analysing past data. It involves techniques such as drill-down analysis, data discovery, data mining, and correlation analysis. Diagnostic analytics helps uncover root causes and factors contributing to specific outcomes or events. It can be used for troubleshooting issues, identifying opportunities for improvement, and gaining insights into past performance. Diagnostic analytics is categorized into discovering and alerting potential issues before they occur and querying and drilling down into data for deeper insights.

### Descriptive Data Analytics

Descriptive analytics involves summarizing historical data to describe past events or trends. It addresses fundamental questions such as “how many, when, where, and what.” Descriptive analytics forms the basis of reporting and business intelligence tools, providing insights into historical performance and trends. It includes ad hoc reporting, which is generated on-demand to answer specific business questions, and canned reports, which are pre-designed reports containing information on specific subjects. Descriptive analytics helps organizations understand past performance and make informed decisions based on historical data.

## Reasons for data analytics – for business

1. **Informed Decision-Making:** Data analytics enables organisations to make informed decisions based on evidence and insights derived from data. By analysing past trends, current patterns, and potential future scenarios, decision-makers can identify opportunities, mitigate risks, and optimize strategies.
2. **Understanding Customer Behaviour:** Data analytics helps businesses understand customer preferences, behaviours, and trends. By analysing customer data, organizations can personalise marketing campaigns, improve product offerings, and enhance customer experiences to increase satisfaction and loyalty.
3. **Operational Efficiency:** Data analytics can optimise operational processes by identifying inefficiencies, streamlining workflows, and reducing costs. By analysing operational data, organizations can identify areas for improvement, automate repetitive tasks, and optimise resource allocation to improve efficiency and productivity.

## Reasons for data analytics – general

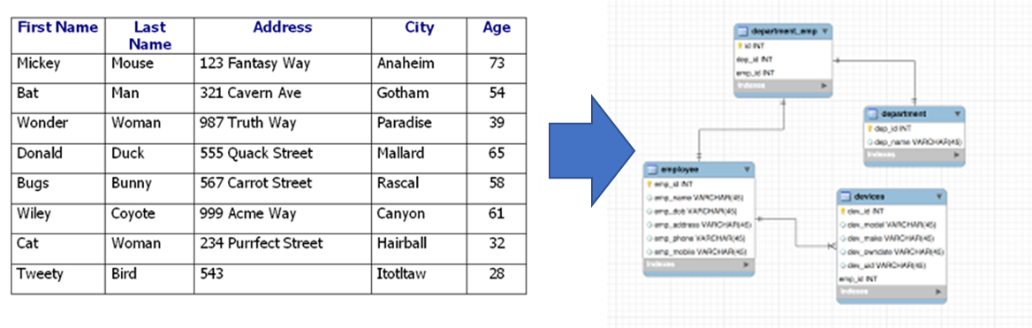
1. **Public Health:** Data analytics is employed in public health to monitor disease outbreaks, track population health trends, and identify risk factors for various diseases. It enables health agencies to allocate resources efficiently, implement targeted interventions, and prevent the spread of infectious diseases.
2. **Education and Student Performance:** Data analytics is used un education to analyse student performance data, identifying learning patterns, and personalise instruction to meet individual student needs. It helps educators assess student progress, identify at-risk students, and implement interventions to improve learning outcomes.
3. **Urban Planning and Infrastructure Development:** Data analytics is applied in urban planning to analyse demographic data, traffic patterns, and infrastructure usage to optimise city planning and development. It helps urban planners make informed decisions about transportation, housing, and public services to improve quality of life and sustainability in cities.

# A Brief History of Data Analytics

We have been analysing data since the dawn of human civilization, in Sumeria, where the country of Iraq is now located, we have discovered one of the first ever databases inscribed onto a clay tablet containing lists of ploughmen employed by the state. The wages of these ploughmen are calculated from this raw data thus the discipline of data analytics was born. Inscriptions on clay tablets were of course eventually replaced by paper and in the ninth century Sumeria is also where algebra and the decimal system were invented, both of which made the calculations and the structure of data more efficient. Data has been collected and analysed since ancient times for various purposes such as censuses, planning activities, agriculture, taxation and trade (Dercyk, 2020).



Prior to computers and machines it used to take over seven years for the United States Census Bureau to finish collecting data for their final census count and report, a tabulating machine was used at the end of the nineteenth century which made use of punch cards to speed up the census process to the point where it could be finished in a year and a half (Foote, 2021). The emergence of computing in the mid twentieth century has enabled the capabilities of data analytics to expand exponentially. Businesses and organisations started using computers to process and analyse data, mainframe systems were developed, and static reports were born, these static reports could be read but not edited once it was completed by the programmer, basic search and custom filtering was introduced but still tied to whatever the programmer thought was useful to present. Data within these static reports was often presented as a single list, the relational model was introduced, and networks of tables were utilised allowing for dynamic data collection (Codd, 1970).

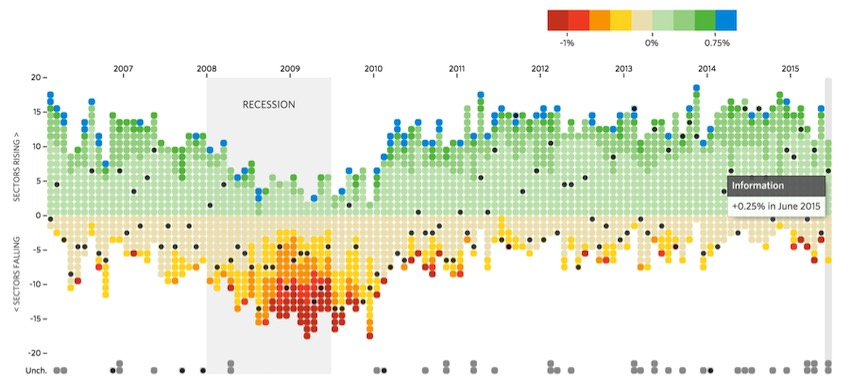
With the introduction of relational models in databases, the need for a dedicated standardised logical language arose, SQL (Structured Query Language) was designed in the 1970s at IBM by Donald D. Chamberlin and Raymond F. Boyce and offered a standard for managing and manipulating databases, in 1979 the first commercially available relational database management system, Oracle, implemented the language of SQL. The language was extremely concise and did not allow any room for errors thus the formal job of the trained data analyst was born. Relational databases and SQL allowed for data analysis on demand and is still widely used today even though they are unfortunately quite rigid (Dercyk, 2020).

Data mining began in the 1990s, it was the process of discovering patterns within large data sets, it was a direct result of the evolution of relational database technology and the leaps and reductions in cost in the field of data storage, allowing businesses and organisations to store more data and also analyse it quickly and efficiently to the point where a business could create models to predict the needs and behavioural patterns of their customers (Foote, 2021).

With the advent of the personal computer and the internet’s rise in popularity in the 1990s, these relational databases could not keep up with the demand of the time, the constant stream of data and the variety of data types from many sources led to non-relational databases also known as NoSQL (Not Only SQL), which was coined by Carlo Strozzi in 1998 but the modern concept of NoSQL truly began around 2009 when big data companies such as Amazon, Google and Facebook faced challenges managing massive amounts of data (Lacefield, 2018).

Big Data is a term coined by Roger Magoulas when he was describing the overwhelming volume of data being stored and processed in the mid-2000s, it was impossible to deal with using the corporate tools available. In 2005, Apache Hadoop was developed as an open-source framework which could process both structured and unstructured data from many different digital sources, it was followed up by Apache Spark and Apache Cassandra in the late 2000s, furthering the ability to process big data. These frameworks bypassed the need for all large data to be structured which would have been impossible given the size and complexity of the data (Dercyk, 2020).

Data Visualisation’s history can be traced back to ancient times when mankind used maps, charts and diagrams to represent information, however in the modern day, enabled by a surge in computing power after the rise of Big Data and the size and complexity of what was being stored grew the demand for Data Visualisation. Businesses and organisations had to be able to make sense of what was being stored often in the form of easily understandable visual representations such as charts and graphs (Dercyk, 2020).



# Why Use/Necessary: Prediction, Data Availability

Data analytics has become an indispensable tool in the modern world, driven by the need for informed decision-making, accurate predictions, and effective strategies across various sectors. The significance of data analytics lies in its ability to transform vast and complex datasets into actionable insights, providing a competitive edge and operational efficiency.

A group of people in a room with a large screen with graphs and charts

Description automatically generated

## Importance of Data Analytics

At the heart of modern strategic planning lies data analytics, enabling organizations to navigate the vast ocean of available data to unearth actionable insights. By systematically analyzing data, businesses can determine patterns, trends, and anomalies, transforming raw data into strategic intelligence. This transition from intuition-based to data-driven decision-making enhances accuracy and efficiency, mitigating risks, and capitalizing on opportunities with precision (Provost & Fawcett, 2013).

**A ship sailing on waves in front of a lighthouse

Description automatically generated**

## Prediction

The ability to forecast future trends and behaviours is a pivotal application of data analytics. In sectors like finance, healthcare, and retail, predictive analytics is a strategic imperative, informing decisions from market investments to patient care protocols and inventory management (Bose, 2009). Predictive analytics also extends to risk assessment, enabling organizations to identify potential pitfalls and strategize pre-emptively, safeguarding their interests.

Furthermore, predictive analytics underpins the shift toward customer-centric business models, enabling personalization and enhancing customer engagement by anticipating needs and preferences (Peppers & Rogers, 1997).

## Data Availability: The Fuel for Analytics

The proliferation of big data in the digital era has revolutionized the landscape of data analytics. The abundance of data, coupled with advanced analytical tools, has democratized the ability to glean insights, making analytics accessible to a broader range of businesses and sectors.

* **Technological Advancements**: The surge in data generation is matched by technological advancements that facilitate the efficient processing, storage, and analysis of vast datasets. This synergy between data availability and technology empowers organizations to harness the power of analytics, irrespective of their size or industry (Kshetri, 2014).
* **Real-Time Data Processing**: The value of data is often time-sensitive, particularly in fast-paced environments where decisions must be made swiftly and accurately. Real-time data processing allows organizations to act on insights almost instantaneously, adjusting to market dynamics, responding to emerging trends, and addressing challenges as they arise, providing a competitive edge in a rapidly evolving landscape (Bifet & Kirkby, 2009).
* **Strategic Decision-Making**: With the wealth of data at their disposal, organizations can make more informed, strategic decisions. Data availability enables a comprehensive view of business operations, market conditions, and customer insights, leading to better-aligned strategies and outcomes.
* **Innovation and Growth**: The accessibility of diverse data sets also fuels innovation, driving the development of new products, services, and business models. By leveraging data from various sources, companies can uncover unique insights, identify new opportunities, and foster growth in an increasingly data-driven world.

The digital age's big data explosion has made analytics indispensable for extracting actionable insights and competitive intelligence. The democratization of data, propelled by technological advancements, allows entities of all sizes to harness the power of analytics. Real-time data processing, critical in an era where promptness is as crucial as accuracy, provides organizations with a definitive edge, enabling rapid responses to market changes or emergent challenges.

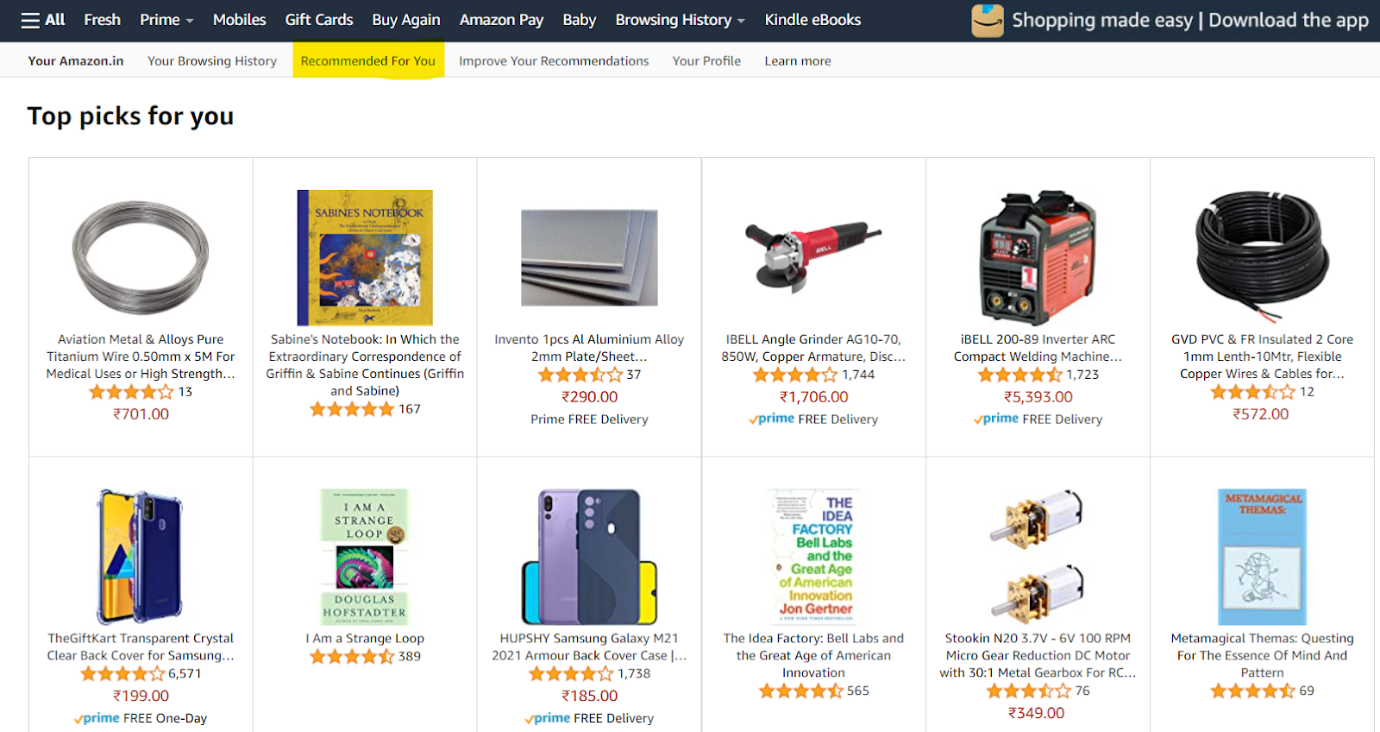
## Applications Across Diverse Fields

* **Healthcare**: Data analytics aids in disease predictive modelling, patient outcome analysis, and treatment optimization. During the COVID-19 pandemic, analytics played a vital role in tracking infection rates and guiding public health policies.
* **Finance**: Financial institutions use analytics for risk assessment, fraud detection, and customer segmentation, enhancing security and operational efficiency.
* **Marketing**: Analytics helps understand customer preferences and tailor marketing strategies, boosting engagement and ROI.
* **Retail**: Retailers employ analytics for inventory management and demand forecasting, improving customer satisfaction and sales.

## Real-World Example: Amazon's Use of Data Analytics

### Amazon's recommendations

Amazon's recommendation engine exemplifies data analytics in action. By analyzing customer behaviour and preferences, Amazon personalizes product suggestions, boosting user experience and sales, and illustrating predictive analytics' tangible benefits in enhancing customer satisfaction and revenue growth.



### Walmart's Inventory Management

Walmart's analytics-driven inventory management optimizes stock levels based on predictive data analysis, enhancing efficiency and customer satisfaction.

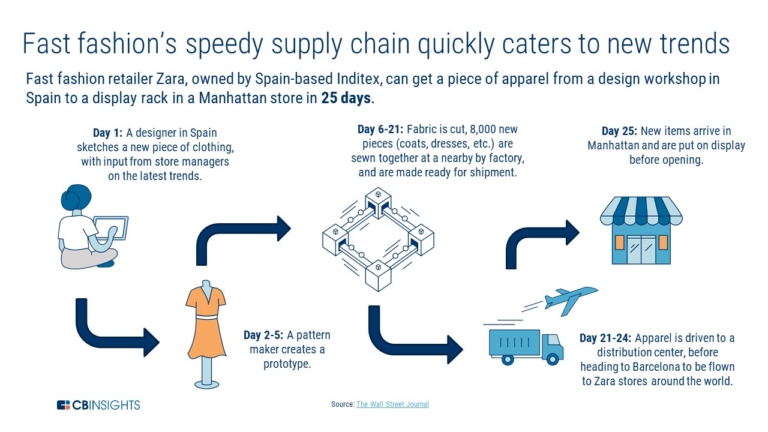
### Uber's Dynamic Pricing

Uber employs analytics for dynamic pricing, adjusting real-time prices to balance driver supply and user demand, demonstrating analytics' strategic application in service industries.



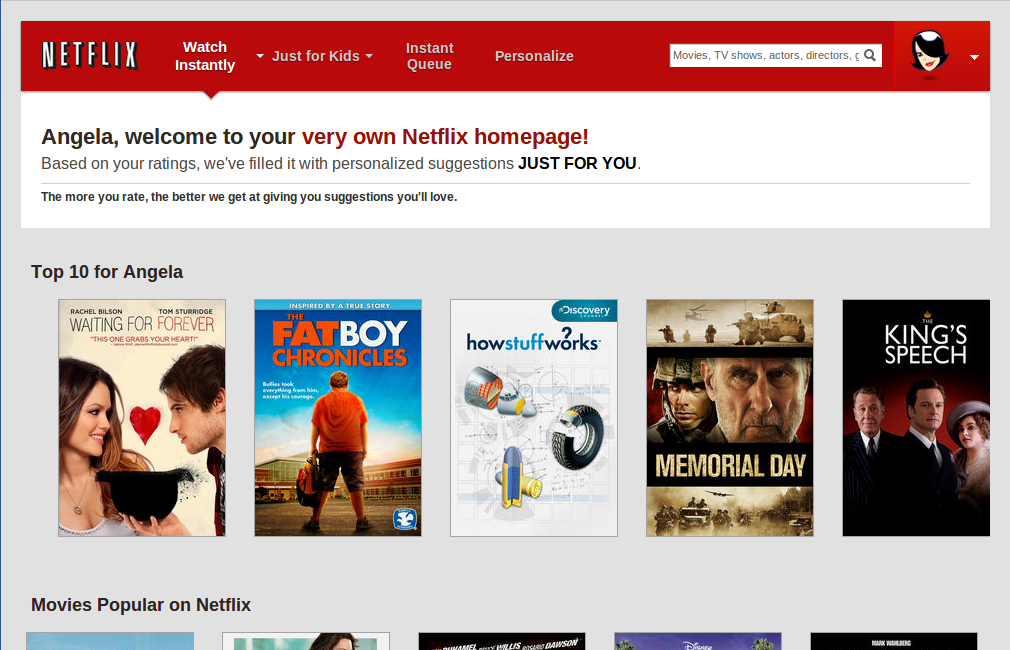
### Zara's Fast Fashion Strategy

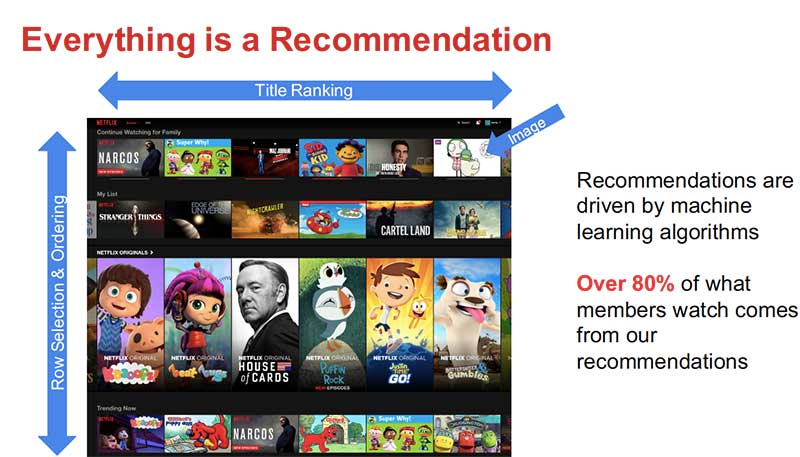
Zara uses analytics to adapt rapidly to market demands, showcasing its pivotal role in the retail sector's competitive strategy.



### Netflix's Personalized Recommendations

Netflix's recommendation engine, powered by analytics, personalizes viewer content, enhancing satisfaction and retention, and highlighting analytics' impact on entertainment.





# What fields is Data Analytics applied to?

If we were to write down all of the industries in which data analytics is applied to, we would not have enough paper to do it. Almost every industry nowadays uses it. Many businesses use it to make informed decisions, manufacturers use it to look at sales data to analyse which designs to retire and which to keep, administrators may use it to look at inventory data to check what they need to order.

All of these applications improve the efficiency of workers and the accuracy of what they produce, as well as saving money and other resources. These can give a huge advantage in the market for small companies. (Data Analytics, no date)

## How is it applied?

There are four main types of data analysis techniques, which are descriptive, diagnostic, predictive and prescriptive. (Data Analytics: What It Is, How It’s Used, and 4 Basic Techniques, no date)

Descriptive analysis is used to understand what has happened in the past and why it happened. Some examples of this would be sale performance, fraud detection and product demand forecasts.

Diagnostic analysis ask why things happened, investigates what events lead to an event and answers why that event occurred. This is very useful when we want to prevent something from happening again.

Predictive analysis is used to predict what will happen in the future, and it uses existing data to achieve just this. This is very wise and past behaviour can be used to generate a good guess regarding what would happen under different circumstances. Some examples include customer pricing or retail sales forecasting.

Prescriptive analysis is predictive analysis taken further, made to design actions to take in the future based on past data and trends. It is very useful when we are tyring to optimise resources or look for new business opportunities. Some examples include launching a new product line or send a targeted ad to particular customers.

These four ways to apply data analytics are very useful in a lot of industries and purposes and is used by many companies nowadays. It is a critical part of business success. It also helps tremendously in improving efficiency by identifying weaknesses in their business models. Making better decisions is one of the most popular uses, understanding what has happened in the past and what can happen in the future can give a huge advantage to businesses. (‘Data Analytics in Business: A Complete Overview - Caltech’, 2024)

In the area of business management it is also crucial to perform data analytics. Some examples are to gain customer insights. gain competitive advantage and make informed decisions. The following are some examples of how it is used (Data Analytics: What It Is, How It’s Used, and 4 Basic Techniques, no date; Data Analytics, no date)

In customer service it follows some critical tasks :

* Providing customers with personalised content and precisely customised recommendations
* Identifying common complaints from customers about certain items or services
* Provide self-service options to reduce support costs.
* Improve issue resolution by understanding customer history and needs.
* Predicting future client purchases - Automating payment and fraud detection processes.

For marketing and sales the following are used:

* Evaluating advertising and marketing campaigns
* Choosing the best product combination for a customer
* Obtaining the best price for a product or bundle
* Identifying customers who are most likely to respond to an offer
* Discovering new markets for existing goods and services

For healthcare the following are used:

* Forecasting patient admissions to ensure optimum resource allocation.
* Predicting disease outbreaks using demographic and environmental data.
* Helping physicians make diagnosis and treatment decisions by assessing patient data and recommending best practices.
* Alerting to probable hazardous medication interactions or allergies.
* Identifying high-risk patient populations and initiating preventive interventions.
* Tracking health patterns and outcomes across communities to help shape public health strategies.
* Improving hospital operations and resource utilisation by analysing patient flows and anticipating demand

# Payment fraud analytics

Payment fraud occurs when a fraudster obtains sensitive payment information from customers or businesses and uses it for financial gain. This includes but is not limited to stealing credit card details individually or through data breach of a company which stores such information.

In 2022, U.S. businesses faced a staggering loss of $1.59 billion due to payment fraud(Sift, 2024). It is therefore in the interest of businesses to invest in ways to mitigate these losses.

Data analytics is being used extensively and increasingly to do just time in a number of interrelated ways.

Patterns in payment behaviour can be collected by financial institutions themselves since they record the information as they process it. There are of course enormous amounts of data to be processed. There is also a very small number of actual examples of fraudulent behaviour which is comparable to finding a needle of undefined shape and size in a haystack(Boiarskaia, Albert and Lee, 2019). That said, details such as transaction amount, geographic location, time and frequency of transactions, preferred merchant types, and type of purchase can be recorded by financial institutions interested in examining normal behaviour and fraudulent behaviour patterns.

Unsupervised machine learning techniques are currently used to identify patterns that deviate from normal behaviour without labelled data.

Profiles for individual credit card holders can be created by studying their normal payments(Budd, 2016).

Supervised machine learning can be used to train a detection model using the historical data available to payment processors. This in turn can be used when processing payments to permit or deny transactions in real-time.

In this way, anything that falls outside the norm for an individual card holder, such as a sudden increase in transaction frequency, large transactions, or transactions in unfamiliar locations, may result in prohibited transactions.

Below is an example of a decision tree which can be used to monitor, permit or deny payments made by individual card owners.

A diagram of a network

Description automatically generated

Each node in the decision tree asks the question, is this behaviour typical of an individual customer? In each case, if the answer is no, the payment will not proceed.

Of course, this is a simplistic view. More accurately, nodes would be assigned an aggregating value depending on their outcome, which would be weighed against a predefined threshold. If the final value is greater than the threshold, the payment will not proceed. It might look something like the graphic below (Fraud.net, 2024).



As mentioned before, also a very small amount of data containing actual fraudulent behaviour exists. To discern normal behaviour patterns experts will define with a set of rules based on what fraud normally looks like. This has led to a rule-based system consisting of predefined criteria determining whether transactions will be allowed or denied.

As new methods of committing fraud emerge, constant analysis is required so that detection systems evolve to meet current demands. Continuous analysis of current data takes place to try to stay ahead of fraudsters. At the same time, considerable collaboration between financial institutions, technology organisations and others also contribute to an ever-evolving rules list.

This has resulted in innovations like two-factor and biometric authentication being used to avoid fraud. For instance, relatively recently many institutions have begun using push notifications on customers mobile devices to confirm the validity of payments.

By combining these approaches, financial institutions can create robust credit card fraud detection systems that effectively identify and prevent fraudulent activities while minimizing false positives. The integration of advanced technologies, such as artificial intelligence and machine learning, further enhances the accuracy and efficiency of these systems.

# Conclusion

Data analytics' necessity and utility are clear across sectors, offering predictive insights, informed decision-making, and effective data utilization. Real-world examples from Amazon, Walmart, Uber, Zara, and Netflix underscore data analytics' relevance and impact, providing valuable insights for organizations leveraging this tool.

# **References**

2U Inc., 2023. Master's in Data Science with edX. [Online] Available at: <https://www.mastersindatascience.org/learning/what-is-data-analytics/> [Accessed 1 March 2024].

Cambridge Spark, 2023. Cambridge Spark. [Online] Available at: <https://www.cambridgespark.com/info/7-benefits-of-data-analytics-in-higher-education> [Accessed 1 March 2024].

CompTIA Inc., n.d. CompTIA. [Online] Available at: <https://www.comptia.org/content/guides/what-is-data-analytics#DataAnalyticsguide> [Accessed 1 March 2024].

CompTIA Inc., n.d. CompTIA. [Online] Available at: <https://www.comptia.org/content/articles/how-is-data-analytics-used-in-health-care#:~:text=Data%20analytics%20in%20health%20care,organizations%20and%20boost%20business%20intelligence>. [Accessed 1 March 2024].

KeyStone MasterStudies, 2018. KeyStone MasterStudies. [Online] Available at: <https://www.masterstudies.com/articles/four-ways-data-analysis-can-improve-society> [Accessed 1 March 2024].

Oracle, 2024. Oracle. [Online] Available at: <https://www.oracle.com/ie/business-analytics/data-analytics/> [Accessed 1 march 2024].

Codd, E. F. (1970). A Relational Model of Data for Large Shared Data Banks. Information Retrieval Vol 13, 377-387.

Dercyk, B. (2020, September 15). A brief, brief history of Data Analytics. Retrieved from LinkedIn: https://www.linkedin.com/pulse/brief-history-data-analytics-deryck-brailsford-%E5%AD%99%E5%BE%B7%E7%91%9E

Foote, K. D. (2021, September 20). A Brief History of Analytics. Retrieved from Dataversity: https://www.dataversity.net/brief-history-analytics/

Lacefield, J. (2018, August 29). The Evolution of NoSQL. Retrieved from Datastax: https://www.datastax.com/blog/evolution-nosql

Bose, R. (2009). Advanced analytics: opportunities and challenges. Industrial Management & Data Systems, 109(2), 155-172.

Bifet, A., & Kirkby, R. (2009). Data stream mining: A practical approach. The University of Waikato.

Kshetri, N. (2014). Big data’s impact on privacy, security and consumer welfare. Telecommunications Policy, 38(11), 1134-1145.

Peppers, D., & Rogers, M. (1997). Enterprise one to one: Tools for competing in the interactive age. Currency Doubleday.

Provost, F., & Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.

Data Analytics (no date) Corporate Finance Institute. Available at: https://corporatefinanceinstitute.com/resources/data-science/data-analytics/ (Accessed: 4 March 2024).

‘Data Analytics in Business: A Complete Overview - Caltech’ (2024), 27 February. Available at: https://pg-p.ctme.caltech.edu/blog/data-analytics/data-analytics-in-business (Accessed: 4 March 2024).

Data Analytics: What It Is, How It’s Used, and 4 Basic Techniques (no date) Investopedia. Available at: https://www.investopedia.com/terms/d/data-analytics.asp (Accessed: 4 March 2024).

Boiarskaia, E., Albert, N. and Lee, D. (2019) Detecting Financial Fraud at Scale with Decision Trees and MLflow on Databricks, Databricks. Available at: <https://www.databricks.com/blog/2019/05/02/detecting-financial-fraud-at-scale-with-decision-trees-and-mlflow-on-databricks.html> (Accessed: 4 March 2024).

Budd, J. (2016) ‘Modelling credit card usage for individual card-holders’.

Fatima, A. (2022) ‘Data Analysis for detecting Credit Card Fraud’, Medium, 1 October. Available at: <https://aliya-fatima.medium.com/data-analysis-for-detecting-credit-card-fraud-f55d33c091b3> (Accessed: 8 February 2024).

Fraud.net (2024) Fraud Score, Fraud.net. Available at: <https://fraud.net/d/fraud-score/> (Accessed: 4 March 2024).

Gautam, A. (2023) Data Analytics Case Studies: Success Stories from Various Industries. Available at: <https://www.linkedin.com/pulse/data-analytics-case-studies-success-stories-qyhvf> (Accessed: 8 February 2024).

Sa’adah, S.A. and Hartanto, A. (2023) ‘FRAUD DETECTION USING DATA ANALYTICS IN THE CUSTOMS CONTEXT: A CASE STUDY OF UNDER INVOICING IMPORTATION FRAUD IN Indonesia’, 16.

Sift (2024) How to evaluate payment fraud analytics and measure risk, How to evaluate payment fraud analytics and measure risk. Available at: <https://sift.com/blog/how-to-evaluate-payment-fraud-analytics-and-measure-risk> (Accessed: 4 March 2024).