A Primer in Spatial Data

- The future of data

- Why is spatial data important?

- Maximizing the value of data with FME

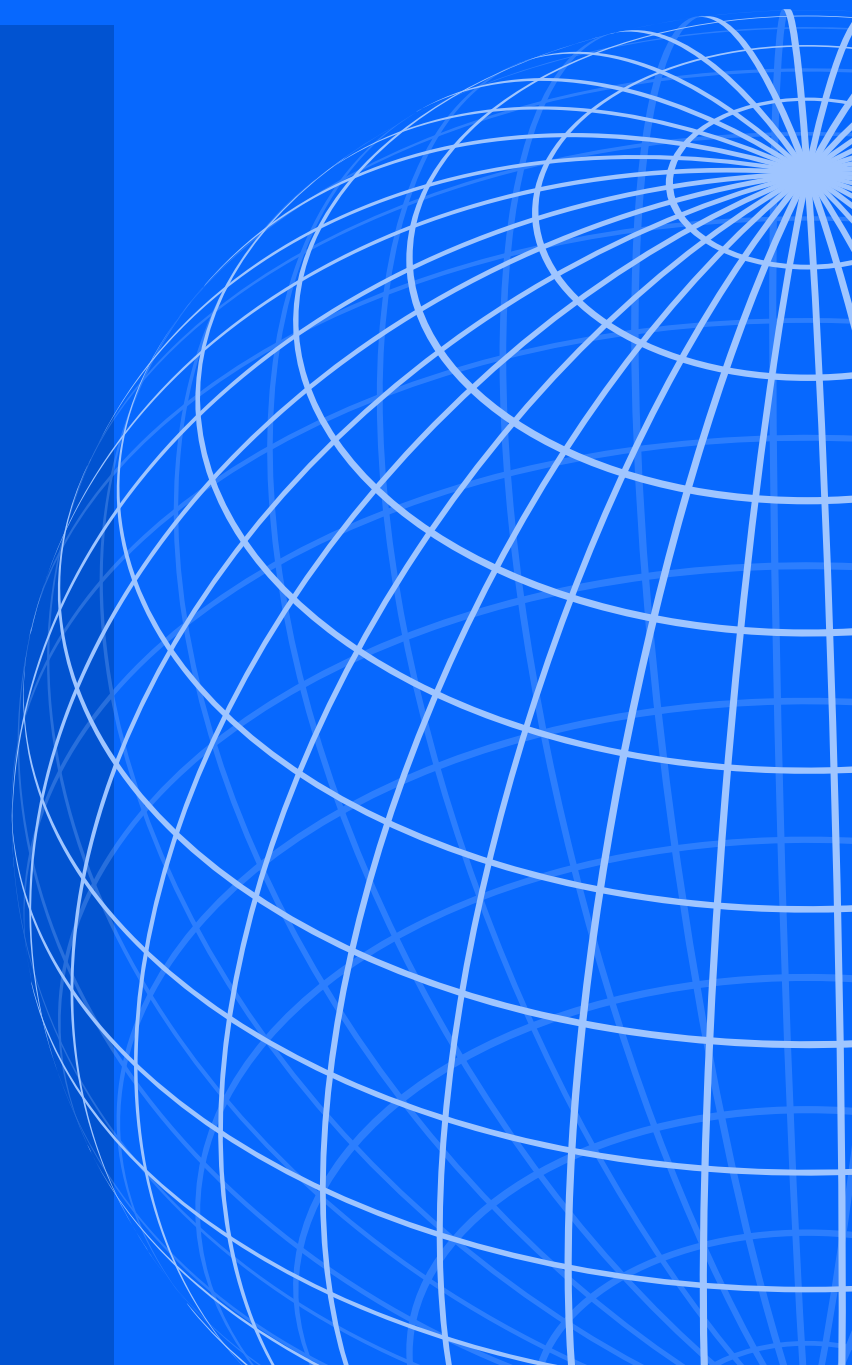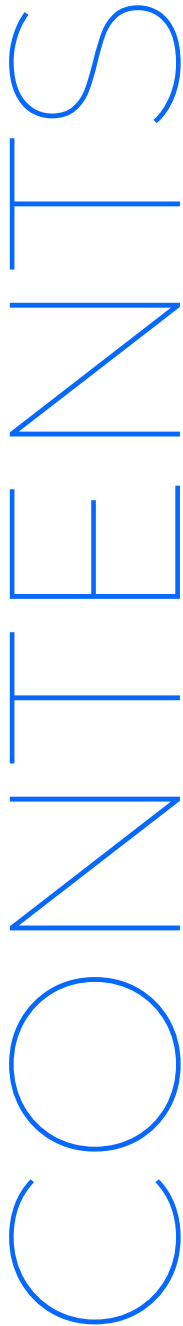# Table of Contents

CONTENTS

01. A Brief History of Data

'Data' comes from the latin word that means "a fact given or granted". The collection of data is part of what makes us human. The first evidence of data goes back to 19,000 BC where notched baboon bones (Ishango bone) were used as primitive storage devices (tally sticks). What has changed constantly over the years is how we store, collect, use and process data.

Computers were invented to curate, create, collect, compute and cleanse data. It is no surprise that since the dawn of the information era, the volume of data has grown exponentially. We live in the *data age*, where everything is recorded, measured, and stored. This trend is increasing with more sensors, devices, and applications generating more data than ever before. Data touches every aspect of our lives and society.

Data offers insights and a better understanding of our world. In business, companies that harness data make smarter decisions. These companies have a huge and unique strategic advantage.

Today, data is powering the innovations of tomorrow.

02. The 4 V's of Data

As data continues to evolve and become a bigger and bigger part of our world, it is important to understand and leverage data to make better decisions. The top challenges that organizations must understand and address to maximize the value of data are: data velocity, data variety, data veracity and data volume. These are the 4 V's of data.

Data Volume and Data Velocity

Today, organizations must handle large volumes of data that arrive at different data velocities. Data velocity is simply the speed at which data is delivered and processed. High velocity or real-time data allows organizations to make immediate decisions. Data with lower velocity is often processed in "batch processing" at a later time. Data of all velocities is important. Volumes of data are growing and the more that we process and use allows us to make decisions that will grow our businesses or improve our processes. This provides organizations with an opportunity to make better real-time and trend based decisions.

Data Variety

Like data volume and data velocity, **data variety** is increasing. Traditional Enterprise Integration tools are dedicated to tabular and traditional business data related to financial performance and customer information. However, new types of data are emerging and becoming more important. They are driving new classes of cost-effective decisions that enable organizations to serve their customers better.

The 4 V's of Data
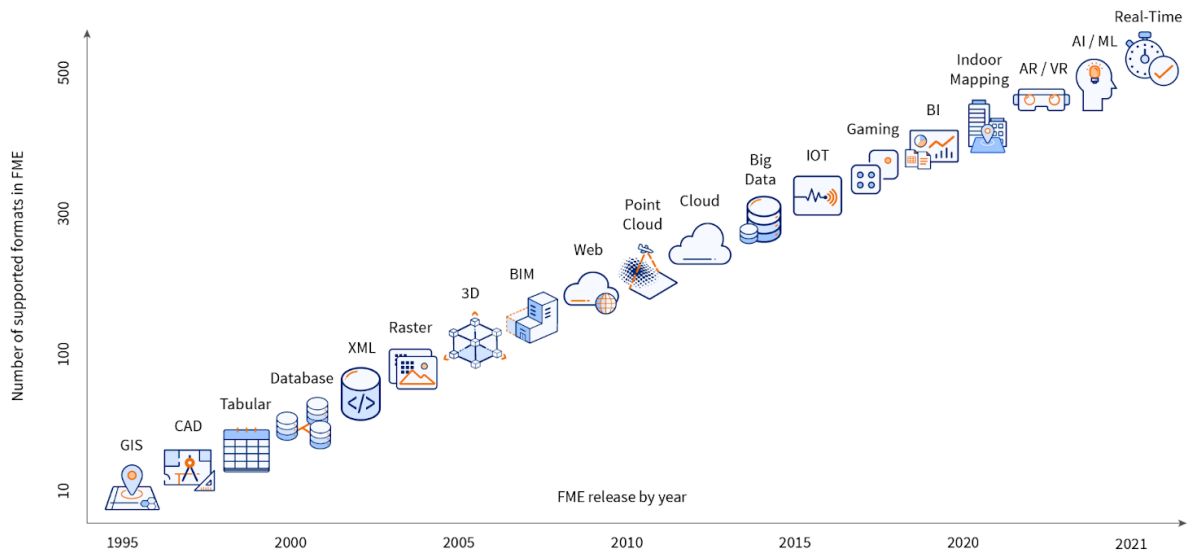
Data Veracity

Veracity is another word for **data quality**. Data quality is a constant battle for all organizations as there is always data that is incorrect. This incorrect data may come from a faulty sensor, or may simply go out of date. Hence, organizations need to ensure data quality and monitor it to continue to improve data quality. Remember, bad data and bad decisions are costly.

03. What Does the Future Hold for Data?

The explosive growth of data is relentless. As the rate of data volume and variety increases, organizations face the challenge of data veracity. How do they make sense of all this data? How accurate, complete, suitable or reliable is this data? Is this data real-time? Organizations that ensure their systems are scalable and efficient to improve the quality of their data will benefit the most. Bad data is expensive. It leads to misinformed decisions and sometimes irreparable consequences. The cost of poor data quality is an estimated $3.1 trillion a year to US businesses alone.

Uses for data also continue to grow with new technologies like Artificial Intelligence, Machine Learning, Augmented Reality (AR) and Virtual Reality. This growth is driven by a thirst to maximize large volumes of data, especially real-time data.

Future Applications and Data

When it comes to the future of data and applications it is clear that some of the more exciting new applications will require the following:

- Higher data volumes to drive applications to be the smartest they can be. This provides their users with more value.
- More data variety which will enable a richer analysis or experience for application users.
- Higher data velocity to improve accuracy. New applications will be built leveraging the growing availability of high speed data. Other solutions will require data of multiple velocities in order to deliver.
- Better data veracity will continue to be important as data volumes of all varieties and velocities continue to increase.

Some of the most exciting new applications and solutions today require a variety of data types to be effective. Here are examples of types of applications that are going to grow in the coming years:

Augmented Reality

Augmented Reality puts data in context with the real world. This requires spatial data and will drive the evolution of not just spatial data but data in general. While AR is in its infancy, this is an area to watch. It is predicted to change the way some industries, especially those with physical assets, interact with data. At Safe, when we talk about AR, we talk about how it enables users to truly experience their data.

The best AR systems will understand and be able to work with 3D data, raster data, traditional data and most importantly, spatial data. Data variety makes the AR experience more pleasing, effective and can convey more information than just an object in place. Spatial data, in particular, is a big part of the real world and is recognized more and more as important to business with leading vendors like Amazon, Google, Snowflake, and others adding support albeit sometimes limited to their offerings.

Machine Learning and Artificial Intelligence

ML/AI needs data, in volumes. ML/AI "training" requires huge amounts of data to learn the solution space. Data veracity is important as well to train. Once the data is trained, the system will be expected to work with data of a higher velocity to provide timely information that users can make sense of. Ultimately, the goal of ML/AI is to deliver more accurate results, in a more timely and cost effective manner than could be done by an expert.

04. Advantages of Real-Time Data Integration
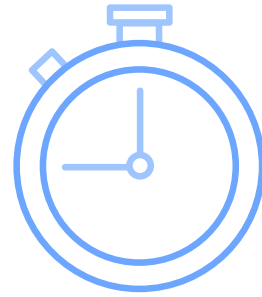
Here are five reasons and examples that illustrate why integrating real-time data is important. It is worth pointing out that for many of the real-time examples below, "location" is a key part of many of them, whether explicit or implicit.

Customer Experience

Improving customer experience is a goal of any organization. Traditional data processing is great for identifying trends and understanding trends. However, organizations that process real-time or high velocity data have the opportunity to improve customer experience by making decisions about what is happening "now". Immediacy is key.

There are many examples. Any organization that is providing services to clients can use real-time data. This monitors, detects and addresses problems before a client is even aware of it. This is particularly important for companies that deliver utilities such as water, electricity, and communications. They can use the real-time knowledge of their fleet to improve their responsiveness and solutions when they have to visit a client to fix a system. With real-time knowledge, clients get their problems fixed sooner and companies can tailor their response to improve customer experience. This is a win-win situation.

Situational Awareness

Understanding unfolding situations is important in many instances. During emergencies like earthquakes, floods, or fires, knowledge of what is happening makes a significant difference. For example, live updates of fire location, wind direction, and the location of crews makes battling the fire more effective. This immediate information also saves lives as the situation develops.

Monitoring Systems & Improving Facility Maintenance

Today, it is common for equipment to have sensors. Sensors constantly share the state of the equipment. Large amounts of data is produced by these systems. Using this data can greatly improve the efficiency of managing a facility. This often enables organizations to know, address and repair a piece of equipment or any other faults before it fails based on the sensor data. This also ensures equipment that needs to be shut down do so before any catastrophic damage to the equipment or the surrounding environment occurs.

For example, motors are everywhere. Simply reporting the temperature of a motor in real-time indicates the health of the motor. Hence, real-time data can make a real impact. Regardless, immediate knowledge gives the operators the chance to fix the problem before any error occurs. This can save time and the costs that are associated with damage or downtime.

Batch Processing for Lower Velocity Data

While processing real-time data is great for immediate decisions, organizations do need to also store the data. Organizations may also need to summarize data to support analysis at a later stage. This can be done in batches called "batch processing".

Batch processing enables organizations to identify important trends that cannot be determined in real-time. Thus organizations cannot ignore traditional data-driven batch processing based decisions.

Ideally, data integration systems can support both real-time and batch processing. Unfortunately, there are few data integration tools that support both high velocity real-time data streams and batch processing. This forces organizations to have different solutions. This incurs many procurement expenses and training that go with that.

05. Why is Spatial Data Important?

Leading companies are always looking to increase and improve their business efficiencies. The ones that are experiencing high growth tend to capture information about location, or spatial data.

Spatial data is more than maps. Maps are a great way to display spatial data and gain understanding about the world around us. However, that is just scratching the surface of spatial data. Spatial data provides us with a whole new way to understand relationships between people, places, things, and events.

Therefore, spatial data support is being added to more systems by more vendors than ever before. It is evident that adding spatial data offers businesses an edge in the marketplace. Spatial data is not only driving new decisions but also enabling organizations to increase productivity and customer service in new ways. Spatial data is found in all data. Today, location data originates from many sources including but not restricted to events, devices or location of people. Virtually everything has a spatial context.

The uses of spatial data are vast. There are over a <u>thousand common use cases!</u> spatial data is useful and its use in research and in business continues to expand as shown in this graph.

Source: sciencedirect.com
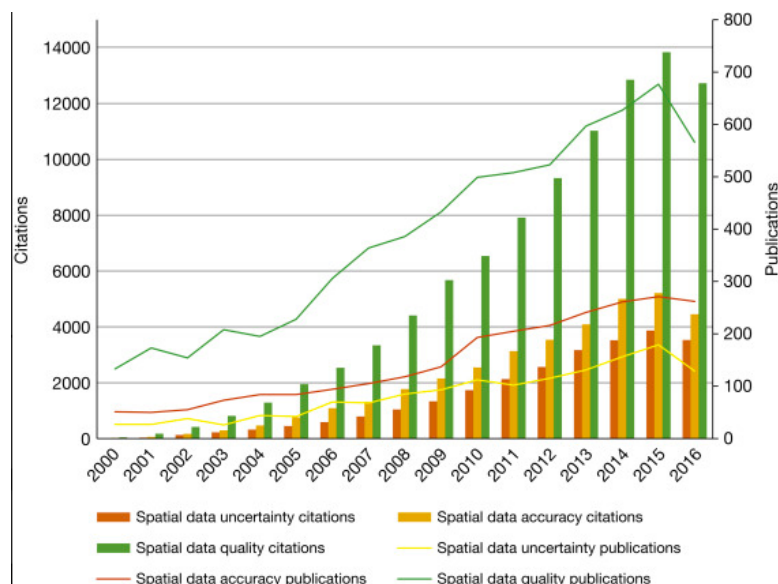
Spatial data is rich and comes in many different forms. Here are some of the most common types of spatial data and what they can be used to represent:

Vector Data

Vector data can be 2D or 3D and represents simple points, lines, and areas that represent locations on or relative to the surface of the earth. Supporting vector data greatly expands the decision making capabilities of any tool making vector data the obvious first choice when spatial data is being added to any tool.

Raster Data

Raster data is data that is stored in a 2D grid called pixels. Each pixel within a raster has 1 or more values. The value(s) in a pixel vary and represent a unit of measurement. Raster data typically represent imagery; taken by a satellite, an airborne craft, or a land based camera, but can represent any grid based measure such as noise, temperature, or air quality.

Point Cloud Data

A point cloud is a 3D set of data points. Each 3D point may have many associated values. Point clouds are collected by lidar or other 3D scanners. Point clouds can be used for many things such as creating 3D models of portions of cities or buildings, for quality inspection of infrastructure. Point cloud datasets can be massive.

3D Data

3D data is used to represent 3D objects. These objects could be buildings or terrain or infrastructure. There are many ways that 3D objects can be represented. Many new technologies rely on location accurate representation of 3D data, especially Augmented Reality and Virtual Reality. 3D data is also critical to building a digital twin which is used to improve efficiencies thru the use of a digital model. Common use cases are improving facilities management, managing underground assets and predicting maintenance and performance of 3D assets.

Indoor Mapping Data

We spend 90% of our time indoors. Yet, the mapping and location data that has been collected historically has been for the outdoors. Recently, effort has started to go into mapping the indoors so that users have a seamless experience. As we spend much of our time indoors, there is a great opportunity to understand how indoor spaces are used, and where they are congested. The graph below shows the relationship between spatial data and the availability of devices that use it. Mobile devices and indoor mapping go hand in hand to improve the lives of users.
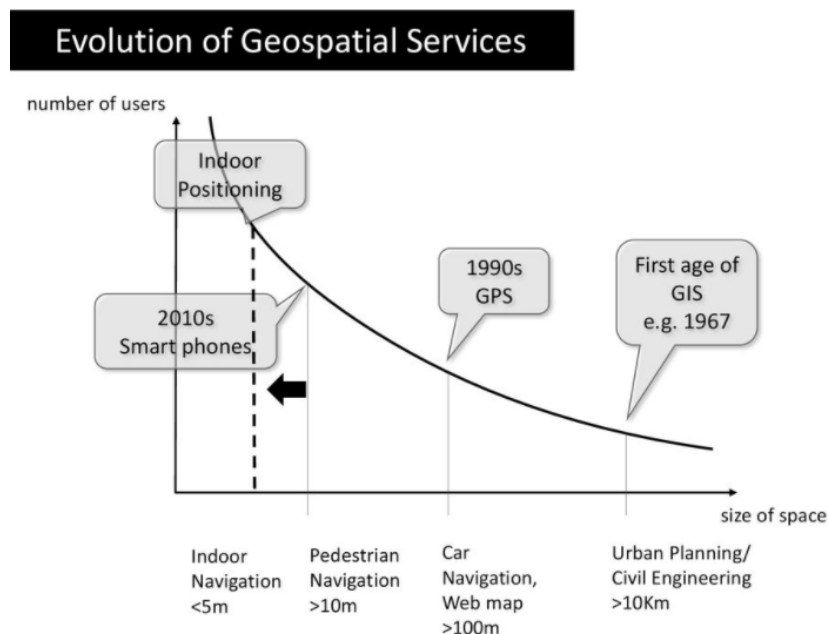
Source: Standard Indoor Spatial Data Model, Published by James Hudson
https://slideplayer.com/slide/14030563/

06. Understanding Spatial Data

By analyzing spatial data, we can learn more about why certain spatial relationships exist. Why are some stores or business locations more popular than others? Spatial data can be used in every industry. Spatial data helps organizations make decisions that were not possible before. Adopt spatial and location data practices to better understand human behavior and to make smart data-driven decisions sooner than later! Our understanding of spatial data and its importance was recognized by Gartner in the Magic Quadrant for Data Integration Tools. The growing importance of spatial data shows no sign of abating. Spatial data is our strength and a growing number of clients of all sizes are making the switch to our FME Data Integration Platform. What is our solution?

Maximizing the Value of Data with FME

The FME (aka. Feature Manipulation Engine) platform is a data integration tool. FME consists of support for hundreds of formats and applications as well as data transformation tools. This allows users to build and automate custom integration workflows **without** having to write code. Whether your data challenges have to do with spatial data, data storage, or business intelligence, FME is here to help you solve them.

Case Study: COVID-19, FME & Bluedot

Every day, BlueDot monitors 200 infectious diseases around the world. Seven days before the World Health Organization (WHO) announced the emergence of a novel coronavirus, BlueDot had already sent out the first warning of COVID-19. How did they do it?

Challenged with a huge array of data sources, BlueDot leveraged FME to automate complex data and application integration workflows. BlueDot also builds enterprise integration patterns, and synchronizes processed data via FME with their products. FME provides an enterprise integration solution that brings BlueDot's data into their global early warning system for disease outbreaks. Emergency responses have never had a greater need for the ability to detect and prepare for disease outbreaks. Before COVID-19, BlueDot was already helping governments, businesses, and healthcare organizations safeguard lives and livelihoods with their Early Warning System. This system uses FME to integrate an array of data and feed it to AI for analysis. It detected COVID-19 a week before it hit mainstream news.

However, only tracking the disease is not enough to prepare for outbreaks. BlueDot also tracks datasets regarding the movement of people, like flight schedules and historical travel data. With FME, these two kinds of datasets are integrated and fed into their data warehouse (AWS), where artificial intelligence takes over to analyze the data and detect credible information.

Case Study: COVID-19, FME & Bluedot

BlueDot forecasted what was coming, and once the coronavirus hit the news, their data intake exploded. They knew they needed an immediate strategy to process it all. When COVID-19 entered BlueDot's system, it dramatically increased their data processing needs. Using FME, they could use as many engines as needed across multiple machines to manage peaks in processing demands. Enterprise integration with FME made this possible.

BlueDot is able to monitor 200 infectious diseases around the world on a daily basis in order to identify disease activities and contextualize disease outbreaks. By automating web scraping for disease related data, they were able to issue an alert through their Insights platform about COVID-19 seven days earlier than the World Health Organization. FME also enables them to deliver data and analytics insights to their clients' preferred cloud data storage locations, such as Amazon S3, ArcGIS Online and Snowflake.

With FME, BlueDot empowers responses to infectious disease risks around the world using human and artificial intelligence, helping governments, businesses, and healthcare organizations safeguard lives and livelihoods.

[insert chart]
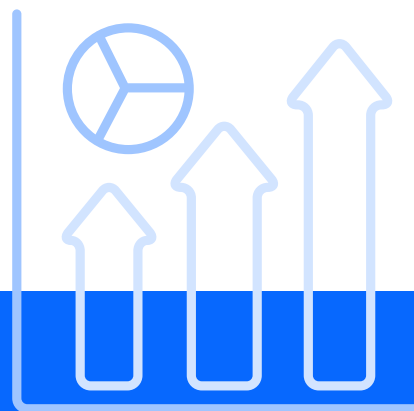
07. Change the World and Your Organization with Data Today

BlueDot's detection of COVID-19 is just one example of maximizing data with FME. There are many other ways to make sense of data and maximize data with FME. You can find many others at safe.com/customers and check out our other resources.

As data, including spatial data, increases in importance so does data and enterprise integration. Data integration produces a single, unified view of an organization's data. This bird's eye view provides actionable insights based on all of the organization's data assets, no matter the source or format. Enterprise integration builds on this and uses multiple integration approaches (including API management, application integration and messaging) to leverage enterprise services and assets and connect them. This enables organizations to seamlessly integrate and unify an organization's capabilities across various data and application environments.

Data and enterprise integration makes sense of data and helps organizations connect different applications, transfer and transform data easily, ensure data quality (or veracity) and the list goes on. Making your data, including spatial data, work for you helps organizations to optimize business processes, decisions and performance by:

1. Driving Sales with Customer Location (by understanding customer behaviors, delivering personalized promotions & driving store traffic and purchases).
2. Improving Operations with Asset Location (by tracking asset location & condition, predicting maintenance & improving workforce coordination).
3. Increasing Situational Awareness with Event Location (by tracking people, places & resources, improving disaster response and recovery & improving workplace safety).

Gain value from your data and change the world or your organization with data, bit by bit or byte by byte, today!

Safe Software (aka. "Safe") — the creators and developers of FME, the data integration platform with the best support for spatial data. Safe Software and its team of over 150 partners worldwide are dedicated to helping over 10,000 organizations around the world discover the power that their data holds.

FACTS AND FIGURES ABOUT
Spatial Data and Why You Need a Spatial Data Integration Tool Now
By Safe Software

IMPORTANCE OF
SPATIAL DATA

95% of business executive agree that spatial data are important in achieving desired business results today. Retailers' use of spatial data, in particular, has more than doubled in 2020 compared to 2019, rising from 10% to 23%.
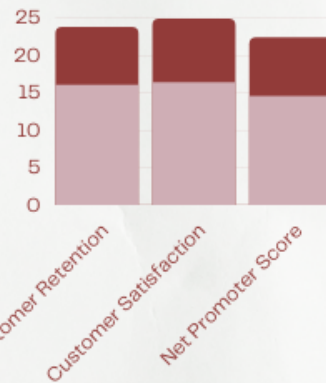
Benefits of Spatial Data

According to BCG.com, businesses that used spatial data and were location intelligence leaders in their industries reported improvements that were 1.3 to 2.0 times as great as those achieved by followers in customer experience and sales performance. The interesting thing is, these higher returns **did no**t vary significantly by industry, company size, or digital adoption - spatial data is for every business.
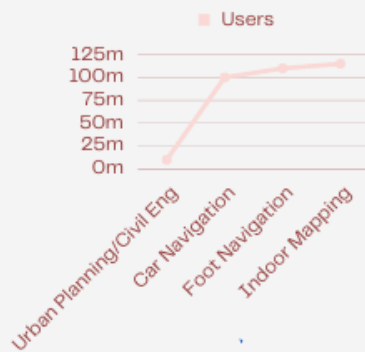
How Does Spatial Data Affect the Sales Funnel?

Spatial Data helps improve sales performance and sales funnel efficiency.

How Does Spatial Data Affect Customer Experience?

Spatial Data can help improve the customer experience.

USERS AND USES OF SPATIAL DATA

Both uses and users of spatial data continue to increase and show no sign of slowing down...

THE FUTURE OF SPATIAL DATA   91%

91% of business executives predict that it will be even more essential in 3-5 years! This is why you need a comprehensive spatial data integration tool, like **FME**, to **maximize the value of your data**.

References   https://www.bcg.com/publications/2021/leveraging-location-intelligence-across-industries
https://gis.usc.edu/blog/how-gis-is-taking-advantage-of-big-data/
https://slideplayer.com/slide/14030563/

Try FME (#1 Data Integration Tool for Spatial Data) for Free Today at https://www.safe.com/fme/trial/

Resources

Website
https://www.safe.com/

Spatial Data
https://www.safe.com/why/spatial-data/
https://www.safe.com/blog/2022/1/7-emerging-data-enterprise-integration-trends-2022/
https://www.safe.com/blog/2021/11/use-spatial-data/
https://www.safe.com/what-is/spatial-data/
https://gisgeography.com/gis-applications-uses/

FME
https://www.safe.com/fme/
https://www.safe.com/fme/trial/

Customer Stories
https://www.safe.com/customers/