

Analysis and visualization after the data pipeline

Previously, you learned that the data pipeline is an essential tool for creating a systematic flow of data through various stages from collection to visualization. Each part of the data pipeline is crucial for effectively using data, and gaining actionable insights. In this reading, you'll learn more about the key components of the data pipeline, and review a scenario of how data is queried at the end of a pipeline.

From collection to visualization

Effective data management is pivotal in addressing various business needs. Each step in the data journey can help you gain valuable business insights. Before data is visualized, it must go through three essential phases:

1. First, in the collection phase, data is gathered from various sources, including but not limited to: internal or external databases, or Internet of Things (IoT) devices, like smartphones, tablets, sensors, and trackers, or online collection channels.
2. Second, in the processing phase, data is cleaned, enriched, and transformed into a format that can be analyzed and used for desired outcomes determined by the business.
3. Finally, in the storage phase, processed data is stored in a structured, unstructured, or semi-structured way to be accessed for further analysis.

Querying data for analysis

When data reaches the store phase, it's retrieved or queried for analysis through different querying languages, like Structured Query Language (SQL). As an analyst, you'll use these languages to retrieve specific datasets from storage, so that they can be used to focus analysis, develop reports, and generate insights that can influence strategic business decisions. For example, an analyst for an e-commerce business might query the latest data on product sales for a particular region. They might use this information to guide product restock and marketing decisions.

Bringing data to light

After data is queried, it's visualized using various tools and platforms. Visualization is not just the act of generating charts and graphs; it is a crucial step in making data accessible and

comprehensible to stakeholders who will utilize data-driven insights to guide business strategies.

For example, an e-commerce business collects data from user interactions on their platform. As an analyst, you clean and organize the data as part of the process phase. Then, the data is stored in a cloud database. You notice a potential trend in customer purchases, and use SQL to extract the necessary data from storage. After you've completed analysis, you use a visualization tool like Looker to visualize this data, revealing the purchasing trend in a line graph for leadership to make strategic decisions about products and marketing.

Visualizations make insights accessible

As an analyst, it's your responsibility to present insights in an accessible way. Consider your audience and desired business outcomes. As a reminder, your audience won't have access to the data. Take care to present visualizations in a way your audience can understand. Accessible visualizations are easy to read, showcase the main point clearly, and are not distracting in design or color palette.

When you create visualizations, adhere to the Web Content Accessibility Guidelines (WCAG). You can implement accessibility features, like text alternatives for non-text content and background colors with sufficient contrast ratios. Make sure all users, including those with disabilities, can access and understand your visualizations. You can adapt your content to be distinguishable, while also providing various ways to access your content. Aim to provide an inclusive experience for everyone in your audience.

Key takeaways

Understanding the fundamental journey of data from collection to visualization is important for deriving meaningful business insights. As a reminder, data travels through three stages: collection, processing, and storage. Ultimately, data is queried for analysis and visualization. SQL is just one tool you'll use to extract data for meaningful analysis. Visualizations make data insights accessible and actionable for decision-makers.

Resources for more information

Further enhance your understanding of SQL, and explore different visualization tools with these resources:

- This resource provides a high-level overview of SQL, its purpose, and how to use it, along with additional resources for learning SQL: [Introduction to SQL](#)
- This resource provides an overview of data visualization, why it's important, and a list of the top data visualization tools available to professionals according to Harvard Business School: [Top Data Visualization Tools for Business Professionals](#)

- This resource provides a customizable, quick reference to Web Content Accessibility Guidelines, like WCAG 2 requirements, success criteria, and techniques: [How to Meet WCAG \(Quick Reference\)](#)