

Get started with data modeling, schemas, and databases

Choose the right database

How data moves

Data-processing with Dataflow

Organize data in BigQuery

▶

Video: Gather information from stakeholders

3 min

⌕

Reading: Merge data from multiple sources with BigQuery

10 min

✔

Practice Quiz: Activity: Set up a sandbox and query a public dataset in BigQuery

4 questions

⌕

Reading: Unify data with target tables

10 min

✔

Practice Quiz: Activity: Create a target table in BigQuery

3 questions

⌕

Reading: Activity Exemplar: Create a target table in BigQuery

20 min

⌕

Reading: Case study: Wayfair - Working with stakeholders to create a pipeline

20 min

Review: Data models and pipelines

[Optional] Review Google Data Analytics Certificate content

Case study: Wayfair - Working with stakeholders to create a pipeline

Working with stakeholders while designing and iterating on a pipeline system is an important strategy for ensuring that the BI systems you put in place answer their business needs. In this case study, you'll discover how the BI team at e-commerce home retailer [Wayfair](#), headquartered in Boston, Massachusetts, works with their stakeholders throughout a project to create a pipeline system that works for them.

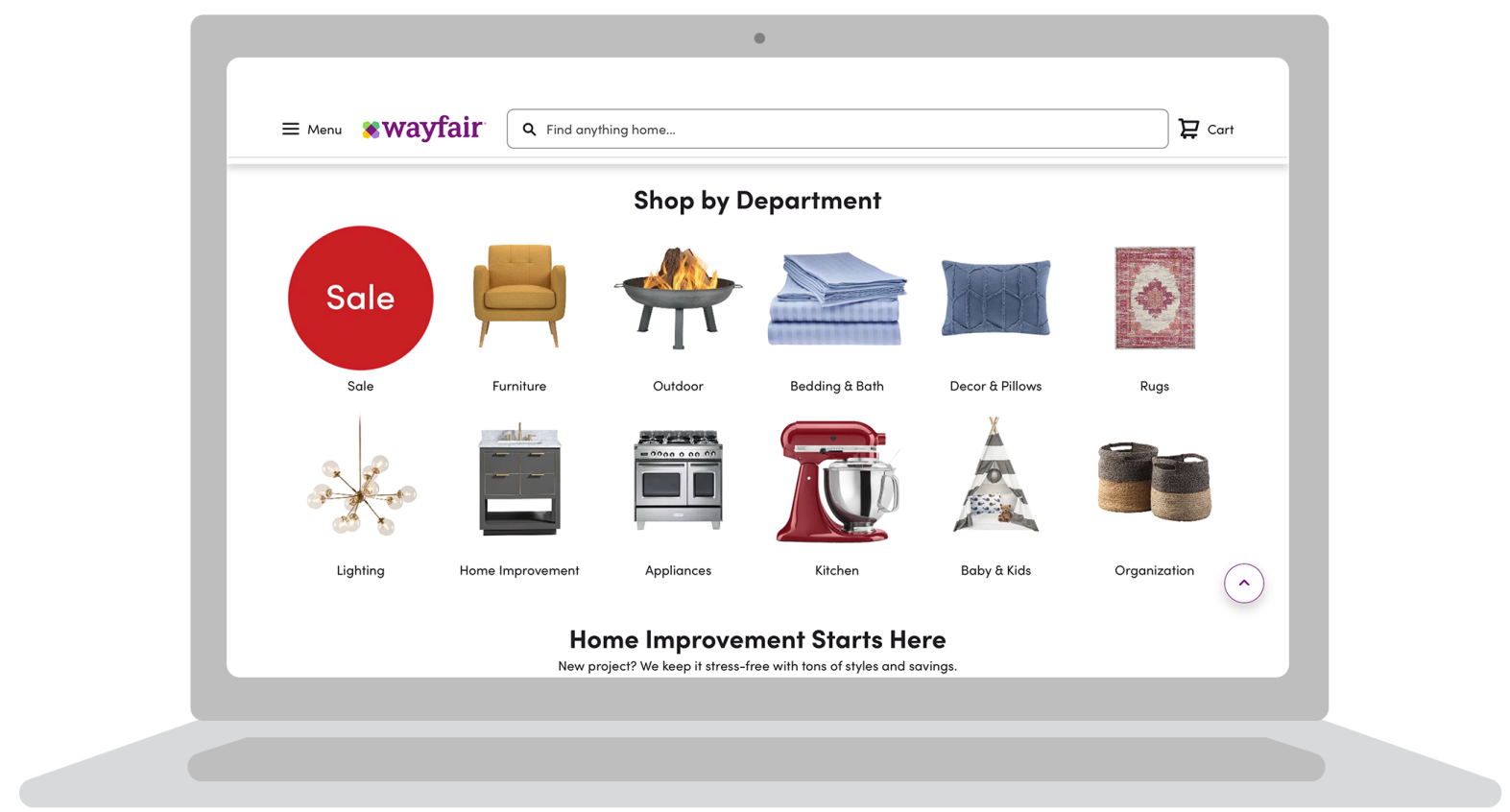


Company background

Longtime friends Niraj Shah and Steve Conine started the online-only company in 2002 after deciding they wanted to offer a larger selection of choices to customers—more than could fit in a brick-and-mortar space. They started the company as a collection of more than 200 e-commerce stores, each selling separate categories of products. In 2011, the company combined these sites to establish wayfair.com.



Wayfair is now one of the world's largest home retailers. The company's goal is to help everyone, anywhere, create their feeling of home. It empowers customers to create spaces that reflect who they are, what they need, and what they value.



The challenge

The Wayfair pricing ecosystem includes thousands of different inputs and outputs across a full catalog of products, which change multiple times a day. All of these inputs and outputs are being generated in different ways from different sources. Because of this, the BI team and other data professionals who needed to access pricing data were having trouble locating, querying, and interpreting the complete dataset. This led to incomplete and often inaccurate insights that weren't useful for decision makers.

To address this, the BI team decided to design and implement a new pipeline system to consolidate all the data stakeholders needed. They also needed to consider a few additional challenges with their pipeline system:

- Monitoring and reporting around these processes would need to be included in the design to track and manage errors.
- Data would need to be clean before it could be shared with downstream users.
- Due to the variety of data types being joined, the BI team also needed to better understand the data relationships so they could accurately consolidate the data.
- Training sessions would be required to help educate users on how to best access and use the new datasets.

These unique challenges meant that it was especially important for the BI team to work closely with stakeholders while developing their new system to address their needs and create something that worked across multiple teams.

The approach

Given the massive amount of data within the system, it was important for the BI team to step back and work with stakeholders to really understand how they were using the data currently. That included understanding the business problems they were trying to solve, the data they were already using and how they were accessing it, and the data they wanted to use but couldn't access yet.

Once they had communicated with stakeholders, the team was able to design a pipeline that achieved three key goals:

- All the required data could be made available and easy to understand and use
- The system was more efficient and could make data available without delays
- The system was designed to scale as the dataset expanded vertically and horizontally to support future growth

After this initial design was completed, the system was presented to stakeholders for review to ensure they understood the system and that it met all of their needs. This project required collaboration across a variety of stakeholders and teams:

- Software engineers:** The software engineer team were the primary owners and generators of data, so they were key to understanding the current state of the data and helped make it accessible for the BI team to work with.
- Data architects:** The BI team consulted with data architects to ensure that the pipeline design was all-encompassing, efficient, and scalable so the BI team could handle the amount of data being ingested by the system and ensure that downstream users would have access to the data as the system was being scaled.
- Data professionals:** As the core users, these teams provided the use cases and requirements for the system so that the BI team could ensure that the pipeline addressed their needs. Because each of their respective teams' needs were different, it was important to ensure the system design and data included was wide enough to account for all of those needs.
- Business stakeholders:** As the end users of the insights generated by the entire pipeline, the business stakeholders ensured all development work and use cases were rooted with clear business problems to ensure what the BI team built could be immediately applied to their work.

Communicating with all of the stakeholders throughout the design process ensured that the Wayfair BI team created something useful and long-lasting for their organization.

The results

The final pipeline that the BI team implemented achieved a variety of key goals for the entire organization:

- It enabled software engineering teams to publish data in real-time for the BI team to use.
- It consolidated the different data components into one unified dataset for ease of access and use.
- It allowed the BI team to store different data components in their own individual staging layers.
- It included additional processes to monitor and report on the system's performance to inform users where failures were occurring and enable quick fixes.
- It created a unified dataset that users could leverage to build metrics and report on data.

The greatest benefit of this pipeline solution was that Wayfair now had the ability to provide accurate information in one place for users, eliminating the need to join different sources themselves. This meant that the team could promote more accurate insights for stakeholders and get rid of costly ad-hoc processes.

The response cross-team was very positive. The director of analytics at Wayfair said that this was revolutionary for their team's daily work because they had information on retail price, cost inputs, and product status in the same place for the first time. This was a huge benefit for their processes and to help them handle their data in a more intelligent way.

Conclusion

A significant benefit that business intelligence provides an organization is that it makes the systems and processes more efficient and effective for users across the organization; basically, BI makes everyone's jobs a little easier. Ensuring that the BI team is tightly aligned with the business stakeholders and other teams is critical to their success. Without great partnership, problems can't be solved correctly.

Mark as completed

Like Dislike Report an issue