

Evaluation and Analysis of Data Management Tools

Based on *Bytebase*

Assignment #5

Group Number: 4

Data Curation

Lanzhou University

|  |
| --- |
| 2\_Anzelin\_320230941661\_安泽麟 |
| 2\_Chenyanyu\_320230941700\_陈彦妤 |
| 1\_Chenyujia\_320230941710\_陈雨佳 |
| 2\_SunYinuo\_320230942370\_孙一诺 |
| 1\_Zhugeyi\_320230942820\_朱哿依 |

Assignment #5

Evaluation and Analysis of Data Management Tools Based on *Bytebase*

**Introduction**

In today's data-driven world, data management tools play a vital role in ensuring the integrity, availability and security of data. As the demand for data quality and consistency continues to increase for enterprises, research institutions and development teams, efficient data management solutions are becoming an essential technology.

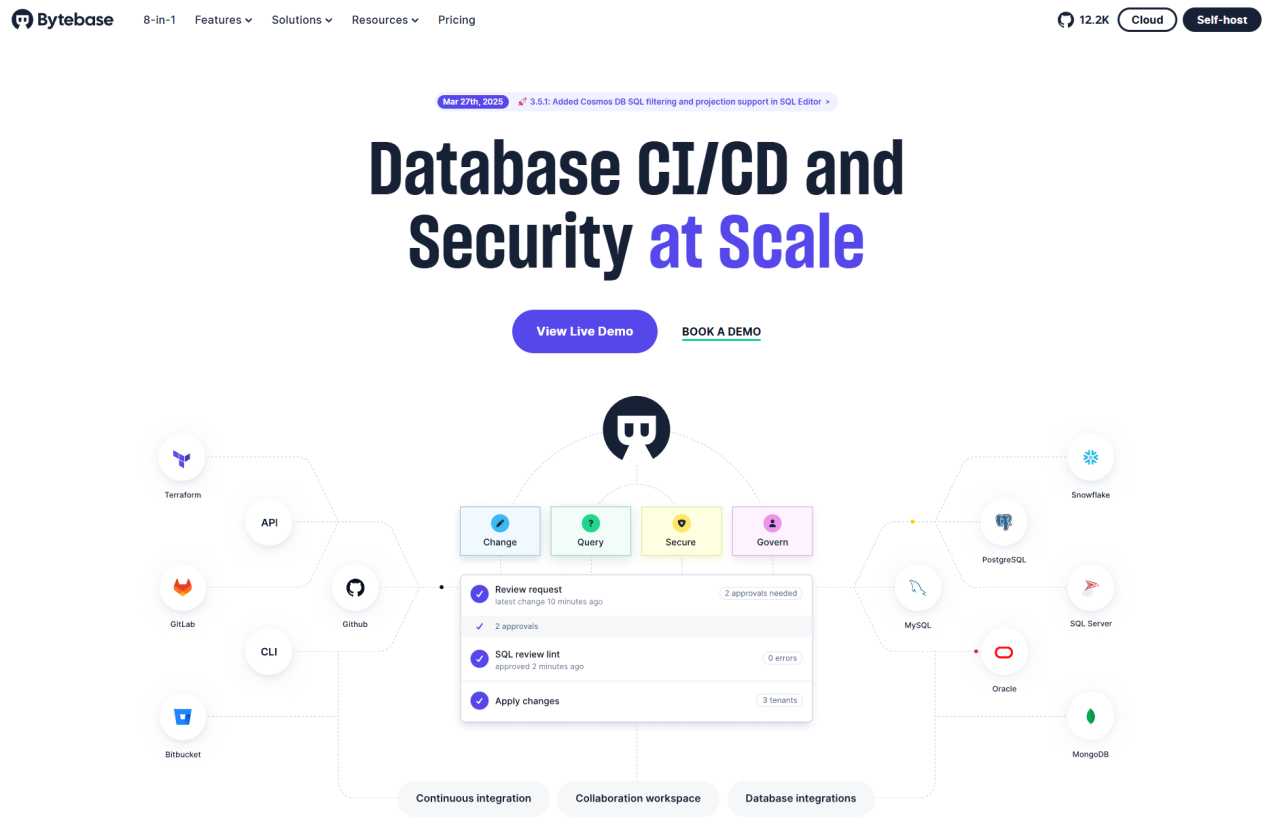


Figure 1

The welcome page of Bytebase’s official website.

Note. Screenshot taken from Bytebase (https://www.Bytebase.com).

*Copyright [2025] by Bytebase.*

The object of our analysis is Bytebase, an 8-in-1 open source tool focused on database development lifecycle management, which aims to improve the control and security of data management by simplifying the database change process. Its core functions include database change review, version control, team collaboration and SQL review to reduce human error in database operation and improve the transparency of data changes.

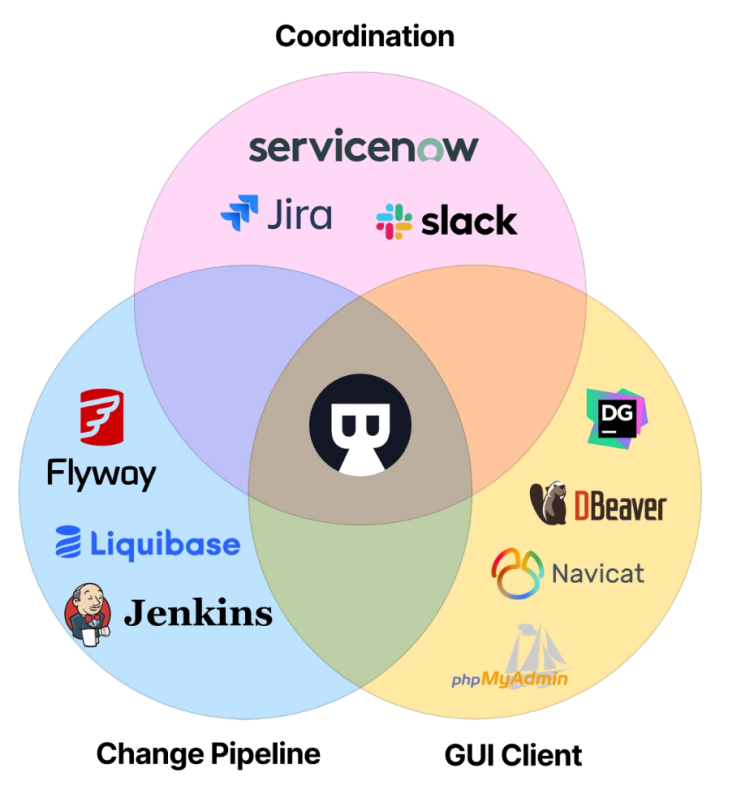


Figure 2

Bytebase's All-in-One Integration: Coordination, Change Pipeline, and GUI Client.

Note. Screenshot taken from Bytebase (https://www.Bytebase.com/docs/introduction/what-is-Bytebase/).

*Copyright [2025] by Bytebase.*

We will discuss its original design intention, target user group, functional features, technical limitations, development status, cost, installation and use difficulty, etc., to determine whether it can meet the long-term data management needs.

**Tool Analysis**

**Intended Use of the Tool**

**Design Purpose**

Bytebase is primarily designed for database change management and versioning to simplify and standardize the process of database schema changes. Traditional database management often relies on manual SQL changes, which can easily lead to problems such as data consistency. Bytebase makes database changes more controlled and transparent and reduces the risk of human data corruption by providing visual change processes, change review mechanisms, and team collaboration.

**Applicability**

Bytebase is designed for database change management as a middleware that connects users to the database, especially for teams that require rigorous auditing and versioning, such as DevOps, DBAs, and development teams, to reduce database risk. It provides a GUI mode that provides a collaborative workspace for teams to manage development tasks for all database systems; It also supports an API (headless mode) that acts as a backend for database operations and integrates with the team's development process.These features are detailed in Table 1.

Table 1

Comparison of the two modes of Bytebase.

|  |  |  |
| --- | --- | --- |
| **Pattern** | **GUI (Graphical User Interface)** | **API (Headless Mode)** |
| Interactive method | Web interface | REST API |
| Applicable subjects | DBA, developers | DevOps, Automation Platform |
| Main use | Visual Management of Database Changes | Automation changes, CI/CD processes |
| Whether additional development is required | No additional development required, ready to use out of the box. | Requires API calls and development |
| Advantage | Intuitive and easy to use, facilitating collaboration | Flexible and scalable, supports automation |

In addition, Bytebase supports rights management, team collaboration, and SQL specification auditing, but it is not a complete data governance or data warehouse management tool that can replace Apache Atlas or IBM InfoSphere MDM. Overall, Bytebase is suitable for DevOps teams and SMBS, but for complex data governance and convergence needs, it needs to be used in conjunction with other tools.

**Target User Groups**

Bytebase's core users are technology-driven teams, including development teams (reliant on automated SQL review), DevOps engineers (requiring CI/CD integration), and DBAs (ensuring security through permission control and multi-environment management). It suits compliance-driven industries like finance and healthcare, scales well for SMEs and global teams, but non-technical users require technical team collaboration, creating an adoption barrier.

Table 2

Analysis of Industrial Demand.

|  |  |  |
| --- | --- | --- |
| **Industry** | **Key Needs** | **Scalability** |
| Finance | Compliance, audit logs | Global team support |
| Healthcare | Sensitive data masking | Multi-environment |
| SMEs | Cost efficiency | Open-source friendly |

**Features and Limitations**  
**Features**

During database changes, Bytebase reduces human errors through automated SQL syntax checks and risk prevention (e.g., blocking DROP TABLE operations). All changes support version tracking and quick rollback. For multi-team collaboration, it provides isolated management of dev/test/prod environments, preventing critical data mishandling via access controls, while deeply integrating with GitHub/GitLab for automated change scripts. Built-in sensitive data masking and comprehensive audit logs meet compliance requirements in finance/healthcare.

***Limitations***

Bytebase primarily supports relational databases like MySQL/PostgreSQL, with limited compatibility for non-relational databases (e.g., MongoDB). It lacks built-in data cleansing/integration capabilities, requiring external tools. Long-term archiving relies on manually configured backups with insufficient automation. Additionally, the open-source version offers basic features, while enterprise needs require subscription-based paid services, creating cost barriers for resource-limited teams.

Table 3

Features and Limitations of Bytebase.

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Capability** | **Gap** |
| **DB Support** | MySQL, PostgreSQL | NoSQL (e.g., MongoDB) |
| **Backup** | Manual config | No auto-archiving |
| **Data Tools** | Needs external ETL | — |
| **Cost** | Free (OSS) | $$$ Enterprise plans |

**Development Status and Future Prospects**

**Development Stage**

Bytebase is currently in a mature and usable stage, with continuous iterative updates.

The release timeline is as follows:

Table 4

Release Timeline of Bytebase.

|  |  |
| --- | --- |
| **Year** | **Stage** |
| 2021 | Open Source Version Release(Apache 2.0 License) |
| 2022 | Launch of enterprise-grade features (such as audit logs, LDAP/SSO integration) |
| 2023 | Official release of the commercial version (BytebasePro), with support for more databases (such asMongoDB, experimental support for Redis) |
| 2024 | Enhanced CI/CD integration (such as GitHub Actions. GitLab CI plugins) |

Open Source Version Release(Apache 2.0 License)

Launch of enterprise-grade features (such as audit logs, LDAP/SSO integration)

Official release of the commercial version (BytebasePro), with support for more databases (such asMongoDB, experimental support for Redis)

Enhanced CI/CD integration (such as GitHub Actions. GitLab CI plugins)

Future Prospects  
The long-term reliability of Bytebase depends on a clear business model (open-source version + commercial version), adaptation to cloud-native trends (supporting Kubernetes deployment), and ecosystem expansion (supporting multiple databases). However, the open-source version lags in functionality, with some advanced features limited to the commercial version, which could affect adoption by community users. At the same time, although tools like Liquibase and Flyway are more mature in CI/CD integration, Bytebase remains competitive due to its user-friendly UI/UX.

**Costs**

**Monetary Costs**

Table 5

Monetary Costs of Bytebase.

|  |  |  |
| --- | --- | --- |
| **Versions** | **Pricing Model** | **Use Case** |
| Community Edition | Completely free (Apache 2.0) | Small teams, individual developers, Proof of Concept(PoC) |
| Commercial Edition | Subscription-based, per instance/user | Enterprise-level requirements (e.g., audit logs, SSO) |

#### Human Resource Costs

#### Bytebase has a low learning curve, with team members needing only about 1-2 days of training to adapt to the work order approval process (similar to Git MR). In terms of deployment, Docker setup is straightforward (within 30 minutes), but achieving high availability in a production environment requires configuring Kubernetes or load balancing.

#### **Hidden Costs**

#### When migrating to Bytebase from traditional tools (like phpMyAdmin), it is necessary to rewrite some SQL scripts to adapt to Bytebase's change management process. Regarding integration, although integrating with existing CI/CD tools (such as Jenkins) may require additional debugging, Bytebase’s OpenAPI helps reduce the complexity of integration.

**Prerequisites for Software and Hardware**

**Operating System**

Supports Linux, macOS, and Windows (via Docker or direct installation).

Can also be deployed via Kubernetes.

If using Docker deployment, Docker (≥20.10) or Docker Compose must be installed.

**Hardware Configuration**

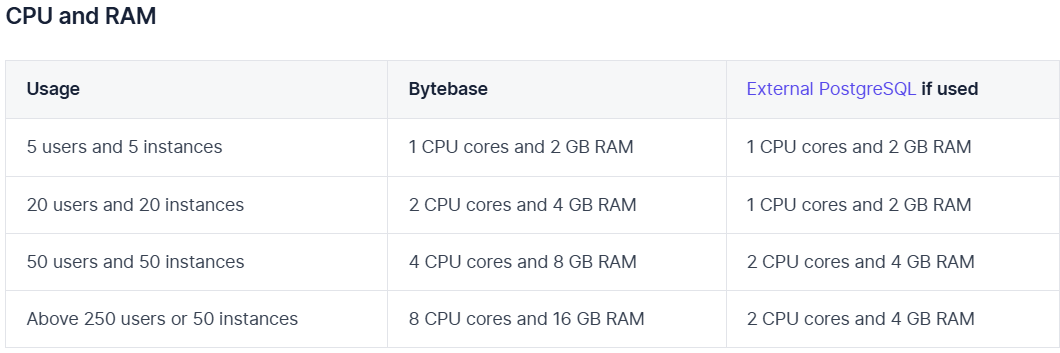


Figure 3

Recommended hardware specs for Bytebase and PostgreSQL.

Note. From Self-host Bytebase, by Bytebase, n.d. (https://Bytebase.cc/docs/faq/).

*Copyright [2025] by Bytebase.*

**Network Environment**

Requires access to the target database instance, supporting both local and cloud environments (e.g., AWS, GCP, Azure).

**Ease of Installation and Use**

**User Experience:**

**Simple:** Offers multiple installation methods (Docker, Kubernetes, binary files). The Docker method requires only one command to run.

**Easy to Use:** Provides a clear web UI, supporting visual database management (schema changes, backups, permission control, etc.).

**Collaboration Features:** Supports Git-like workflow (development → review → release), making it suitable for team collaboration.

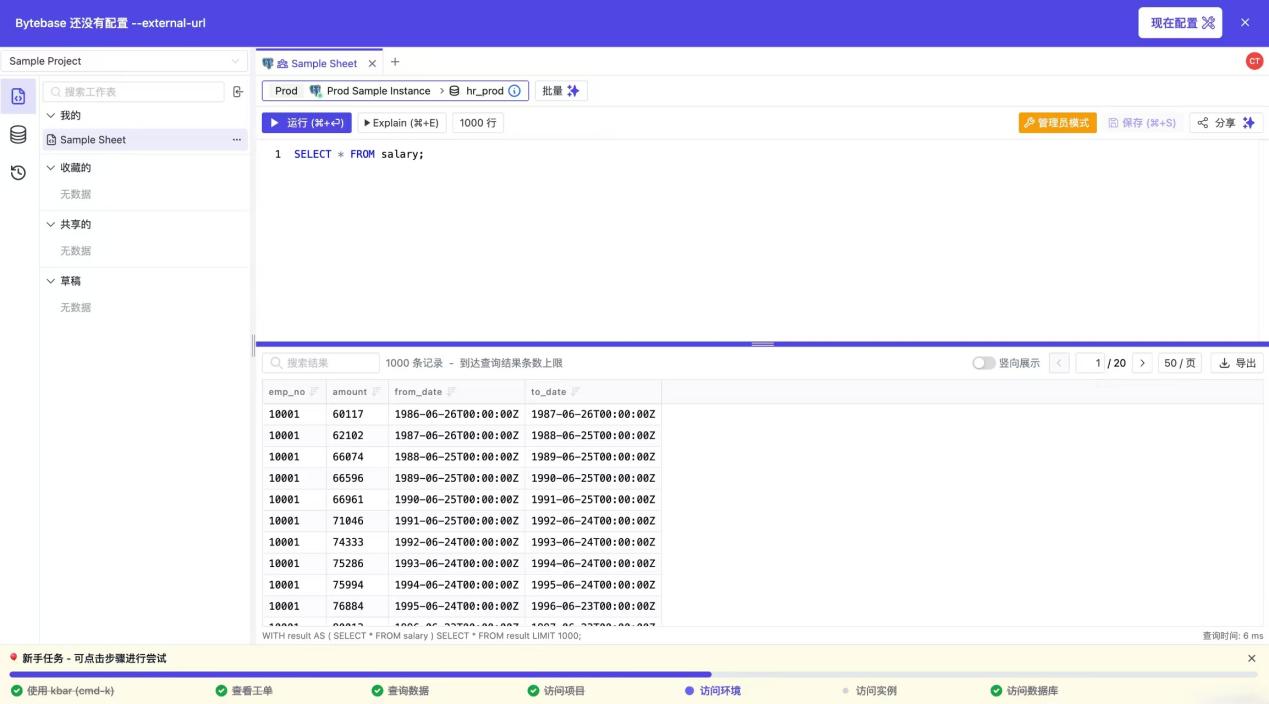


Figure 4

Person Usecase about Querying Data from Dataset.

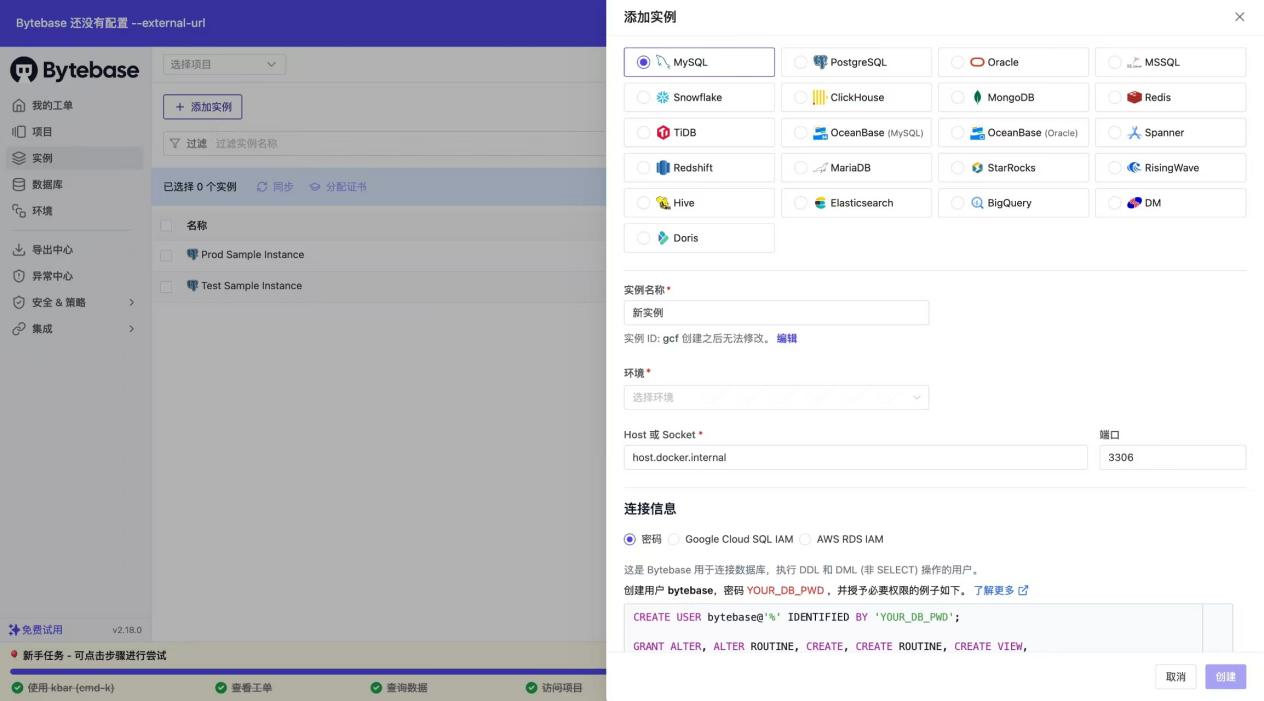
.

Figure 5

Person Usecase about Adding A New Dataset.

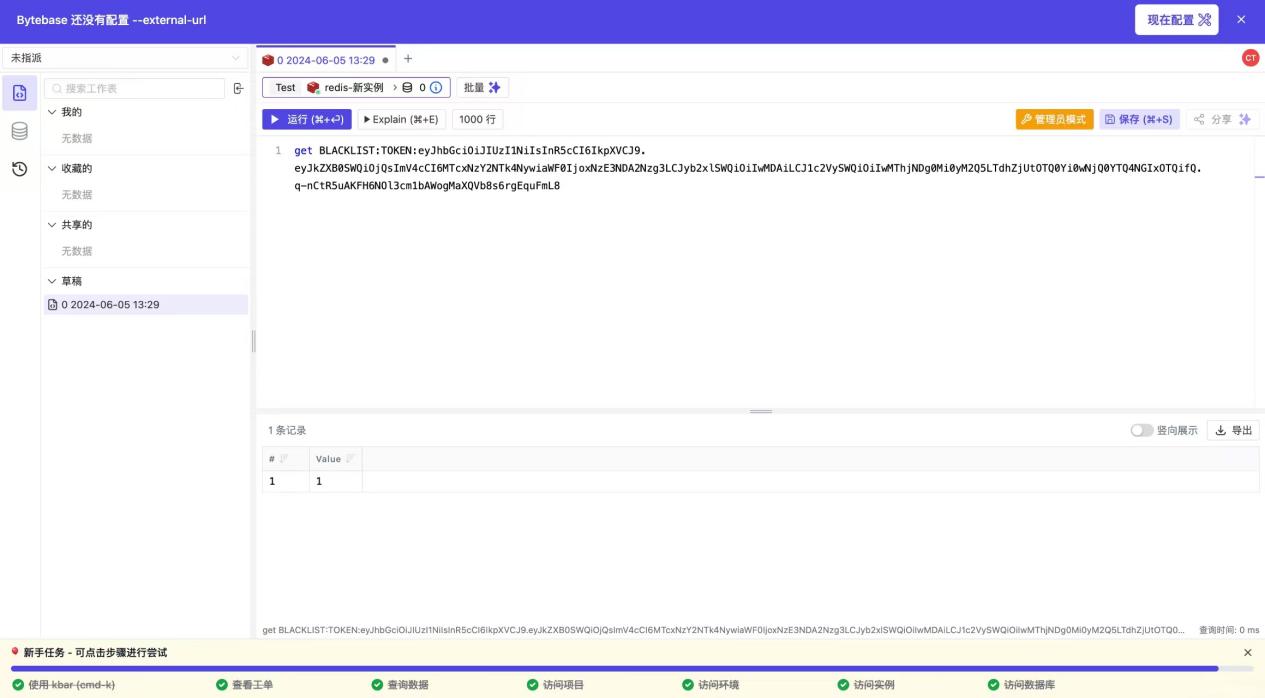


Figure 6

Person Usecase about Executing Instructions

**Conclusion and Recommendations**

Bytebase offers efficient solutions in database change management, version control, and environment management. It particularly shines when handling large-scale datasets. Bytebase's automated deployment and integration capabilities make it a powerful assistant for team collaboration. However, its high usage cost and high hardware requirements pose a significant barrier for small and medium-sized enterprises.

To enhance Bytebase's market competitiveness, it is recommended to reduce costs, optimize the user experience, and develop a more lightweight version to expand the user base. Through these improvements, Bytebase is expected to occupy a more important position in the field of data management.

**Reference**

Bytebase. (n.d.). What is Bytebase? Bytebase Documentation. Retrieved March 31, 2025, from [https://www.Bytebase.com/docs/introduction/what-is-Bytebase/](https://www.bytebase.com/docs/introduction/what-is-bytebase/)

Bytebase. (n.d.). Self-host Bytebase. Bytebase Documentation. Retrieved Marc-h 31, 2025, from [https://Bytebase.cc/docs/get-started/self-host/](https://bytebase.cc/docs/get-started/self-host/)

Bytebase. (n.d.). FAQ. Bytebase Documentation. Retrieved March 31, 2025, f-rom [https://Bytebase.cc/docs/faq/](https://bytebase.cc/docs/faq/)

CSDN Blog CSDN Blog. (2023). Database change and version control man-agement tool --Bytebase installation and deployment . CSDN Blog. https://blog.csdn.net/tomxjc/article/details/141725824 [Retrieved October 15, 2023]

What's Worth Buying What's Worth Buying. (2025). Bytebase: It has to be said, this database management and DevOps tool is amazing! . What's Worth Buying. https://post.m.smzdm.com/p/a8p0go27/ [Retrieved October 15, 2023]

Zhihu Column Zhihu Column. (2023). Bytebase, an open-source review tool for managing MySQL . Zhihu Column. https://zhuanlan.zhihu.com/p/619562023 [Retrieved October 15, 2023]