Зонова Анна, ИИ5-21М

Apache Superset — открытое программное обеспечение для исследования и визуализации данных, ориентированное на большие данные. Разработан сотрудником Airbnb Максимом Бошменом, который также является одним из создателей Apache Airflow. Вошёл в программу Apache Incubator в 2017 году.

Superset не поставляется в комплекте с возможностью подключения к базам данных. Основным шагом при подключении Superset к базе данных является **установка подходящих драйверов базы данных** в вашей среде.

Расширенный интерфейс cli позволяет импортировать и экспортировать источники данных из YAML и в него. Источники данных включают базы данных.

Сам по себе Superset не имеет уровня хранения для хранения ваших данных, а вместо этого соединяется с существующей базой данных или хранилищем данных, поддерживающей SQL.

Характеристики установки Superset зависят от количества пользователей и их активности, а не от размера ваших данных. Администраторы Superset в сообществе сообщили, что 8 ГБ ОЗУ и 2 виртуальных ЦП достаточны для запуска экземпляра среднего размера. Для разработки Superset, например компиляции кода или создания образов, вам может потребоваться больше мощности.

Необходимо контролировать использование ресурсов и увеличение или уменьшение их по мере необходимости. Использование Superset имеет тенденцию производить всплески, например, если все участники собрания одновременно загружают одну и ту же панель мониторинга.

Метаданные приложения Superset не требуют для хранения очень большой базы данных, хотя файл журнала со временем увеличивается.

**Публичный REST API** Superset соответствует спецификации OpenAPI.

Достоинства интеграции метаграфового источника данных и BI-системы

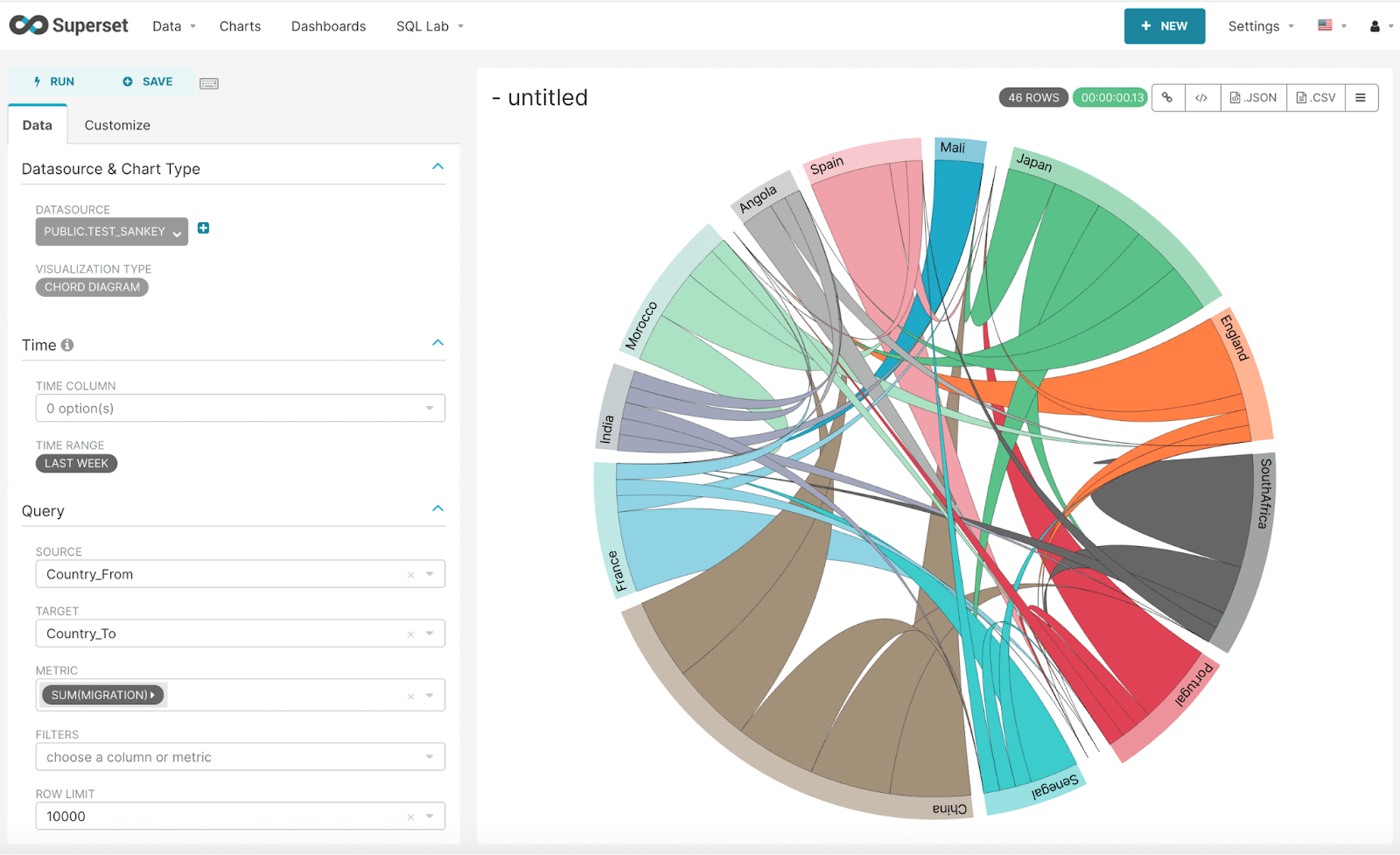
Недостатки интеграции метаграфового источника данных и BI-системы Apache Superset:

1. Интеграция метаграфовых источников данных может потребовать дополнительной настройки и конфигурации, что может быть сложно для пользователей без специфических знаний и навыков

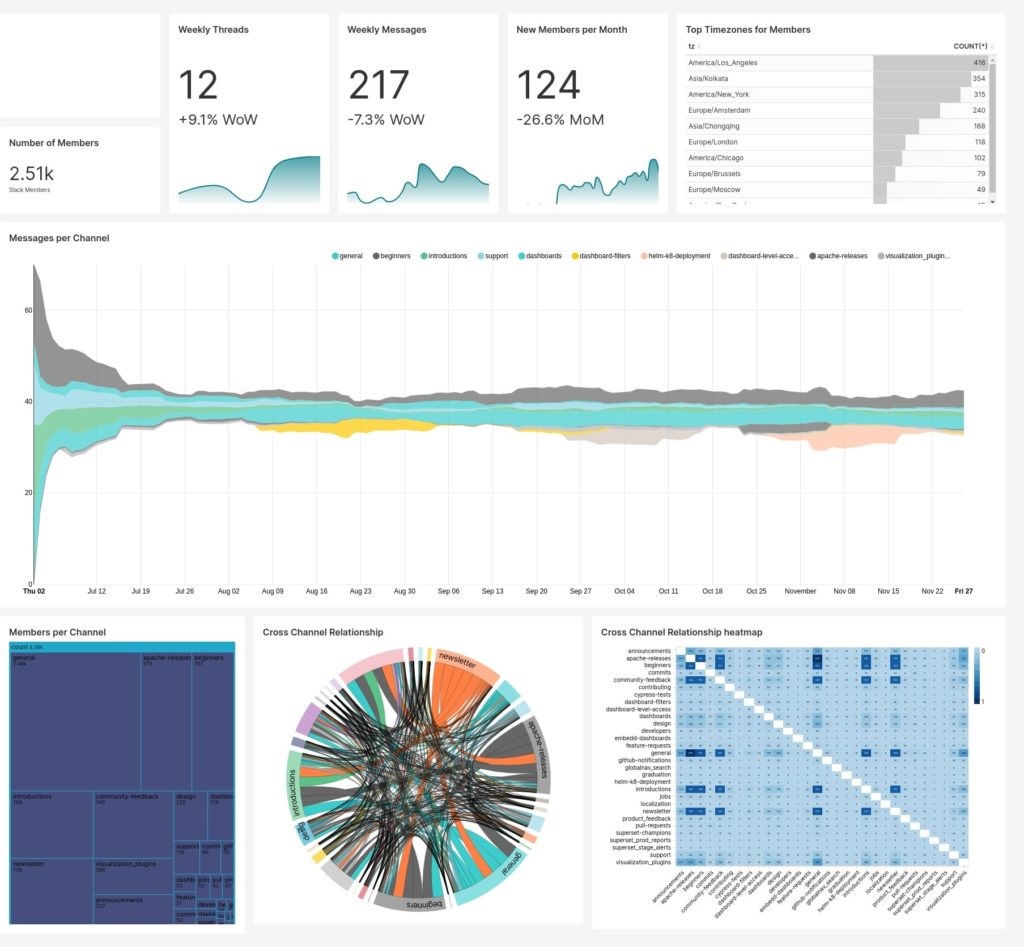
2. Подключение дополнительных источников данных может увеличить уязвимость системы к возможным угрозам безопасности данных.

3. Использование метаграфовых источников данных может привести к большему объему данных, что может повлечь за собой дополнительные нагрузки на систему при их обработке.

Однако Apache Superset поддерживает визуализацию 3-D объектов, что может быть удобным при отображении отчётов вершин метаграфа:



*Рисунок 10. Варианты визуализации в Apache Superset*



*Рисунок 11. Варианты визуализации в Apache Supe**rset*

# Вывод

В ходе исследования методов интеграции генераторов отчетов и метаграфовых источников данных, с фокусом на BI-системах Grafana, Metabase и Apache Superset, было установлено, что различные подходы предоставляют разные уровни гибкости, производительности и удобства использования. На основе проведенного анализа можно сделать вывод о необходимости учитывать специфику каждой BI-системы при выборе метода интеграции, а также о важности тщательного планирования и настройки процесса интеграции для достижения оптимальных результатов.