

Tìm hiểu lý thuyết về Apache Superset

1. Apache Superset là gì?

- Apache Superset là một ứng dụng web thông minh kinh doanh hiện đại, sẵn sàng cho doanh nghiệp. Nó nhanh, nhẹ, trực quan và được tải với các tùy chọn giúp người dùng thuộc tất cả các bộ kỹ năng dễ dàng khám phá và hình dung dữ liệu của họ, từ biểu đồ hình tròn đơn giản đến biểu đồ không gian địa lý deck.gl rất chi tiết.

- Superset là nền tảng đám mây và được thiết kế để có tính khả dụng cao. Nó được thiết kế để mở rộng ra các môi trường lớn, phân tán và hoạt động rất tốt bên trong các thùng chứa. Mặc dù bạn có thể dễ dàng kiểm tra ổ đĩa Superset trên một thiết lập khiêm tốn hoặc đơn giản trên máy tính xách tay của mình, nhưng hầu như không có giới hạn về việc mở rộng nền tảng.

2. Apache Superset cung cấp những gì?

Giao diện trực quan để trực quan hóa tập dữ liệu và tạo bảng điều khiển tương tác

Một loạt các hình ảnh trực quan đẹp mắt để giới thiệu dữ liệu của bạn

Trình tạo trực quan hóa không cần mã để trích xuất và trình bày tập dữ liệu

Một SQL IDE đẳng cấp thế giới để chuẩn bị dữ liệu để hiển thị, bao gồm một trình duyệt siêu dữ liệu phong phú

Một lớp ngữ nghĩa nhẹ cho phép các nhà phân tích dữ liệu nhanh chóng xác định các thứ nguyên và chỉ số tùy chỉnh

Hỗ trợ out-of-the-box cho hầu hết các cơ sở dữ liệu nói tiếng SQL

Bộ nhớ đệm và truy vấn không đồng bộ liền mạch, trong bộ nhớ

Một mô hình bảo mật có thể mở rộng cho phép cấu hình các quy tắc rất phức tạp về việc ai có thể truy cập các tính năng và bộ dữ liệu của sản phẩm.

Tích hợp với các phụ trợ xác thực chính (cơ sở dữ liệu, OpenID, LDAP, OAuth, REMOTE_USER, v.v.)

Khả năng thêm các plugin trực quan hóa tùy chỉnh

Một API để tùy chỉnh có lập trình

Một kiến trúc gốc đám mây được thiết kế từ đầu cho quy mô

3. Các cơ sở dữ liệu được hỗ trợ

- Amazon Athena
- Amazon Redshift
- Apache Drill
- Apache Druid
- Apache Hive
- Apache Impala
- Apache Kylin
- Apache Pinot
- Apache Spark SQL
- BigQuery
- ClickHouse
- CockroachDB
- Dremio
- Elasticsearch
- Exasol
- Google Sheets
- Greenplum
- IBM Db2
- MySQL
- Oracle
- PostgreSQL
- Presto
- Snowflake
- SQLite
- SQL Server
- Teradata
- Vertica
- Hana

- Ngoài ra Superset cũng hoạt động tốt với các dịch vụ như NewRelic, StatsD và DataDog, đồng thời có khả năng chạy khối lượng công việc phân tích dựa trên hầu hết các công nghệ cơ sở dữ liệu phổ biến.