

**BỘ THÔNG TIN VÀ TRUYỀN THÔNG  
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN  
THÔNG**



**BÁO CÁO THỰC TẬP CƠ SỞ**

**BÁO CÁO TÌM HIỂU VỀ ETL VÀ DATA  
PIPELINE**

**Giảng viên hướng dẫn : Kim Ngọc Bách**

**Họ và tên : Ngô Vũ Minh Quý**

**Mã sinh viên : B22DCVT427**

**Lớp : E22CQCN02-B**

## **MỤC LỤC**

- I. Giới thiệu về Looker
- II. Ứng dụng của Looker trong phân tích dữ liệu hiện đại
- III. Cài đặt và triển khai Looker
- IV. Các thao tác cơ bản với Looker
- V. Đánh giá và kết luận

## I. Giới thiệu về Looker

Looker là một nền tảng Business Intelligence (BI) hiện đại được Google Cloud mua lại vào năm 2019. Looker cung cấp giải pháp phân tích dữ liệu mạnh mẽ, cho phép người dùng trực quan hóa dữ liệu, tạo báo cáo và dashboard mà không cần phải viết mã SQL phức tạp.

Điểm nổi bật của Looker là sử dụng **ngôn ngữ LookML (Looker Modeling Language)** – một ngôn ngữ mô tả dữ liệu giúp trừu tượng hóa logic phân tích, từ đó tách biệt phần hiển thị và phần truy vấn dữ liệu.

### Các đặc điểm nổi bật của Looker:

- **Truy vấn trực tiếp trên cơ sở dữ liệu:** Không cần ETL sang hệ thống khác, Looker truy vấn trực tiếp trên các nguồn dữ liệu lớn như BigQuery, Redshift, Snowflake,...
- **Ngôn ngữ LookML:** Cho phép định nghĩa các mô hình dữ liệu, tránh lặp lại truy vấn SQL và giúp người dùng phi kỹ thuật có thể dễ dàng sử dụng.
- **Tùy biến linh hoạt:** Dashboard, biểu đồ, KPI có thể cá nhân hóa hoặc chia sẻ dễ dàng trong nhóm.
- **Tích hợp mạnh mẽ:** Hỗ trợ API, webhooks và tích hợp với Slack, Google Sheets, và các dịch vụ đám mây.

## II. Ứng dụng của Looker trong phân tích dữ liệu hiện đại

### 1. Lý do lựa chọn Looker

Looker được lựa chọn phổ biến trong các doanh nghiệp hiện đại nhờ vào khả năng:

- **Truy vấn nhanh, hiệu quả:** Looker chỉ thực hiện truy vấn khi cần,

giảm tải hệ thống.

- **Quản lý phân quyền chi tiết:** Cho phép kiểm soát truy cập dữ liệu theo vai trò, đảm bảo tính bảo mật.
- **Tự động hóa báo cáo:** Hỗ trợ gửi báo cáo định kỳ qua email hoặc webhook.
- **Khả năng mở rộng:** Phù hợp cho cả startup nhỏ và doanh nghiệp lớn với hàng tỷ dòng dữ liệu.

## 2. Ứng dụng thực tế của Looker

- **Phân tích hành vi khách hàng:** Xây dựng các dashboard theo dõi hành vi người dùng theo thời gian thực.
- **Theo dõi hiệu suất kinh doanh:** Tạo các báo cáo KPI theo từng phòng ban hoặc sản phẩm.
- **Giám sát dữ liệu vận hành:** Phân tích lỗi hệ thống, thời gian downtime và đưa ra cảnh báo.
- **Tích hợp vào sản phẩm:** Looker có thể nhúng (embed) vào phần mềm SaaS hoặc website để cung cấp báo cáo trực tiếp cho người dùng cuối.

## III. Cài đặt và triển khai Looker

Looker không được cung cấp miễn phí và thường được triển khai dưới dạng **SaaS (Software-as-a-Service)** qua Google Cloud, tuy nhiên quá trình triển khai có thể mô phỏng như sau:

### Bước 1: Đăng ký và truy cập

- Truy cập: <https://looker.com>
- Đăng ký tài khoản thử nghiệm qua Google Cloud hoặc liên hệ tư vấn.

## Bước 2: Kết nối cơ sở dữ liệu

- Chọn nguồn dữ liệu (MySQL, PostgreSQL, BigQuery, Snowflake,...).
- Nhập thông tin kết nối: host, user, password, schema.

## Bước 3: Tạo mô hình dữ liệu với LookML

- Định nghĩa **view** và **model** bằng LookML để ánh xạ tới bảng dữ liệu thực.
- Khai báo các **dimension**, **measure**, **filter** và **explore**.

## Bước 4: Truy vấn và xây dựng dashboard

- Sử dụng giao diện kéo thả để tạo báo cáo.
- Lưu lại dashboard, chia sẻ hoặc lên lịch gửi email tự động.

# IV. Các thao tác cơ bản với Looker

## 1. Tạo view và model với LookML

```
view: orders {  
  dimension: order_id {  
    type: number  
    sql: ${TABLE}.order_id ;;  
  }  
  
  dimension: created_date {  
    type: date  
    sql: ${TABLE}.created_at ;;  
  }  
  
  measure: total_revenue {  
    type: sum  
    sql: ${TABLE}.total_amount ;;  
  }  
}
```

## 2. Truy vấn dữ liệu

- Sử dụng “Explore” để chọn bảng, lọc dữ liệu, nhóm và tạo biểu đồ.
- Looker sẽ tự động tạo truy vấn SQL tương ứng và gửi đến CSDL.

## 3. Tạo dashboard và báo cáo

- Thêm các biểu đồ: cột, đường, donut, bảng.
- Cấu hình filter tương tác (ví dụ: chọn ngày/tháng, người dùng).
- Lên lịch gửi định kỳ qua email, webhook, hoặc Google Sheets.

## 4. Quản lý quyền truy cập

- Gán quyền xem/sửa với từng dashboard.
- Phân quyền theo vai trò (admin, developer, viewer).

# V. Đánh giá và kết luận

Qua quá trình tìm hiểu về Looker, em nhận thấy rằng đây là một công cụ cực kỳ mạnh mẽ và phù hợp với nhu cầu phân tích dữ liệu hiện đại. Khác với các nền tảng BI truyền thống yêu cầu ETL và lưu trữ dữ liệu riêng biệt, Looker truy vấn trực tiếp vào các nguồn dữ liệu lớn hiện có, giúp giảm thiểu độ trễ và chi phí lưu trữ.

Looker đặc biệt hiệu quả trong các tổ chức cần báo cáo thường xuyên, đa tầng, và nhiều người dùng. Tính năng mô hình hóa dữ liệu bằng LookML giúp tách biệt logic nghiệp vụ và giao diện trình bày, từ đó dễ bảo trì và mở rộng hệ thống phân tích.

Một điểm mạnh vượt trội khác là khả năng tích hợp với hệ sinh thái Google Cloud và các dịch vụ CI/CD, giúp doanh nghiệp dễ dàng xây dựng các hệ thống phân tích tự động, linh hoạt và dễ mở rộng theo nhu cầu.

Trong bối cảnh dữ liệu trở thành tài sản cốt lõi của doanh nghiệp, việc thành thạo Looker sẽ mang lại nhiều lợi thế cạnh tranh. Với sinh viên ngành CNTT, kỹ năng làm việc với Looker không chỉ là điểm cộng trong mắt nhà tuyển dụng mà còn giúp nâng cao tư duy dữ liệu, vốn rất cần thiết trong thời đại số hóa.

Việc tìm hiểu và làm chủ Looker là một bước đi cần thiết để tiếp cận với các vị trí như **Data Analyst**, **BI Developer**, hoặc **Data Engineer**, đồng thời mở rộng cơ hội nghề nghiệp trong các công ty công nghệ, tài chính, và thương mại điện tử.