

Lab 03: Trực quan hóa dữ liệu với POWERBI

Võ Nhật Tân

1 Quy định chung

- Bài làm được thực hiện theo nhóm tối thiểu 2 người và tối đa 3 người.
- Thành viên không tham gia sẽ không có điểm bài tập này.
- Bài làm phải tuân thủ theo yêu cầu đề án.
- Các nguồn tài liệu tham khảo (nếu có) cần ghi đầy đủ trong báo cáo ở mục Tài liệu tham khảo.
- Các công cụ hỗ trợ trong việc code, viết báo cáo như ChatGPT, Github Copilot, ... chỉ dùng như một công cụ tham khảo; nội dung phải được kiểm tra và chỉnh sửa phù hợp với bài toán đặt ra. Nếu phát hiện sử dụng công cụ AI để sinh nội dung quá nhiều, hoặc nội dung không phù hợp/sai lệch thì sẽ trừ tối đa 50% số điểm (tùy mức độ).
- Bài giống nhau sẽ **0 điểm môn học**

2 Giới thiệu đề án

2.1 Giới thiệu đề án

Power BI là công cụ, mạnh mẽ do Microsoft phát triển giúp kết nối, xử lý và trực quan hóa dữ liệu từ nhiều nguồn khác nhau, hỗ trợ tạo dashboard, biểu đồ tương tác và chia sẻ báo cáo trực tuyến.

Trong đề án này, các nhóm sẽ tự chọn dữ liệu và sử dụng Power BI để phân tích, trực quan hóa cũng như tạo các dashboard nhằm khám phá các xu hướng, mối quan hệ và đưa ra những thông tin hữu ích.

2.2 Nhiệm vụ đề án

Trong đề án này, các bạn sẽ thực hiện các nhiệm vụ sau:

1. Tìm hiểu công cụ PowerBI

- Giới thiệu về PowerBI
- Tìm hiểu các tính năng của PowerBI kèm các ví dụ minh họa sử dụng các mẫu dataset có sẵn, không liên quan đến dataset chính của bài

2. Xác định mục tiêu và lựa chọn dataset phù hợp với bài phân tích

- Xác định rõ mục tiêu phân tích. Các mục tiêu cần phải liên quan đến một chủ đề.
- Lựa chọn **một dataset phù hợp nhất** với các mục tiêu đã đề ra. Các dataset phải đủ lớn để phân tích.
- Các bộ dữ liệu có thể lấy từ nguồn có uy tín hoặc có thể sử dụng dữ liệu của hai bài thực hành trước để thực hiện tiếp. Khi sử dụng dữ liệu, cần phải giới thiệu và trích nguồn dữ liệu vào trong bài báo cáo, đồng thời giải thích vì sao chọn bộ dữ liệu đó.
- Mỗi nhóm cần xác định ít nhất [**Số lượng thành viên**] * **3** câu hỏi để phân tích và trực quan.

3. Phân tích dữ liệu và tạo báo cáo

- Cần xem xét các mối quan hệ giữa các trường để giải quyết câu hỏi.
- Chọn lựa nhiều dạng biểu đồ khác nhau, đánh giá sự phù hợp và thể hiện các biểu đồ này trong PowerBI. Nhóm sử dụng đa dạng các biểu đồ để làm rõ mục tiêu đã đề ra.
- Sử dụng màu sắc để thể hiện dữ liệu, giải thích ý nghĩa các màu và tại sao mình sử dụng màu như vậy. Sau khi bổ sung màu, nhóm có rút ra thêm ý nghĩa gì không?
- Cần tạo dashboard để bài phân tích trở nên thuyết phục hơn.

4. Nhận xét và đưa ra kết luận dựa trên các biểu đồ

3 Yêu cầu đề án

- Nhóm chỉ sử dụng môi trường PowerBI để trực quan và tạo dashboard; các thư viện ngoài thì cần ghi rõ trong báo cáo lý do sử dụng.
- Có thể sử dụng các thuật toán máy học trong việc phân tích và thể hiện trên biểu đồ.
- Viết báo cáo ngắn gọn và đầy đủ toàn bộ quá trình thực hiện; kết quả và nhận xét. Báo cáo tối đa là 24 trang (không bao gồm trang bìa, mục lục và tài liệu tham khảo). Nội dung báo cáo bao gồm:
 - Thông tin nhóm: tên nhóm, mssv...
 - Mức độ hoàn thành tổng thể của mỗi yêu cầu.
 - Mức độ hoàn thành của từng thành viên.
 - Chi tiết các bước thực hiện (kèm hình ảnh), thuật toán, chạy ví dụ, nhận xét. Trình bày đơn giản, có hình minh họa.
- Trong file PowerBI cần thể hiện như một storyboard, rõ ràng dễ hiểu và hợp lý, ngoài báo cáo, trình bày trong PowerBI là một tiêu chí để cộng điểm.

4 Quy định nộp bài

- Nhóm sẽ cử đại diện 1 người nộp bài.
- Bài nộp là một file nén đặt tên là **[MSSV_1, MSSV_2, MSSV_3].zip** sẽ gồm:
 1. Thư mục chứa dữ liệu/Link tới dữ liệu
 2. File PowerBI: **[MSSV_1, MSSV_2, MSSV_3].pbix**
 3. File báo cáo: **[MSSV_1, MSSV_2, MSSV_3].pdf**
- Trong trường hợp dữ liệu quá nặng, các bạn chỉ upload dữ liệu lên server ngoài như Google Drive, ..., nộp link và giữ link public ít nhất trong 2 năm; còn file PowerBI và báo cáo vẫn nộp trên Moodle.

5 Liên hệ

Mọi thắc mắc trong quá trình thực hiện vui lòng gửi mail về vntan.work@gmail.com