

<b>Module/môn: SQL Cơ bản</b>	<b>Số hiệu assignment:</b>	<b>Điểm (60%)</b>
<b>Người điều phối: Nguyễn Nghiệm</b>	<b>Ngày ban hành:</b>	
<i>Bài assignment này đòi hỏi sinh viên phải dùng khoảng 50h làm để hoàn thành. Giảng viên hướng dẫn sinh viên thực hiện trong suốt thời gian học tại lớp và đánh giá 2 lần trong quá trình học và 1 lần bảo vệ trước hội đồng.</i>		
<i>Mục tiêu: Sinh viên dựa vào đặc tả yêu cầu để lập kế hoạch thực hiện, viết báo cáo và hoàn thành các yêu cầu của assignment.</i>		

## CẢNH BÁO:

**Gian lận** là hình thức lấy bài làm của người khác và sử dụng như là mình làm ra. Hình thức đó bao gồm những hành động như: copy thông tin trực tiếp từ trang web hay sách mà không ghi rõ nguồn tham khảo trong tài liệu; gửi bài assignment làm chung như là thành quả cá nhân; copy bài assignment của các sinh viên khác cùng khóa hay khác khóa; ăn trộm hay mua bài assignment của ai đó và gửi lên như là sản phẩm mình làm ra. Những sinh viên bị nghi ngờ gian lận sẽ bị điều tra và nếu phát hiện là có gian lận thì sẽ phải chịu các mức phạt theo quy định của Nhà trường.

Mọi tài nguyên copy hay điều chế từ bất cứ nguồn nào (VD: Internet, sách) phải được đặt trong cặp dấu nháy kép và in nghiêng, với thông tin tham khảo đầy đủ về nguồn tài liệu.

Bài làm của bạn sẽ được đưa vào phần mềm kiểm tra gian lận. Mọi hình thức cố tình đánh lừa hệ thống phát hiện gian lận sẽ bị coi là Vi phạm quy định thi cử.

## QUY ĐỊNH NỘP BÀI ASSIGNMENT:

- ✓ Một bản mềm kết quả bài làm assignment của bạn phải được upload trước nửa đêm (giờ địa phương) vào ngày hạn nộp. Quá hạn nộp hệ thống sẽ khóa lại và sinh viên không còn quyền nộp bài.
- ✓ Phiên bản upload lên cuối cùng sẽ được chấm điểm. Sinh viên có quyền upload đè file nhiều lần trước khi hết hạn nộp.
- ✓ Tất cả những file tài liệu văn bản phải để ở dạng file gốc chứ không file dạng file được xuất ra từ định dạng khác (ví dụ pdf được xuất từ doc). Không được gửi tài liệu văn bản dưới dạng ảnh chụp.
- ✓ Đối với bài assignment này bạn cũng phải đưa các bằng chứng hay sản phẩm khác vào trong file nén dạng zip.
- ✓ Kích thước file cần tuân thủ theo giới hạn trên hệ thống nộp bài (thông thường là <50M).
- ✓ Hãy đảm bảo các file được upload lên không bị nhiễm virus (điều này có thể dẫn đến file bị hệ thống xóa mất) và không đặt mật khẩu mở file. Nếu vi phạm những điều này, bài coi như chưa được nộp.
- ✓ Hãy chú ý xem thông báo sau khi upload để chắc chắn bài của bạn đã được nộp lên hệ thống chưa.
- ✓ Bạn không phải gửi lại file đề bài của assignment (file này).

## QUY ĐỊNH ĐÁNH GIÁ BÀI ASSIGNMENT:

- ✓ Sinh viên không có bài assignment trên hệ thống sẽ bị 0 điểm bài assignment.
- ✓ Sau hạn nộp bài một tuần, sinh viên nộp muộn có quyền nộp đơn kiến nghị xin được chấp nhận gia hạn nộp. Hội đồng Nhà trường sẽ xét duyệt từng trường hợp. Nếu kiến nghị không được chấp nhận, bài giữ nguyên điểm 0. Nếu quá một tuần không có kiến nghị thì bài cũng sinh viên không nộp mặc nhiên nhận điểm 0.
- ✓ Ngay cả trường hợp bài của sinh viên bị phát hiện gian lận sau khi có điểm, sinh viên sẽ không được công nhận bài đó và chịu mức kỷ luật như quy định của Nhà trường.

## ASSIGNMENT XỬ LÝ DỮ LIỆU

<b>Mục tiêu cụ thể</b>	Sau bài assignment này , sinh viên có khả năng: <ul style="list-style-type: none"> <li>✓ Ứng dụng được quy trình xử lý dữ liệu</li> <li>✓ Kỹ năng làm sạch dữ liệu thô</li> <li>✓ Các kỹ thuật phân tích và xử lý dữ liệu</li> <li>✓ Trực quan hóa dữ liệu</li> <li>✓ Khả năng phân tích và dự báo dữ liệu</li> </ul>
<b>Các công cụ cần có</b>	<ul style="list-style-type: none"> <li>✓ MS. Office 2013+</li> </ul>
<b>Tài nguyên</b>	Thư mục tài nguyên đi kèm

## ĐẶT VẤN ĐỀ

Hiện tại dữ liệu của shop quần áo tại công ty Thiên Phát được thu thập về ở dạng file Excel đang gặp khá nhiều lỗi như: chính tả, khoảng trống thừa, lỗi định dạng tiền, ngày nhập hàng và xuất không đúng định dạng... và đặc biệt dữ liệu được sắp xếp rất lộn lộn, không tuân theo một quy tắc, quy chuẩn...như vậy không thể giúp công ty phân tích, tổng hợp và thống kê, gây ra rất nhiều khó khăn và thách thức trong việc đưa ra các chiến lược phát triển công ty trong thời gian tới. Bạn hãy xử lý dữ liệu trên để giúp công ty phân tích, thống kê và tổng hợp dữ liệu kinh doanh mình một cách chính xác và từ đó có những quyết định kinh doanh hiệu quả.

## YÊU CẦU THỰC HIỆN

**Y1.** Làm sạch dữ liệu

**Y2.** Chuyển đổi dữ liệu

**Y3.** Phân tích, thống kê và tổng hợp dữ liệu

**Y4.** Trực quan hoá dữ liệu

**Y5. Phân tích và dự báo dữ liệu**

## HƯỚNG DẪN THỰC HIỆN

✓ **GIẢI ĐOẠN 1: Hoàn thành yêu cầu Y.1**

- Sinh viên sử dụng file Excel QuanLy\_Shop\_QuanAo.xlsx trong thư mục tài nguyên đính kèm.
- Tiếp theo sinh viên áp dụng các hàm đã học để làm sạch dữ liệu cho file Excel trên như:
  - 1. Trim():** Xoá khoảng trống trước và sau text.
  - 2. Home -> Find & Select -> Go to Special...-> Blanks...** để xử lý ô trống.
  - 3. Chuyển giá trị Text sang số.**
  - 4. Data -> Remove Duplicates:** Xoá dữ liệu trùng nhau
  - 5. Chuyển định dạng Text từ thường sang Hoa cho các tiêu đề**
  - 6. Data - Text to column -> Delimited -> Other ->:** Tách nhanh dữ liệu
  - 7. Find và replace:** Tìm và thay thế giá trị lỗi

.....

✓ **GIẢI ĐOẠN 2: Hoàn thành yêu cầu Y.2 và Y.3**

- **Y.2: Các bước thực hiện**
  - 1) Lên ý tưởng, khảo sát nghiệp vụ
  - 2) Xác định các yếu tố đầu vào
  - 3) Xác định các yếu tố đầu ra (báo cáo, dashboard)
  - 4) Áp dụng các kỹ thuật chuẩn hoá để tổ chức lại dữ liệu thành dữ liệu có cấu trúc có quan hệ với nhau.
- **Y.3: Phân tích, thống kê và tổng hợp dữ liệu**
  - 1) Nhóm dữ liệu theo ngày tháng năm.
  - 2) Nhóm dữ liệu theo một khoảng giá trị.
  - 3) Tạo bảng Dashboard.
  - 4) Tạo báo cáo doanh thu.

.....

✓ **GIẢI ĐOẠN FINAL: Hoàn thành yêu cầu Y.4 và Y.5**

- **Y.4: Sử dụng đồ thị để:**
  - 1) Lập biểu đồ hình Pie trong không gian 3 chiều của cả công ty theo từng sản phẩm.
  - 2) Vẽ đồ thị trong không gian 2 chiều để thể hiện 3 đường biến đổi của của tình hình nhập, xuất và tồn kho của công ty theo từng tháng.
  - 3) Lập biểu đồ hình cột với trục tung là doanh thu và trục hoành là các sản phẩm.

phẩm.

○ **Y.5: Phân tích và dự báo**

- 1) Sử dụng hàm TREND để thực hiện dự đoán doanh thu 3 tháng tiếp theo dựa vào các tháng trước đó.
- 2) Sử dụng hàm FORECAST để thực hiện dự đoán doanh thu của năm tiếp theo dựa vào các năm trước đó.

## **MÔ TẢ SẢN PHẨM PHẢI NỘP**

Đóng gói tất cả theo định dạng <Tên nhóm/ mã sinh viên>\_ Mã Môn\_Assignment.zip (ví dụ: PS12345\_DAT101\_Assignment.zip). Nộp bài lên LMS theo yêu cầu của giảng viên

## **THANG ĐÁNH GIÁ**

<b>A</b> <b>(70%-100%)</b>	- Hoàn thành các yêu cầu từ Y1 đến Y5 - Nhóm bảo vệ trả lời được các câu hỏi của giảng viên
<b>B</b> <b>60%-69%</b>	- Hoàn thành các yêu cầu từ Y1 đến Y3 - Nhóm bảo vệ trả lời được các câu hỏi của giảng viên
<b>C</b> <b>50%-59%</b>	- Hoàn thành các yêu cầu từ Y1 đến Y2 - Nhóm bảo vệ trả lời tương đối các câu hỏi của giảng viên
<b>TRƯỢT</b> <b>(&lt; 50%)</b>	- Hoàn thành dưới 02 yêu cầu - Nhóm bảo vệ không trả lời được câu hỏi của giảng viên