

BÀI 8. SẮP XÉP VÀ LỌC DỮ LIỆU

(Thời lượng: 2 tiết)

A - Mục đích, yêu cầu

- Hiểu được nhu cầu sắp xếp, lọc dữ liệu;
- Biết các bước cần thực hiện sắp xếp, lọc dữ liệu.

B - Những điểm cần lưu ý và gợi ý dạy học

- a) Để bắt đầu bài này GV cần giới thiệu cho HS biết nhu cầu sắp xếp và lọc dữ liệu trong một danh sách dữ liệu là một nhu cầu thường xuyên trong cuộc sống, nhất là trong quản lí, ví dụ sắp xếp danh sách lớp học theo họ tên; sắp xếp theo kết quả học tập, sắp xếp các doanh nghiệp theo tên (theo ngành nghề, theo doanh thu),... Đây là những tính năng rất quan trọng và hiệu quả của chương trình bảng tính. HS cũng cần biết rằng bất cứ phần mềm ứng dụng nào có liên quan đến xử lí dữ liệu cũng đều phải có các tính năng này.

Có thể HS mới chỉ tiếp xúc với các bảng dữ liệu gồm quá ít hàng dữ liệu (5-50 hàng), do vậy việc so sánh và chọn các dữ liệu lớn nhất hay nhỏ nhất tương đối dễ dàng. Tuy nhiên, khi danh sách gồm nhiều trăm hàng dữ liệu thì việc so sánh và sắp xếp nhanh sẽ trở thành vấn đề khó khăn. Để minh họa điều này, GV nên chuẩn bị sẵn một trang tính với số hàng có dữ liệu đủ nhiều sao cho không thể hiển thị hết trên một trang màn hình. Giới thiệu trang tính này, GV giải thích cho HS biết được những khó khăn khi sắp xếp các hàng của bảng dữ liệu này. GV nên yêu cầu HS trả lời các câu hỏi trong phần mở đầu của bài học trong SGK (hoặc các câu hỏi tương tự do giáo viên đề xuất), từ đó HS sẽ nhận biết những khó khăn trong việc xử lý dữ liệu trên trang tính, đặc biệt là những trang tính có nhiều hàng, nhiều cột dữ liệu. Qua đó đặc biệt nhấn mạnh đến lợi điểm của việc sử dụng chương trình bảng tính để tính toán và xử lý dữ liệu.

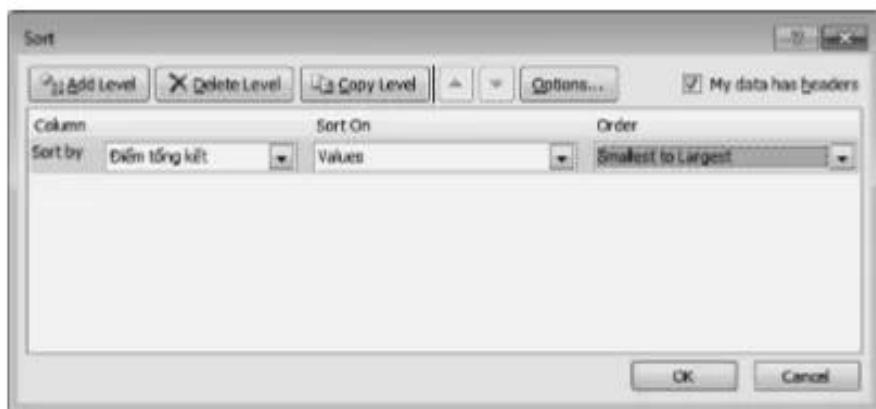
- b) Thao tác sắp xếp dữ liệu trên trang tính khá đơn giản, trước hết chọn một ô trong cột dữ liệu cần sắp xếp, sau đó nháy lệnh tương ứng trong nhóm **Sort & Filter** trên dài lệnh **Data** để sắp xếp theo chiều tăng dần hay giảm dần. Khi sắp xếp dữ liệu theo chiều tăng dần hay giảm dần của một cột nào đó thì dữ liệu tương ứng trên các hàng được di chuyển theo.

Một câu hỏi tự nhiên xuất phát từ nhu cầu thực tế là với ví dụ minh họa bảng hình trong SGK, liệu có thể sắp xếp HS theo tên (không chỉ theo họ và tên) được không? Đây là vấn đề liên quan đến việc tổ chức dữ liệu trên trang tính. Đây là một nội dung ẩn trong SGK. Việc thiết kế và tổ chức dữ liệu trên trang tính một cách hợp lí có vai trò rất quan trọng trong việc khai thác dữ liệu đó một cách có hiệu quả. Chúng ta đã biết, khi sao chép một công thức từ một ô tính sang một ô khác, địa chỉ tương đối trong công thức ở ô đích được điều chỉnh thích hợp. Tổ chức dữ liệu một cách hợp lí không chỉ phục vụ việc sắp xếp mà còn giúp cho việc thực hiện các tính toán nhanh và hiệu quả hơn. Nội dung tổ chức dữ liệu hợp lí trên trang tính cần phải được giới thiệu và rèn luyện kỹ năng cho HS trong suốt chương, ngay từ những bài học đầu tiên thông qua các bài tập và các bài thực hành. Vì vậy GV nên dành thời gian hướng dẫn HS thảo luận mở rộng nội dung này để HS hiểu được vai trò của việc tổ chức dữ liệu.

Với trang tính nói trên, để có thể sắp xếp dữ liệu tên HS thì cần tách cột *họ* và *tên* thành hai cột: cột *họ* và *tên đệm* và cột *tên*.

c) Đối với HS THCS, SGK chỉ giới thiệu cách *sắp xếp nhanh*. Dưới đây giới thiệu các thao tác sắp xếp chuẩn bằng hộp thoại Sort để GV tham khảo.

- Nháy chuột chọn một ô trong cột có dữ liệu cần sắp xếp của danh sách dữ liệu.
- Chọn lệnh Sort trong nhóm Sort & Filter trên tab lệnh Data để hiển thị hộp thoại Sort.
- Trên hộp thoại Sort được hiển thị ngay sau đó, thực hiện tiếp các thao tác được minh họa trên hình dưới đây. Sau đó chọn các tùy chọn sau:



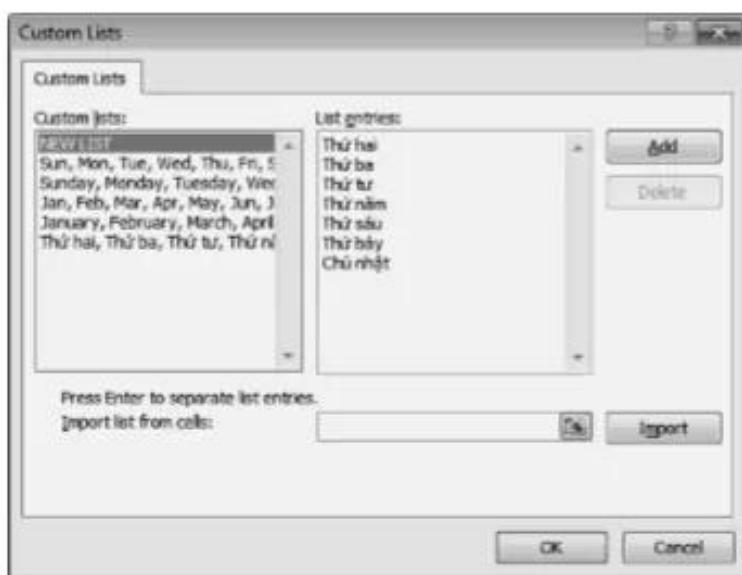
- Chọn tiêu đề của cột có dữ liệu cần sắp xếp trong ô Sort by.
- Chọn thứ tự sắp xếp tăng (Smallest to Largest) hay giảm (Largest to Smallest) trong ô Order.

Tùy chọn ngầm định trong ô Sort on là Values. Ta không chỉ có thể sắp xếp theo giá trị dữ liệu mà còn theo màu nền của ô, màu chữ,... Khi đó cần chọn tùy chọn thích hợp trong ô này.

Theo ngầm định, chương trình bảng tính chỉ sắp xếp dữ liệu văn bản theo thứ tự bảng chữ cái tiếng Anh. Để sắp xếp tên theo đúng thứ tự bảng chữ cái tiếng Việt cần phải tạo ra một danh sách thứ tự tự tạo (Custom List) bằng cách sau đây:

- Mở bảng chọn File, chọn Options để hiển thị hộp thoại Excel Options. Sau đó mở trang Advanced của hộp thoại và chọn Edit Custom Lists phía dưới mục General để hiển thị hộp thoại Custom Lists.

- Nháy chọn **NEW LIST** trong khung **Custom Lists** sau đó nhập các dữ liệu làm tiêu chí sắp xếp trong khung **List entries**. Nhấn phím **Enter** sau mỗi mục.
- Nháy **Add** khi kết thúc danh sách. Danh sách dữ liệu làm tiêu chí sắp xếp sẽ được đưa vào khung **Custom lists**.
- Cuối cùng nháy **OK** hai lần.



Để sắp xếp theo thứ tự tự tạo, trên hộp thoại **Sort** cần nháy chọn **Custom lists** trong ô **Order** để hiển thị hộp thoại **Custom Lists** và chọn thứ tự tự tạo thích hợp để sắp xếp.

- d) Ví dụ về bảng thành tích SeaGames, một thực tế là bảng này có thể được sắp xếp theo cả ba loại huy chương, trước hết là huy chương vàng, đến huy chương bạc và cuối cùng là huy chương đồng. GV có thể hướng dẫn HS chọn lệnh **Sort** trong nhóm **Sort & Filter** trên dài lệnh **Data** để hiển thị hộp thoại **Sort** và nháy chọn **My data has headers**. Sau đó chọn **Vàng** trong ô **Sort by** và nháy chọn **Largest to Smallest**, tiếp theo nháy **Add Level** và chọn **Bạc** trong ô **Then by** rồi chọn **Largest to Smallest**, cuối cùng nháy **Add Level** và chọn **Đồng** trong ô **Then by** thứ hai và chọn **Largest to Smallest**. Thao tác này sẽ thực hiện việc sắp xếp thứ tự theo chiều giảm dần số huy chương vàng, nếu số huy chương vàng bằng nhau sẽ sắp xếp tiếp theo chiều giảm dần của số huy chương bạc và nếu cùng số huy chương bạc sẽ sắp xếp tiếp theo chiều giảm

dẫn của số huy chương đồng. Tuy nhiên, đây không phải là nội dung bắt buộc đối với HS, GV chỉ nên giới thiệu các thao tác này khi điều kiện cho phép và phù hợp với khả năng hiểu bài của HS.

- e) Việc lọc dữ liệu thực chất là hiển thị dữ liệu thỏa mãn yêu cầu đặt ra và ẩn những dữ liệu không thỏa mãn yêu cầu. Tuy nhiên, việc lọc dữ liệu không làm thay đổi thứ tự các hàng, không sắp xếp lại dữ liệu. Nội dung chỉ yêu cầu HS thực hiện được việc lọc dữ liệu với lệnh **AutoFilter** và sử dụng các nút mũi tên xuất hiện ở đầu cột dữ liệu để lựa chọn hiển thị các hàng theo điều kiện nào đó.

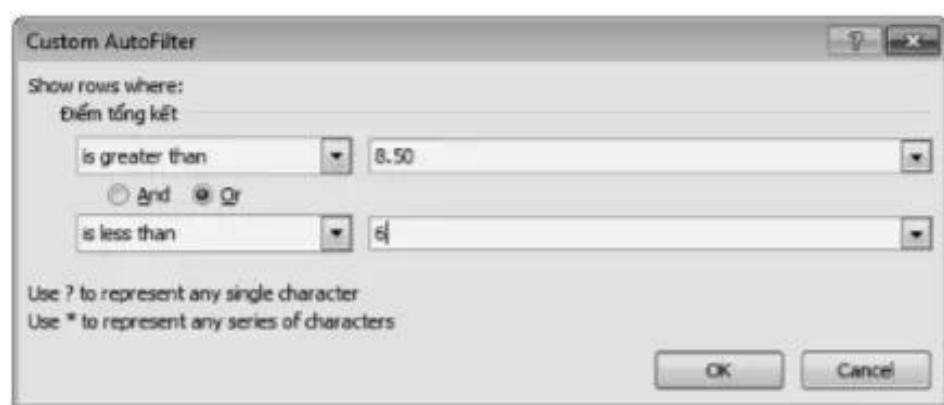
Với lựa chọn (**Top 10**) sau khi chọn **Number Filters** (hay **Text Filters**) với những GV đã quá quen thuộc với phiên bản XP cần lưu ý rằng với phiên bản Excel XP lựa chọn này lọc ra *các hàng có giá trị thuộc nhóm các giá trị dữ liệu lớn nhất hay nhỏ nhất*. Khác với các phiên bản trước, trong phiên bản Excel 2010 tác dụng của lệnh là lọc *số hàng nhất định* có giá trị dữ liệu lớn nhất hay nhỏ nhất (khác với phiên bản Excel XP). Ví dụ, không phải lọc *các HS có điểm trung bình thuộc nhóm 3 điểm cao nhất* mà là lọc *3 HS (3 hàng) có điểm trung bình cao nhất*. Nếu có 3 HS cùng đạt điểm cao nhất bằng nhau thì chỉ ứng với *một giá trị dữ liệu* trong cột *điểm trung bình*. Kết quả lọc cũng có thể cho số HS nhiều hơn 3 vì có thể có nhiều HS cùng điểm trung bình, chẳng hạn có 2 HS cùng đạt điểm cao nhất và 5 HS cùng đạt điểm cao thứ hai thì lọc không chỉ cho kết quả chính xác 3 HS mà là 7 HS. Việc lọc *các hàng có giá trị thuộc nhóm các giá trị dữ liệu lớn nhất hay nhỏ nhất* như trong các phiên bản trước có thể được thực hiện một cách đơn giản hơn bằng việc nháy chọn các *các giá trị dữ liệu* theo yêu cầu sau khi nháy nút mũi tên ở tiêu đề cột dữ liệu.

GV có thể giới thiệu thêm lọc kết hợp hai tiêu chí (chọn mục **Custom Filter...**) nhưng không nhất thiết phải đi sâu khai thác nội dung này. Dưới đây giới thiệu cách sử dụng tùy chọn **Custom Filter...** để GV tham khảo.

Sau khi chọn **Number Filters** (hay **Text Filters**) và chọn tùy chọn **Custom Filter...**, hộp thoại **Custom AutoFilter** hiện ra để giúp tạo ra các tiêu chí lọc phức hợp.

Hộp thoại **Custom AutoFilter** chứa hai cặp ô giống nhau. Các bước thực hiện như sau:

1. Chọn quan hệ cho tiêu chí thứ nhất trong ô bên trái.
2. Chọn hoặc nhập giá trị cho quan hệ thứ nhất trong ô bên phải tiêu chí.
3. Chọn **And** (và) hoặc **Or** (hoặc).
4. Chọn quan hệ cho tiêu chí thứ hai trong ô bên trái ở ngay dưới.
5. Chọn hoặc nhập giá trị cho quan hệ thứ hai trong ô bên phải tiêu chí.
6. Nhấn **OK**.



Không thực hiện các bước 3, 4 và 5 nếu chỉ cần dùng một quan hệ để lọc.

Các quan hệ làm tiêu chuẩn hay được sử dụng:

Quan hệ	Ý nghĩa
equals	Bằng
does not equal	Không bằng
is greater than	Lớn hơn
is greater or equal to	Lớn hơn hoặc bằng
is less than	Nhỏ hơn
is less than or equal to	Nhỏ hơn hoặc bằng
begins with	Bắt đầu bằng
ends with	Kết thúc bằng

Nếu ở bước 3 chọn **And**, dữ liệu được lọc phải thỏa mãn đồng thời cả hai tiêu chuẩn được định nghĩa, còn nếu chọn **Or** thì dữ liệu được lọc chỉ cần thỏa mãn một trong hai tiêu chuẩn.

Để thuận tiện cho HS, GV nên dịch một số thuật ngữ tiếng Anh hoặc gợi ý HS sử dụng phương pháp thử và sai để tìm hiểu ý nghĩa của các thuật ngữ tiếng Anh.

Hướng dẫn trả lời câu hỏi và bài tập

3. Nếu có nhiều hơn 3 hàng có cùng giá trị trong cột tương ứng ta có thể nhận được kết quả nhiều hơn ba nước. Ví dụ có hai nước Xinh-ga-po và Thái Lan có tổng số huy chương là 259 và 247, trong khi Việt Nam và Ma-lai-xi-a có cùng tổng số huy chương là 186 thì kết quả lọc là 4 (mặc dù yêu cầu lọc là **Top 3**).