

C. DẠNG PHÂN TÍCH TỔNG HỢP

Bài tập 1: Mở file dữ liệu Bai tap tong hop 1.xlsx và thực hiện các yêu cầu sau:

1. Tại sheet **Dữ liệu** hãy làm sạch dữ liệu theo yêu cầu:
 - 1.1. Làm sạch khoảng trắng thừa tại cột **Họ và tên**, cột **Trường**
 - 1.2. Xóa dòng dữ liệu trống
 - 1.3. Xóa dữ liệu trùng lặp
 - 1.4. Tách dữ liệu từ cột **Họ và tên** vào cột **Họ đệm** và cột **Tên** tương ứng.
 - 1.5. Tạo Data Validation cho cột **SAT**, với giá trị tại cột này chỉ nằm trong khoảng từ 400 tới 1600 khoanh tròn màu đỏ các giá trị không hợp lệ. Hiện thị cảnh báo Warning với Title là "Vui lòng nhập lại", và Error message là "Nhập giá trị từ 400 tới 1600". Với dữ liệu sai nhập lại giá trị 450 cho các dữ liệu sai này.
2. Sắp xếp dữ liệu tăng dần tại cột **Tên**, tăng dần tại cột **Họ đệm**, và giảm dần tại cột **SAT**.
3. Tại sheet **Dữ liệu**, định dạng dữ liệu kiểu Table với tên là **DuLieuSAT** cho vùng dữ liệu từ ô A2 tới ô giá trị cuối cùng của cột **Thời gian đọc báo giờ/tuần**.
4. Tại sheet **Dữ liệu**, định dạng có điều kiện cho Top 10 với kiểu Red text.
5. Sử dụng hàm:
 - 5.1. Tại sheet **Dữ liệu**, dùng hàm công thức cho tại ô "?" để tìm kiếm điểm **SAT** theo họ tên nhập vào trong ô B1.
 - 5.2. Tại sheet **Summary**, hãy thực hiện các yêu cầu sau:
 - Tính điểm trung bình SAT trung bình của thí sinh theo từng trường.
 - Đếm số lượng thí sinh theo từng trường.
 - Tính thời gian đọc báo trung bình theo từng trường.
6. Tạo **Pivot Table** cho **Bảng DuLieuSAT** của sheet **Dữ liệu** trong sheet mới. Đặt tên sheet mới này là **Pivot**.
 - 6.1. Hãy tạo Pivot table có tên Pivot1 với yêu cầu thống kê số lượng thí sinh và tỉ lệ % sinh viên từng trường so với tổng số thí sinh. Định dạng các giá trị lấy hai số phần thập phân.
 - 6.2. Copy Pivot 1 và dán bên cạnh đổi tên thành Pivot 2 đó sửa lại để thống kê điểm **SAT** trung bình theo từng trường và thời gian đọc báo trung bình theo từng trường. Định dạng các giá trị lấy hai số phần thập phân.
 - 6.3. Copy Pivot 1 và dán bên cạnh Pivot 2, đổi tên thành Pivot 3. Sau đó sửa lại để thống kê để thống kê điểm **SAT** trung bình theo từng **Major** và thời gian đọc báo trung bình theo từng **Major**.
 - 6.4. Chèn slicer là **Trường** định dạng 3 cột dùng cho cả ba pivot, chèn slicer **Major** định dạng 1 cột dùng cho Pivot 3.
 - 6.5. Đổi tên các trường trong các Pivot cho phù hợp.
7. Tạo Pivot Chart

Tại sheet **Pivot**, trên cơ sở 3 Pivot vừa với tạo hãy:

- Với dữ liệu Pivot 1, tạo Pivot chart kiểu Clustered Column-Line on Secondary Axis.
- Với dữ liệu Pivot 2, tạo Pivot chart kiểu Clustered Column-Line on Secondary Axis.
- Với dữ liệu Pivot 3, tạo Pivot chart kiểu Clustered Column-Line on Secondary Axis.

8. Nhận xét: Hãy hoàn thành một số nhận xét về dữ liệu sau khi thống kê tại sheet Nhận xét.

9. Định dạng dữ liệu tại các sheet cho đẹp mắt.

Bài tập 2. Mở dữ liệu Bai tong hop 2.xlsx và thực hiện các yêu cầu sau:

1. Tại sheet Dữ liệu

1.1. Làm sạch các lỗi spelling tại cột **Country**.

1.2. Làm sạch các dòng trống.

1.3. Đồng bộ dữ liệu ở cột tiêu đề dạng viết hoa kí tự đầu tiên từng chữ.

1.4. Đồng bộ dữ liệu ở cột **Address** dạng viết hoa chữ cái đầu tiên.

1.5. Loại bỏ tất cả các khoảng trắng thừa.

1.6. Sao lưu 1 bản dữ liệu sau đó loại dữ liệu trùng lặp.

1.7. Tách dữ liệu ở cột **Cust_name** thành hai cột **FirstName** và **LastName** dùng Flass fill hoặc Text to column.

1.8. Định dạng dữ liệu kiểu bảng với tên là **DuLieu** cho vùng dữ liệu từ ô A2 cho tới hàng chứa dữ liệu cuối cùng của cột Y.

2. Tại sheet Dữ liệu: Sắp xếp dữ liệu tăng dần theo **Order_Ship_Date** và giảm dần theo **Order_Value**

3. Sử dụng hàm:

3.1. Tại sheet **Dữ liệu**, tại ô ? Trên dòng 2, hãy áp dụng hàm vlookup (hoặc hàm khác) để tìm kiếm **Order_Value** theo tên khách hàng **Cust_Name** nhập vào.

3.2. Tại **Order_Type** hãy dùng công thức điền dữ liệu dựa và giá trị cột **Order_Value** như sau:

- Nếu **Order_Value** có giá lớn hơn 300 thì **Order_Type** là "Large",
- Ngược lại nếu **Order_Value** có giá trị lớn hơn 200 thì **Order_Type** là "Medium",
- Ngược lại thì **Order_Type** là "Small".

3.3. Tại sheet **Summary**

- Tính số lần đặt hàng bị Cancelled theo từng khách hàng
- Tính tổng giá trị hàng bị Cancelled theo từng khách hàng

4. Định dạng có điều kiện cho cột **Order_Value**: Theo kiểu định dạng là Icon sets loại 4 Arrow

5. Tạo Pivot Table cho table DuLieu:

5.1. Tạo Pivot thống kê số lượng kiểu T_Type và tổng giá trị hàng đã đặt **Order_Value** theo từng khách hàng.

5.2. Chèn Slicer T_Type

5.3. Sắp xếp dữ liệu cho cột tổng giá trị hàng hóa theo thứ tự từ lớn tới nhỏ.

5.4. Định dạng dữ liệu kiểu Currency cho tổng giá trị hàng theo đơn vị tiền tệ là \$

6. Tạo Pivot Chart: Tạo Pivot chart kiểu Bar cho Pivot Table đã tạo.

7. Nhận xét: Hoàn thành dữ liệu tại sheet **Nhận xét**

8. Định dạng các sheet dữ liệu cho đẹp mắt.

Bài tập 3: Mở file Bai tap tong hop 4.xlsx và thực hiện các yêu cầu:

1. Tại sheet **Home Values**, làm sạch dữ liệu:

1.1. Loại bỏ lỗi Spelling.

1.2. Đồng bộ dữ liệu viết hoa kí tự đầu tiên của từng chữ cột Region.

1.3. Loại bỏ các khoảng trắng thừa.

1.4. Loại bỏ các dữ liệu trùng lặp.

1.5. Định dạng dữ liệu kiểu tiền tệ \$ cho các dữ liệu của các năm 1970, 1980, 1990, 2000, 2010.

2. Dùng công thức điền dữ liệu vào các cột sau:

2.1. Dùng công thức điền dữ liệu trung bình giá nhà các năm vào cột Average Value.

2.2. Dùng công thức xếp hạng giá nhà năm 2010 theo từng bang (state) tại cột 2010 Rank.

2.3. Dùng công thức để nhận định giá nhà là "High", "Moderate", "Low" trong cột **Home Pricing** biết:

- Nếu giá nhà năm 2010 lớn hơn 225000 thì **Home Pricing** là "High",
- Ngược lại nếu giá nhà năm 2010 lớn hơn 150000 thì **Home Pricing** là "Moderate",
- Ngược lại **Home Pricing** là "Low".

3. Tại Sheet **Summary**, hãy tính trung bình giá nhà theo từng Region của từng năm.

4. Tại sheet Home Values, hãy định dạng có điều kiện các dữ liệu 1970, 1980, 1990, 2000, 2010 theo điều kiện như sau:

Hãy định dạng theo rule là Format all cells based on their values, với Format Style là 3-Color Scale trong đó Minimum là giá trị thấp nhất có màu đỏ, Midpoint là Percentile giá trị là 50, Maximum là giá trị cao nhất."

5. Chèn Sparkline theo kiểu Line cho giá theo từng năm của từng State với Maker color bất kì.

6. Tạo Pivot Table trong sheet mới thể hiện giá trị thống kê giá trung bình theo từng năm 1970, 1980, 1990, 2000, 2010 của từng Region.

6.1. Chèn Slicer để lọc Region.

6.2. Định dạng dữ liệu kiểu tiền tệ \$ cho giá trung bình từng năm của các vùng.

6.3. Đổi tên sheet này là Pivot.

7. Tạo Pivot Chart kiểu Column cho Pivot Table vừa tạo.

8. Tại sheet **Nhận xét** hãy hoàn thành nhận xét về kết quả thống kê dữ liệu này.

9. Định dạng các bảng dữ liệu cho đẹp mắt.