



**FPT POLYTECHNIC**



---

[caodang.fpt.edu.vn](http://caodang.fpt.edu.vn)

---

## **NHẬP MÔN XỬ LÝ DỮ LIỆU**

---

### **BÀI 5: XỬ LÝ DỮ LIỆU**

# MỤC TIÊU

- ◎ CƠ SỞ DỮ LIỆU TRONG EXCEL
- ◎ TÌM KIẾM THÔNG TIN
- ◎ SẮP XẾP VÀ LỌC DỮ LIỆU
- ◎ SỬ DỤNG PIVOT TABLE



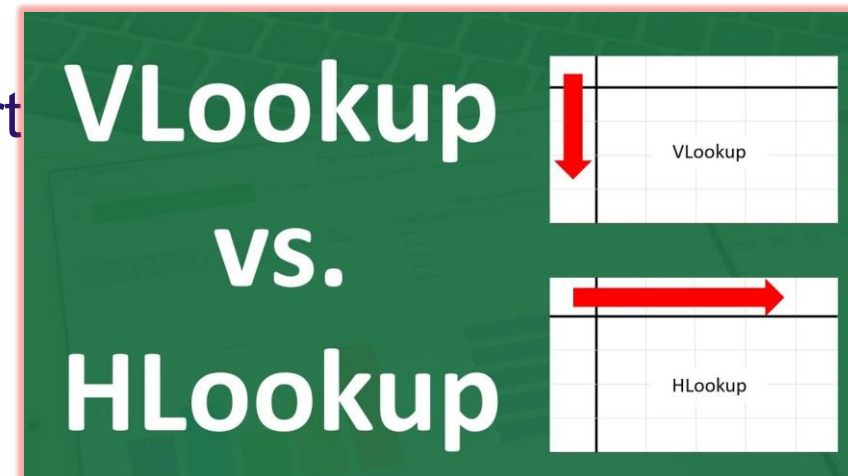
- ❑ Giới thiệu xử lý dữ liệu
- ❑ Cơ sở dữ liệu trong Excel
- ❑ Tìm kiếm thông tin
- ❑ Sắp xếp và lọc dữ liệu
- ❑ Sử dụng PivotTable và PivotChart



# PHẦN I: XỬ LÝ DỮ LIỆU

## ❑ Xử lý dữ liệu là gì ?

- ❖ Gồm các tình huống **phân tích, tìm kiếm, thống kê và tổng hợp** tự động với các hàm liên quan tới số và chữ.
- ❖ Xử lý dữ liệu số: abs, sqrt, sin, cos, round, int, if, count, sum, sumif, countif..
- ❖ Xử lý dữ liệu văn bản: left, right, mid, len, find, search, ... dùng kết hợp với hàm if và các hàm tìm kiếm Vlookup, Hlookup.
- ❖ Thống kê và lọc dữ liệu với PivotTable, PivotChart



## Cơ sở dữ liệu là gì?

- ❖ Là tập hợp thông tin được tổ chức theo kiểu danh sách có thể tìm kiếm, lọc, xóa những dòng dữ liệu thỏa mãn một tiêu chuẩn một cách nhanh chóng.
- ❖ Mỗi danh sách gồm:
  - **Field (trường):**
    - ✓ Hàng đầu tiên của danh sách chứa tên trường.
    - ✓ Tên của các trường phải là dạng chuỗi và khác nhau.
  - **Record (bản ghi):**
    - ✓ Các hàng còn lại chứa thông tin về một đối tượng trong danh sách

## ❑ Các hàm cơ sở dữ liệu

❖ Cấu trúc tổng quát: **Dfunction(Database, Field, Criteria)**

- ✓ **Database:** Là một cơ sở dữ liệu dạng danh sách của Excel
- ✓ **Field:** là tên của trường
- ✓ **Criteria:** là vùng điều kiện cần thiết mà hàm phải thỏa mãn đều tác động lên trường dữ liệu đã chỉ ra bởi Field.
- ✓ Hàm cơ sở dữ liệu sẽ trả lại kết quả tính toán phù hợp với những ràng buộc được chỉ ra ở vùng điều kiện.

## □ Các hàm cơ sở dữ liệu:

- ❖ Tính tổng các số trong một cột của cơ sở dữ liệu thỏa mãn điều kiện.  
**=DSUM(Database,Field, Criteria)**
- ❖ Tính trung bình cộng các số trong một cột của cơ sở dữ liệu thỏa mãn điều kiện.  
**=DAVERAGE(Database,Field, Criteria)**
- ❖ Đếm các ô chứa số trong một cột của cơ sở dữ liệu thỏa mãn điều kiện.  
**=DCOUNT(Database,Field, Criteria)**
- ❖ Đếm các ô chứa dữ liệu trong một cột của cơ sở dữ liệu thỏa mãn điều kiện.  
**=DCOUNTA(Database,Field, Criteria)**
- ❖ Hàm tìm giá trị lớn nhất/nhỏ nhất của các ô trên một cột của cơ sở dữ liệu.  
**=DMAX(Database,Field, Criteria)**  
**=DMIN(Database,Field, Criteria)**



VÍ DỤ:

[illegible]

## □ Hàm Vlookup:

- ❖ **VLOOKUP**(Giá trị tìm, vùng cần tìm, cột cần lấy, kiểu tìm): trả lại giá trị ở "cột cần lấy" của "Giá trị tìm" trong "vùng cần tìm".
  - **Giá trị tìm**: một trong những giá trị ở cột đầu tiên của vùng cần tìm.
  - **Vùng cần tìm**: nằm ở bảng phụ ngoài vùng cơ sở dữ liệu.
  - **Cột cần lấy**: là thứ tự cột cần lấy trong vùng cần tìm
  - **Kiểu tìm**: 0 tìm chính xác hoặc 1 tìm tương đối.

## ❑ Hàm Hlookup:

- ❖ HLOOKUP(Giá trị tìm, vùng cần tìm, cột cần lấy, kiểu tìm): trả lại giá trị ở "dòng cần lấy" của "Giá trị tìm" trong "vùng cần tìm".
  - *Giá trị tìm*: một trong những giá trị ở dòng đầu tiên của vùng cần tìm.
  - *Vùng cần tìm*: nằm ở bảng phụ ngoài vùng CSDL
  - *Dòng cần lấy*: là thứ tự dòng cần lấy trong vùng cần tìm
  - *Kiểu tìm*: 0 tìm chính xác hoặc 1 tìm tương đối.

📄 Ví dụ Vlookup:

E6		fx =VLOOKUP(C6,\$A\$1:\$B\$3,2,0)*D6								
	A	B	C	D	E	F	G	H	I	J
1	A	200000								
2	B	100000								
3	C	50000								
4										
5	TT	Họ tên	Loại	Số công	Tiền					
6	1	Tuấn	A	28	5600000					
7	2	Anh	C	25	1250000					
8	3	Hùng	A	17	3400000					
9	4	Thanh	B	27	2700000					
10										
11	<p>Tính lương cho ba loại công lao động khác nhau, biết rằng số ngày công cho từng loại là:          Loại A: 200000 đồng/công, Loại B: 100000 đồng/công, Loại C: 50000 đồng/công          Tiền lương = số ngày công x tiền công một ngày</p>									

[illegible]

## □ Sắp xếp dữ liệu:

- ❖ *Sắp xếp dữ liệu* là hoán đổi vị trí các hàng để giá trị dữ liệu trong một hay nhiều cột được sắp theo thứ tự tăng dần hay giảm dần.

### Chú ý:

- ❖ *Thứ tự cột có dữ liệu kiểu kí tự là thứ tự theo bảng chữ cái tiếng anh.*
- ❖ *Dữ liệu trong các cột khi sắp xếp phải có cùng kiểu (kí tự, số, ...)*

## □ Sắp xếp dữ liệu:

❖ Các bước thực hiện sắp xếp dữ liệu.

The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The 'Filter' button in the 'Sort & Filter' group is highlighted with a red box and the number 2. The 'Data' tab itself is highlighted with a red box and the number 1. Below the ribbon, a formula bar shows the formula  $=IF(((G3+H3)/2+I3)>10,10,G3+H3)/2+I3$ . The main area displays a table titled 'BẢNG KẾT QUẢ THI CỦA CÁC LỚP' (Exam Results Table of Classes). The table has columns: 'Tên' (Name), 'Pha' (Phase), 'Ngày Sinh' (Date of Birth), 'Lớp' (Class), 'Lý Thuyết' (Theory), and 'Th' (Test). The 'Pha' column header is highlighted with a red box and the number 3. The table contains 10 rows of data.

	Tên	Pha	Ngày Sinh	Lớp	Lý Thuyết	Th
	Kim	TRUE	3/13/1982	D	8.0	
	Thu	FALSE	8/21/1980	C	6.5	
nh	Loan	TRUE	3/22/1981	D	4.5	
	Duy	FALSE	7/1/1979	B	8.0	
c	Hà	TRUE	9/3/1976	B	6.5	
	Yến	TRUE	7/4/1971	B	4.0	
	Mai	TRUE	9/12/1974	C	8.0	
	Bình	FALSE	1/6/1971	D	6.0	
	Minh	FALSE	8/31/1977	C	7.0	



## Lọc dữ liệu:

❖ Là chọn và chỉ hiển thị các hàng thỏa mãn các tiêu chuẩn nhất định nào đó.

### **Bước 1: Chuẩn bị (mở chế độ lọc):**

- Nháy chuột => chọn 1 ô trong vùng có dữ liệu cần lọc.

- Mở lệnh Data => chọn lệnh Filter  nhóm **Sort & Filter**

### **Bước 2: Lọc**

-Nháy chuột vào biểu tượng mũi tên  trên hàng tiêu đề cột có giá trị dữ liệu cần lọc.

-Nháy chuột chọn các giá trị dữ liệu cần lọc trên danh sách hiện ra  
chọn **OK**.



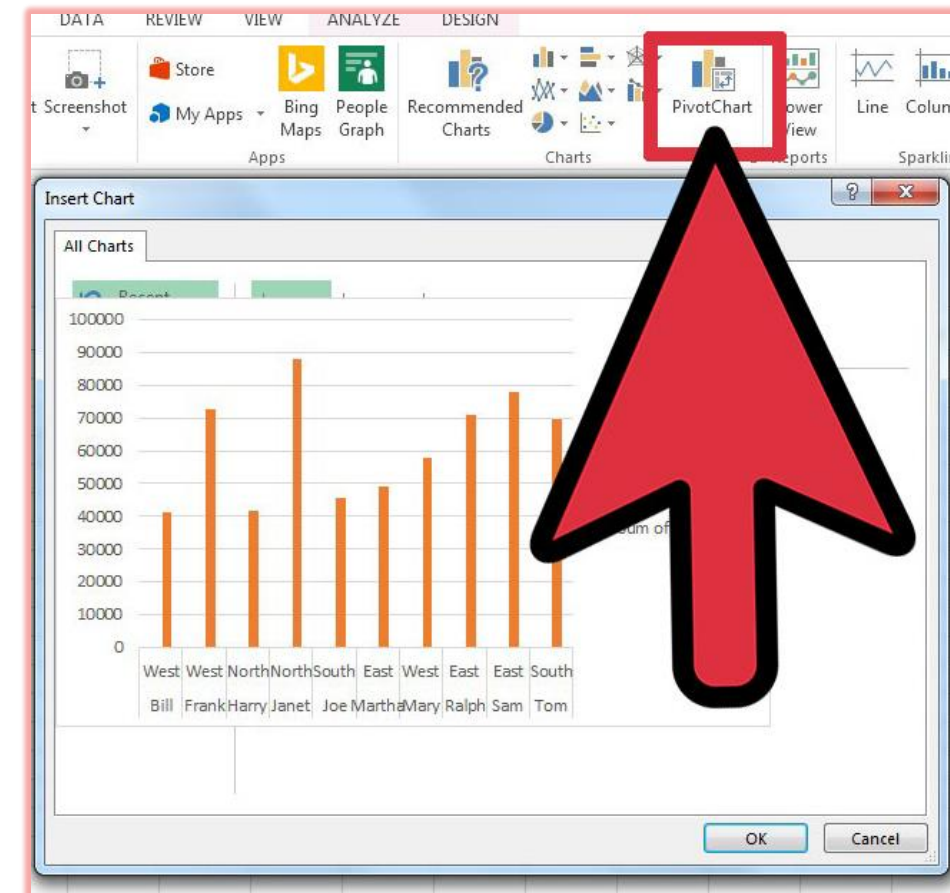


## PHẦN II: XỬ LÝ DỮ LIỆU (TT)

---

## ❑ Giới thiệu PivotTable

- ❖ PivotTable là một cách sắp xếp lại dữ liệu bảng tính từ đó có thể phân tích dữ liệu trong nhiều cách khác nhau.
- ❖ Hiển thị dữ liệu dạng biểu đồ từ dữ liệu của bảng **PivotTable**.



## ❑ Tạo PivotTable

❖ Giao diện của PivotTable gồm 02 phần:

- PivotTable Field List dialog box
- PivotTable pane

The screenshot shows an Excel spreadsheet with a PivotTable summarizing student loan amounts by class level and major. The PivotTable is located in the range A3:F16. The PivotTable Field List task pane is open on the right side of the screen.

**PivotTable Data:**

Row Labels	Column Labels	Junior	Senior	Sophomore	Grand Total	
Freeshman		5136	4059		9197	
Biology		11868			11868	
Chemistry		9803	1411	3244	14458	
Computers		1531	1400	0	3021	
Economics			1490	3823	1433	6745
Education		5000	3425	0021	1240	10292
English		234	676	1373		2283
Finance		1895	8375	450	4309	15029
History			3318	1433	0240	13091
Mathematics		1450			1675	3125
Music		1232		3300		4532
Grand Total		38151	21343	19901	21147	100542

**Annotations:**

- PivotTable tabs:** Points to the PivotTable Tools ribbon tabs (Options and Design).
- Column Labels down arrow:** Points to the dropdown arrow next to the Column Labels header in the PivotTable.
- Row Labels area with field:** Points to the Row Labels area in the PivotTable Field List task pane, which currently shows 'Major'.
- Row Labels down arrow:** Points to the dropdown arrow next to the Row Labels header in the PivotTable.
- View PivotTable in PivotTable Pane:** Points to the 'View PivotTable' button in the PivotTable Field List task pane.

## ❑ Tạo PivotTable

### ❖ Các bước thực hiện:

1. Click vào một ô trong bảng dữ liệu trên bảng tính hoặc chọn một vùng dữ liệu cho **PivotTable**
2. Ở nhóm **Tables** trên thẻ **Insert** tab, click vào nút
3. Click nút **Select a table or range** kiểm lại vùng tham chiếu trong ô **Table/Range box**
4. Click nút **New Worksheet** hoặc nút **Existing worksheet**
5. Click nút **OK**
6. Click hộp kiểm (check box) tại trường (field) muốn thêm vào PivotTable
7. Nếu cần sắp xếp lại, kéo (drag) các trường bỏ sang các ô khác nhau trong mục layout section

## Ví dụ:

- ❖ Cho một danh sách các tác phẩm nghệ thuật tranh vẽ, hãy thống kê tổng giá trị thẩm định của các tác phẩm nhóm theo điều kiện ?

Museum - Microsoft Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW Team Sign in

A1 : ArtID

	A	B	C	D	E	F	G	H
1	ArtID	Artist	Title	Date Acquired	Category	Condition	Location	Appraised Value
2	1	Mogan	Red Rock Mountain	3/19/2005	Painting	Excellent	East Pavilion	\$ 18,000
3	2	Novarre	Offerings	5/16/2005	Painting	Excellent	East Pavilion	\$ 10,000
4	3	Chico	Spring Flowers	3/20/2004	Sculpture	Excellent	East Pavilion	\$ 2,400
5	4	Roman	Seeking Shelter	10/8/2005	Sculpture	Excellent	Courtyard	\$ 52,000
6	5	DiGrigoro	The Hang	7/16/2004	Painting	Excellent	East Pavilion	\$ 8,000
7	6	Ibe	House Remembered	8/16/2004	Sculpture	Good	East Pavilion	\$ 700
8	7	Zischke	Homage to the Ancestors	7/17/2004	Textile	Excellent	East Pavilion	\$ 1,200
9	8	Gilhooly	End of the Path	8/16/2004	Sculpture	Excellent	East Pavilion	\$ 1,900
10	9	Guys	Amen	12/13/2003	Sculpture	Excellent	East Pavilion	\$ 3,000
11	10	Swartz	Untitled (two figures)	9/8/2007	Sculpture	Excellent	East Pavilion	\$ 800
12	11	Dill	Eve	7/16/2004	Sculpture	Excellent	East Pavilion	\$ 975
13	12	Udinotti	Man on horseback	11/10/2007	Painting	Good	East Pavilion	\$ 8,000
14	13	Aserty	Superstitions	1/10/2009	Painting	Excellent	Courtyard	\$ 78,000
15	14	McGraw	Plenty	2/19/2005	Sculpture	Excellent	West Pavilion	\$ 500
16	15	Mclver	Punch	3/19/2005	Painting	Excellent	East Pavilion	\$ 10,000
17	16	Schenck	Untitled	1/7/2006	Painting	Excellent	East Pavilion	\$ 6,000
18	17	Bindner	Brittlecone	1/10/2004	Sculpture	Excellent	East Pavilion	\$ 1,300
19	18	Blum	Mountain Scene	1/10/2004	Sculpture	Excellent	East Pavilion	\$ 2,500
20	19	Reed	The White Heart	8/14/2005	Painting	Excellent	East Pavilion	\$ 9,300

Art Collection Sheet3

READY 100%

❑ Ví dụ:

❖ Thực hiện các bước tạo PivotTable:

The screenshot displays a Microsoft Excel window titled "Museum - Microsoft Excel". The ribbon includes FILE, HOME, INSERT, PAGE LAYOUT, FORMULAS, DATA, REVIEW, VIEW, Team, and ANALYZE. The active cell is A3, and the formula bar shows "Row Labels".

The PivotTable is located in the range A3:D8. The data is as follows:

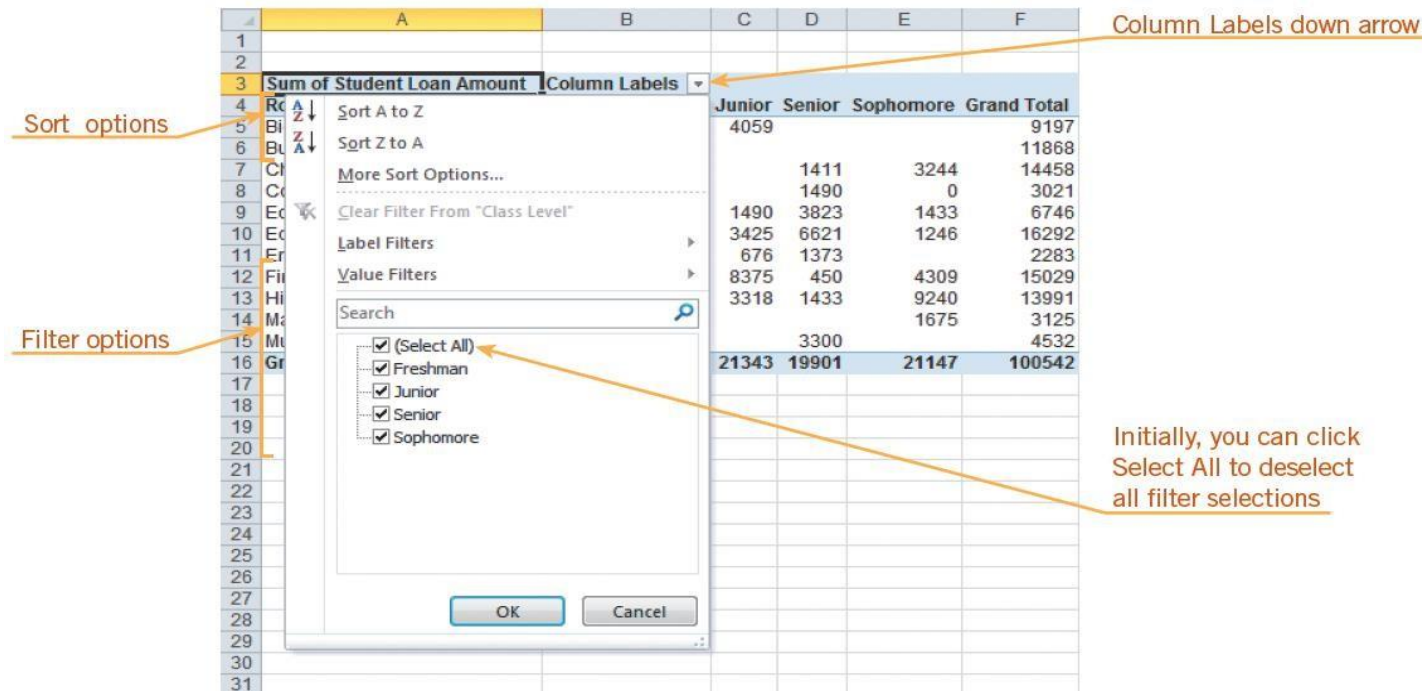
Row Labels	Sum of Appraised Value
Excellent	966712
Fair	51692
Good	84569
Poor	12650
<b>Grand Total</b>	<b>1115623</b>

The PivotTable Fields task pane on the right shows the following configuration:

- Choose fields to add to report:**
  - ☐ Date Acquired
  - ☐ Category
  - ☒ Condition
  - ☐ Location
  - ☒ Appraised Value
- Drag fields between areas below:**
  - FILTERS:** (Empty)
  - COLUMNS:** (Empty)
  - ROWS:** Condition
  - VALUES:** Sum of Appraised ...
- ☐ Defer Layout Update
- UPDATE** button

## ❑ Sắp xếp và lọc dữ liệu PivotTable:

- ❖ Tùy chọn sắp xếp và lọc dữ liệu sẽ được hiển thị khi click vào:
  - ✓ Nút mũi tên xuống **Column Labels** hoặc Nút mũi tên xuống **Row Labels**
- ❖ Chọn **sort option** để sắp xếp dữ liệu theo thứ tự tăng hay giảm dần.



Column Labels down arrow

Sort options

Filter options

Initially, you can click Select All to deselect all filter selections

Class Level	Junior	Senior	Sophomore	Grand Total
4059				9197
		1411	3244	11868
		1490	0	3021
1490	3823	1433		6746
3425	6621	1246		16292
676	1373			2283
8375	450	4309		15029
3318	1433	9240		13991
		1675		3125
	3300			4532
<b>21343</b>	<b>19901</b>	<b>21147</b>		<b>100542</b>

## ❑ Hiệu chỉnh PivotTable:

- ❖ Có thể sắp xếp lại dữ liệu sau khi tạo PivotTable.
- ❖ Để thay đổi vị trí dòng và cột:
  - ✓ Hiển thị lại hộp thoại **PivotTable Field List**.
  - ✓ Dịch chuyển thứ tự các trường (field) trong hộp thoại.
- ❖ Có thể bổ sung thêm mức chi tiết của trường.
- ❖ Khi thêm các mức chi tiết của trường, nút **Outline** xuất hiện trong PivotTable.



## Hiệu chỉnh PivotTable:

- PivotTable với hai mức chi tiết (levels of Row Label detail).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2													
3	Sum of Student Loan Amount	Column Labels											
4	Row Labels	Junior	Senior	Grand Total									
5	Music		3300	3300									
6	Caldwell		1450	1450									
7	Hernandez		1850	1850									
8	History		3318	1433	4751								
9	Aannestad		2500		2500								
10	Feraco		775	775									
11	Prager		658	658									
12	Rosenberg		818	818									
13	Finance		8375	450	8825								
14	Bennett		0		0								
15	Hargus		1358		1358								
16	Jackson			450	450								
17	Johnson		775		775								
18	Laporte		3333		3333								
19	Lightford		1336		1336								
20	Smith		1573		1573								
21	English		676	1373	2049								
22	Nader			1373	1373								
23	Thompson		676		676								
24	Education		3425	6621	10046								
25	Ascot		0		0								
26	Lopez			2640	2640								
27	Middleton		3425		3425								
28	Prescott			462	462								
29	Selinger			3519	3519								
30	Economics		1490	3823	5313								
31	Coats			200	200								
32	Cortez			623	623								
33	Gattis		1490		1490								
34	Velasquez			3000	3000								

PivotTable Field List

Choose fields to add to report:

☒ Last Name  
☐ First Name  
☒ Class Level  
☒ Major  
☐ Grade Point Average  
☒ Student Loan Amount

Report Filter

Column Labels  
Class Level

Row Labels  
Major  
Last Name

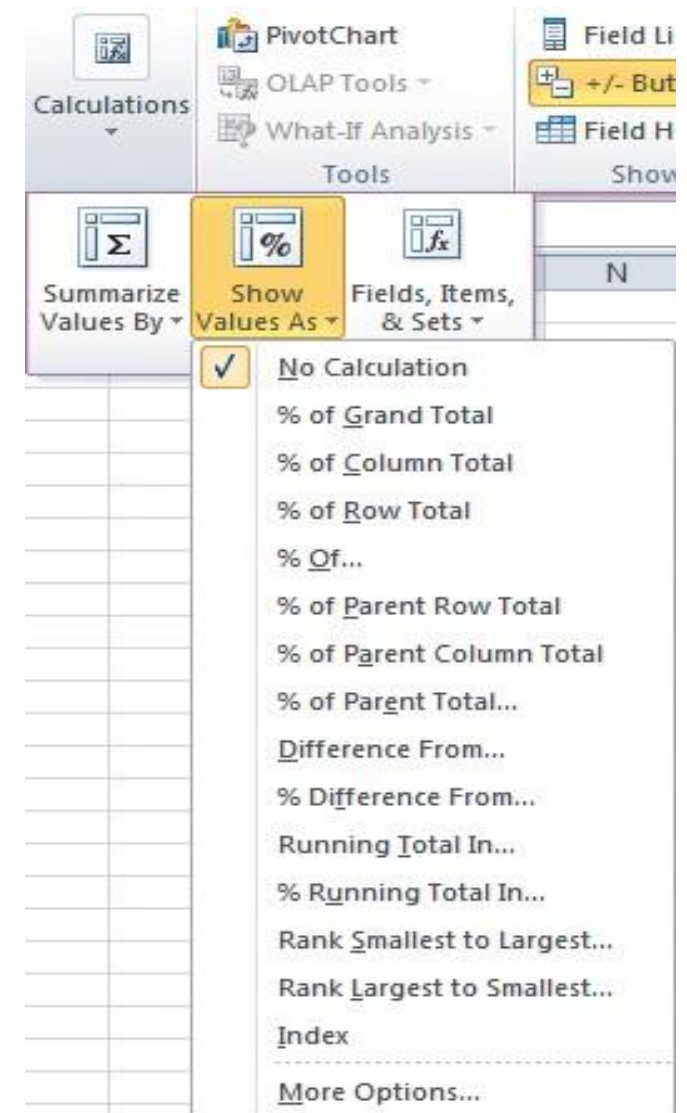
Σ Values  
Sum of Student Loan Amount

Defer Layout Upd...

Update

## ❑ Thay đổi định dạng và giá trị tính toán:

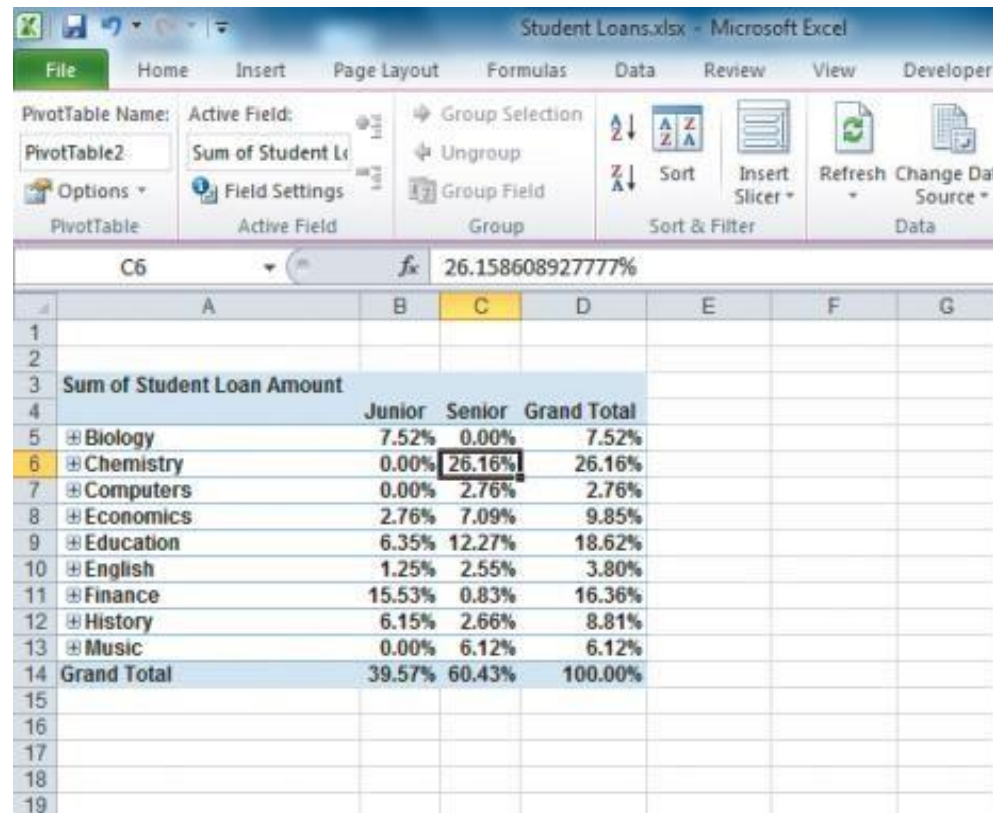
- ❖ Có thể thay đổi định dạng của dữ liệu PivotTable
  - ✓ **Ví dụ:** hiển thị giá trị số dạng dollar, %, hoặc thêm dấu phân cách phần ngàn,...
- ❖ Show Values As list
  - ✓ Cho phép thay đổi cách thức giá trị được hiển thị hoặc tính toán.



## ❑ Cập nhật dữ liệu:

❖ Để cập nhật dữ liệu PivotTable phản ánh sự thay đổi dữ liệu bảng tính:

- ✓ Click nút **Refresh**, ở nhóm **Data** trên thẻ **Options** tab.

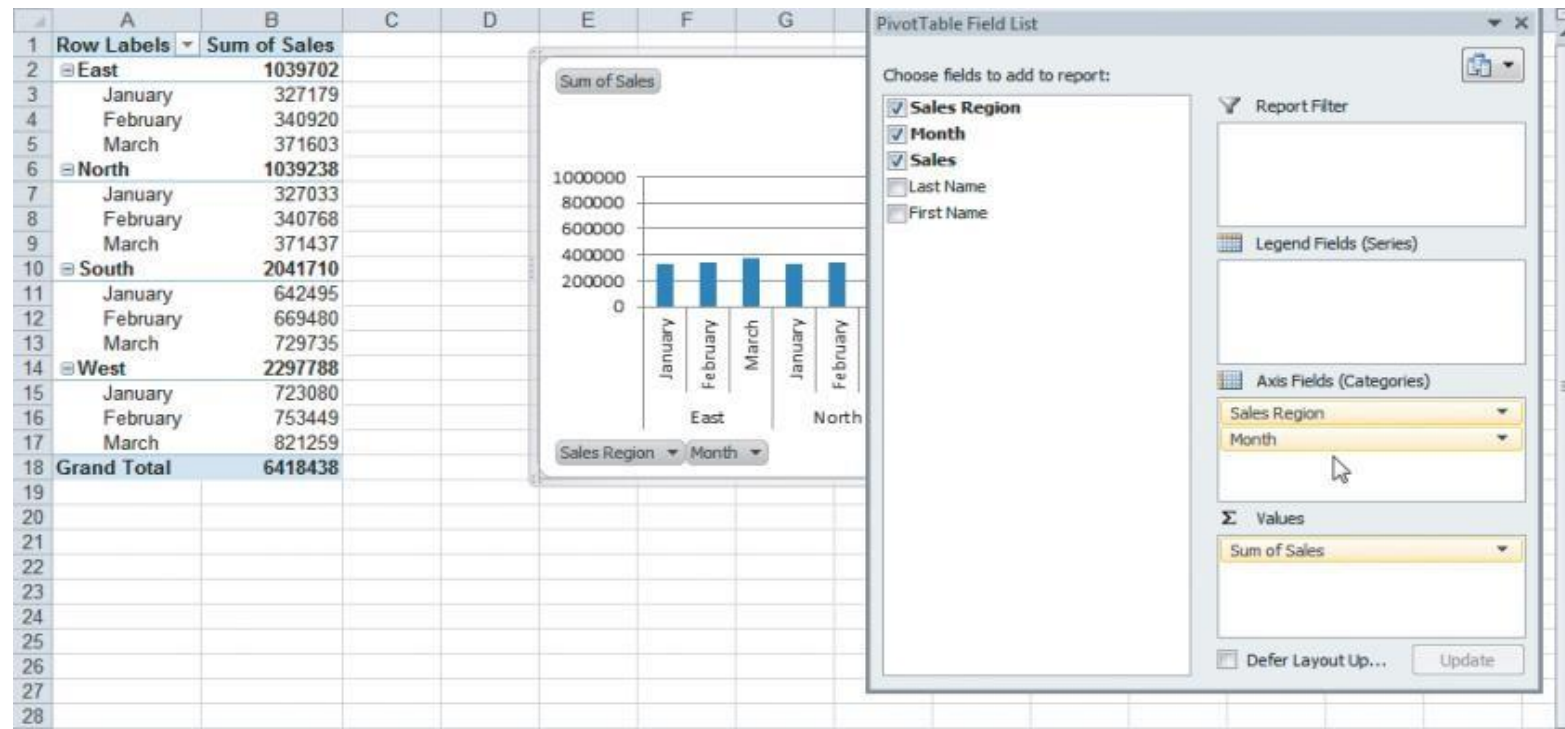


The screenshot shows the Microsoft Excel interface with the 'Student Loans.xlsx' file open. The 'Options' tab is active, and the 'Refresh' button is highlighted in the 'Data' group. The PivotTable below shows the 'Sum of Student Loan Amount' by subject.

	Junior	Senior	Grand Total
Biology	7.52%	0.00%	7.52%
Chemistry	0.00%	26.16%	26.16%
Computers	0.00%	2.76%	2.76%
Economics	2.76%	7.09%	9.85%
Education	6.35%	12.27%	18.62%
English	1.25%	2.55%	3.80%
Finance	15.53%	0.83%	16.36%
History	6.15%	2.66%	8.81%
Music	0.00%	6.12%	6.12%
Grand Total	39.57%	60.43%	100.00%

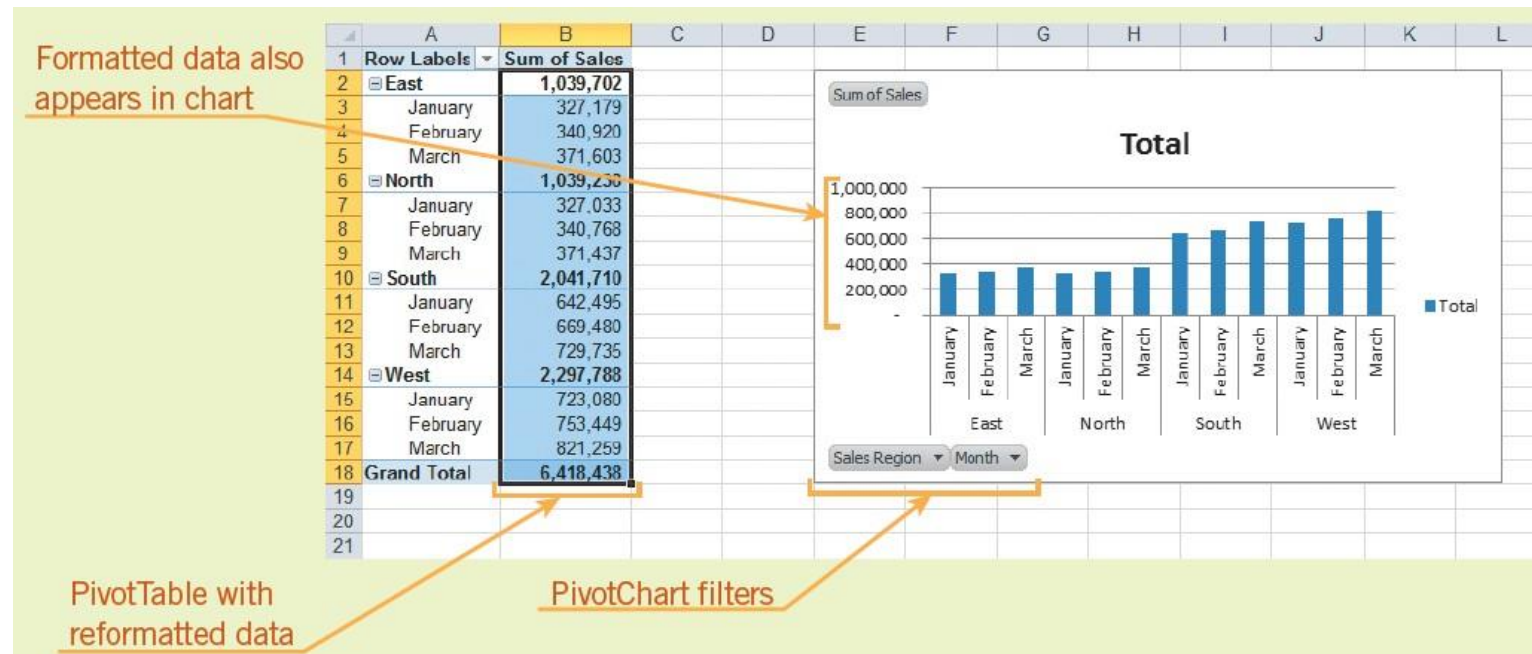
## PivotChart:

- ❖ **PivotChart** là một cách trình bày dữ liệu PivotTable một cách **trực quan**.
- ❖ Khi tạo PivotChart với dữ liệu bảng tính thì PivotTable cũng tạo cùng lúc.
- ❖ **PivotTable** và **PivotChart** được link với nhau.



## PivotChart:

- ❖ Khi lọc dữ liệu trên PivotChart thì dữ liệu cũng được lọc trên PivotTable.
- ❖ Bộ lọc Chart filters xuất hiện trên chính biểu đồ.



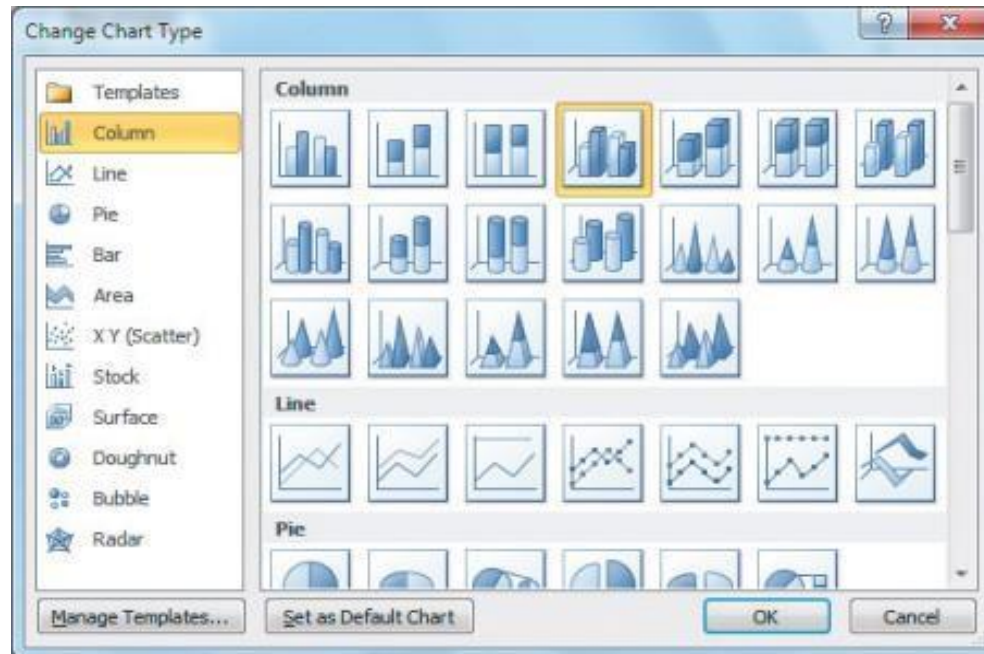


## ❑ PivotChart:

❖ PivotChart có thể được chỉnh sửa như biểu đồ Excel.

**Ví dụ:** có thể thay đổi kiểu chart type và chart styles.

❖ Khi **PivotChart** được chọn thì các thẻ bổ sung thêm sẽ xuất hiện trên **Ribbon**





- ☑ Các kỹ thuật xử lý dữ liệu quan trọng như tìm kiếm, sắp xếp và lọc dữ liệu.
- ☑ Phân tích thống kê và lọc dữ liệu với PivotTable, PivotChart.



FPT POLYTECHNIC

**Thank you**