

SIT103/SIT772: Database Fundamentals



9.1P: Data Analysis & Visualization using Excel

Overview

In this task, you will learn to use Microsoft Excel for simple data analysis and visualization using Pivot Tables/Charts.

Pivot Tables/Charts are useful built-in data-analysis feature in Excel that help to summarize/visualize data. They allow us to build an interactive dashboard. They are called a Pivot because fields can be moved within the table/chart to create different types of summary lists, providing a “pivot.” They offer flexible and intuitive analysis of data.

Although the data that appear in Pivot Tables look like any other worksheet, the data in the data area of the Pivot Table cannot be directly entered or changed. The Pivot Table is linked to the source data; the output in the cells of the table are read-only data. The formatting (number, alignment, font, etc.) can be changed as well as a variety of computational options such as SUM, AVERAGE, MIN, and MAX.

Tasks To Do

In this task, you are expected to summarize and visualize data in the “sales.xlsx” file available in the task resources .zip file. The data file looks like this.

Order_ID	Order_Date	Cust_Name	Salesman_Name	Prod_Name	Prod_Category	Prod_Sales
4	9-Apr-15	International Paper	Summer Payne	Crucial	Storage	92684
4	9-Apr-15	International Paper	Summer Payne	Western Digital WDS500G1B0B	Storage	7044.36
4	9-Apr-15	International Paper	Summer Payne	Asus Z10PE-D16 WS	Mother Board	35249.25
4	9-Apr-15	International Paper	Summer Payne	G.Skill Ripjaws V Series	Storage	95338.6
4	9-Apr-15	International Paper	Summer Payne	MSI X99A XPOWER GAMING TITANIUM	Mother Board	19799.4
4	9-Apr-15	International Paper	Summer Payne	Corsair Dominator Platinum	Storage	57599.2
4	9-Apr-15	International Paper	Summer Payne	Intel Xeon E5-2690 V4	CPU	191471.04
4	9-Apr-15	International Paper	Summer Payne	Corsair Dominator Platinum	Storage	114549.21
41	11-May-17	Emerson Electric	Summer Payne	Western Digital WD10EZEX	Storage	3830.4
41	11-May-17	Emerson Electric	Summer Payne	Intel Xeon E5-2660 V3	CPU	106577.86
41	11-May-17	Emerson Electric	Summer Payne	Intel Core i7-7820X	CPU	58372.5
41	11-May-17	Emerson Electric	Summer Payne	PNY SSD7CS1311-120-RB	Storage	2841.02
41	11-May-17	Emerson Electric	Summer Payne	Asus SABERTOOTH X99	Mother Board	25951.61
41	11-May-17	Emerson Electric	Summer Payne	MSI X99A XPOWER GAMING TITANIUM	Mother Board	36958.88
41	11-May-17	Emerson Electric	Summer Payne	Samsung MZ-75E250B/AM	Storage	10592.88
41	11-May-17	Emerson Electric	Summer Payne	Corsair Dominator Platinum	Storage	23249.69
41	11-May-17	Emerson Electric	Summer Payne	Crucial CT525MX300SSD4	Storage	9210.39
82	3-Dec-16	Jabil Circuit	Summer Payne	Kingston SA400S37/120G	Storage	7588.62

You are required to create a **Pivot Table/Chart** to see **Total sales per product category over years with**

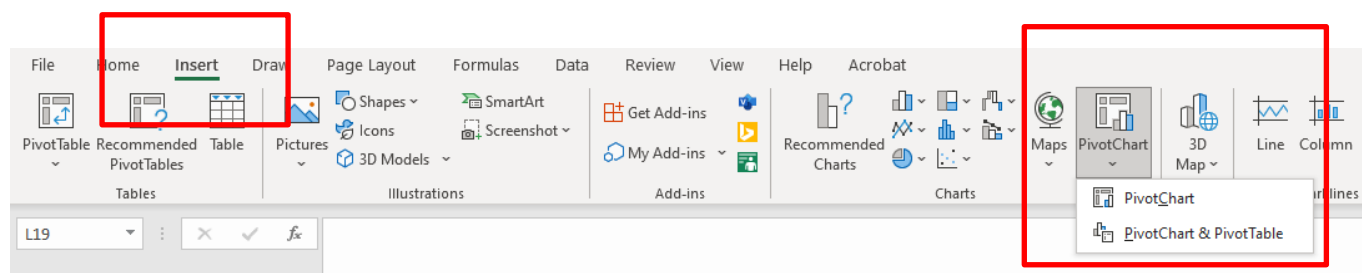
ability to filter per salesman and customer.

You can watch the following YouTube videos for more information.

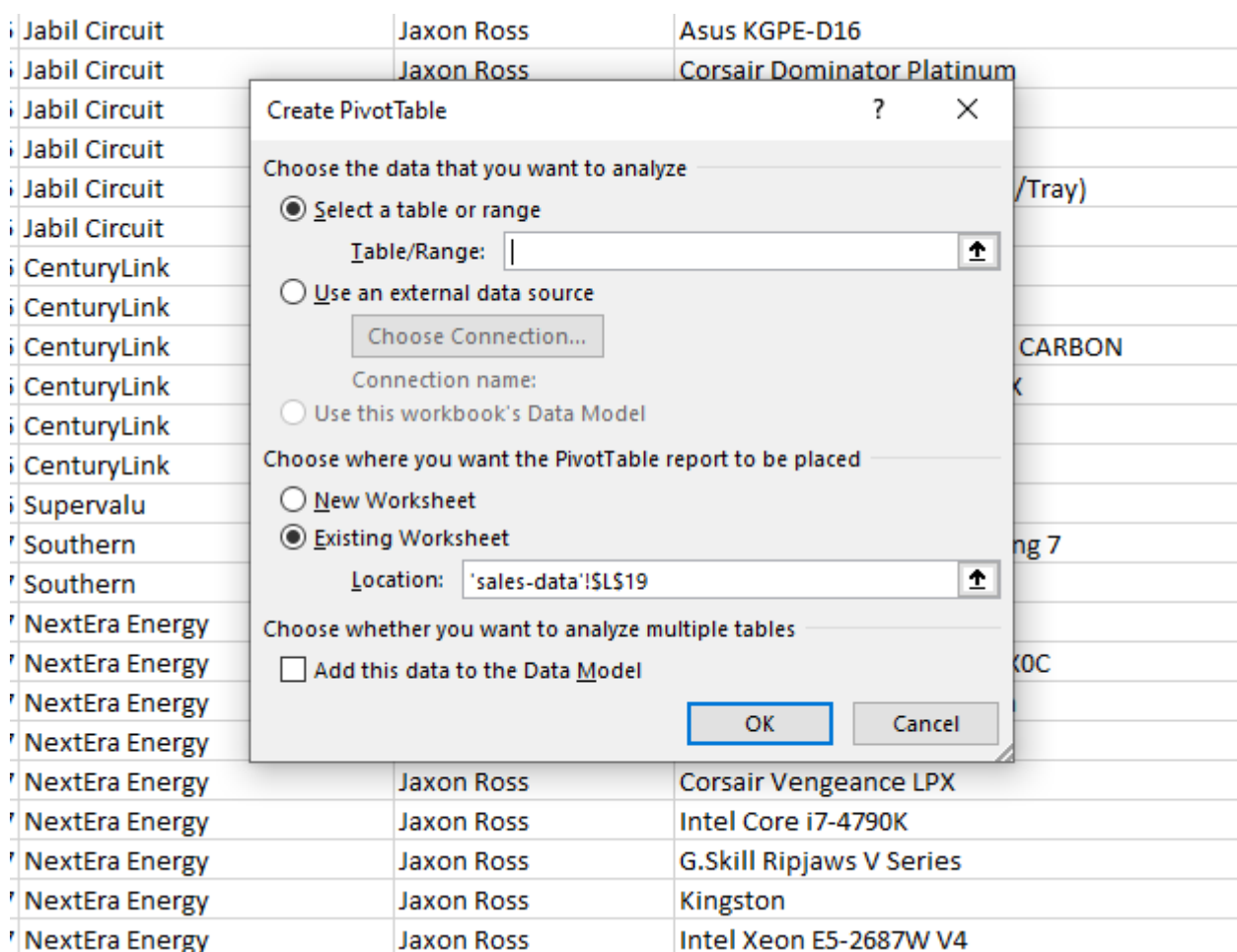
https://www.youtube.com/watch?v=9NUjHBNWe9M&ab_channel=ExcelCampus-Jon

https://www.youtube.com/watch?v=g530cnFfk8Y&ab_channel=ExcelCampus-Jon

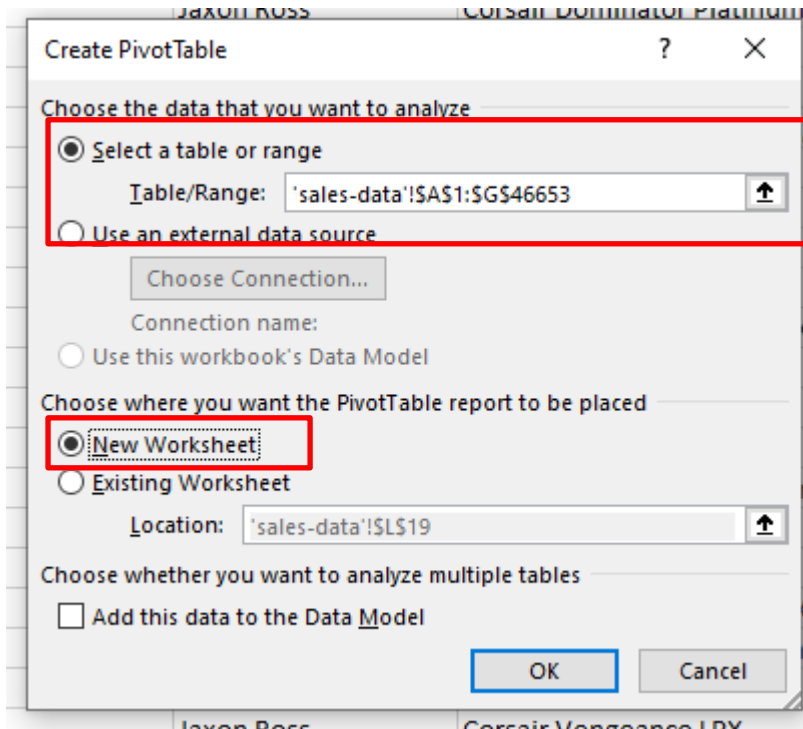
1. Open the 'sales.xlsx' file.
2. Go to "Insert" menu and then "PivotChart".



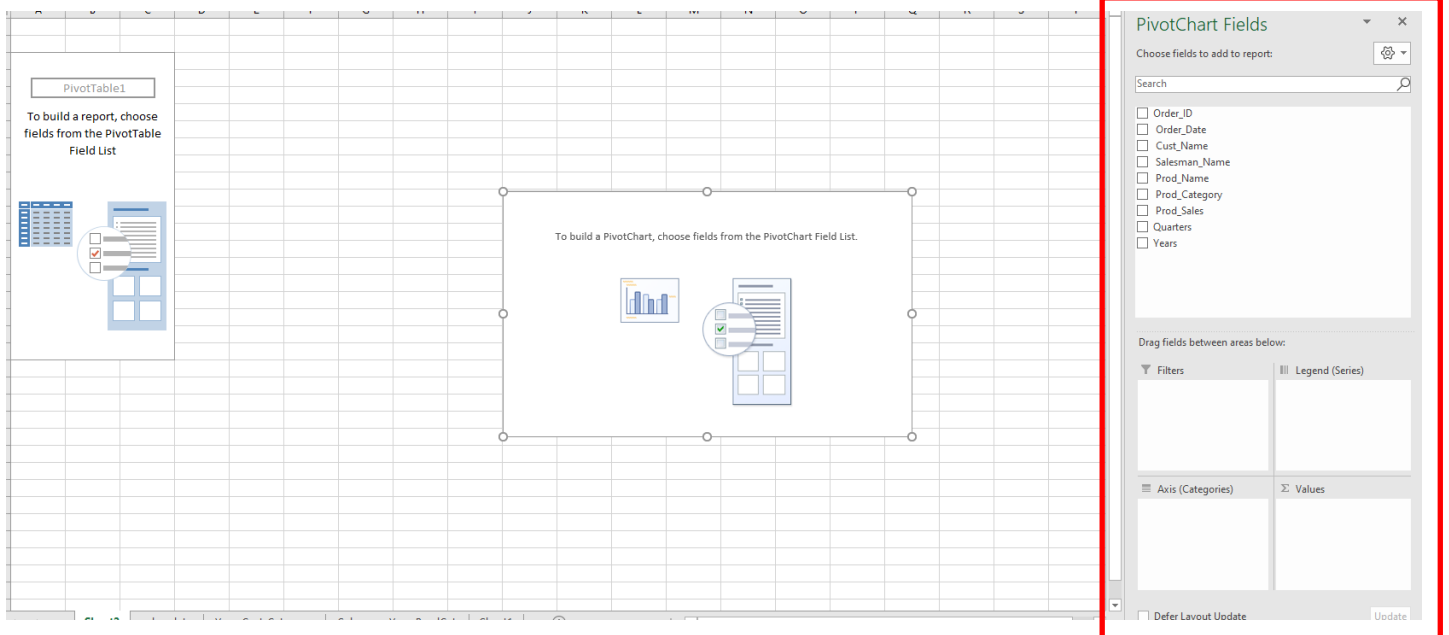
3. Choose "PivotChart & PivotTable", the following dialogue box will appear



4. For Table/Range, select the entire range of data including the column header line in your data worksheet and choose “New Worksheet” to place PivotTable/Chart in a new worksheet.



5. You will have a PivotTable/Chart template appear in a new worksheet. Select anywhere within the PivotTable/Chart area, you will see PivotChart Fields and Areas on the right.



6. Drag and drop fields in the Pivot areas as shown in the following screen shoot:

You can see ‘Years’ and ‘Quarters’ fields in the list which are not in the original data. They will appear if you check/select ‘Order_Date’ once. Excel recognizes the data in this field as ‘Date’ and gives you the option to apply ‘Year’ and ‘Quarters’ filters.

PivotChart Fields

Choose fields to add to report:



Search



- ☐ Order_ID
- ☐ Order_Date
- ☒ **Cust_Name**
- ☒ **Salesman_Name**
- ☐ Prod_Name
- ☒ **Prod_Category**
- ☒ **Prod_Sales**
- ☐ Quarters
- ☒ **Years**

Drag fields between areas below:

Filters

Salesman_Name ▼

Cust_Name ▼

Legend (Series)

Prod_Category ▼

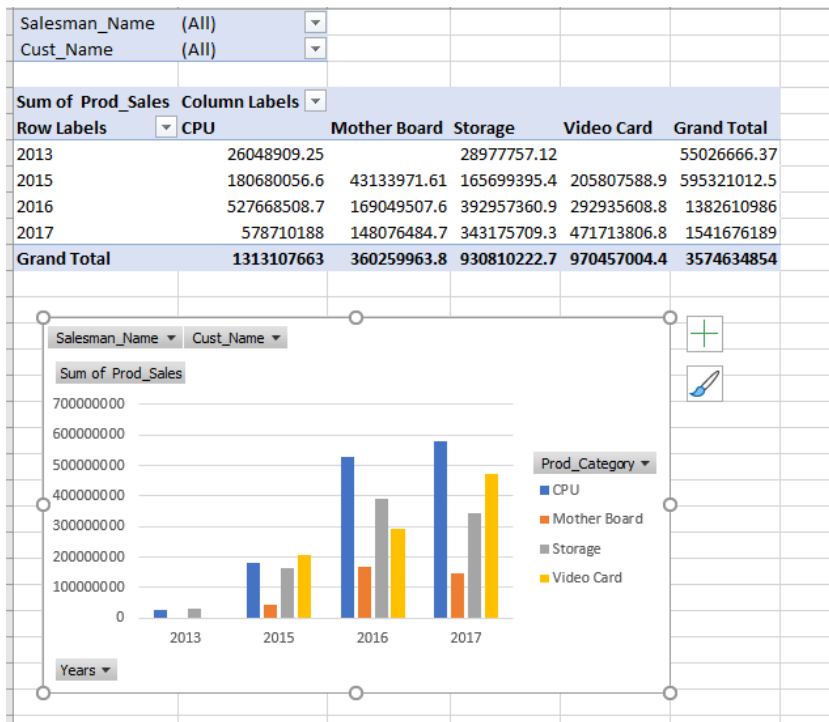
Axis (Categories)

Years ▼

Values

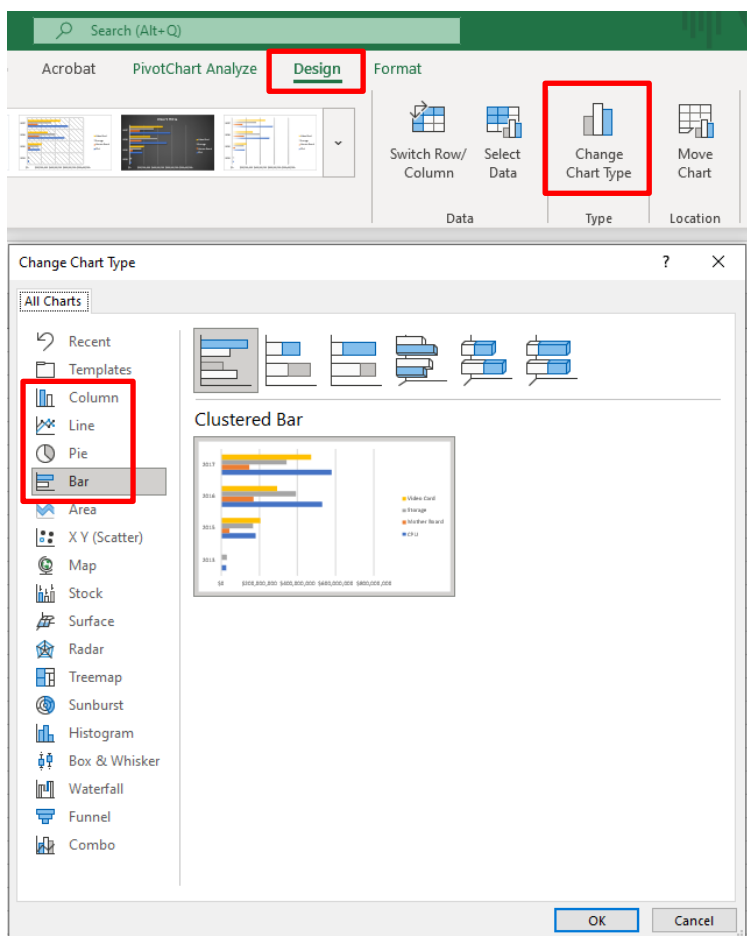
Sum of Prod_Sales ▼

7. You will see total sales for each product category over years.



8. You can use filters to visualize data for particular salesman(s), Customer(s), year(s), Prod_Category(ies).

9. You can change Chart Type as:

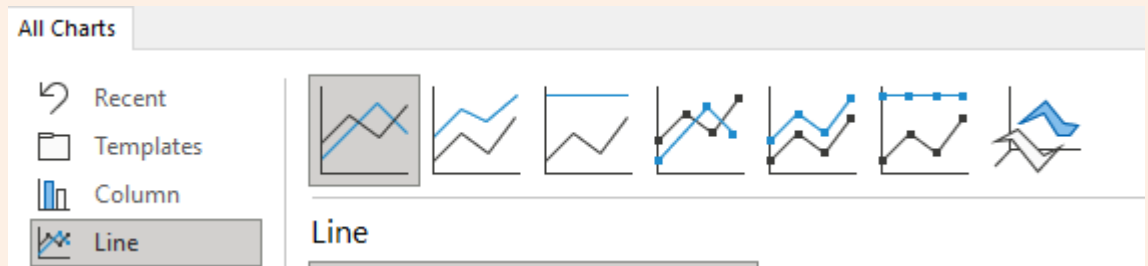


Try other fields, chart types, and play with it. Practice PivotTable and PivotChart, they are simple but very useful tools that comes with standard Excel package that allow us to build a simple dashboard for data visualization.

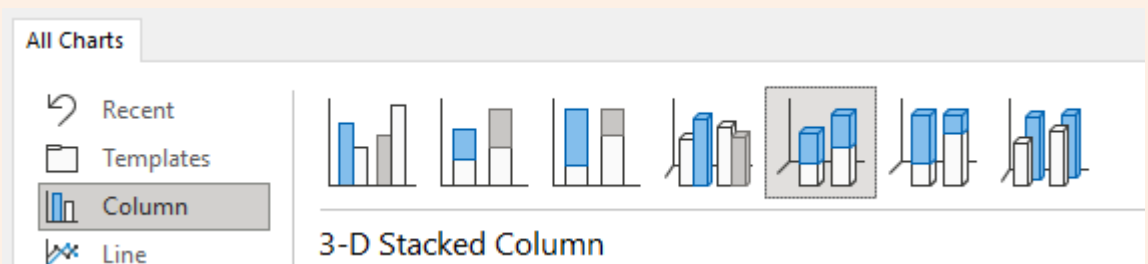
Submission Requirements:

Submit one PDF/WORD file with the following graphs:

1. Line plot of all four product categories over all years with total sale by all customers and all salesmen. This will show the trend of sales over years for each product category.



2. 3-D Stacked Column plot of sales of Mother Board and CPU product categories by the salesman named "Ellis Washington" only.



ADDITIONAL REQUIREMENT FOR SIT772 STUDENTS ONLY

3. Choose any real-world business (e.g., Coles, Qantas, eBay, Facebook, Netflix, etc.). Those who are aiming for a 'C' or above and did the Mini-Project (Task 4.2C and 7.2C) can use the same company you selected for the project if you want. In 750-1000 words, discuss the importance of data and Business Intelligence (BI) for your chosen company. Your discussion must include: (i) what the company is; (ii) what is its main business; (iii) what data they collect about what; and (iv) how they use the collected data and BI tools to have competitive advantage in terms of revenue and/or productivity. Note that (iv) is the most important part in this additional task and it is expected that 75-80% of your word limit will be used to answer this part.

Submission Due

The due for each task has been stated via its OnTrack task information dashboard.