Table of Contents

Define Data Analysis

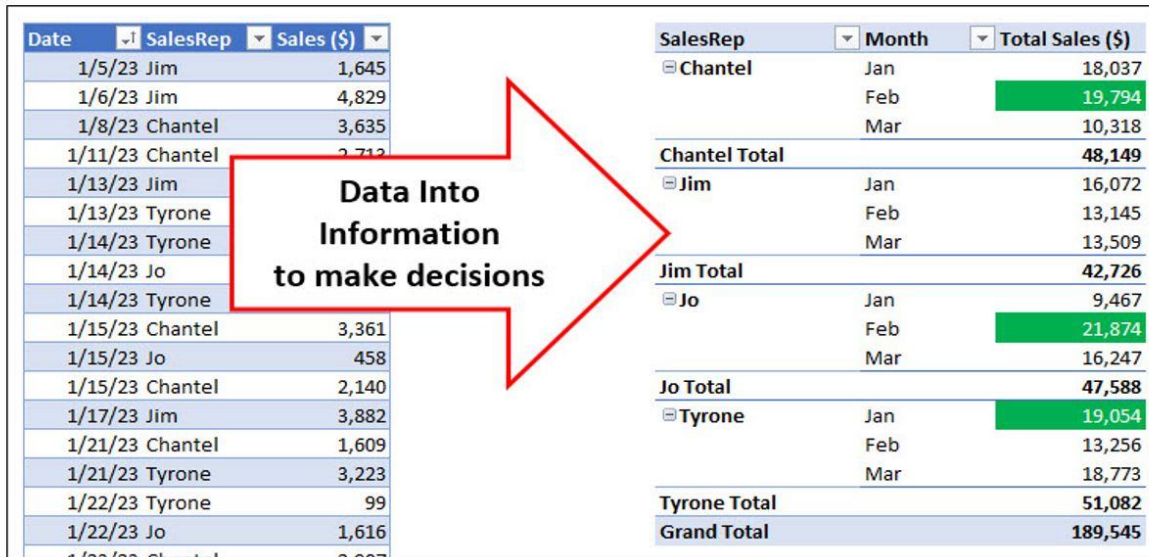**Data Analysis** = Convert raw data into useful information to gain insight and make decisions.

Synonyms: Data Analytics, Analytics, Business Intelligence, Data Analysis

Goal: help make data-driven decisions, which tend to be more accurate & help to achieve goals more consistently

Examples:

1) Sales data in a table converted into a Monthly SalesRep Sales Report. What are top 3 sales amounts?
   Q: If a bonus is given for the 3 biggest monthly sales amounts, who gets bonus? A: Chantel, Jo, Tyrone.

Date	SalesRep	Sales ($)
1/5/23	Jim	1,645
1/6/23	Jim	4,829
1/8/23	Chantel	3,635
1/11/23	Chantel	2,713
1/13/23	Jim	
1/13/23	Tyrone	
1/14/23	Tyrone	
1/14/23	Jo	
1/14/23	Tyrone	
1/15/23	Chantel	3,361
1/15/23	Jo	458
1/15/23	Chantel	2,140
1/17/23	Jim	3,882
1/21/23	Chantel	1,609
1/21/23	Tyrone	3,223
1/22/23	Tyrone	99
1/22/23	Jo	1,616

Data Into Information to make decisions

SalesRep	Month	Total Sales ($)
Chantel	Jan	18,037
	Feb	19,794
	Mar	10,318
Chantel Total		**48,149**
Jim	Jan	16,072
	Feb	13,145
	Mar	13,509
Jim Total		**42,726**
Jo	Jan	9,467
	Feb	21,874
	Mar	16,247
Jo Total		**47,588**
Tyrone	Jan	19,054
	Feb	13,256
	Mar	18,773
Tyrone Total		**51,082**
Grand Total		**189,545**

2) Sales data in a Table converted into a Monthly Sales Trends Chart. What is the trend?
   Q: What were sales trends over last six months? A: Mostly up.

Date	SalesRep	Sales ($)
5/10/23	Chantel	1,605
4/29/23	Tyrone	3,677
3/29/23	Jim	
6/20/23	Jo	
5/3/23	Cha	
4/8/23	Cha	
1/14/23	Tyr	
1/21/23	Chantel	
1/13/23	Jim	679
1/6/23	Jim	4,829
4/15/23	Chantel	2,981
6/5/23	Tyrone	6,269
6/18/23	Tyrone	6,948
6/1/23	Tyrone	7,224

Data Into Information to gain insight

Monthly Sales ($) Trends

Jan	Feb	Mar	Apr	May	Jun
62,630	68,069	58,847	112,048	111,955	186,718

Define Data and Table

Raw Data = data stored in its smallest form in a cell

Not Raw Data:

Date, Person, Sales
01/05/2023, Jim, $1,645.01

Raw Data:

Date	Person	Sales
1/5/2023	Jim	$1,645.01

Proper Data Set = Data Set = Table

Table is made up of:

Field = column in table

Field Name= name at top of field that describes what data goes into field

Record = one row in table

Table requirements in Excel:

1) Field names in first row

2) Records of related data in subsequent rows

3) Empty cells or Excel Row/Column Headers, all the way around table

Date	Person	Sales		
1/5/2023	Jim	1,645	←	Fields
1/6/2023	Jim	4,829	←	Records
1/8/2023	Chantel	3,635	←	
1/11/2023	Chantel	2,713	.	
1/13/2023	Jim	679	.	
1/13/2023	Tyrone	1,527	.	
1/14/2023	Tyrone	3,152		
1/14/2023	Jo	3,554		
1/14/2023	Tyrone	3,410		

Survey Data		
Yes	←	Fields
Yes	←	Records
No	←	
Yes	.	
No	.	
No	.	
No		
Somewhat		
Yes		

Some of the Data Analysis Tools in Excel

Excel Table feature = Convert proper data sets to a Table Object that can expand and contract and auto fill formulas.
Sort = Sort data A-Z (Ascending) or Z-A (Descending).
Filter = Show or extract records based on conditions, criteria and logical tests.
PivotTable = Drag and drop summary report tool that makes calculations based on conditions, criteria and logical
Excel Charts = Visualize data with charts and graphs.
Conditional Formatting = Visualize data based on conditions, criteria and logical tests
Worksheet formulas = To create helper fields (like with XLOOKUP function) or make conditional calculations (like with SUMIFS or COUNTIFS functions) or create full reports (like with LET, SUMIFS, HSTACK or VSTACK functions).
Flash Fill = One-time data cleaning tool
Power Query = Import data into worksheet, PivotTable or Data Model. Clean & transform data, tables and other related data objects.
Relationships = Creates relationships between tables (Substitute for XLOOKUP). Method of creating related tables to make PivotTable reporting more efficient.
Power Pivot Data Model and Data Model PivotTable = Used when you need Excel worksheet reports and visualizations and you have large data (about 50,000 rows or more), calculations that are hard to do with a PivotTable, or you have related tables of data.
Power BI Desktop and Data Model Visualizations = Used when you need Power BI reports and visualizations (more varied than in Excel) or you want to publish and share data analysis results online.
Power BI Online = share data analysis results online.

Excel Table Feature

- If you have a proper data set, you can convert your proper data set to an Excel Table by selecting one cell in the proper data set, click on the Table button in Table group in Insert Ribbon Tab, or just use keyboard: Ctrl + T.
    - The advantage to using the Excel Table feature is that when new rows or columns are added to the Excel Table, all objects, such as PivotTable, Charts, Formulas, or other features can be refreshed, and the new data will be incorporated into the object.
    - You can name your Excel Table: select one cell in Excel Table, click the Table Design Ribbon Tab, Properties group. You can NOT use spaces in the Table Name.
- Add new records to Excel Table by typing or pasting new data in the first row below the Excel Table.
- If new rows and columns are not added to Excel Table, you will need to change option settings for Excel Tables: File menu, Options button, Proofing tab on left, "AutoCorrect Options" button, check "Include new rows and columns in table".

PivotTable feature

- What PivotTables do:
  - Create Summary Reports that contain calculations with Conditions or Criteria.

- Summary of how to create PivotTable:
  - Click in one cell in Proper Data Set
  - Insert Ribbon Tab, Tables group, PivotTable button.
    - Keyboard = Alt, N, V, T.
  - From Field List, drag field name to Rows area or Columns area or Filter area. These are the conditions/criteria for the calculation in the Values area of the PivotTable.
    - Fields in the Rows area add a condition/criterion to the row
    - Fields in the Columns area add a condition/criterion to the column
    - Fields in the Filter area add a condition/criterion to the entire PivotTable.
  - From Field List drag the field you would like to make a calculation on to values area.
    - Number fields default to a SUM calculation (add numbers)
    - Text fields default to a COUNTA calculation (count non-empty cells)
  - With a cell selected in the PivotTable, click on PivotTable Tools Design Ribbon Tab, go to the Layout group, click drop- down for Report Layout and then click on "Show in Tabular Form" or "Show in Outline Form". These Layouts shows the Field Names in your Report. To set the layout default for all PivotTable, click on File menu, Options button, in the Excel Options dialog box, click the Data tab on the left, click the Edit Default Layout button, then in Report Layout check box, select "Show in Tabular Form" or "Show in Outline Form".
  - To add Number Formatting to the Values area of the PivotTable, click in one cell in the Values area of the PivotTable, Right-click the cell and click on "Number Format…", then in the Number Formatting dialog box select the Number Formatting that you would like and then click OK.
  - Slicers can be added to the PivotTable to add a condition/criterion to the entire PivotTable, similar to a field in the Filter area. To insert a Slicer into a PivotTable, click in one cell in the PivotTable Report, then go to the PivotTable Tools Analyze Ribbon Tab, then in the Filter Group, click the Insert Slicer button
  - If you want to change the summary calculation in the Values area, right-click the Values area in the PivotTable Report, point to one of these two options:
    - "Summarize Values By" and then select an aggregate calculation such as "Average", or "Max", or "More Options".
    - "Show Values As" and then select a calculation such as "% of Column Total", "Running Total", "Difference From" or other calculations.

- o Standard PivotTable Cached Data
  - When you create a Standard PivotTable, Excel creates a copy of the source data and stores it in the Pivot Cache.
  - The Pivot Cache is stored in Excel's memory.
  - This is why the PivotTable does not update when source data changes.
  - If source data changes, you can right-click the PivotTable and click Refresh, or use the Refresh button in the Data Ribbon Tab.
  - You can't see Cached PivotTable Data, but that's the data the PivotTable references when you build your PivotTable or change a Slicer selection or move rows/columns around.
  - When we group dates in a PivotTable, the Pivot Cache is where this action takes place. The Pivot Cache save the action of grouping by dates so that when we use the date column in other places, the date column will remain grouped.
- o When we use a Slicer or change the conditions and criteria in the PivotTable, these actions are interacting with the Pivot Cache of data, not the original source data.

Excel Charts

- What do Charts do?
- o Visually portray quantitative data (number data).
- o Give a quick impression of the number data.
- o Create a picture that can communicate more quickly than just the numbers alone.
- o Charts allow you to see patterns or trends that you may not be able to see if you are looking at just the number data.
- o Allows you to make relative comparisons more quickly than if you are using a table

- Types of Charts
- o Column: Use to compare differences across categories. Height of column conveys number.
- o Bar: Use to compare differences across categories. Length of bar conveys number.
- o Stacked Column/Bar: Good for displaying crosstabulation, emphasis on horizontal axis categories.
- o Clustered Column/Bar: Good for displaying crosstabulation, emphasis on legend categories.
- o Line: Use to show trend for a number variable over a category such as time.
- o X-Y Scatter: Used to show relationship between two number variables (x and y variables).

- Rules for Charts:
  o No Chart Junk.
  o Eliminate all chart elements that do not help to communicate the message.

- Formatting Charts:
  o You can add and remove chart elements by using the **Green +** on the right side of chart.
  o To format a chart element, select element and use the keyboard Ctrl + 1 to open the Format Chart Element task pane.

- Link Labels to Cells
  o Click on Chart Title, type an equal sign, click on cell with title, hit Enter.

Conditional Formatting

- Built-in Conditional Formatting
  o Select cells and apply a conditional format from the Styles group in the Home Ribbon
  o Each cell is evaluated to TRUE or FALSE. TRUE = Formatting applied. FALSE = Formatting not applied.
  o To edit rule, go to Manage Rules
- When you apply Conditional Formatting to a PivotTable, you can use the smart tag that pops up after you apply the conditional formatting to change where the formatting is applied:
  o **Selected cells:** only selected cells.
  o **All cells showing SUM Sales($):** this will apply the formatting to all cells including the totals and grand totals (not usually what you want).
  o **All cells showing SUM Sales($) for Month:** this will apply the formatting to all cells except the totals and grand totals (usually what you want).

Data Analysis Example in Video:

Goal 1: Create monthly sales rep report

Goal 2: Visualize trends in month sales

Chart Title: Company Monthly Sales ($) Trends

Date	SalesRep	Sales ($)		SalesRep	Months	Sum of Sales ($)		Months	Sum of Sales ($)
01/05/2023	Jim	1645.01		⊟Chantel	Jan	18,036.68		Jan	62,629.60
01/06/2023	Jim	4828.54			Feb	19,794.25		Feb	68,068.69
01/08/2023	Chantel	3634.56			Mar	10,317.84		Mar	58,846.96
01/11/2023	Chantel	2713.46			Apr	39,529.88		Apr	112,047.83
01/13/2023	Jim	679.07			May	20,023.19		May	111,955.33
01/13/2023	Tyrone	1526.73			Jun	42,592.27		Jun	186,718.15
01/14/2023	Tyrone	3151.76		Chantel Total		150,294.11		Grand Total	600,266.56
01/14/2023	Jo	3554.02		⊟Jim	Jan	16,071.94			
01/14/2023	Tyrone	3409.51			Feb	13,144.82			
01/15/2023	Chantel	3361.29			Mar	13,509.22			
01/15/2023	Jo	458.09			Apr	22,758.20			
01/15/2023	Chantel	2140.04			May	24,567.18			
01/17/2023	Jim	3882.28			Jun	51,211.06			
01/21/2023	Chantel	1608.7		Jim Total		141,262.42			
01/21/2023	Tyrone	3223.12		⊟Jo	Jan	9,467.30			
01/22/2023	Tyrone	98.57			Feb	21,873.70			
01/22/2023	Jo	1616.32			Mar	16,247.12			
01/23/2023	Chantel	2907.42			Apr	25,026.29			
01/25/2023	Tyrone	2384.97			May	38,757.67			
01/25/2023	Jim	4326.08			Jun	33,445.38			
01/25/2023	Jim	554.62		Jo Total		144,817.46			
01/26/2023	Tyrone	1904.62		⊟Tyrone	Jan	19,053.68			
01/26/2023	Jim	156.34			Feb	13,255.92			
01/27/2023	Chantel	883.34			Mar	18,772.78			
01/29/2023	Jo	3838.87			Apr	24,733.46			
01/30/2023	Chantel	787.87			May	28,607.29			
01/30/2023	Tyrone	3354.4			Jun	59,469.44			
02/01/2023	Jo	4467.02		Tyrone Total		163,892.57			
02/01/2023	Jim	4434.06		Grand Total		600,266.56			
02/02/2023	Jim	1970.14							

SalesRep

Chantel

Jim

Jo

Tyrone

Company Monthly Sales ($) Trends