ACC575: Data Analytics for Accounting
LN2. Introduction to Excel - Part II

Jaeyoon Yu, Ph.D.
Central Michigan University

August 31, 2025

Table of Contents

Table of Contents

PivotTable

- A PivotTable is a table that summarizes data from a larger dataset.
- PivotTables are useful for summarizing and analyzing data.
- PivotTables are also useful for creating charts and graphs.

Ex. Sales order data [Lab3-1&3-2]

Description:

- The data includes 100+ sales order data.
- One order may include multiple products.

	A	B	C	D	E	F	G	H	I
1	Sales_Order_ID	Sales_Order_Date	Sales_Employee_ID	Customer_ID	Sales_Order_Quantity_Sold	product_description	Product_Sale_Price	customer_st	customer_city
2	20001	11/1/2024	1006	2056	10	Pale Ale	95	KS	Wichita
3	20005	11/1/2024	1006	2012	8	Pale Ale	95	TN	Memphis
4	20017	11/7/2024	1006	2043	1	Pale Ale	95	TX	Houston

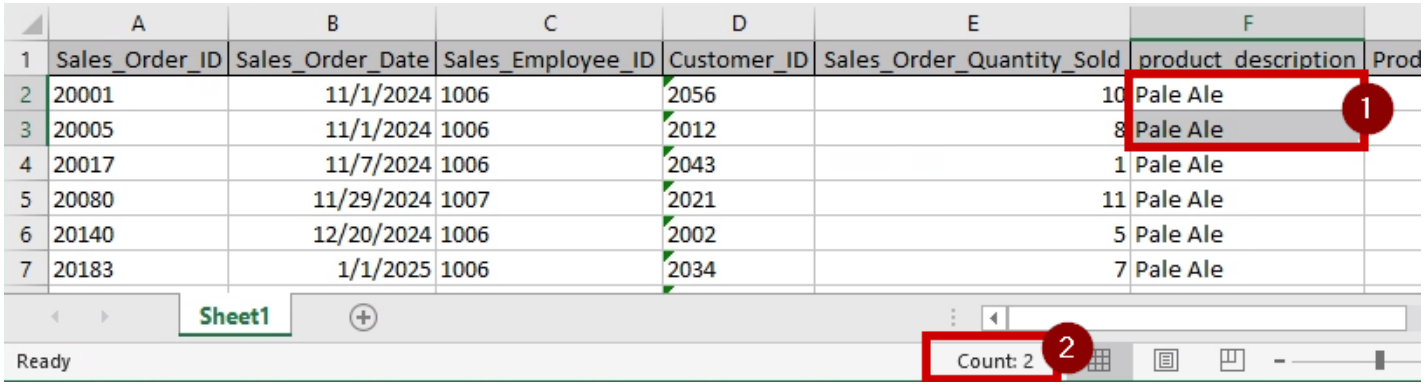
Field Name	**Description**
Sales_Order_ID	Unique identifier for each individual Sales Order
Sales_Order_Date	Date each Sales Order was placed
Sales_Employee_ID	Unique identifier for the employee who was responsible for taking the Sales Order
Customer_ID	Unique identifier for the Customer who placed the Sales Order
Sales_Order_Quantity_Sold	Quantity of each product sold on the transaction
Product_Description	Description of the product sold
Product_Sale_Price	Price of each product sold on the transaction
Customer_St	State in which the customer lives
Customer_City	City in which the customer lives

Required:

1. Inspect the data.
2. Calculate Sum of **Sales_Order_Quantity_Sold**, Count of **Sales_Order_Date**, Count of **Product_Description**.
3. Which product was sold the most? Make a pivot table.
4. Which product has the highest average of quantity sold? Make a pivot table.
5. How many times was each product sold? Make a pivot table.
6. Visualize the pivot table.
7. Calculate sum of Quantity Sold by State. Make a pivot table and chart.

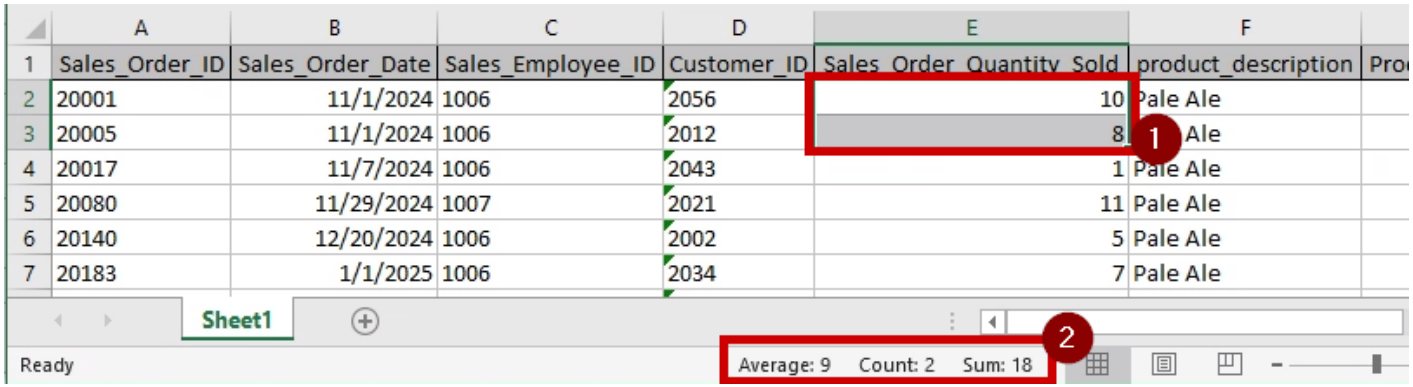
1. Basic inspections:
   1. Navigate the excel file.
   2. **ctrl + arrow keys** to move between cells.
   3. how many rows and columns are there?
   4. Any missing values or a pattern of missing values?
   5. Fix any data type issues, if any.

Select a couple of cells in each column and check the bottom right of the excel sheet.



	A	B	C	D	E	F
1	Sales_Order_ID	Sales_Order_Date	Sales_Employee_ID	Customer_ID	Sales_Order_Quantity_Sold	product_description
2	20001	11/1/2024	1006	2056	10	Pale Ale
3	20005	11/1/2024	1006	2012	8	Pale Ale
4	20017	11/7/2024	1006	2043	1	Pale Ale
5	20080	11/29/2024	1007	2021	11	Pale Ale
6	20140	12/20/2024	1006	2002	5	Pale Ale
7	20183	1/1/2025	1006	2034	7	Pale Ale

Figure: Text



	A	B	C	D	E	F
1	Sales_Order_ID	Sales_Order_Date	Sales_Employee_ID	Customer_ID	Sales_Order_Quantity_Sold	product_description
2	20001	11/1/2024	1006	2056	10	Pale Ale
3	20005	11/1/2024	1006	2012	8	Pale Ale
4	20017	11/7/2024	1006	2043	1	Pale Ale
5	20080	11/29/2024	1007	2021	11	Pale Ale
6	20140	12/20/2024	1006	2002	5	Pale Ale
7	20183	1/1/2025	1006	2034	7	Pale Ale

Figure: Numeric

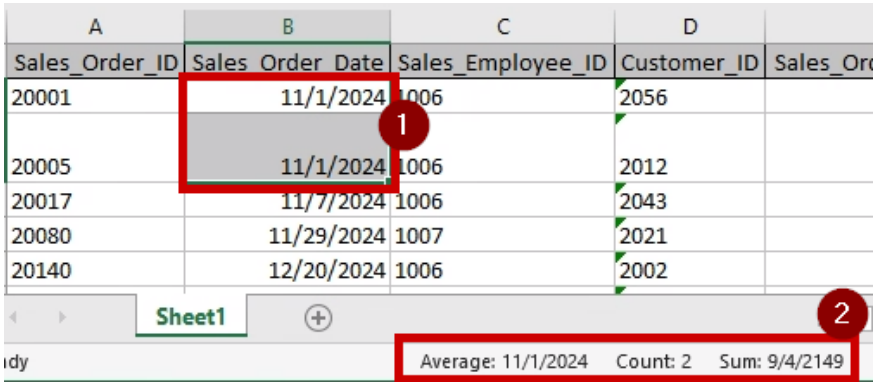


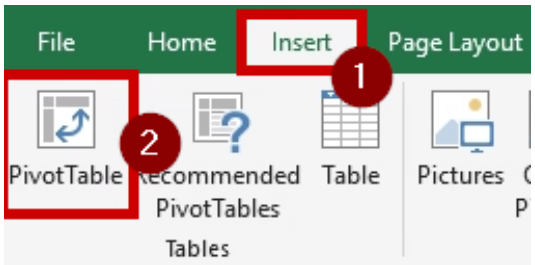
Figure: Date

Notes on date types:

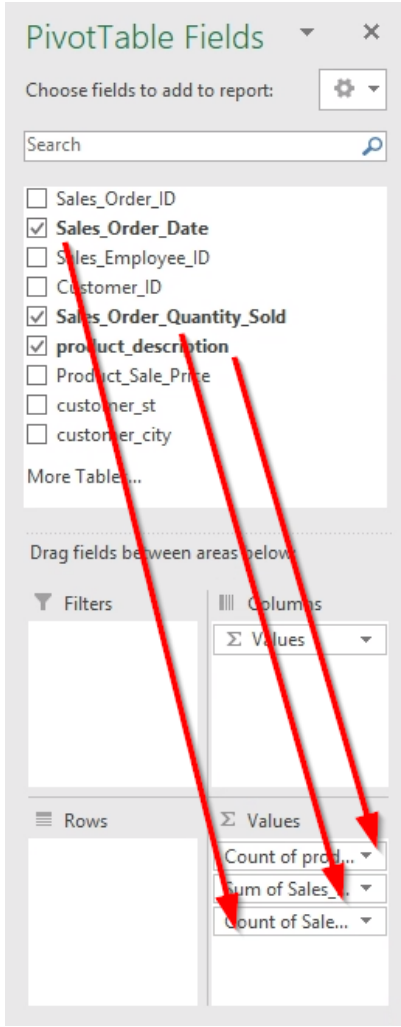
- Excel automatically recognizes most date formats.
- If Excel doesn't recognize a date, it will be treated as text.
- Excel associates a date with a number.
  1. 1900-01-01 is associated with 1.
  2. 1900-01-02 is associated with 2.
  3. 1900-01-03 is associated with 3.
- Due to this, Excel can calculate the number of days between two dates.

2. Calculate Sum of **Sales_Order_Quantity_Sold**, Count of **Sales_Order_Date**, Count of **Product_Description**.

- Click the **Insert** tab.
- Click the **PivotTable** button.
- Select the **range of cells** you want to pivot.
- Click **OK**.
- Drag the selected cells to **Values field**.



How to insert a PivotTable

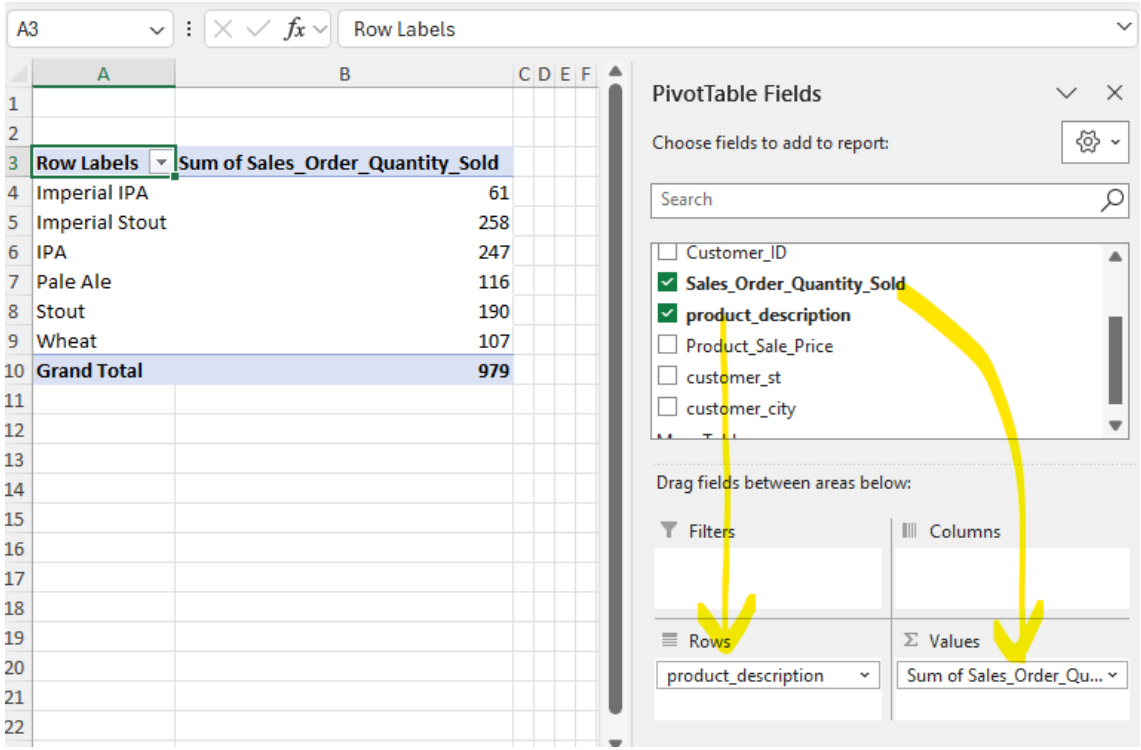


PivotTable Fields

Sum of Sales_Order_Quantity_Sold	Count of Sales_Order_Date	Count of product_description
979	149	149

Does the output have any insight?

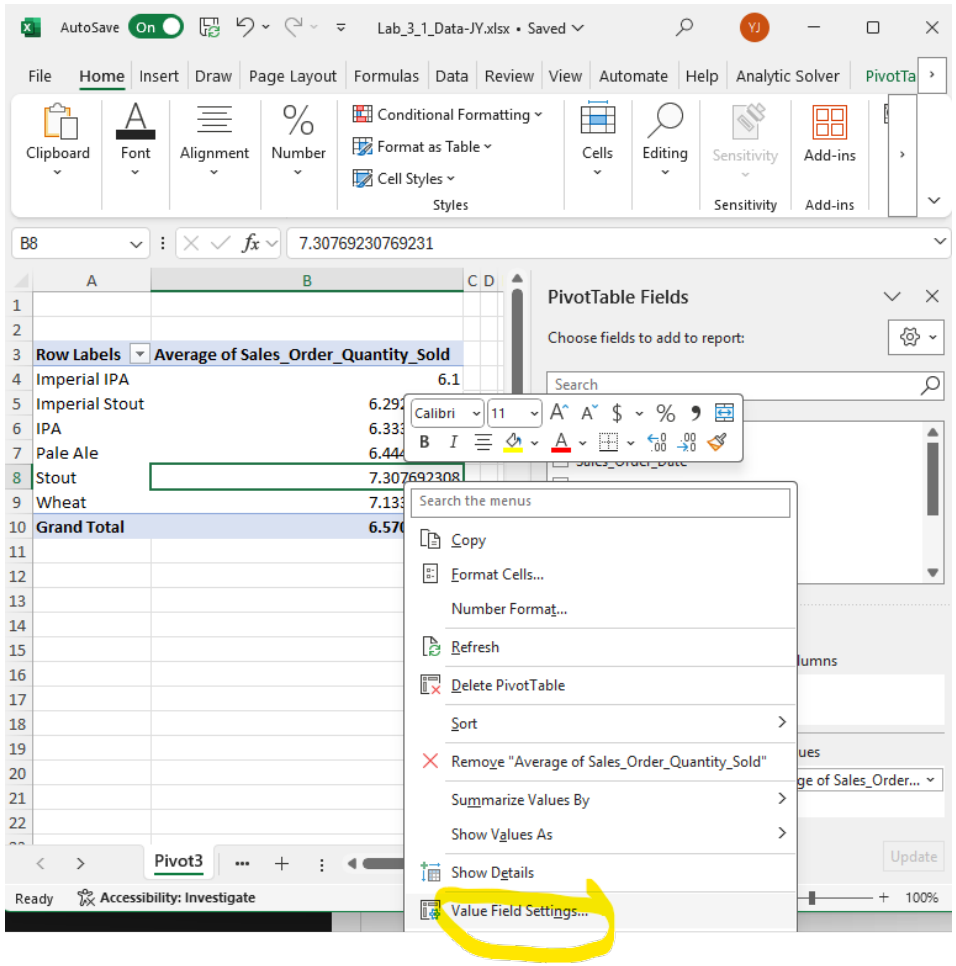
3. Which product was sold the most? Make a pivot table.
   1. Drag the **Product_Description** field to **Rows field**.
   2. Drag the **Sales_Order_Quantity_Sold** field to **Values field**.



Does the output have any insight?

4. Which product has the highest average of quantity sold? Make a pivot table.

1. Copy the previous pivot table to a new sheet.
2. Right click on the **Sales_Order_Quantity_Sold** and choose **Value Field Settings**.
3. Choose **Average** and click **OK**.

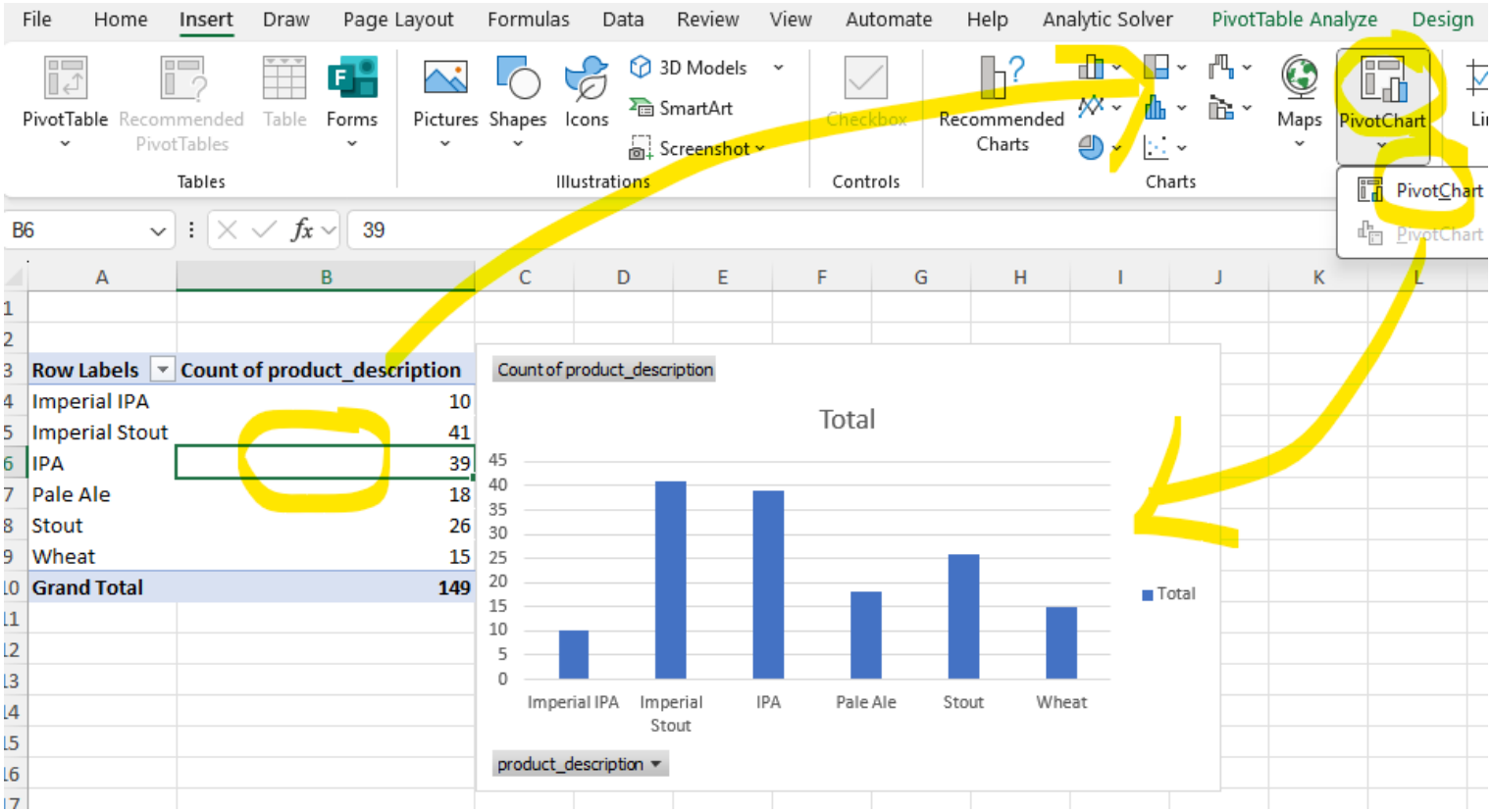


5. How many times was each product sold? Make a pivot table.
   1. Make a new pivot table from the data.
   2. Drag the **Product_Description** field to **Rows field**.
   3. Drag the **Product_Description** field to **Values field**.

Row Labels	Count of product_description
Imperial IPA	10
Imperial Stout	41
IPA	39
Pale Ale	18
Stout	26
Wheat	15
Grand Total	**149**

6. Visualize the pivot table.

1. Have your active cell in the pivot table.
2. Insert > PivotChart.



7. Calculate sum of Quantity Sold by State. Make a pivot table and chart.

Row Labels	Sum of Sales_Order_Quantity_Sold
AR	144
IA	83
KS	98
LA	35
MO	135
NE	96
OK	134
TN	43
TX	211
Grand Total	**979**

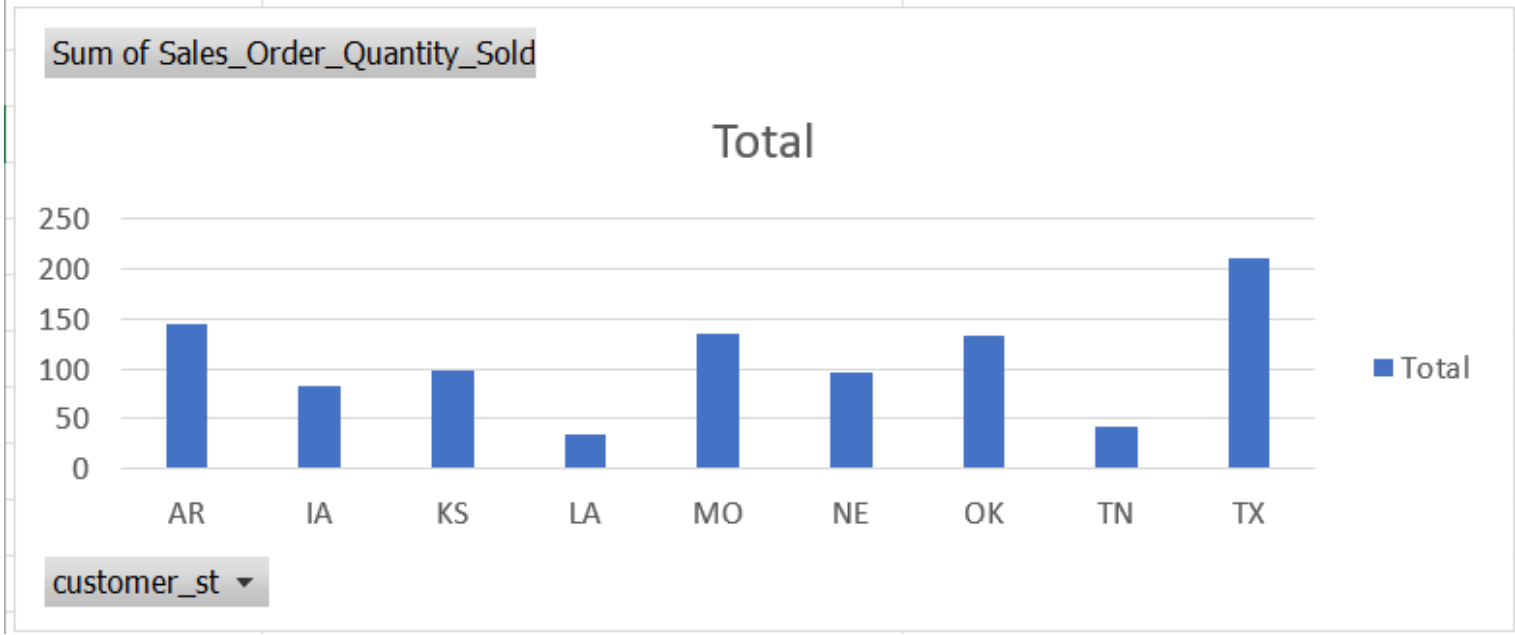


Table of Contents

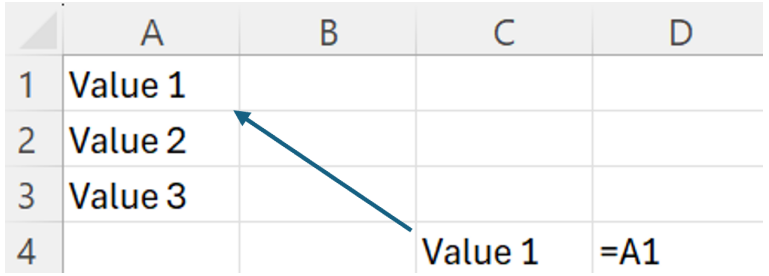
Relative vs. Absolute References

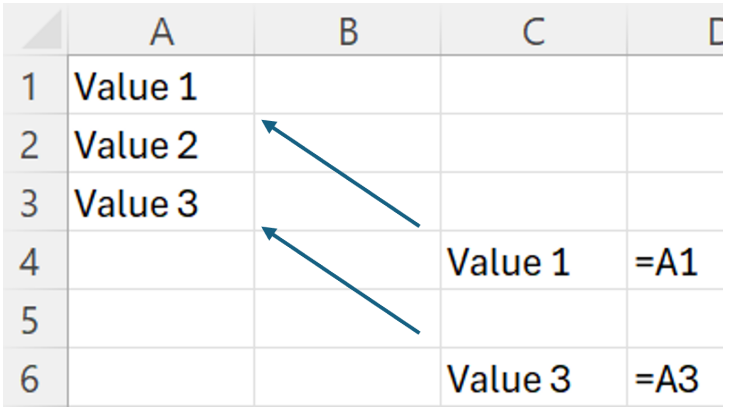
- Relative references change when copied to another cell.
- Absolute references do not change when copied to another cell.

Relative vs. Absolute References

Relative Reference

- Cell C4 (relatively) refers to the cell A1.
- Copy the cell C4 to the cell C6.
- What is the value of the cell C6?

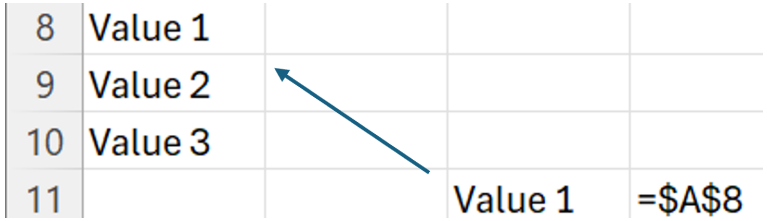


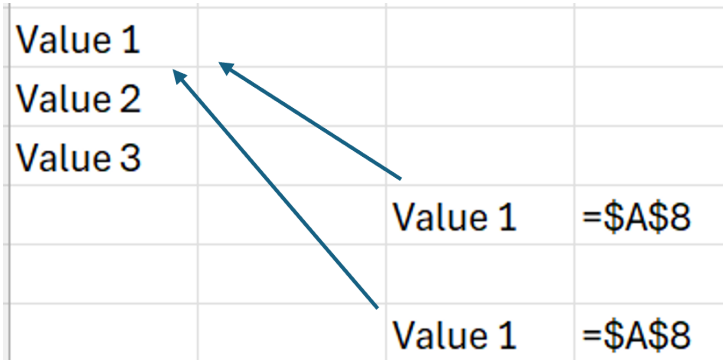


Relative vs. Absolute References

Absolute Reference

- Cell C4 (absolutely) refers to the cell A1.
- Copy the cell C4 to the cell C6.
- What is the value of the cell C6?





Relative vs. Absolute References

Differences between the A1, A1, $A1, and A$1:

- **A1**: Relative reference.
- **A1**: Absolute reference.
- **$A1**: Absolute column, relative row.
- **A$1**: Relative column, absolute row.

Table of Contents

Functions

- Functions are predefined formulas that perform specific calculations.
- Functions are useful for performing complex calculations.
- Functions are also useful for performing calculations on a range of cells.

Benford's Law

Benford's Law

- The first digit of a number is not uniformly distributed.
- The first digit of a number is more likely to be 1 than 2.
- The first digit of a number is more likely to be 2 than 3.
- The first digit of a number is more likely to be 3 than 4.
- Some, but not all, follow Benford's Law.

Ex. Benford's Law [Lab7-4]

Required:

1. From the Sales data, get the 1st digit of each sales amount.
2. Count the frequency of each 1st digit and its percentage.
3. Calculate the expected percentage of each 1st digit under Benford's Law.
4. Compare 2. with 3 using a combo chart.

	A	B
1	**Order_Number**	**Sales**
2	00001	129.33
3	00002	234.85
4	00003	934.83
5	00004	485.48

1. Get the **1st digit** of each sales amount using left().

Order_Number	Sales	First_Digit
00001	129.33	1
00002	234.85	2
00003	934.83	9
00004	485.48	4
00005	384.58	3

2. Count the frequency of each 1st digit in a new sheet using countif().

1. Insert a new sheet.
2. Type 1-9 in the first column.
3. Use countif() to count the frequency of each 1st digit.
4. Use sum() to calculate the total frequency.

	A	B
1	First_Digit	Count
2	1	46.00
3	2	42.00
4	3	27.00
5	4	20.00
6	5	16.00
7	6	21.00
8	7	10.00
9	8	10.00
10	9	8.00
11		200.00

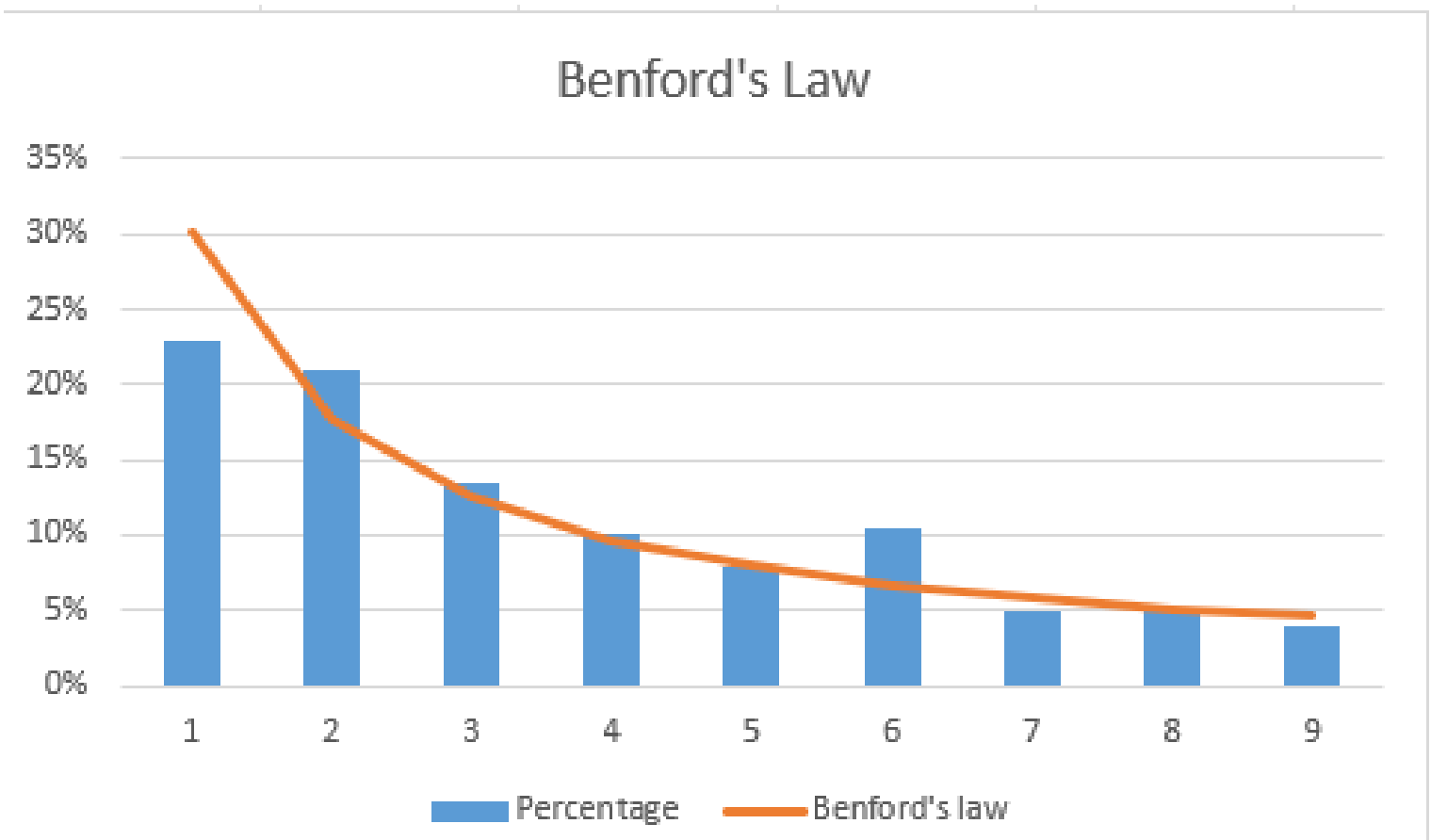
3. Calculate the expected frequency of each 1st digit under Benford's Law.

1. Percentage column: Count of each 1st digit / Total count.
2. Benford's law column: Use the formula.
3. Get the difference between them.

First_Digit	Count	Percentage	Benford's law	Difference
1	46.00	23%	30%	-7%
2	42.00	21%	18%	3%
3	27.00	14%	12%	1%
4	20.00	10%	10%	0%
5	16.00	8%	8%	0%
6	21.00	11%	7%	4%
7	10.00	5%	6%	-1%
8	10.00	5%	5%	0%
9	8.00	4%	5%	-1%

Q. Is there anything suspicious?

4. Compare 2. with 3. using combo chart.



Q. Does visualization help?

Financial Statement Analysis: Horizontal and Vertical

- **Horizontal analysis**: Compare the current year with the previous year.
- Example:
  - Revenues in 2024 is 100 mil. Revenues in 2023 is 90 mil.
  - Revenues increased by 10 mil (or 11.1%).

- **Vertical analysis**: Compare the current year with the total.
- Example:
  - Revenues is 100 mil. Cost of goods sold is 70 mil.
  - Out of 100, 70% is cost; 30% is profit.

Ex. Financial Statement Analysis [Lab6-2&6-3]

Descriptions

1. 2-year income statement of AAL.

Required:

1. Perform a horizontal analysis.
2. Perform a vertical analysis.

Q. Any quick insights?

	A	B	C
1	American Airlines Group Inc. (AAL)		
2	Consolidated Statements of Operations		
3	*Period End*	**2019**	**2020**
4		10-K	10-K
5	**Operating revenues:**	($ millions)	($ millions)
6	Passenger	$42,010	$14,518
7	Cargo	$863	$769
8	Other	$2,888	$2,048
9	Total operating revenues	$45,761	$17,335
10	**Operating expenses:**		
11	Aircraft fuel and related taxes	$7,526	$2,581
12	Salaries, wages and benefits	$12,600	$10,955
13	Regional expenses	$7,518	$4,676
14	Maintenance, materials and repairs	$2,380	$1,583
15	Other rent and landing fees	$2,055	$1,536
16	Aircraft rent	$1,326	$1,341
17	Selling expenses	$1,602	$513
18	Depreciation and amortization	$1,982	$2,040
19	Special items, net	$635	($657)
20	Other	$5,090	$2,991
21	Total operating expenses	$42,714	$27,559
22	Operating income	$3,047	($10,224)

Horizontal Analysis

Steps:

1. Type the title of column D.
2. Type in **D6: C6 - B6**.
3. Drag **D6** to the rest of the column.

	A	B	C	D
1	American Airlines Group Inc. (AAL)			
2	Consolidated Statements of Operations			
3	*Period End*	**2019**	**2020**	Change from
4		10-K	10-K	2019 to 2020
5	**Operating revenues:**	($ millions)	($ millions)	
6	Passenger	$42,010	$14,518	($27,492)
7	Cargo	$863	$769	
8	Other	$2,888	$2,048	
9	Total operating reven	$45,761	$17,335	

	A	B	C	D
1	American Airlines Group Inc. (AAL)			
2	Consolidated Statements of Operations			
3	*Period End*	**2019**	**2020**	Change from
4		10-K	10-K	2019 to 2020
5	**Operating revenues:**	($ millions)	($ millions)	
6	Passenger	$42,010	$14,518	($27,492)
7	Cargo	$863	$769	($94)
8	Other	$2,888	$2,048	($840)
9	Total operating reven	$45,761	$17,335	($28,426)

Steps:

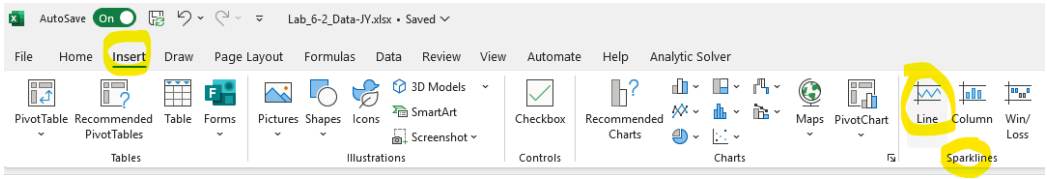
1. Copy the formatting of column C to column D.
    1. Select column C.
    2. Home > Clipboard > Format Painter.
    3. Select column D.
2. Make a new column (E) for the percentage change.
3. Type in **E6: D6 / B6**.
4. % formatting: Home > Number > Percentage.

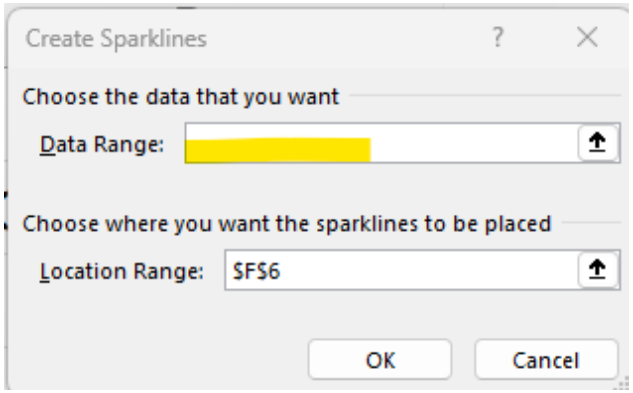
	A	B	C	D
1	American Airlines Group Inc. (AAL)			
2	Consolidated Statements of Operations			
3	*Period End*	**2019**	**2020**	**Change from**
4		10-K	10-K	2019 to 2020
5	**Operating revenues:**	($ millions)	($ millions)	
6	Passenger	$42,010	$14,518	($27,492)
7	Cargo	$863	$769	($94)
8	Other	$2,888	$2,048	($840)
9	Total operating reven	$45,761	$17,335	($28,426)

	A	B	C	D	E
1	American Airlines Group Inc. (AAL)				
2	Consolidated Statements of Operations				
3	*Period End*	**2019**	**2020**	**Change from**	% Change
4		10-K	10-K	2019 to 2020	
5	**Operating revenues:**	($ millions)	($ millions)		
6	Passenger	$42,010	$14,518	($27,492)	-65.4%
7	Cargo	$863	$769	($94)	-10.9%
8	Other	$2,888	$2,048	($840)	-29.1%
9	Total operating reven	$45,761	$17,335	($28,426)	-62.1%

Steps:

1. Sparkline chart:
    1. Move to cell **E6**.
    2. Insert > Sparklines > Line.
    3. Select **B6:C6** for Data Range in Create Sparkline dialog.
    4. Drag the sparkline in cell **E6** to the rest of the column.





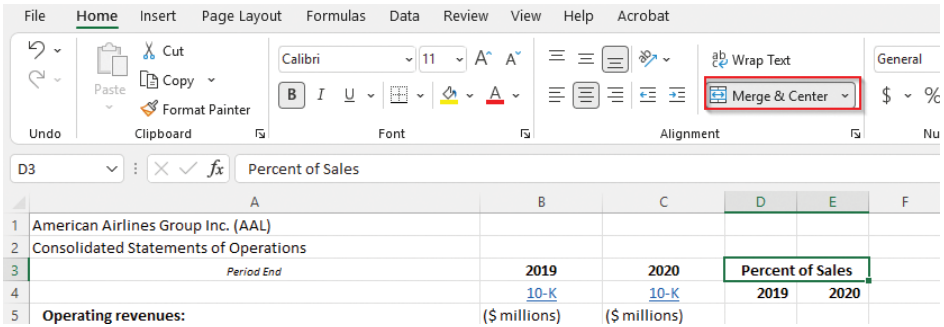
	A	B	C	D	E	F
1	American Airlines Group Inc. (AAL)					
2	Consolidated Statements of Operations					
3	*Period End*	**2019**	**2020**	**Change from**	**% Change**	
4		10-K	10-K	2019 to 2020		
5	**Operating revenues:**	($ millions)	($ millions)			
6	Passenger	$42,010	$14,518	($27,492)	-65.4%	
7	Cargo	$863	$769	($94)	-10.9%	
8	Other	$2,888	$2,048	($840)	-29.1%	
9	Total operating reven	$45,761	$17,335	($28,426)	-62.1%	

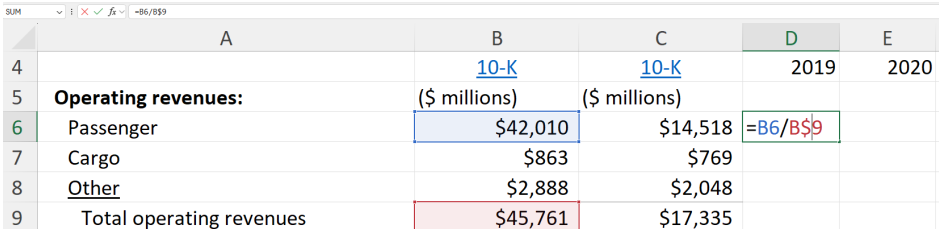
Vertical Analysis

Letś use **absolute** and **relative** references properly to convert numbers in **columns B and C** to percentages (rlative to total revenues) in **columns D and E**.

Steps:

1. Merge **D3** and **E3**: Home > Alignment > Merge & Center.
2. Type headers in **D3**, **D4**, and **E4**.
3. In **D6**: Type in **B6 / B$9**. (We fix row 9 but not column B in **B$9**.)
4. Drag **D6** to the rest of the column.
5. Also copy the column **D** to **E**. (In columns **D** and **E**, denominator is always the total revenue (in row 9).)





Steps:

1. Use percent format for columns D and E.
2. Add sparklines in column F.

	A	B	C	D	E	F
4		10-K	10-K	2019	2020	
5	**Operating revenues:**	($ millions)	($ millions)			
6	Passenger	$42,010	$14,518	91.8%	83.7%	
7	Cargo	$863	$769	1.9%	4.4%	
8	Other	$2,888	$2,048	6.3%	11.8%	
9	Total operating revenues	$45,761	$17,335	100.0%	100.0%	
10	**Operating expenses:**					
11	Aircraft fuel and related taxes	$7,526	$2,581	16.4%	14.9%	
12	Salaries, wages and benefits	$12,600	$10,955	27.5%	63.2%	
13	Regional expenses	$7,518	$4,676	16.4%	27.0%	
14	Maintenance, materials and repairs	$2,380	$1,583	5.2%	9.1%	
15	Other rent and landing fees	$2,055	$1,536	4.5%	8.9%	
16	Aircraft rent	$1,326	$1,341	2.9%	7.7%	
17	Selling expenses	$1,602	$513	3.5%	3.0%	
18	Depreciation and amortization	$1,982	$2,040	4.3%	11.8%	
19	Special items, net	$635	($657)	1.4%	-3.8%	
20	Other	$5,090	$2,991	11.1%	17.3%	
21	Total operating expenses	$42,714	$27,559	93.3%	159.0%	
22	Operating income	$3,047	($10,224)	6.7%	-59.0%	
23	**Nonoperating income (expense):**					
24	Interest income	$515	$337	1.1%	1.9%	
25	Interest expense, net	($1,109)	($1,171)	-2.4%	-6.8%	
26	Other, net	$152	$155	0.3%	0.9%	
27	Total nonoperating expense, net	($442)	($679)	-1.0%	-3.9%	
28	Income before income taxes	$2,605	($10,903)	5.7%	-62.9%	
29	Income tax provision (benefit)	$633	($2,453)	1.4%	-14.2%	

Table of Contents

Formatting

- Formatting is the process of changing the appearance of data.
- Formatting is useful for making data more readable and visually appealing.
- Formatting is also useful for making data more informative.

Back to Sales order data [Lab3-1&3-2]

- Highlight the product with the highest quantity sold in green.
- Highlight the product with the lowest quantity sold in red.

Back to Benford's Law [Lab7-4]

- Highlight with red if the difference is greater than 0.05.
- Highlight with green if the difference is less than 0.05.
- Replace 0.05 with 0.03