📊 Day 11 — Excel for Data Analysis

📌 Objective

The objective of this exercise is to analyze sales data using Microsoft Excel / Google Sheets by applying data cleaning techniques, essential formulas, PivotTables, and basic visualizations to extract meaningful business insights.

📂 Dataset Used

- **File name:** `cleaned_sales_data.csv`
- **Source:** Cleaned sales dataset generated during earlier data cleaning steps
- **Key Columns:**
    - OrderID
    - Region
    - Product
    - Category
    - Sales
    - Quantity

🧹 1. Data Cleaning in Excel

Before analysis, the dataset was reviewed and cleaned directly in Excel.

Steps Performed:

- Checked for **duplicate rows**

- Verified **no blank cells** in critical columns
- Ensured **Sales** column is formatted as numeric
- Ensured **Quantity** column contains integer values
- Verified column names are consistent and readable

Why Excel for Cleaning?

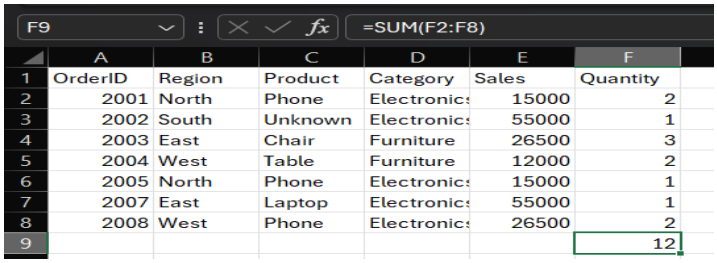
Excel allows fast visual inspection of data, making it ideal for quick validation and minor cleaning tasks before deeper analysis.

	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2

🧮 2. Excel Formulas Used

a) Total Sales

`=SUM(F2:F8)`



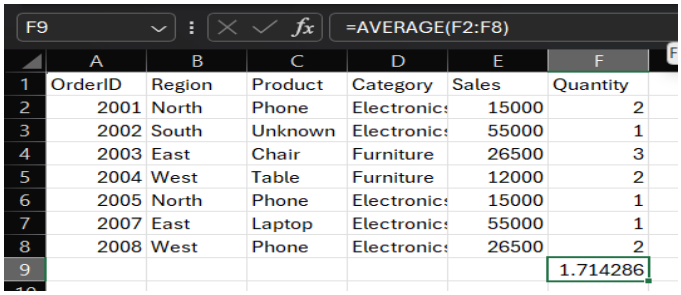
	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2
9						12

Calculates total revenue from all orders.

---

b) Average Sales

`=AVERAGE(F2:F8)`



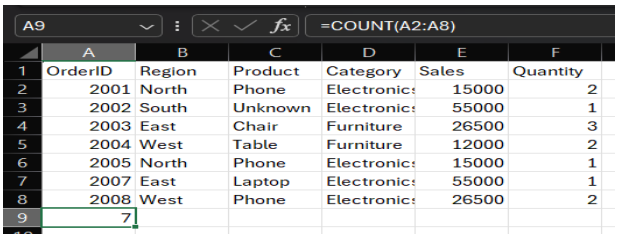
	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2
9						1.714286

Finds the average sales value per order.

c) Count of Orders

```
=COUNT(A2:A8)
```

Counts total number of orders.



	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2
9	7					

---

d) Conditional Count (Sales > 20000)

```
=COUNTIF(F2:F8, ">20000")
```

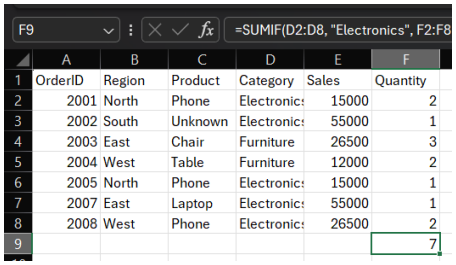
Counts number of high-value orders.

---

e) Conditional Sum (Electronics Sales)

```
=SUMIF(D2:D8, "Electronics", F2:F8)
```

Calculates total sales generated from Electronics category.



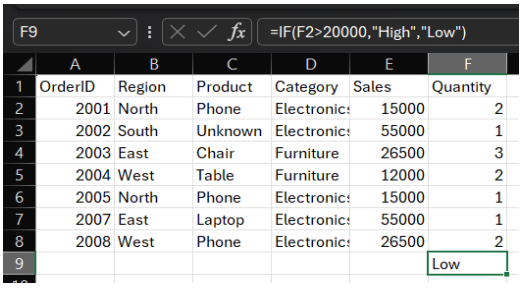
	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2
9						7

---

f) Logical Condition Example

```
=IF(F2>20000,"High","Low")
```

Classifies orders as High or Low value based on sales amount.



	A	B	C	D	E	F
1	OrderID	Region	Product	Category	Sales	Quantity
2	2001	North	Phone	Electronics	15000	2
3	2002	South	Unknown	Electronics	55000	1
4	2003	East	Chair	Furniture	26500	3
5	2004	West	Table	Furniture	12000	2
6	2005	North	Phone	Electronics	15000	1
7	2007	East	Laptop	Electronics	55000	1
8	2008	West	Phone	Electronics	26500	2
9						Low

📊 3. Pivot Table Analysis

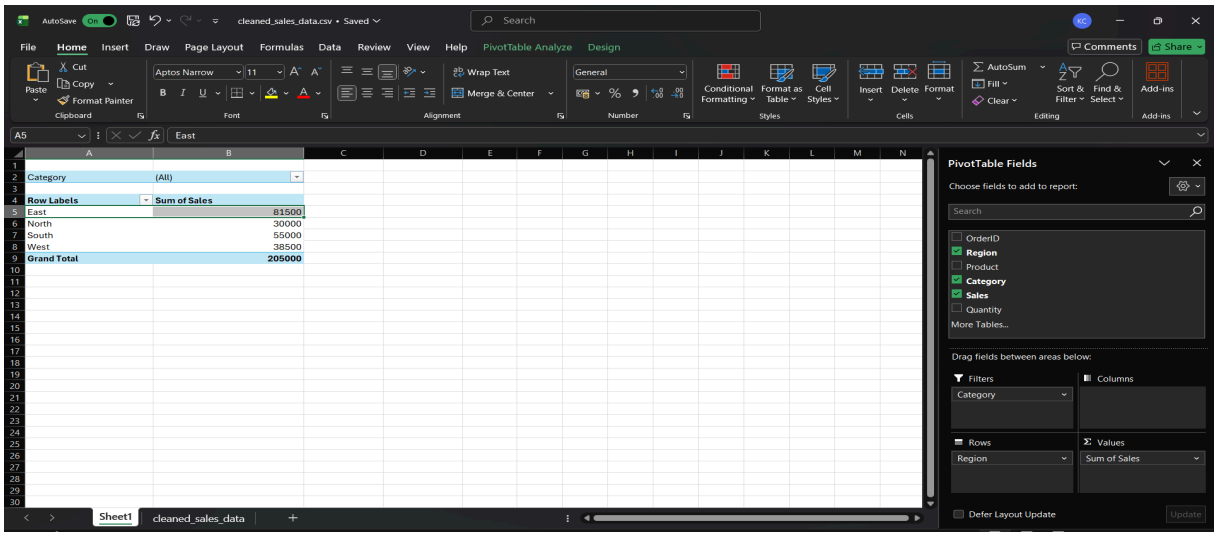
Pivot Table Configuration:

- **Rows:** Region
- **Values:** Sum of Sales
- **Filters:** Category

This PivotTable provides a quick comparison of regional sales performance and allows filtering by product category.

Insights from Pivot Table:

- Certain regions consistently outperform others
- Electronics category contributes higher revenue than Furniture
- PivotTables enable rapid business-level decision making



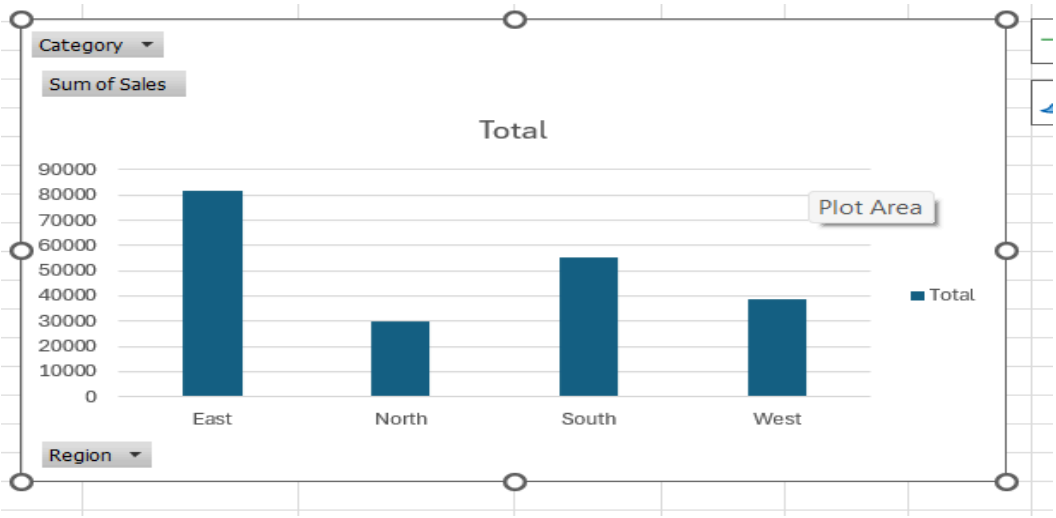
📈 4. Excel Charts

Charts Created:

1. **Bar Chart** — Total Sales by Region
2. **Column Chart** — Total Sales by Category

Best Practices Followed:

- Clear axis labels
- Minimal colors
- No unnecessary visual clutter



🧠 5. Business Insights

- The **Electronics category** generates the highest revenue overall.
- **Laptop sales** contribute significantly despite lower order frequency.
- Some regions consistently underperform and may need strategic attention.
- High sales value does not always correlate with high quantity sold.

🏁 Conclusion

Excel proves to be a powerful and efficient tool for quick data analysis, especially for business stakeholders. Using formulas, PivotTables, and charts, meaningful insights can be derived rapidly without complex programming.

This exercise demonstrates practical Excel skills that are highly relevant for data analyst internships and entry-level roles.