Excel for Data Analysis Cheat Sheet

⊙ Type	Data science masterclass

🟢 Excel for Data Analysis Cheat Sheet

1. Data Cleaning

- **Key Tools:**
    - `Remove Duplicates` : Clean duplicate records (`Data` → `Remove Duplicates`).
    - **Text to Columns**: Split data based on delimiters (e.g., commas, spaces).
    - **Find & Replace**: `Ctrl + H` to clean unwanted characters or errors.
    - **TRIM**: `=TRIM(cell)` removes unnecessary spaces.
    - **CLEAN**: Removes non-printable characters.
    - **Error Handling**: `=IFERROR(value, "alternative")` for clean error handling.
    - **Advanced Filters**: Filter data by criteria (`Data` → `Advanced Filter`).

2. Data Transformation

- **Text Functions:**
    - **Extract Data:**
        - `LEFT(text, num_chars)`
        - `RIGHT(text, num_chars)`
        - `MID(text, start, num_chars)`
    - **Text Formatting:**

- ■ `UPPER(text)`

  - ■ `LOWER(text)`

  - ■ `PROPER(text)`

- **Date Functions:**

  - `=NOW()` → Current date and time.

  - `=EDATE(start_date, months)` → Add/subtract months.

  - `=WORKDAY(start_date, days)` → Exclude weekends.

- **Transform Data with Power Query:**

  - **Connect Data Sources**: Import data from external files (CSV, databases, etc.).

  - **Combine Queries**: Merge or append datasets.

  - **Unpivot Columns**: Reshape data into a columnar structure for analysis.

3. Data Analysis Functions

- **Statistical Functions:**

  - Central Tendency: `=AVERAGE(range)` , `=MEDIAN(range)` , `=MODE(range)` .

  - Variance: `=VAR.S(range)` (sample), `=VAR.P(range)` (population).

  - Standard Deviation: `=STDEV.S(range)` , `=STDEV.P(range)` .

- **Logical Functions:**

  - `=IF(condition, value_if_true, value_if_false)`

  - `=AND(condition1, condition2)`

  - `=OR(condition1, condition2)`

  - Nested `IF` : Combine multiple conditions.

- **Lookup & Reference:**

  - Vertical/Horizontal Lookup: `=VLOOKUP` , `=HLOOKUP` .

  - Advanced Lookup: `=INDEX + MATCH` , `=XLOOKUP` .

4. Data Visualization

- **Charts and Graphs:**

    - Recommended Charts: `Insert` → `Recommended Charts`.

    - Popular Types:

        - Bar/Column Chart → Compare data categories.

        - Line Chart → Trends over time.

        - Scatter Plot → Relationship analysis.

        - Pie Chart → Distribution analysis.

- **Conditional Formatting:**

    - Data Bars, Color Scales, and Icon Sets for insights at a glance.

    - Example Rule: Highlight sales > 10,000 (`New Rule` → Greater Than).

- **Sparklines:** Add mini-charts in cells (`Insert` → `Sparklines`).

5. Advanced Analysis Tools

- **PivotTables:**

    - Summarize and analyze large datasets.

    - Add **Slicers** for easy filtering.

    - Use Calculated Fields for advanced metrics.

- **What-If Analysis:**

    - **Goal Seek:** Find input values to achieve a specific outcome.

    - **Scenario Manager:** Analyze different scenarios (e.g., Best/Worst Case).

    - **Data Tables:** Test multiple variables simultaneously.

- **Solver:**

    - Optimize results based on constraints (e.g., minimize cost).

- **Power Pivot:**

- Create complex data models and relationships between tables.
- DAX formulas for advanced calculations (e.g., `SUMX`, `CALCULATE`).

6. Keyboard Shortcuts

Action	Shortcut
Select Entire Column	`Ctrl + Space`
Select Entire Row	`Shift + Space`
Create Table	`Ctrl + T`
AutoSum	`Alt + =`
Insert New Sheet	`Shift + F11`
Open Filter Dropdown	`Alt + Down Arrow`
Freeze Panes	`Alt + W + F + F`

7. Pro Tips for Efficient Analysis

1. **Use Named Ranges**: Easier to reference ranges (Formulas → Define Name).
2. **Data Validation**: Create dropdowns for cleaner data entry (`Data` → `Data Validation`).
3. **Combine Functions**: Use nested formulas like `=IF(AND(A1>10, B1<5), "Yes", "No")`.
4. **Backup Data**: Always work on a copy to avoid overwriting raw data.
5. **Organize Sheets**: Color tabs, add comments, and use a clean layout.

Summary: Excel for Data Analysis

1. **Data Cleaning**: Remove duplicates, trim spaces, and handle errors.
2. **Data Transformation**: Use text, date, and lookup functions. Power Query is a game-changer.
3. **Descriptive Analysis**: Averages, medians, and variance functions are your go-to.

4. **Visualization**: Master charts, conditional formatting, and sparklines.

5. **Advanced Tools**: PivotTables, Solver, and What-If Analysis for deeper insights.

6. **Shortcuts**: Memorize common shortcuts to save time.

7. **Pro Tools**: Leverage Power Query and Power Pivot for advanced data handling.