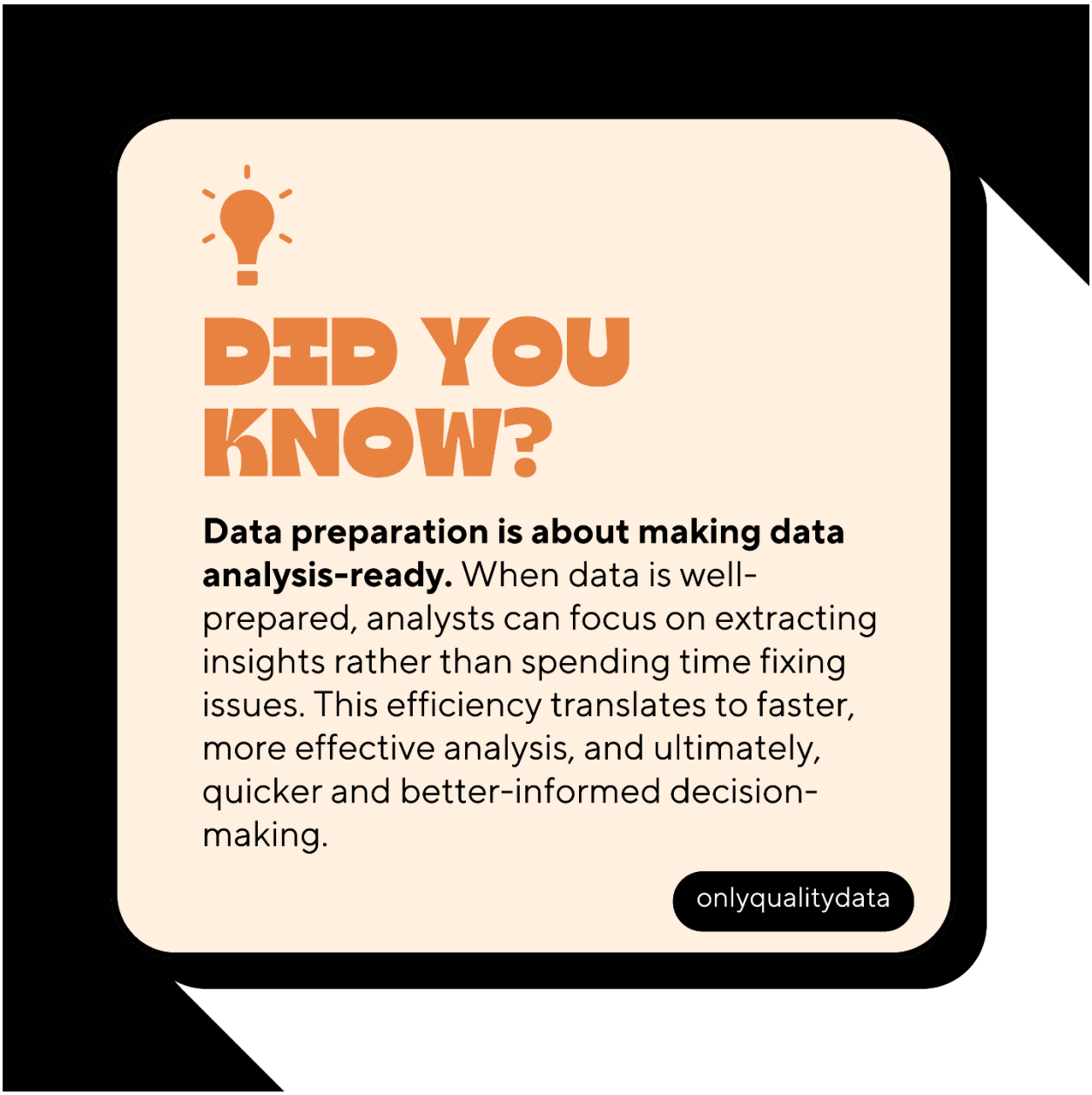
OUR AIM IS TO PROMOTE OUTSANDING DATA TRANSFORMATION AND ENGINEERING, THUS FACILITATING THE GENERATION OF VALUABLE INSIGHTS, AND OPENING A WORLD OF LIMITLESS POSSIBILITIES WITH DATA LEAD DECISION MAKING.

Authored by Enoch Nunekpeku
https://www.linkedin.com/in/enoch-nunekpeku/



Authored by Enoch Nunekpeku
https://www.linkedin.com/in/enoch-nunekpeku/

PREPARING AND TRANSFORMING ZENOMART SALES DATA IN EXCEL POWER QUERY.

Phase 1

1. **Download the ZenoMart Data.** This can be accessed from the Google Classroom LMP.
2. **Organize Your Workspace:**

   • Create a new folder on your computer in a convenient location and name it ZenoMart-Sales-Data-Cleaning-Project.

   • Move all the downloaded files into this new folder
3. **Review the ZenoMart Data:**
   - Open the Dataset: Begin by loading the dataset into Excel.
   - Scan the Dataset: Quickly glance through the entire dataset to get a sense of its structure, including the number of columns, rows, and any immediately noticeable patterns or inconsistencies.
   - Identify Key Columns: Identify columns that seem relevant for data cleaning, such as IDs, dates, numerical values, or categories.

4. Understand Data Types

- **Categorize Data Types:** Note the data types (e.g., text, numbers, dates) of each column. This will help in identifying potential type mismatches or errors in data.
- **Look for Mixed Data Types:** Identify any columns where data types are mixed (e.g., numbers and text in the same column), which could indicate data entry errors or inconsistencies.

Authored by Enoch Nunekpeku
https://www.linkedin.com/in/enoch-nunekpeku/

5. **Check for Missing Data**

- **Identify Missing Values:** Determine which columns have missing data. Look for cells with blanks, NULL, or other indicators of missing data (e.g., "N/A").
- **Assess Impact:** Consider how missing data in key columns might affect the analysis and plan for strategies to handle or impute these missing values during data cleaning.

6. Understand the Context

- **Familiarize with Domain Knowledge:** If possible, acquire some basic domain knowledge relevant to the dataset to help you understand what constitutes normal or abnormal data.
- **Consult Documentation:** If there is any documentation or metadata available for the dataset, review it to understand the purpose of each column and the expected data.

7. Document Observations

- **Keep a Record:** As you review the data, document any observations, potential issues, or questions. This record will be valuable during the cleaning process in Power Query.

8. Prepare for Power Query

- **Plan the Cleaning Process:** Based on your review, plan the steps you'll need to take in Power Query, such as filtering, removing duplicates, replacing or filling missing values, and transforming data types. It will be helpful to write these steps done.
- **Identify Transformations:** Consider any necessary transformations, like splitting columns, unpivoting data, or merging datasets, that might be required during the cleaning process.

9. Set Expectations

- **Acknowledge Limitations:** Recognize that the data review is a preliminary step, and some issues may only become apparent during the data cleaning process in Power Query.

Authored by Enoch Nunekpeku
https://www.linkedin.com/in/enoch-nunekpeku/

- **Stay Flexible:** Be prepared to adjust your approach as new information and challenges arise during data cleaning.

By following these steps, you'll ensure that you thoroughly understand the dataset's structure, content, and potential issues before beginning the data cleaning process in Power Query.

Phase 2: Data Cleaning and Transformation Steps:

1. Change the Customer's Date of Birth to a Proper Date Format. Next, extract the Customer's Age from the Customer's Date of Birth.
    - *Excel formula tip: =DATEDIF(Cell Coordinate, TODAY(), "Y").*
2. Now give the Raw Dataset a Table format. Call the data table **ZenoMartDt**
3. Now let's Clean and Transform the Data in Power Query. Select the Data Tab and Click on From Table/Range.
4. Extract the first name and last name from the Email Column using Text with Delimiters. Begin the First and Last Name with an Upper Case. Next, Merge the first and last name columns and rename the column to Full Name.
5. Quick one: Do you have any tips for a more efficient way to extract first and last names from the email column?
6. Use choose column to remove the Email Column.
7. Trim the Product Name Column to Remove Extra Spaces.

8. Change the Purchase Date to a Proper Date Format. Next, extract the Year from the Purchase Date and Rename the New Column as the Year of Purchase.
9. Use the Conditional Column technique to change the 1s to Yes and 0s to No in the Loyalty Card Column. Now Rename the New Column as Have Loyalty Card.
10. Extract the Region from the Location Column. Now rename the new column as Region and the Location Column as Country.
11. Give the following Columns a Currency Data Type: Unit Price, Cost Price, & Selling Price.
12. Now using Custom Column under the Add Column tab, Calculate the CostPerOrder and SalesPerOrder. Next, Calculate the Profit.
13. When you are done with the Data Cleaning and Preparation, Close and Load the Clean Data to a New Sheet in the Existing Worksheet.