Data Cleaning (chipotle.tsv):

Columns:

- Order Id
- Quantity
- Item Name
- Choice Description
- Item Price

In this task we are working on tsv file so, before diving into the questions, here's a brief introduction to working with TSV files in **Pandas**:

A TSV (Tab-Separated Values) file is a plain text file where each line represents a data record, and values are separated by tabs. In Pandas, you can use the **read_csv** function with the sep parameter set to **'\t'** to read TSV files.

Optional: Converting TSV to CSV in pandas:
If needed, you can convert a TSV file to CSV using the to_csv function with the sep parameter set to ','. Now, let's proceed with the questions for data cleaning using the Chipotle dataset:

Reading TSV File:

Use the **pd.read_csv** function with sep='\t' to read the Chipotle TSV file into a pandas DataFrame.

1. Missing Values:

Question: Check for missing values in each column (Order ID, Quantity, Item Name, Choice Description, Item Price). How should missing values be handled?

2. Data Types:

Question: Verify the data types of each column. Do they align with their expected types, and should any adjustments be made?

3. Duplicated Entries:

Question: Identify and handle duplicated entries in the dataset. How might duplicates impact analysis, and what is the appropriate action?

4. Quantity and Item Price:

Question: Examine the Quantity and Item Price columns. Are there any inconsistencies or anomalies that need correction?

5. Choice Description:

Question: Analyze the Choice Description column. How should choices be handled, especially when there are multiple descriptions for a single item?

6. Handling Special Characters:

Question: Check for special characters in text-based columns (e.g., Item Name, Choice Description). How can these be addressed for consistency?

7. Order Id Integrity:

Question: Cross-reference the Order ID column for integrity. Are there any irregularities or patterns that need validation?

8. Item Name Standardization:

Question: Standardize the Item Name column. Are there variations that can be unified for better analysis?

9. Quantity and Price Relationships:

Question: Investigate the relationships between Quantity and Item Price. Are there cases where adjustments need to be made for accurate analysis?

10. Data Integrity Check:

Question: Perform a data integrity check by ensuring that quantities and prices align with the corresponding items and descriptions.

11. Converting to CSV:

Optional Question: If needed, convert the cleaned dataset to a CSV file using the to_csv function with sep=','.

12. Handling Categorical Data:

Question: For categorical columns (e.g., Item Name), consider encoding or transforming them into a format suitable for analysis.

13. Consistent Quantity and Price Units:

Question: Ensure consistency in units for Quantity and Item Price. Should any conversions or adjustments be made for uniform analysis?

These questions guide you through the process of cleaning and preparing the Chipotle dataset using pure pandas, addressing various aspects of data quality and consistency. Feel free to explore the dataset.