**CLEANING DATA**

| **Data cleaning checklist** | **Cleaning methods** |
| --- | --- |
| * ~~Size of Data set~~ * ~~Number of categories~~ * ~~Unformatted data~~ * ~~The different data types~~ * ~~Blanks(missing values)~~ * ~~Extra Spaces~~ * ~~Misspells(length)~~ * ~~Duplicates~~ * ~~Date formats~~ | * Sort and Filter (filter by condition-is empty) * Transposing (copy-paste special-paste transposed) * Trim (data-trim whitespace) * ChangeCase (add on) * Format (Clear Formatting) * Conditional formatting * Remove Duplicates(data-remove) * Dates (format-number-date) * Split,Concatenate,Trim * Left,Right,Mid * Countif * Len * Pivot tables * Vlookup (data to look up, ‘where to look’! ,Range, column, false)($ before to lock) * Find and Replace * Plotting (outliers, skewed data) * Copy-Paste Special-values only * Importrange |

**SECOND ROUND**

* **~~Sources of errors: Did you use the right tools and functions to find the source of the errors in your dataset?~~**
* **~~Null data: Did you search for NULLs using conditional formatting and filters?~~**
* **~~Misspelled words: Did you locate all misspellings?~~**
* **~~Mistyped numbers: Did you double-check that your numeric data has been entered correctly?~~**
* **~~Extra spaces and characters: Did you remove any extra spaces or characters using the TRIM function?~~**
* **~~Duplicates: Did you remove duplicates in spreadsheets using the Remove Duplicates function or DISTINCT in SQL?~~**
* **~~Mismatched data types: Did you check that numeric, date, and string data are typecast correctly?~~**
* **~~Messy (inconsistent) strings: Did you make sure that all of your strings are consistent and meaningful?~~**
* **~~Messy (inconsistent) date formats: Did you format the dates consistently throughout your dataset?~~**
* **~~Misleading variable labels (columns): Did you name your columns meaningfully?~~**
* **~~Truncated data: Did you check for truncated or missing data that needs correction?~~**
* **~~Business Logic: Did you check that the data makes sense given your knowledge of the business?~~**