A comprehensive list of the changes made in the project files:

1. Download dependencies
2. Download and install openpyxl (library)
3. Brief description of dataframe using .info()
4. Getting the columns within the dataframe
5. Segmentation of column values within dataframe, like module 1 challenge:

* Reference (column segmentation)

<https://pandas.pydata.org/docs/reference/api/pandas.Series.str.split.html#pandas.Series.str.split>

1. getting the unique values within columns and creation of lists

* Reference (unique values)

<https://pandas.pydata.org/docs/reference/api/pandas.api.extensions.ExtensionArray.unique.html#pandas.api.extensions.ExtensionArray.unique>

* Reference (lists)

<https://pandas.pydata.org/docs/reference/api/pandas.api.extensions.ExtensionArray.tolist.html#pandas.api.extensions.ExtensionArray.tolist>

1. Creation of f string to depict category and sub-category (will use a separate cell to ensure depicted values are correct)
   * <https://www.geeksforgeeks.org/formatted-string-literals-f-strings-python/?ref=header_outind>
2. Using LEN function to determine number of distinct variables within category
   * <https://pandas.pydata.org/docs/reference/api/pandas.Series.str.len.html#pandas.Series.str.len>