synthetic_data_generation

August 18, 2025

1 Synthetic Data Generation

- Name: M. H. Bethmi Kisalka
- Student ID: s224028677

```python
[1]: # import libraries
     import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     import seaborn as sns
```

```python
[2]: # load one file
     file_name = 'ScrappedData.2025_04_02_Coles.csv'
     data = pd.read_csv(file_name)
     data.head()
```

```
[2]:                        _id  product_code category  \
     0  67ed14d6a8bdc17e3ddceed5       8371390   Easter
     1  67ed14d6a8bdc17e3ddceed6       7473849   Easter
     2  67ed14d6a8bdc17e3ddceed7       5726070   Easter
     3  67ed14d6a8bdc17e3ddceed9       4885191   Easter
     4  67ed14d6a8bdc17e3ddceedb       3378370   Easter

                                             item_name best_price  \
     0       Coles Hot Cross Buns Traditional Fruit | 6 Pack      $3.0
     1               Coles Hot Cross Buns Choc Chip | 6 Pack      $3.0
     2  Coles Hot Cross Buns Traditional Fruit Mini | …      $3.0
     3  Cadbury Dairy Milk Easter Chocolate Eggs Bag |…     $6.70
     4     Coles Hot Cross Buns Apple & Cinnamon | 6 Pack      $3.0

       best_unit_price item_price      unit_price price_was special_text  \
     0   $0.73 per 1ea      $4.40   $0.73 per 1ea       NaN      SPECIAL
     1   $0.73 per 1ea      $4.40   $0.73 per 1ea       NaN      SPECIAL
     2   $0.49 per 1ea      $4.40   $0.49 per 1ea       NaN      SPECIAL
     3  $5.88 per 100g      $6.70  $5.88 per 100g       NaN          NaN
     4   $0.73 per 1ea      $4.40   $0.73 per 1ea       NaN      SPECIAL

       promo_text                                              link
```

```
0   2 for $6   https://www.coles.com.au/product/coles-hot-cro…
1   2 for $6   https://www.coles.com.au/product/coles-hot-cro…
2   2 for $6   https://www.coles.com.au/product/coles-hot-cro…
3        NaN   https://www.coles.com.au/product/cadbury-dairy…
4   2 for $6   https://www.coles.com.au/product/coles-hot-cro…
```

[117]:
```python
# get column names
data.columns
```

[117]:
```
Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
       'best_unit_price', 'item_price', 'unit_price', 'price_was',
       'special_text', 'promo_text', 'link'],
      dtype='object')
```

[116]:
```python
# get data shape
data.shape
```

[116]: (140, 12)

[7]:
```python
# check for duplicates
data.duplicated().sum()
```

[7]: 0

[8]:
```python
# check for missing values
data.isna().sum()
```

[8]:
```
_id                0
product_code       0
category           0
item_name          0
best_price         0
best_unit_price    0
item_price         0
unit_price         0
price_was        116
special_text      66
promo_text       115
link               0
dtype: int64
```

[21]:
```python
# get unique product items
data[['product_code','item_name']].value_counts()
```

[21]:
```
product_code  item_name
8371390       Coles Hot Cross Buns Traditional Fruit | 6 Pack
3
2686693       Coles Deli Thawed Australian Cooked Black Tiger Pra … approx.
```

```
250g        3
4885191          Cadbury Dairy Milk Easter Chocolate Eggs Bag | 114g
2
1786160          Red Rock Deli Honey Soy Chicken Potato Chips | 165g
2
8940564          Bibigo Korean Corndog Potato Crispy | 400GRAM
2
        ..
3301820          Castello Double Cream Truffle Cheese | 150g
1
3292718          Kiddylicious Gluten Free Strawberry & Banana Smoothie Melts | 6g
1
3272845          Castello Double Cream Brie Chilli | 150g
1
3029086          Castello Creamy Havarti | 200g
1
9633786          Cadbury Selections Large Easter Chocolate Eggs Bag | 545g
1
Length: 128, dtype: int64
```

[7]: `data.columns`

[7]:
```
Index(['Unnamed: 0', '_id', 'product_code', 'category', 'item_name',
       'best_price', 'best_unit_price', 'item_price', 'unit_price',
       'price_was', 'special_text', 'promo_text', 'link', 'brand', 'item_size',
       'unit_weight', 'similarity_score'],
      dtype='object')
```

[67]: `data1.isna().sum()`

[67]:
```
Unnamed: 0             0
_id                   0
product_code          0
category              0
item_name             0
best_price            0
best_unit_price   77053
item_price            0
unit_price         2237
price_was         77641
special_text      69680
promo_text        59274
link                  0
brand                 0
item_size         66683
unit_weight        2238
similarity_score      0
```

```
dtype: int64
```

```python
[2]: file_name = 'merged_data.csv'
     data = pd.read_csv(file_name, index_col=False)
     data.head()
```

C:\ProgramData\Anaconda3\lib\site-
packages\IPython\core\interactiveshell.py:3165: DtypeWarning: Columns (9) have
mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,

```
[2]:    Unnamed: 0                      _id  product_code category  \
     0            0  67ed14d6a8bdc17e3ddceed5       8371390   Easter
     1            1  67ed14d6a8bdc17e3ddceed6       7473849   Easter
     2            2  67ed14d6a8bdc17e3ddceed7       5726070   Easter
     3            3  67ed14d6a8bdc17e3ddceed9       4885191   Easter
     4            4  67ed14d6a8bdc17e3ddceedb       3378370   Easter

                                              item_name  best_price  \
     0        Coles Hot Cross Buns Traditional Fruit | 6 Pack         3.0
     1                Coles Hot Cross Buns Choc Chip | 6 Pack         3.0
     2  Coles Hot Cross Buns Traditional Fruit Mini | …         3.0
     3  Cadbury Dairy Milk Easter Chocolate Eggs Bag |…         6.7
     4       Coles Hot Cross Buns Apple & Cinnamon | 6 Pack         3.0

        best_unit_price  item_price  unit_price price_was special_text promo_text  \
     0             0.73         4.4        0.73       NaN      SPECIAL   2 for $6
     1             0.73         4.4        0.73       NaN      SPECIAL   2 for $6
     2             0.49         4.4        0.49       NaN      SPECIAL   2 for $6
     3             5.88         6.7        5.88       NaN          NaN        NaN
     4             0.73         4.4        0.73       NaN      SPECIAL   2 for $6

                                              link    brand item_size  \
     0  https://www.coles.com.au/product/coles-hot-cro…    Coles    6 Pack
     1  https://www.coles.com.au/product/coles-hot-cro…    Coles    6 Pack
     2  https://www.coles.com.au/product/coles-hot-cro…    Coles    9 Pack
     3  https://www.coles.com.au/product/cadbury-dairy…  Cadbury      114g
     4  https://www.coles.com.au/product/coles-hot-cro…    Coles    6 Pack

        unit_weight  similarity_score
     0          1ea          0.198373
     1          1ea          0.198373
     2          1ea          0.198373
     3         100g          0.214731
     4          1ea          0.198373
```

```python
[3]: df = data.drop(columns=['Unnamed: 0'], errors='ignore')
```

```
[4]: df.columns
```

```
[4]: Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
             'best_unit_price', 'item_price', 'unit_price', 'price_was',
             'special_text', 'promo_text', 'link', 'brand', 'item_size',
             'unit_weight', 'similarity_score'],
            dtype='object')
```

1.0.1 Plot by category

```
[24]: data['category'].unique()
```

```
[24]: array(['Easter', "Footy's Back", 'Bonus Credit Products',
             'Meat & Seafood', 'Fruit & Vegetables', 'Dairy, Eggs & Fridge',
             'Bakery', 'Deli', 'Pantry', 'Dietary & World Foods',
             'Chips, Chocolates & Snacks', 'Drinks', 'Frozen', 'Household'],
            dtype=object)
```

```python
[3]: plt.figure(figsize=(12, 6))
     ax = sns.countplot(data=data, x='category', order=data['category'].
      ↪value_counts().index)

     # Rotate x-axis labels vertically
     plt.xticks(rotation=90)

     for p in ax.patches:
         count = int(p.get_height())
         ax.text(p.get_x() + p.get_width() / 2, p.get_height() + 0.5,  # x, y␣
      ↪position
                 str(count), ha='center', va='bottom', fontsize=9)

     plt.title('Product Count by Category')
     plt.xlabel('Category')
     plt.ylabel('Count')
     plt.tight_layout()
     plt.show()
```

Product Count by Category


```
[31]: # get brand per category
      items_per_category = {}
      categories = list(data['category'].unique())
      for category in categories:
          items_per_category[category] = data[data['category'] ==␣
       ↪category]['item_name'].tolist()
```

```
[33]: items_per_category['Easter']
```

```
[33]: ['Coles Hot Cross Buns Traditional Fruit | 6 Pack',
       'Coles Hot Cross Buns Choc Chip | 6 Pack',
       'Coles Hot Cross Buns Traditional Fruit Mini | 9 Pack',
       'Cadbury Dairy Milk Easter Chocolate Eggs Bag | 114g',
       'Coles Hot Cross Buns Apple & Cinnamon | 6 Pack',
       'Cadbury Selections Large Easter Chocolate Eggs Bag | 545g',
       'Greens Deluxe Jam & Cream Lamington Cake Roll | 300g',
       'Cadbury Dairy Milk Chocolate Easter Egg Hunt Carton | 306g',
       'Coles Lamb Leg Roast Half | approx. 1.74kg',
       'Coles Deli Thawed Australian Cooked Black Tiger Pra … approx. 250g']
```

1.1 Extract Brand

```
[47]: # extract the brand names
      def extract_brand_name(items_per_category, categories):
          brand_names = []

          for category in categories:
              for item in items_per_category[category]:
                  item_words = item.split()[:2]
```

```
                brand_guess = ' '.join(item_words)

                if brand_guess not in brand_names:
                    brand_names.append(brand_guess)

    return brand_names
```

[48]: `brand_names = extract_brand_name(items_per_category, categories)`

[49]: `brand_names`

[49]: ['Coles Hot',
 'Cadbury Dairy',
 'Cadbury Selections',
 'Greens Deluxe',
 'Coles Lamb',
 'Coles Deli',
 'Mcvities Digestives',
 'Red Rock',
 'Ferrero Rocher',
 'Peckish Gluten',
 'Cadbury Favourites',
 "Four'N Twenty",
 'Maybelline Colossal',
 'Maybelline Lash',
 'Maybelline Sky',
 'Sirena Tuna',
 'Maybelline Falsies',
 'Coles No',
 'Coles Boneless',
 'Continental Stock',
 'Coles RSPCA',
 'Mount Barker',
 'Continental Chicken',
 'Coles Pink',
 'Coles Shepard',
 'Coles Raspberries',
 'Coles Bananas',
 'Fresh Grapes',
 'Coles Strawberries',
 'Coles Blueberries',
 'Coles Seedless',
 'Coles Continental',
 'Brownes Dairy',
 'Fresh Fodder',
 'The Good',
 'Danone Ultimate',

```
'Danone The',
'Devondale Salted',
'Devondale Extra',
'Lurpak Butter',
'Tip Top',
'Noshu 95%',
'Mr Kipling',
'Castello Aged',
'Castello Double',
'Castello Creamy',
'Castello Blue',
'Kenilworth Roast',
'KCF Lime',
'Kenilworth Sweet',
"Leggo's Bolognese",
'Wicked Kimchi',
'Val Verde',
'Nestle Kit',
'Nutella Hazelnut',
'Dolmio Extra',
'Wicked Burger',
'Wicked Sliced',
'Tegel Takeouts',
'Nuttelex Lite',
'Healthy Baker',
'Old El',
'Eclipse Spearmint',
'Nuttelex Original',
'Bibigo Korean',
'Kiddylicious Gluten',
"Arnott's Tiny",
'Smiths Snack',
'Sunbites Snack',
'Smiths Fun',
"Smith's Crinkle",
'Coles Spring',
'Illy Classico',
'Pepsi Max',
'Fever-Tree Premium',
'Illy Espresso',
'Solo Thirst',
'Pepsi Cola',
'Schweppes Lemonade',
'Sunkist Orange',
'Birds Eye',
'Coles Finest',
'Pickers Mini',
```

```
    'Coles Frozen',
    'Comfort Ultra-Care',
    'OMO Ultimate',
    'Easter Bunny',
    'OMO Laundry',
    'Tide Laundry',
    'Comfort Sensitive',
    'OMO Touch',
    'Surf Laundry',
    'OMO Active']
```

```python
[34]: # extract the actual brand names from the brand names
actual_brand_names = ['Coles', 'Cadbury', 'Greens', 'Mcvities Digestives', 'Red␣
 ↪Rock', 'Ferrero Rocher', 'Peckish',
                      "Four'N Twenty", 'Maybelline', 'Sirena', 'Continental',␣
 ↪'Mount Barker', 'Brownes', 'Fresh Fodder', 'The Good Farmer',
                      'Danone', 'Devondale', 'Lurpak', 'Tip Top', 'Noshu', 'Mr␣
 ↪Kipling', 'Castello', 'Kenilworth', "Leggo's", 'Wicked',
                      'Val Verde', 'Nestle', 'Nutella', 'Dolmio', 'Tegel␣
 ↪Takeouts', 'Nuttelex', 'Healthy Baker', 'Old El Paso', 'Eclipse',
                      'Bibigo', 'Kiddylicious', "Arnott's", "Smith's",␣
 ↪'Sunbites', 'Illy','Pepsi', 'Fever_Tree', 'Solo Thirst Crusher',
                      'Schweppes', 'Sunkist', 'Birds Eye', 'Pickers', 'Comfort',␣
 ↪'OMO', 'Surf', 'Tide','Smiths']
```

```python
[52]: # match brand name
def match_brand(item_name, actual_brand_names):
    if not isinstance(item_name, str):
        return "Unknown"

    item_name_lower = item_name.lower()

    for brand in actual_brand_names:
        if brand.lower() in item_name_lower:
            return brand  # return the original-cased brand
    return "Unknown"
```

```python
[6]: data['brand'] = data['item_name'].apply(match_brand)
```

```python
[7]: data[['item_name', 'brand']]
```

```
[7]:                                          item_name       brand
     0      Coles Hot Cross Buns Traditional Fruit | 6 Pack      Coles
     1            Coles Hot Cross Buns Choc Chip | 6 Pack      Coles
     2   Coles Hot Cross Buns Traditional Fruit Mini | …      Coles
     3   Cadbury Dairy Milk Easter Chocolate Eggs Bag |…    Cadbury
     4        Coles Hot Cross Buns Apple & Cinnamon | 6 Pack      Coles
```

```
     ..                                                        …            …
135  Sunkist Orange Soft Drink Cans Multipack 375mL…       Sunkist
136  Pickers Mini Italian Herb Mozzarella Sticks | …      Pickers
137      Birds Eye Frozen Fish Fingers 40 Pieces | 1kg   Birds Eye
138              Coles Frozen Fruit Blueberries | 500g       Coles
139  OMO Active Clean 3 in 1 Laundry Capsules 50 Wa…         OMO

[140 rows x 2 columns]
```

1.2 Extract item size using regex

- Extract sizes:
  - Pack
  - g
  - kg
  - each
  - ml
  - l etc

```python
[56]: import re

      def extract_clean_item_size(item_name):
          if not isinstance(item_name, str):
              return None

          marker_match = re.search(r'(?:\|\s*|approx\.?\s*|…\s*)([\d]+\.?
      ↪\d*\s*[a-zA-Z]+)', item_name, re.IGNORECASE)
          if marker_match:
              return marker_match.group(1).strip()

          fallback_match = re.search(r'(\d+\.?\d*\s*(?:
      ↪g|kg|ml|mL|L|GRAM|each|pack|pk|x\s*\d+\s*[a-zA-Z]+))$', item_name, re.
      ↪IGNORECASE)
          if fallback_match:
              return fallback_match.group(1).strip()

          return None
```

```python
[9]: data['item_size'] = data['item_name'].apply(extract_clean_item_size)   # extract␣
     ↪item size from item name
```

```python
[10]: data[['item_name', 'item_size']]
```

```
[10]:                                           item_name item_size
      0      Coles Hot Cross Buns Traditional Fruit | 6 Pack    6 Pack
      1              Coles Hot Cross Buns Choc Chip | 6 Pack    6 Pack
      2   Coles Hot Cross Buns Traditional Fruit Mini | …    9 Pack
```

```
3      Cadbury Dairy Milk Easter Chocolate Eggs Bag |…       114g
4          Coles Hot Cross Buns Apple & Cinnamon | 6 Pack      6 Pack
..                                                    …         …
135  Sunkist Orange Soft Drink Cans Multipack 375mL…    30 pack
136  Pickers Mini Italian Herb Mozzarella Sticks | …        230g
137      Birds Eye Frozen Fish Fingers 40 Pieces | 1kg        1kg
138              Coles Frozen Fruit Blueberries | 500g        500g
139  OMO Active Clean 3 in 1 Laundry Capsules 50 Wa…    50 pack

[140 rows x 2 columns]
```

```python
for index, row in data[['item_name', 'item_size']].iterrows():
    print(f"Item: {row['item_name']}  |  Size: {row['item_size']}")
```

```
Item: Coles Hot Cross Buns Traditional Fruit | 6 Pack  |  Size: 6 Pack
Item: Coles Hot Cross Buns Choc Chip | 6 Pack  |  Size: 6 Pack
Item: Coles Hot Cross Buns Traditional Fruit Mini | 9 Pack  |  Size: 9 Pack
Item: Cadbury Dairy Milk Easter Chocolate Eggs Bag | 114g  |  Size: 114g
Item: Coles Hot Cross Buns Apple & Cinnamon | 6 Pack  |  Size: 6 Pack
Item: Cadbury Selections Large Easter Chocolate Eggs Bag | 545g  |  Size: 545g
Item: Greens Deluxe Jam & Cream Lamington Cake Roll | 300g  |  Size: 300g
Item: Mcvities Digestives Biscuits Milk Chocolate | 266g  |  Size: 266g
Item: Mcvities Digestives Biscuits Dark Chocolate | 266g  |  Size: 266g
Item: Red Rock Deli Potato Chips Sweet Chilli And Sour Cream | 165g  |  Size:
165g
Item: Red Rock Deli Honey Soy Chicken Potato Chips | 165g  |  Size: 165g
Item: Cadbury Dairy Milk Chocolate Roll Pack | 55g  |  Size: 55g
Item: Maybelline Colossal Big Shot Volume Express Blackest Mascara | 9.7mL  |
Size: 9.7mL
Item: Maybelline Lash Sensational Firework Mascara | 28g  |  Size: 28g
Item: Maybelline Sky High Mascara Cosmic Black | 7.2mL  |  Size: 7.2mL
Item: Maybelline Lash Sensational Very Black Waterproof Masc … 9.5mL  |  Size:
9.5mL
Item: Coles Lamb Leg Roast Half | approx. 1.74kg  |  Size: 1.74kg
Item: Coles No Added Hormone Beef 5 Star Extra Trim Mince | 500g  |  Size: 500g
Item: Coles Pink Lady Apples Medium | approx. 200g  |  Size: 200g
Item: Brownes Dairy Milk Full Cream | 2L  |  Size: 2L
Item: Fresh Fodder Greek Style Tzatziki Dip | 200g  |  Size: 200g
Item: Tip Top The One White Toast Bread | 700g  |  Size: 700g
Item: Noshu 95% Sugar Free Iced Carrot Cakes | 180g  |  Size: 180g
Item: Mr Kipling Chocolate Slices 6 Pack | 155g  |  Size: 155g
Item: Mr Kipling Lemon Slices 6 Pack | 165g  |  Size: 165g
Item: Castello Aged Havarti | 200g  |  Size: 200g
Item: Castello Double Cream Brie Chilli | 150g  |  Size: 150g
Item: Sirena Tuna in Oil Italian Style | 95g  |  Size: 95g
Item: Leggo's Bolognese with Chunky Tomato Garlic & Herbs Pasta Sauce | 500g  |
Size: 500g
Item: Wicked Kimchi | 460g  |  Size: 460g
```

Item: Val Verde Passata Cooking Sauce | 700mL | Size: 700mL
Item: Nestle Kit Kat Breakfast Cereal | 330g | Size: 330g
Item: Tegel Takeouts Korean Style Frozen Chicken Chunks Ruffle … 400GRAM |
Size: 400GRAM
Item: Nuttelex Lite Spread Tub | 500g | Size: 500g
Item: Healthy Baker Flour Plain | 1kg | Size: 1kg
Item: Old El Paso Hard N Soft Taco Kit Mexican Style | 350g | Size: 350g
Item: Eclipse Spearmint Sugar Free Mints Tin | 40g | Size: 40g
Item: Cadbury Dairy Milk Chocolate Easter Egg Hunt Carton | 306g | Size: 306g
Item: Arnott's Tiny Teddy Multipack Variety Biscuits 15 Pack | 375g | Size:
375g
Item: Smiths Snack Mix Variety Multipack 20 Pack | 395g | Size: 395g
Item: Cadbury Dairy Milk Easter Chocolate Eggs Bag | 114g | Size: 114g
Item: Coles Spring Water 24X600mL | 24 pack | Size: 24 pack
Item: Illy Classico Instant Coffee | 95 GRAM | Size: 95 GRAM
Item: Pepsi Max No Sugar Cola Soft Drink Cans Multipack 375mL x 30 Pack | 30
pack | Size: 30 pack
Item: Tegel Takeouts Korean Style Frozen Chicken Chunks Ruffle … 400GRAM |
Size: 400GRAM
Item: Bibigo Korean Waffles Sweet Red Bean Flavoured | 300g | Size: 300g
Item: Comfort Ultra-Care Fabric Conditioner Sky Blue | 900mL | Size: 900mL
Item: OMO Ultimate 3 in 1 Laundry Capsules 50 Washes | 50 pack | Size: 50 pack
Item: Cadbury Dairy Milk Chocolate Easter Egg Hunt Carton | 306g | Size: 306g
Item: Ferrero Rocher Chocolate Gift Box 16 Pack | 200g | Size: 200g
Item: Peckish Gluten Free Fancies Aged Cheddar & Chives Flavoured Ri … 90g |
Size: 90g
Item: Sirena Tuna in Oil Italian Style | 95g | Size: 95g
Item: Maybelline Lash Sensational Sky High Waterproof Mascar … 6mL | Size:
6mL
Item: Coles Lamb Whole Leg Roast | approx. 2.8kg | Size: 2.8kg
Item: Coles Boneless Pork Leg Roast | approx. 2.1kg | Size: 2.1kg
Item: Continental Stock Pot Beef | 112g | Size: 112g
Item: Coles RSPCA Approved Chicken Thigh Fillets Large Pack | approx. 1.2kg |
Size: 1.2kg
Item: Coles Shepard Avocados | 1 each | Size: 1 each
Item: Coles Raspberries | 125g | Size: 125g
Item: Tip Top Cafe Bread Raisin Toast | 650g | Size: 650g
Item: Noshu 95% Sugar Free Banana Bread Slices | 240g | Size: 240g
Item: Mr Kipling Angel Slice 6 pack | 165g | Size: 165g
Item: Castello Creamy Havarti | 200g | Size: 200g
Item: Castello Double Cream Truffle Cheese | 150g | Size: 150g
Item: Castello Blue Cheese | 150g | Size: 150g
Item: Leggo's Bolognese With Red Wine Pasta Sauce | 500g | Size: 500g
Item: Nutella Hazelnut Chocolate Spread | 750g | Size: 750g
Item: Nuttelex Original Spread | 500g | Size: 500g
Item: Sunbites Snack Crackers Mix 2 Multipack 12 Pack | 288g | Size: 288g
Item: Red Rock Deli Honey Soy Chicken Potato Chips | 165g | Size: 165g
Item: Fever-Tree Premium Mixers Light Meditteranean T … 4 pack | Size: 4

pack

Item: Birds Eye Deli Sea Salt & Rosemary Seasoned Potato Chips | 600g | Size: 600g

Item: Easter Bunny Ears Assorted | 1 Each | Size: 1 Each

Item: OMO Laundry Liquid Sensitive Front & Top | 4L | Size: 4L

Item: Tide Laundry 3 In 1 Pods Original | 16 pack | Size: 16 pack

Item: Comfort Sensitive Fabric Conditioner Pure White | 2L | Size: 2L

Item: Coles Lamb Leg Roast Half | approx. 1.74kg | Size: 1.74kg

Item: Maybelline Lash Sensational Blackest Black Mascara | 9mL | Size: 9mL

Item: Mount Barker Free Range Chicken Breast Fillets Skin Off | approx. 500g | Size: 500g

Item: Coles Bananas | approx. 170g | Size: 170g

Item: Fresh Grapes Seedless White Paper | approx. 850g | Size: 850g

Item: Coles Strawberries | 250g | Size: 250g

Item: Coles Blueberries Prepacked | 125g | Size: 125g

Item: Fresh Grapes Seedless Red Paper | approx. 850g | Size: 850g

Item: Coles Seedless Watermelon Cut | approx. 1.8kg | Size: 1.8kg

Item: Coles Continental Cucumbers Loose | 1 each | Size: 1 each

Item: The Good Farmer Gingin Free Range Eggs 12 pack | 700g | Size: 700g

Item: Danone Ultimate Greek Yoghurt Limited Edition 4 pack | 460g | Size: 460g

Item: Danone The Ultimate Black Cherry Yoghurt 4x115g | 4 pack | Size: 4 pack

Item: Fresh Fodder Taramosalata Dip | 200g | Size: 200g

Item: Devondale Salted Butter | 250g | Size: 250g

Item: Devondale Extra Soft Butter Less Fat | 750g | Size: 750g

Item: Danone The Ultimate Golden Honey Yoghurt 4x115g | 4 Pack | Size: 4 Pack

Item: Kenilworth Roast Garlic & Cracked Pepper Cheddar | 165g | Size: 165g

Item: KCF Lime & Cracked Pepper Club Cheddar | 165g | Size: 165g

Item: Dolmio Extra Tomato, Onion & Roast Garlic Pasta Sauce | 500g | Size: 500g

Item: Wicked Burger Pickles | 500g | Size: 500g

Item: Wicked Sliced Pickled Red Onions | 500g | Size: 500g

Item: Coles Hot Cross Buns Traditional Fruit | 6 Pack | Size: 6 Pack

Item: Bibigo Korean Corndog Potato Crispy | 400GRAM | Size: 400GRAM

Item: Bibigo Korean Waffles Sweet Red Bean Flavoured | 300g | Size: 300g

Item: Smiths Fun Mix Variety Multipack 20 Pack | 375g | Size: 375g

Item: Smith's Crinkle Cut Original Potato Chips | 170g | Size: 170g

Item: Illy Espresso Classico Capsules | 10 pack | Size: 10 pack

Item: Illy Classico Ground Coffee | 250g | Size: 250g

Item: Solo Thirst Crusher Original Lemon Soft Drink Cans Multipack … 30 pack | Size: 30 pack

Item: Pepsi Cola Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack | Size: 30 pack

Item: Coles Finest Chocolate & Vanilla Easter Egg | 445 GRAM | Size: 445 GRAM

Item: Bibigo Korean Corndog Potato Crispy | 400GRAM | Size: 400GRAM

Item: Birds Eye Frozen Fish Fillets With Original Lightly Battered O … 425g | Size: 425g

Item: Birds Eye Deli Davidson Plum Ginger Sweet Potato And Red Onion | 600g | Size: 600g

```
Item: OMO Touch of Comfort 3 in 1 Laundry Capsules 50 Washes | 50 pack  | Size:
50 pack
Item: Tide Laundry 3 In 1 Pods Original | 31 pack  | Size: 31 pack
Item: Surf Laundry Liquid Detergent Tropical 80 Washes | 4L  | Size: 4L
Item: Coles Deli Thawed Australian Cooked Black Tiger Pra … approx. 250g  |
Size: 250g
Item: Red Rock Deli Sea Salt & Bal Vinegar Potato Chips | 165g  | Size: 165g
Item: Cadbury Favourites Ultimate Share Boxed Chocolate | 820g  | Size: 820g
Item: Four'N Twenty Frozen Meat Pies 4 Pack | 700g  | Size: 700g
Item: Maybelline Lash Sensational Sky High Washable Masca … 7.2mL  | Size:
7.2mL
Item: Maybelline Falsies Surreal Washable Mascara | 9mL  | Size: 9mL
Item: Maybelline Colossal Washable Mascara Glam Black | 8mL  | Size: 8mL
Item: Coles Deli Thawed Australian Cooked Black Tiger Pra … approx. 250g  |
Size: 250g
Item: Coles RSPCA Approved Chicken Breast Fillets Large Pack | approx. 1.4kg  |
Size: 1.4kg
Item: Continental Chicken Stock Pot | 112g  | Size: 112g
Item: Lurpak Butter Spreadable Slightly Salted | 400g  | Size: 400g
Item: Coles Hot Cross Buns Traditional Fruit | 6 Pack  | Size: 6 Pack
Item: Tip Top The One White Sandwich Bread | 700g  | Size: 700g
Item: Tip Top The One Wholemeal Sandwich Bread | 700g  | Size: 700g
Item: Coles Deli Thawed Australian Cooked Black Tiger Pra … approx. 250g  |
Size: 250g
Item: Castello Double Cream Brie | 150g  | Size: 150g
Item: Kenilworth Sweet Chilli Sour Cream & Bell Peppers Club Che … 165g  |
Size: 165g
Item: Kiddylicious Gluten Free Strawberry & Banana Smoothie Melts | 6g  | Size:
6g
Item: Cadbury Selections Easter Chocolate Eggs Bag | 230g  | Size: 230g
Item: Smith's Crinkle Cut Salt And Vinegar Potato Chips | 170g  | Size: 170g
Item: Schweppes Lemonade Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack  |
Size: 30 pack
Item: Sunkist Orange Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack  |
Size: 30 pack
Item: Pickers Mini Italian Herb Mozzarella Sticks | 230g  | Size: 230g
Item: Birds Eye Frozen Fish Fingers 40 Pieces | 1kg  | Size: 1kg
Item: Coles Frozen Fruit Blueberries | 500g  | Size: 500g
Item: OMO Active Clean 3 in 1 Laundry Capsules 50 Washes | 50 pack  | Size: 50
pack
```

[109]: `data.dtypes`

```
[109]: _id              object
       product_code      int64
       category         object
       item_name        object
```

```
best_price          object
best_unit_price     object
item_price          object
unit_price          object
price_was           object
special_text        object
promo_text          object
link                object
brand               object
item_size           object
dtype: object
```

1.3   Extract price and convert them to float

```python
[11]: # Define the columns to clean
      price_columns = ['best_price', 'item_price']
      unit_price_columns = ['best_unit_price', 'unit_price']

      # Remove "$" and convert to float for simple price columns
      for col in price_columns:
          data[col] = data[col].replace('[\$,]', '', regex=True).astype(float)

      # Extract just the numeric part from unit price columns (e.g., "$0.73 per 1ea"␣
       ↪→ 0.73)
      for col in unit_price_columns:
          price = data[col].str.extract(r'\$?([\d\.]+)').astype(float)

          data[f'unit_weight'] = data[col].str.extract(r'per\s+(\w+)', expand=False)
          data[col] = price
```

```python
[12]: data[['item_name','best_price', 'item_price','best_unit_price', 'unit_price',␣
       ↪→'unit_weight']]
```

```
[12]:                                        item_name  best_price  \
      0        Coles Hot Cross Buns Traditional Fruit | 6 Pack         3.0
      1                 Coles Hot Cross Buns Choc Chip | 6 Pack         3.0
      2      Coles Hot Cross Buns Traditional Fruit Mini | …         3.0
      3      Cadbury Dairy Milk Easter Chocolate Eggs Bag |…         6.7
      4          Coles Hot Cross Buns Apple & Cinnamon | 6 Pack      3.0
      ..                                                   …           …
      135  Sunkist Orange Soft Drink Cans Multipack 375mL…         28.0
      136  Pickers Mini Italian Herb Mozzarella Sticks | …         7.5
      137      Birds Eye Frozen Fish Fingers 40 Pieces | 1kg        11.5
      138              Coles Frozen Fruit Blueberries | 500g         6.5
      139  OMO Active Clean 3 in 1 Laundry Capsules 50 Wa…        30.0

           item_price  best_unit_price  unit_price unit_weight
```

```
0        4.4          0.73        0.73        1ea
1        4.4          0.73        0.73        1ea
2        4.4          0.49        0.49        1ea
3        6.7          5.88        5.88        100g
4        4.4          0.73        0.73        1ea
..       …            …           …          …
135     28.0          2.49        2.49        1l
136      7.5          3.26        3.26        100g
137     11.5         11.50       11.50        1kg
138      6.5         13.00       13.00        1kg
139     30.0          0.60        0.60        1ea

[140 rows x 6 columns]
```

[126]: `data.dtypes`

```
[126]: _id                 object
       product_code         int64
       category            object
       item_name           object
       best_price         float64
       best_unit_price    float64
       item_price         float64
       unit_price         float64
       price_was           object
       special_text        object
       promo_text          object
       link                object
       brand               object
       item_size           object
       unit_weight         object
       dtype: object
```

```python
[13]: # generate similarity score

      def compute_similarity(group):
          prices = group['item_price'].values
          # Mean absolute price difference from others
          similarity_scores = []
          for i, price in enumerate(prices):
              diffs = np.abs(prices - price)
              mean_diff = np.mean(diffs)
              # Convert mean difference to similarity score (inverse relationship)
              similarity = 1 / (1 + mean_diff)  # Avoid division by zero
              similarity_scores.append(similarity)
          group['similarity_score'] = similarity_scores
          return group
```

```
# Apply per category group
data = data.groupby('category', group_keys=False).apply(compute_similarity)
data.head(20)
```

[13]:
```
                         _id  product_code              category  \
0    67ed14d6a8bdc17e3ddceed5       8371390                Easter
1    67ed14d6a8bdc17e3ddceed6       7473849                Easter
2    67ed14d6a8bdc17e3ddceed7       5726070                Easter
3    67ed14d6a8bdc17e3ddceed9       4885191                Easter
4    67ed14d6a8bdc17e3ddceedb       3378370                Easter
5    67ed14d6a8bdc17e3ddceedc       9633786                Easter
6    67ed14d6a8bdc17e3ddceede       9077853                Easter
7    67ed14d6a8bdc17e3ddceedf       5854739           Footy's Back
8    67ed14d6a8bdc17e3ddceee0       5854717           Footy's Back
9    67ed14d6a8bdc17e3ddceee1       1786182           Footy's Back
10   67ed14d6a8bdc17e3ddceee4       1786160           Footy's Back
11   67ed14d6a8bdc17e3ddceee6       5072676           Footy's Back
12   67ed14d6a8bdc17e3ddceeeb       2855600  Bonus Credit Products
13   67ed14d6a8bdc17e3ddceeec       8100086  Bonus Credit Products
14   67ed14d6a8bdc17e3ddceeee       5239330  Bonus Credit Products
15   67ed14d6a8bdc17e3ddceef1       2364438  Bonus Credit Products
16   67ed14d6a8bdc17e3ddceef4       9246065          Meat & Seafood
17   67ed14d6a8bdc17e3ddceefa       8112799          Meat & Seafood
18   67ed14d6a8bdc17e3ddcef04       5111654      Fruit & Vegetables
19   67ed14d6a8bdc17e3ddcef0d       5066120     Dairy, Eggs & Fridge

                                    item_name  best_price  \
0      Coles Hot Cross Buns Traditional Fruit | 6 Pack        3.00
1              Coles Hot Cross Buns Choc Chip | 6 Pack        3.00
2    Coles Hot Cross Buns Traditional Fruit Mini | …        3.00
3    Cadbury Dairy Milk Easter Chocolate Eggs Bag |…        6.70
4        Coles Hot Cross Buns Apple & Cinnamon | 6 Pack        3.00
5    Cadbury Selections Large Easter Chocolate Eggs…       13.00
6    Greens Deluxe Jam & Cream Lamington Cake Roll …        6.00
7    Mcvities Digestives Biscuits Milk Chocolate | …        2.75
8    Mcvities Digestives Biscuits Dark Chocolate | …        2.75
9    Red Rock Deli Potato Chips Sweet Chilli And So…        3.15
10   Red Rock Deli Honey Soy Chicken Potato Chips |…        3.15
11       Cadbury Dairy Milk Chocolate Roll Pack | 55g        1.25
12   Maybelline Colossal Big Shot Volume Express Bl…       21.00
13   Maybelline Lash Sensational Firework Mascara |…       17.50
14    Maybelline Sky High Mascara Cosmic Black | 7.2mL       26.00
15   Maybelline Lash Sensational Very Black Waterpr…       25.00
16          Coles Lamb Leg Roast Half | approx. 1.74kg       26.10
17   Coles No Added Hormone Beef 5 Star Extra Trim …       10.00
18       Coles Pink Lady Apples Medium | approx. 200g        1.70
```

```
19                   Brownes Dairy Milk Full Cream | 2L          4.45

     best_unit_price  item_price  unit_price  price_was  special_text  promo_text  \
0               0.73        4.40        0.73        NaN       SPECIAL    2 for $6
1               0.73        4.40        0.73        NaN       SPECIAL    2 for $6
2               0.49        4.40        0.49        NaN       SPECIAL    2 for $6
3               5.88        6.70        5.88        NaN           NaN         NaN
4               0.73        4.40        0.73        NaN       SPECIAL    2 for $6
5               2.39       13.00        2.39      16.00       SPECIAL         NaN
6               2.00        6.00        2.00        NaN           NEW         NaN
7               1.03        2.75        1.03       5.50           50%         NaN
8               1.03        2.75        1.03       5.50           50%         NaN
9               1.91        3.15        1.91       6.30           50%         NaN
10              1.91        3.15        1.91       6.30           50%         NaN
11              2.27        1.25        2.27       2.50           50%         NaN
12             21.65       21.00       21.65        NaN           NaN         NaN
13             10.36       29.00       10.36        NaN       SPECIAL   2 for $35
14             36.11       26.00       36.11        NaN           NaN         NaN
15             26.32       25.00       26.32        NaN           NaN         NaN
16             15.00       26.10       15.00        NaN           NaN         NaN
17             20.00       10.00       20.00        NaN           NaN         NaN
18              8.50        1.70        8.50        NaN           NaN         NaN
19              2.23        4.45        2.23        NaN     EVERY DAY         NaN

                                            link               brand  \
0   https://www.coles.com.au/product/coles-hot-cro…                Coles
1   https://www.coles.com.au/product/coles-hot-cro…                Coles
2   https://www.coles.com.au/product/coles-hot-cro…                Coles
3   https://www.coles.com.au/product/cadbury-dairy…              Cadbury
4   https://www.coles.com.au/product/coles-hot-cro…                Coles
5   https://www.coles.com.au/product/cadbury-selec…              Cadbury
6   https://www.coles.com.au/product/greens-deluxe…               Greens
7   https://www.coles.com.au/product/mcvities-dige…  Mcvities Digestives
8   https://www.coles.com.au/product/mcvities-dige…  Mcvities Digestives
9   https://www.coles.com.au/product/red-rock-deli…             Red Rock
10  https://www.coles.com.au/product/red-rock-deli…             Red Rock
11  https://www.coles.com.au/product/cadbury-dairy…              Cadbury
12  https://www.coles.com.au/product/maybelline-co…           Maybelline
13  https://www.coles.com.au/product/maybelline-la…           Maybelline
14  https://www.coles.com.au/product/maybelline-sk…           Maybelline
15  https://www.coles.com.au/product/maybelline-la…           Maybelline
16  https://www.coles.com.au/product/coles-lamb-le…                Coles
17  https://www.coles.com.au/product/coles-no-adde…                Coles
18  https://www.coles.com.au/product/coles-pink-la…                Coles
19  https://www.coles.com.au/product/brownes-dairy…              Brownes

    item_size unit_weight  similarity_score
```

```
0      6 Pack        1ea        0.190840
1      6 Pack        1ea        0.190840
2      9 Pack        1ea        0.190840
3       114g        100g        0.203252
4      6 Pack        1ea        0.190840
5       545g        100g        0.125313
6       300g        100g        0.203252
7       266g        100g        0.284091
8       266g        100g        0.284091
9       165g        100g        0.290698
10      165g        100g        0.290698
11       55g        100g        0.223714
12      9.7mL       10ml        0.153728
13       28g        10g         0.140746
14      7.2mL       10ml        0.203874
15      9.5mL       10ml        0.203874
16      1.74kg      1kg         0.067981
17      500g        1kg         0.111483
18      200g        1kg         0.267594
19       2L         1l          0.311042
```

1.4   Merge other data

```python
[14]:  # read other csvs
       file_names = ['ScrappedData.2025_04_03_Coles.csv', 'ScrappedData.
        ↪2025_04_08_Coles.csv', 'ScrappedData.2025_04_09_Coles.csv', 'ScrappedData.
        ↪2025_04_17_Coles_All.csv']
       data_1, data_2, data_3, data_4 = [pd.read_csv(fp) for fp in file_names]
```

```python
[15]:  # merge csv
       merged_df = pd.concat([data_1, data_2, data_3, data_4], ignore_index=True,␣
        ↪sort=False)
```

```python
[19]:  merged_df.columns
```

```python
[19]:  Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
              'best_unit_price', 'item_price', 'unit_price', 'price_was',
              'special_text', 'promo_text', 'link'],
             dtype='object')
```

```python
[35]:  def find_new_brand_names(item_names, actual_brand_names, max_words=3):
           new_brand_names = []

           # Convert actual brand names to lowercase for comparison
           actual_brands_lower = [brand.lower() for brand in actual_brand_names]

           for item in item_names:
```

```python
        item_lower = item.lower()
        # Check if any known brand is already in the item
        if not any(brand in item_lower for brand in actual_brands_lower):
            # Take the first max_words from the item
            brand_guess = ' '.join(item.split()[:max_words]).strip()

            # Add if it's not already captured (case-insensitive)
            if brand_guess.lower() not in [b.lower() for b in new_brand_names]:
                new_brand_names.append(brand_guess)

    return new_brand_names
```

[36]:
```python
item_names = merged_df['item_name'].dropna().tolist()

new_brands = find_new_brand_names(item_names, actual_brand_names)


# updated_brand_list = actual_brand_names + new_brands
```

[37]:
```python
new_brands
```

[37]: ['Sensodyne Repair &',
 'Sensodyne Soft Daily',
 'Sensodyne Complete Care',
 'Tassal Tasmanian Smoked',
 'Tassal Hot Smoked',
 'Fresh Grapes Seedless',
 'Western Star Traditional',
 'Harvey Fresh Vanilla',
 'Pauls Custard Vanilla',
 'Tablelands Dairy Blend',
 'Pauls Custard Snack',
 'Western Star Supersoft',
 'Bega Dairy Natural',
 'Mersey Valley Club',
 'Tasmanian Heritage Triple',
 'Happy Snack Company',
 'La Espanola Extra',
 'Slice X Chupa',
 "Abe's Bagels Everything",
 'Vegemite Spread |',
 'Gravox Traditional Gravy',
 'MasterFoods All Purpose',
 'Carmans Chia Bowls',
 'Liberate Gluten Free',
 'John West Tempters',
 'Bega Tasty Block',

```
'MasterFoods Soy Honey',
'Powerade Berry Ice',
'Powerade Mountain Blast',
'Powerade ION4 Blackcurrent',
'Coca-Cola Zero Sugar',
'Gym Bod Caramel',
'Streets Golden Gaytime',
'Streets Paddle Pop',
'Streets Blue Ribbon',
'Gym Bod Cookies',
'Weis Ice Cream',
'Easter Hunt Basket',
'Canesten Antibacterial and',
'Oral B IO2',
'Libra Goodnights Extra',
'Quick Screen Travel',
'Lady Jayne Detangling',
'Gaviscon Dual Action',
'Libra Pads Ultra',
'Quick Flick Multipurpose',
'Swisspers Cotton Tips',
'Nair Shower Power',
'Sensodyne Sensitivity &',
'Babybel Mini Cheese',
'Sanitarium Weet-Bix Breakfast',
"Campbell's Real Stock",
'Swisspers Baby Organic',
'Aptamil Gold+ Lactose',
'Health Lab Charlottes',
'Health Lab Caramel',
'Kinder Chocolate 16',
'The Natural Confectionery',
'Sakata Rice Cracker',
'Uncle Tobys Chewy',
'Golden Circle Tropical',
'Coca-Cola Classic Soft',
'Powerade Sports Drink',
'Twinings English Breakfast',
'Red Bull Energy',
'Gym Bod Ice',
'Gym Bod Chocolate',
'Gym Bod Vanilla',
'Magnum Ice Cream',
"Ben & Jerry's",
'Easter Bunny Pom',
'Swisse Ub Magnesium',
'Glam By Manicare',
```

'Tamar Valley The',
'Sealord Gluten Free',
'Heaps Normal Half',
'Tegel Take Outs',
'Windex Glass Cleaner',
"Bosisto's Multipurpose Cleaner",
'Finish Ultimate Plus',
'Earth Choice Ultra',
'Finish Ultimate 0%',
'Nilodor Concentrated Deodoriser',
'Radiant Laundry Capsules',
"Bosisto's Bathroom &",
'Epzen Calm Cleanse',
'Vaseline Gluta-Hya Dewy',
'Vaseline Gluta-Hya Pro',
'Murine Relief Eye',
'Great Northern Super',
'Carlton Dry Bottle',
'Gage Roads Single',
"Gage Roads Pinky's",
'Grant Burge 5th',
'Heaps Normal Third',
'Oyster Bay Pinot',
'Heineken Bottle 330mL',
'XXXX Gold Block',
'Sprite Lemonade Soft',
'Coca-Cola Vanilla Soft',
'Blackmores Multivitamins 50+',
'Blackmores Valerian Forte',
'Blackmores Cranberry Forte',
'Don English Style',
'Good Farmer Free',
'Olympian Easter Tsoureki',
'Jarlsberg Cheese Wedge',
'Moro Primero Extra',
'Lindt Easter Gold',
'Soda Stream Gas',
'Connoisseur Classic Vanlla',
'Destination Italy Tiramisu',
'McCain Frozen Baby',
'Haagen-Dazs Ice Cream',
"Kb's Marinara Mix",
'Easter Mug |',
'Morning Fresh Soft',
'Bref Power Active',
'Quilton 3 Ply',
'Quilton White Toilet',

"Swisse Men's 65+",
'Tatachilla White Admiral',
'Victoria Bitter Block',
'Lindt Easter Flower',
'Fanta Orange Soft',
'Harvey Beef Premium',
"Abe's Bagels Thins",
'Moccona Classic Medium',
'Capilano 100% Australian',
'Lindt Lindor Easter',
'Lindt Easter Medley',
'Red Tulip Chocolate',
'Arnotts Tim Tam',
'Coca-Cola Diet Coke',
'Sprite Zero Sugar',
"Blackmores Probiotics+ Women's",
'Jarlsberg Regular Cheese',
'Don Honey Leg',
'Riverview Farms Pork',
'Don Ham On',
'Bertolli Extra Virgin',
'Indomie Fried Mi',
'Gatorade Fast Twitch',
'Fanta Orange Zero',
'Maltesers Milk Chocolate',
'Gatorade Sport Drink',
'Kettle Chilli Potato',
'Blackmores Vitamin B12',
'MT Barker Free',
'Jarlsberg Original Cheese',
'Don Double Smoked',
'Don Italian Style',
'Don Premium Melosi',
'Primo Streaky Bacon',
"Carman's Oat Cups",
'Cocobella Coconut Water',
'Peters Maxibon Vanilla',
'Fropro Protein Beef',
'Duck Fresh Discs',
'Morning Fresh Limited',
'Morning Fresh Dishwashing',
'Easter Egg Jumbo',
'Epzen Soothe Soak',
"Swisse Ultivite Women's",
'Life Space Gummies',
'Corona Bottle 355mL',
'Smirnoff Red Vodka',

'Dr. Oetker Frozen',
'Duck Extra Power',
'Ajax Spray N',
'Kleenex Complete Clean',
'Quilton King Size',
'Ostelin Vitamin D',
'Great Northern Original',
'Jack Daniels Cola',
'Finish Ultimate Dishwashing',
'Handee Ultra Paper',
'Hardys Brave New',
'Cleanskin Vodka 700mL',
'Victoria Bitter Bottle',
'Asahi Super Dry',
'UP&GO Liquid Breakfast',
'MasterFoods Finely Crushed',
'Nescafe Blend 43',
'Cheer Plant Based',
'Keens Curry Powder',
'Golden Crumpet Rounds',
'Golden Crumpet Square',
'Nair Sensitive Shower',
'Morning Fresh Lemon',
'Nair Precision Easiwax',
'Chocolate Easter Bunny',
'Selections Easter Chocolate',
'Easter Gold Bunny',
'Chocolate Easter Hunt',
'Dairy Milk Easter',
'Easter Chocolate Selection',
'Grand Rocher Milk',
'Hot Cross Buns',
'Easter Mini Gold',
'Surprise Chocolate Easter',
'Bubbly Mint Chocolate',
'Milk Chocolate Solid',
'Crunchie Hollow Chocolate',
'Cherry Ripe Hollow',
'Easter Gourmet Milk',
'Collection Boxed Egg',
'Mini Eggs Milk',
'Kitkat Chocolate Easter',
'Easter Bunny Chocolate',
'Bueno Easter Eggs',
'Marvellous Creations Chocolate',
'Surprise Chocolate Giant',
'Dairy Milk Coconut',

```
'Milkybar White Choc',
'Mint Slice Egg',
'Hazelnut Brittle White',
'Creme Egg 6-Pack',
'Peanut Chocolate Speckled',
'Life Savers Fruit',
'Bunny Ears Assorted',
'Milk Chocolate Mini',
'Mug',
'Fresh Tasmanian Salmon',
'Honey Soy Chicken',
'Zero Sugar Soft',
'Classic Soft Drink',
'Fresh West Australian',
'Sweet Gold Potatoes',
'Max No Sugar',
'Potato Chips Sea',
'Lamb Boneless Shoulder',
'Tennessee Whiskey 1L',
'Roses Boxed Chocolates',
'Natural Walnuts',
'Sea Salt &',
'Yellow Label Brut',
'Red Royale Washed',
'Kabana From The',
'Virginian Ham From',
'Simply Thin BBQ',
'UP Rum 1L',
'Cordon Rouge NV',
'Beef Garlic &',
'12YO Single Malt',
'Ultra Low Carb',
'Thawed Basa Fillets',
'Lamb Ribs With',
'Draught Bottle 375mL',
'Whisky 1L',
'Kitchen Caesar Salad',
'Strawberries',
'Lindor Easter Assorted',
'Dairy Milk Chocolate',
'Boneless Pork Leg',
'Easter Flower Gold',
'Humpty Dumpty Easter',
'Thawed Australian Cooked',
'Collection Hazelnut Easter',
'Crunchie Chocolate Easter',
'Dairy Milk Hollow',
```

'Caramello Easter Chocolate',
'Chocolate Easter Bunnies',
'Chocolate & Vanilla',
'Oreo Chocolate Easter',
'Dream Chocolate Easter',
'Easter Hazelnut Dark',
'Arnotts Iced Vovo',
'Milk Chocolate Bunnies',
'Mini Easter Eggs',
'Crispy Chocolate Speckled',
'Kitkat Mini Pouch',
'Easter Fried Eggs',
'Rocklea Road Bunny',
'Surprise Gift Pack',
'Australia Puddles The',
'Easter Tsoureki Gift',
'Easter Brownies',
'Hunt Basket',
'Dry Bottle 330mL',
'Perino Red Grape',
'Gold Block Can',
'The One White',
'No Added Hormone',
'Marlborough Sauvignon Blanc',
'Lamb Loin Chops',
'Crinkle Cut Original',
'Super Dry Bottle',
'Irish Whiskey 700mL',
'Barossa Shiraz 750mL',
'Bottle 330mL',
'Rocher Chocolate Gift',
'Scotch Whisky 700mL',
'4 Leaf Salad',
'Vodka 700mL',
'Lamb Leg Butterflied',
'Australian Roasted &',
'Can 375mL',
'Raffaello 26 Pack',
'Nastro Azzurro 5%',
'Collection Rocher Raffaello',
'White Hungarian Salami',
'Prosecco NV 750mL',
'Pinot Noir Chardonnay',
'Gin 1L',
'Estate Shiraz 750mL',
'Italian Style Pork',
'Semi Dried Tomatoes',

'Lamb Leg Roast',
'Australian Thawed Raw',
'Chocolate Easter Rabbit',
'Clinkers Chocolate Giant',
'Dairy Milk Deluxe',
'Clinkers Chocolate Easter',
'Creme Egg Minis',
'Grand Rocher Dark',
'Gluten Free Hot',
'Caramilk Easter Chocolate',
'Easter Choc Coated',
'Nougat Easter Egg',
'Creme Egg Chocolate',
'Easter Sour Carrots',
'Milk Chocolate Easter',
'Easter Milk Chocolate',
'Raspberry&coconut Slice Egg',
'Kit Kat Mini',
'Britecrawlers & Zappo',
'Deluxe Jam &',
'Blueberries Prepacked',
'Baby Cucumbers',
'Carrots Prepacked',
'Strawberries Premium',
'Chocolate Milk',
'Kent Pumpkin Cut',
'White Bourbon 1L',
'Lamb Shanks 2',
'Milk Chocolate Party',
'Super Dry 3.5',
'RSPCA Approved Free',
'Pistachios Roasted &',
'Bottle 375mL',
'Lindor Assorted Chocolate',
'NV Sparkling Brut',
'Pork Belly Roast',
'Cola Can 375mL',
'Crinkle Cut Salt',
'12YO Scotch Whisky',
'Beef And Thyme',
'Salada Original Crispbread',
'Triple Smoked Leg',
'Brewed Drink Ginger',
'Crinkle Cut Chicken',
'Baby Carisma Potatoes',
'White & Cola',
'Pork Sausage',

'Original Can 375mL',
'Super Crisp Lager',
'Australian Lager Can',
'Original Lager Bottle',
'Roast Salted Macadamias',
'Extra Dry Bottle',
'Butcher Lamb Leg',
'Irish Cream 1L',
'Diet Coke Soft',
'Orange Soft Drink',
'Whole Chicken',
'White Label Bourbon',
'Potato Chips Share',
'Pub Style Extra',
'Australian Garlic Prawns',
'Gentleman Jack Tennessee',
'Black Label 12YO',
'Bourbon & Cola',
'Shiraz 750ml',
'Kitchen Family Ranch',
'Mixed Nuts Dry',
'Kitchen Pasta Salad',
'Mid Strength Bottle',
'Pacific Ale Bottle',
'Irish Cream Liqueur',
'Natural Macadamias',
'Lindor Dark Assorted',
'Hommus Dip',
'Potato Chips Party',
'Traditional Hommus Dip',
'Jumbo Prawn Cutlets.',
'RSPCA Approved Chicken',
'Blue Label Scotch',
'Cream Chocolate Bag',
'Aperitivo 700mL',
'Faith Shiraz 750mL',
'Dry Roasted Cashews',
'Japanese Gin 700mL',
'Blue Cheese',
'Rare Dry Gin',
'XPA Can 375mL',
'Made Easy Slow',
'Cabernet Sauvignon 750mL',
'Cruiser Mixed 275mL',
'Slow Cooked Lamb',
'Lindor Milk Chocolate',
'Dry Can 375mL',

'Mild Hungarian Salami',
'Reserve Sparkling Prosecco',
'Semillon Sauvignon Blanc',
'Dip Hommus Traditional',
'Tasmanian Vodka 1L',
'Chocolate Truffles',
'Club Original Cheddar',
'Zero Sugar Mixers',
'NV Pink 750mL',
'Cruiser Wild Raspberry',
'Pinot Noir 750mL',
'Dollop Cream',
'Baked Pretzels Wheat',
'Selection Pack Sausages',
'Salad Baby Leaf',
'Deli Golden Ale',
'Triple Cream Brie',
'Chicken Mozzarella &',
'Lamb Whole Leg',
'Selections Large Easter',
'Oreo Easter Chocolate',
'Rocher Easter Egg',
'Collection Dark Chocolate',
'Turkish Delight Easter',
'Easter Figurines 6',
'Milk Chocolate Bilby',
'Crunchie Easter Chocolate',
'Mint Easter Chocolate',
'Lindor Easter Mini',
'Easter Chocolate Carrot',
'Mini Eggs Chocolate',
'Speckled Easter Eggs',
'Rocher Milk Chocolate',
'100S & 1000S',
'Caramilk Chocolate Easter',
'Easter Bugs &',
'Gluten Free Chocolate',
'Surprise Chocolate Gift',
'Milk Choc Foiled',
'Hollow Chocolate Bunny',
'Smarties Chocolate Easter',
'Kit Kat Chocolate',
'Milk Chocolate Bunny',
'After Eight Chocolate',
'Chocolate Hazelnut Mini',
'Surprise Maxi Egg',
'Easter Belgian Milk',

'Smarties Mini Eggs',
'Egg Jumbo Chalks',
'Bunny Pom Pom',
'Lamb Cutlets (11x9)',
'Lamb Forequarter Chops',
'Savoy Crackers Original',
'Bottle 355mL',
'Favourites Boxed Chocolate',
'Green Beans Prepacked',
'Red Label Scotch',
'Thawed Barramundi Fillets',
'Corn Chips Supreme',
'Favourites Ultimate Share',
'Red Vodka 700mL',
'Graze Grassfed Beef',
'The Original Beef',
'Celery Sticks Prepacked',
'Original Crackers',
'NV Yellow 750mL',
'Lamb Kebabs',
'Brewed Drink Passionfruit',
'UP & Cola',
'Premium Bottle 375mL',
'Moscato 750mL',
'Sausages Beef Garlic',
'Gin 700mL',
'Club Classic Cheddar',
'Tzatziki Dip',
'Lemonade Soft Drink',
'Turkish Dried Apricots',
'Energised Soft Drink',
'Ginger Beer Can',
'Thawed Cooked Peeled',
'Gluten Free Thinly',
'Corn Chips Mexicana',
'Cheese Brie Wrapped',
'Graze Grass Fed',
'Sugar Free Lemon',
'Classic Beef Burgers',
'Brewed Drink Tropical',
'Sparkling Pinot Noir',
'Premium Sem Sauv',
'Squeezy Sauce Tomato',
'700mL',
'Hot Hungarian Salami',
'Prosciutto',
'Lindor Silver Assorted',

'Made Easy Roast',
'& No Sugar',
'Agrum Blood Orange',
'RSPCA Chicken Portions',
'Deluxe Triple Chocolate',
'Beef Kebabs',
'Kitchen Asian Salad',
'Lemon Lime Zero',
'Cerveza Bottle 355mL',
'Chardonnay Pinot Noir',
'AOC Luberon Rose',
"Max's Shiraz 750mL",
"'The Siding' Cabernet",
'Roasted And Salted',
'Rocher Heart Chocolate',
'Pale Ale Bottle',
'Pink Gin RTD',
'Thick & Juicy',
'Boneless Whole Chicken',
'Boneless Pork Shoulder',
'Extra Bottle 355mL',
'Dry Gin 1L',
'Draught Block Can',
'RSPCA Approved Whole',
'Chocolate Seashells',
'Boneless Pork Loin',
'New York Style',
'Double Jack &',
'Roasted & Salted',
'Gold Bottle 375mL',
'Double Lemon &',
'Kitchen Asian Style',
'Vodka 1L',
'Brewed Drink Guava',
'Cooked Prawns With',
'Original Pale Ale',
'Lemon Lime Bitters',
'Double Black Scotch',
'Lindor 70% Cocoa',
'Pork Cutlets',
'Crinkle BBQ Potato',
'Mediterranean Olives',
'Crinkle Cut Potato',
'Liqueur 700mL',
'White Zero Sugar',
'Lamb Shoulder Butterflied',
'Kitchen Potato Salad',

'Beef Brisket Burgers',
'Pasito Passionfruit Soft',
'Pork Scotch Steaks',
'Antipasto Mix',
'Apple Cider Can',
'Ichiban Bottle 330mL',
'Milk Chocolate Baking',
'Black Double Serve',
'Snacking Carrots',
'Soda Water Bottle',
'Lager Can 375mL',
'Zero Sugar Can',
'Bubbles 750mL',
'Pepperoni Hot Salami',
'Australian Pork &',
'Lamb Garlic And',
'Spiced Rum 1L',
'Peanuts Unsalted',
'Estate Pinot Grigio',
'Brewed Drink Lemon',
'Kitchen Creme Royale',
'12YO Blended Scotch',
'150 Lashes Pale',
'Limoncello Spritz 750mL',
'Prosecco 750mL',
'Black Label Red',
'Zero Sugar Original',
'Australian Salad Family',
'Giant Pitted Kalamata',
'Whiskey 700mL',
'Salt & Pepper',
'Mini Dessert Selection',
'Lindor Pink Assorted',
'Macaron Selection',
'Mixers Tonic Water',
'French Style Cheesecake',
'Cruskits Original Crispbread',
'Prosecco DOC Organic',
'Brut Imperial NV',
'Riesling 750mL',
'Coffee Liqueur 700mL',
'NV 750mL',
'Diet Soft Drink',
'Kitchen American BBQ',
'Red Label Chardonnay',
'Favourites Party Edition',
'Kitchen Vegetables Baby',

'Kitchen Australian Diced',
'Kentucky Straight Bourbon',
'Potato Chips Thai',
'Kalamata Olive Dip',
'Creaming Soda Soft',
'Zero Sugar Classic',
'Zero Sugar Agrum',
'Zero Sugar Cola',
'NV Rose 750mL',
'Kitchen Beetslaw Salad',
'Peanuts Salted',
'Hard Cider Can',
'Lemon & Pepper',
'Artisan Crackers Biscuits',
'Pork Scotch Steak',
'Hot Soppressa Salami',
'Dark Chocolate Baking',
'Chardonnay 750mL',
'Wafers Chocolate',
'Salted Mixed Nuts',
'Milk Chocolate Crunchy',
'Orange Mango Natural',
'Sparkling White 750mL',
'Lamb Burgers With',
'Dip Reduced Fat',
'Pralines Noir Dark',
'Dry Lager Bottle',
'Brewed Drink Sarsaparilla',
'Snakes Lollies Party',
'Praline Master Collection',
'Dairy & Gluten',
'Non Vintage Sparkling',
'Milk Chocolate Orange',
'YB Prosecco 750mL',
'Simply BBQ Pack',
'Crafted Guacamole Dip',
'Lemon Lime Natural',
'Pistachios',
'Kitchen Broccoli &',
'Especial Tequila 700mL',
'Double Barrel Shiraz',
'Quiche Lorraine',
'Silver Tequila 700mL',
'Cheese Apricot &',
'Raffaello Heart T10',
'Classic Water Crackers',
'London Dry Gin',

```
'Chocolate Love Heart',
'Indian Tonic Water',
'Pretzel Twists',
'Tasmanian Gin 700mL',
'Sacred Hill Semillon',
'Zero Sugar Lemon',
'Red Hill Cool',
'Prosecco DOC 750mL',
'Eazy Hazy Can',
'Avocado Dip',
'Timber Spoons',
'Vodka Lemon Squash',
'Sparkling Shiraz 750mL',
'Roast Pumpkin &',
'Pink Gin 700mL',
'Original Cider Can',
'BBQ Aluminium XL',
'Chinotto Drink Cans',
'Refreshing Ale Bottle',
'Crafted Crispy Bacon',
'Bloody Shiraz Gin',
'Swiss Milk Chocolate',
'South Coast Pale',
'Soft Centres Gift',
'Cracked Black Pepper',
'Cheese Spring Onion',
'Draught Can 375mL',
'10YO Single Malt',
'Yarra Valley Pinot',
'Barossa Valley Rose',
'Rose Brut NV',
'Lime Soda Water',
'Cremant Blanc de',
'Sparkling Cuvee Brut',
'Hommus Roasted Pine',
'Blood Orange &',
'Rhubarb & Ginger',
'Party Mix Lollies',
'Brewed Drink Burgundee',
'Double Cream',
'Australian Beef Garlic',
'12YO Double Cask',
'Deli Sea Salt',
'Hard Seltzer Mango',
'Beef And Margaret',
'Green Olives &',
'Dry Ginger Ale',
```

'Rocklea Road',
'Pizza 11" Margherita',
'Barossa Valley Shiraz',
'Apple Cider Bottle',
'Ultra Zero Carb',
'Smooth Classic Hommus',
'Simply Beef BBQ',
'Timber Forks',
'Pink Bubbles 750mL',
'Liqueur 1L',
'Midori Melon Liqueur',
'Quiche Spinach Parmesan',
'Prosecco DOC 18K',
'Blanc De Blancs',
'BBQ Aluminium Large',
'Grissini Rosemary &',
'Pink Gin 1L',
'Graphite Prosecco',
'Soft Drink Orange',
'Marinated Danish Fetta',
'Spritz Bottle 200mL',
'Hot Honey Cheese',
'Taste Collection Buttery',
'Kitchen Carrot &',
'Kitchen Beef Lasagne',
'French Raspberry Tart',
'Apple Lower Sugar',
'15YO Scotch Whisky',
'Bulloak Rose 750mL',
'Prosecco Rose 750mL',
'Tequila Blanco 700mL',
'Hazelnut Dark Chocolate',
'Zero Carb Mid',
'Tasmanian Vodka 700mL',
'Goats Cheese Polenta',
'Basque Cheesecake',
'Pumpkin & Sage',
'Double Cream Camembert',
'Kitchen Grilled Mediterranean',
'Grand Cordon Rose',
'Bertie Cold Pressed',
'Limestone Coast Cabernet',
'Deluxe Red Velvet',
'Lunch Serviettes',
'Tasmanian Bramble Gin',
'French Petits Fours',
"Malt Master's Selection",

'Tradition Brut Vintage',
'Deluxe Vanilla Cupcake',
'Summer Ale Can',
'Crafted Caramelised French',
'Premium Mixers Elderflower',
'Low Sugar Grapefruit',
'Large Patterned Cups',
'Premium Red Wine',
'Kitchen Ratatouille',
'Premium Mixers Light',
'Pale Ale Can',
'The Stag Pinot',
'Pancetta Crumb',
'Kitchen Sesame Prawn',
'Seafood Sauce Creamy',
'Patterned Dinner Plates',
'Pine Lime Mixed',
'NV Yellow Piccolo',
'Wholesome Banana Bread',
'Grapes Seedless White',
'Snack Crackers Mix',
'Fun Mix Variety',
'Original Mini Cookies',
'Bolognese with Chunky',
'Bananas',
'Spring Water 24X600mL',
'Water Multipack 24X350mL',
'Cheese Shredded Tasty',
'Traditional Spreadable Butter',
'Greek Style Natural',
'Weet-Bix Breakfast Cereal',
'Green Kiwifruit',
'Oranges Prepacked',
'Pink Lady Apples',
'Protein Energize Choc',
'Pork & Beef',
'Apple Splash Fruit',
'Australian Still Spring',
'Liquid Breakfast Vanilla',
'Celery Bunch',
'Grain Waves Chips',
'Super Soft Hamburger',
'Liquid Breakfast Choc',
'Tasty Cheese Slices',
'Tri Colour Capsicum',
'Strong & Bitey',
'Bread High Fibre',

'Rindless Pan Sized',
'Red Label Cabernet',
'Original Cheese Block',
'Roses Chocolate Gift',
'Kitchen Sesame Torpedo',
'Fruit Tangle Vodka',
'Large Cups',
'Zero Bottle 330mL',
'Brut Tradition NV',
'Dinner Plates',
'Roasting Pan',
'YB Cuvee NV',
'Corn Relish Dip',
'Premium Mixers Indian',
'Premium Mixers Mediterranean',
"Jumbl'D Cinnamon Scroll",
'Traditional Passionfruit Lemonade',
'Premium Peppercorn Sauce',
'Bulloak Cuvee Blanc',
'Beef Bacon &',
'Signature Rose 750mL',
'Sauce & Dip',
'Patterned Snack Plates',
'Tropical Punch Vodka',
'IPA Can 375mL',
'Patterned Oval Plates',
'Whisky Apple &',
'Portello Vodka Mixed',
'Seafood Sauce',
'Cinnamon Crumble Muffin',
'Lemon Myrtle Tonic',
'Patterned Bowls',
'Snack Mix Variety',
'Raspberries',
'Iceberg Lettuce',
'Fresh Gourmet Tomatoes',
'Dairy Natural Tasty',
'Dairy Stringers Original',
'Greek Yogurt Natural',
'Chickpeas Lightly Salted',
'Dark Choc &',
'Roasted Chickpeas Dark',
'Chocolate Roasted Chickpeas',
'Favva Beans Red',
'Lightly Salted Crunchy',
'Fav-va Beans Lightly',
'Soft Wholemeal Sandwich',

'All Butter Croissants',
'Super Soft Round',
'Peppermint Sugar Free',
'Roasted Seasoned Seaweed',
'Vegemite Scroll',
'Hazelnut Spread With',
'Sultana Bran High',
'Nutri-Grain Protein Breakfast',
'Butter Leaf Blend',
'Hazelnut Chocolate Spread',
'Lamb Strips With',
'Bacon Streaky',
'Greek Style Strawberry',
'Lightly Salted Slightly',
'Pure Honey Squeeze',
'Short Cut Bacon',
'Cup A Soup',
'Milk Chocolate Value',
'Valencia Oranges Medium',
'Leg Shaved Ham',
'Brut NV Champagne',
'Original Corn Chips',
'Reserve Chardonnay Pinot',
'Two Fold Whisky',
'White Chocolate Baking',
'Kitchen Mozzarella Bites',
'Superior Rum 700mL',
'Red Imported Bottle',
'Little Yering Pinot',
'Pinot Gris 750mL',
'Salami Tasting Plate',
'Parmentier Potatoes',
'Roasted Beetroot Hommus',
'Beef Roast With',
'VSOP Cognac 700mL',
'Lindor Strawberries &',
'Pear Cider Bottle',
'Frozen Tiramisu',
'Sauvignon Blanc 750mL',
"Rose d'Anjou 750mL",
'Cheese & Chive',
'Hommus Garnished Zesty',
'Sweet Chilli Pourover',
'Soda & Lime',
'Classic Smooth Hommus',
'Double Jack and',
'Kitchen Broccoli Florets',

'Ginger Beer Soft',
'Frozen Sticky Date',
'Carbon Neutral Beef',
'Bin 128 Shiraz',
'Kitchen Lemon &',
'Japanese Vodka 700mL',
'Ice Double Black',
'Trilogy Sparkling 750mL',
'Traditional Gravy Mix',
'Kitchen Sweet Corn',
'Deli Roast Vegetables',
'Blanco Tequila 750mL',
'Tasmanian Pink Gin',
'Garlic Crackers',
'Brut Vintage 750mL',
'Ready To Roast',
'Brut Rose NV',
'Keyneton Euphonium 750mL',
'Brut Cuvee 750mL',
'Dim Sims',
'Bin 8 Shiraz',
'Frosted Chocolate Brownie',
'Slow Cooked Pulled',
'Apple Sauce',
'Milk Chocolate 8',
'Roast Meat Gravy',
'Oval Plates',
'NV Cava Brut',
'Patterned Timber Cutlery',
'Chorizo Crumb',
'Rainbow Carrots With',
'Free Range Chorizo',
'Raffaello Coconut And',
'Original Water Crackers',
'Prosecco Spumante DOC',
'Artisan Cracker Biscuits',
'Gluten Free Hummous',
"AOC Rose d'Anjou",
'Crisp Dry White',
'French Onion Dip',
'Pear Cider Can',
'BBQ Aluminium Medium',
'Brut NV 750mL',
'18YO Scotch Whisky',
'BBQ Aluminium Small',
'Sparkling Margarita 330mL',
'Brewed Drink Pineapple',

'Church Block Red',
'Garden Spritz 750mL',
'16YO Single Malt',
'Newtowner Can 375mL',
'Original BBQ Briquettes',
'Lindor Cheesecake Chocolate',
'Crackers Checkers',
'Paper Dinner Plates',
'Dutch Gouda',
'Pralines',
'Premium Light Bottle',
'Terre de Providence',
'Sparkling 750mL',
'Squeezy Sauce BBQ',
'Bitter Lemon Soft',
'Barossa Cabernet Sauvignon',
'Organic Pinot Grigio',
'Danish Salami',
'Lindor Milk Salted',
'Zero Sugar Dry',
'Zero Sugar Pink',
'Belgian Chocolate Sea',
'Dark Chocolate Coated',
'Watermelon & Lemon',
'Eden Valley Chardonnay',
'Quiche Leek Bacon',
'Saffron Corn &',
'Cracked Pepper Water',
'Kitchen Carbonara Bites',
'Cider Rose Bottle',
'Deluxe Chocolate Mud',
'Mild Chorizo Salami',
'Variety Pack Diamond',
'Hay Premium Molten',
'Deluxe Orange Poppyseed',
'Premium Mushroom Sauce',
'Refreshr Alcoholic Lemonade',
'Truffled Potato Croquettes',
'Tartare Sauce',
'Aranciata Rossa Drink',
'Fennel Sausage Roll',
'Roasting Pan with',
'Mint Thins',
'Little Kings Grenache',
'Deluxe Raspberry &',
'Small Patterned Cups',
'Timber Knives',

```
        "Jumbl'D Cookie Loaded",
        'Chicken Mushroom &',
        'Deluxe Caramel Mud',
        'Patterned Lunch Serviettes',
        'Tropical Can 330mL',
        'Tempranillo 750mL',
        'Spanish Cava 750mL',
        'Strawberry & Lime',
        'Turkish Delight Vodka',
        'Bowls',
        'Chocolate 16 Treat',
        'Crumpet Rounds Original',
        'Travel Pack',
        'Frozen Fish Fingers',
        'Bananas Mini Pack',
        'Pot Set Greek',
        'High Fibre Wholemeal',
        'Yopro Vanilla Yoghurt',
        'Favva Beans Sea',
        'Kids Fav-va Beans',
        'Roasted Turkey Breast',
        'Mayonnaise',
        'William Pears 1kg',
        'Tempters Tuna Chunks',
        'Kitchen Chicken Caesar',
        'Multipacks Variety 10',
        'Dairy Original Swiss',
        'Vegan Tasty Cheese',
        'Beef & Pork',
        'Super Soft Hot',
        'Kitchen Ancient Grain',
        'Finger Lime Soda',
        'Sunset Tonic',
        'Coloured Lunch Serviettes',
        …]
```

[40]: `unique_brand_names`

```
[40]: ['Sensodyne Repair &',
       'Tassal Tasmanian Smoked',
       'Fresh Grapes Seedless',
       'Western Star Traditional',
       'Harvey Fresh Vanilla',
       'Pauls Custard Vanilla',
       'Tablelands Dairy Blend',
       'Bega Dairy Natural',
       'Mersey Valley Club',
```

'Tasmanian Heritage Triple',
'Happy Snack Company',
'La Espanola Extra',
'Slice X Chupa',
"Abe's Bagels Everything",
'Vegemite Spread |',
'Gravox Traditional Gravy',
'MasterFoods All Purpose',
'Carmans Chia Bowls',
'Liberate Gluten Free',
'John West Tempters',
'Powerade Berry Ice',
'Coca-Cola Zero Sugar',
'Gym Bod Caramel',
'Streets Golden Gaytime',
'Weis Ice Cream',
'Easter Hunt Basket',
'Canesten Antibacterial and',
'Oral B IO2',
'Libra Goodnights Extra',
'Quick Screen Travel',
'Lady Jayne Detangling',
'Gaviscon Dual Action',
'Swisspers Cotton Tips',
'Nair Shower Power',
'Babybel Mini Cheese',
'Sanitarium Weet-Bix Breakfast',
"Campbell's Real Stock",
'Aptamil Gold+ Lactose',
'Health Lab Charlottes',
'Kinder Chocolate 16',
'The Natural Confectionery',
'Sakata Rice Cracker',
'Uncle Tobys Chewy',
'Golden Circle Tropical',
'Twinings English Breakfast',
'Red Bull Energy',
'Magnum Ice Cream',
"Ben & Jerry's",
'Swisse Ub Magnesium',
'Glam By Manicare',
'Tamar Valley The',
'Sealord Gluten Free',
'Heaps Normal Half',
'Tegel Take Outs',
'Windex Glass Cleaner',
"Bosisto's Multipurpose Cleaner",

'Finish Ultimate Plus',
'Earth Choice Ultra',
'Nilodor Concentrated Deodoriser',
'Radiant Laundry Capsules',
'Epzen Calm Cleanse',
'Vaseline Gluta-Hya Dewy',
'Murine Relief Eye',
'Great Northern Super',
'Carlton Dry Bottle',
'Gage Roads Single',
'Grant Burge 5th',
'Oyster Bay Pinot',
'Heineken Bottle 330mL',
'XXXX Gold Block',
'Sprite Lemonade Soft',
'Blackmores Multivitamins 50+',
'Don English Style',
'Good Farmer Free',
'Olympian Easter Tsoureki',
'Jarlsberg Cheese Wedge',
'Moro Primero Extra',
'Lindt Easter Gold',
'Soda Stream Gas',
'Connoisseur Classic Vanlla',
'Destination Italy Tiramisu',
'McCain Frozen Baby',
'Haagen-Dazs Ice Cream',
"Kb's Marinara Mix",
'Morning Fresh Soft',
'Bref Power Active',
'Quilton 3 Ply',
'Tatachilla White Admiral',
'Victoria Bitter Block',
'Fanta Orange Soft',
'Moccona Classic Medium',
'Capilano 100% Australian',
'Arnotts Tim Tam',
'Riverview Farms Pork',
'Bertolli Extra Virgin',
'Indomie Fried Mi',
'Gatorade Fast Twitch',
'Maltesers Milk Chocolate',
'Kettle Chilli Potato',
'MT Barker Free',
'Primo Streaky Bacon',
"Carman's Oat Cups",
'Cocobella Coconut Water',

'Peters Maxibon Vanilla',
'Fropro Protein Beef',
'Duck Fresh Discs',
'Life Space Gummies',
'Corona Bottle 355mL',
'Smirnoff Red Vodka',
'Dr. Oetker Frozen',
'Ajax Spray N',
'Kleenex Complete Clean',
'Ostelin Vitamin D',
'Jack Daniels Cola',
'Handee Ultra Paper',
'Hardys Brave New',
'Cleanskin Vodka 700mL',
'Asahi Super Dry',
'UP&GO Liquid Breakfast',
'Nescafe Blend 43',
'Cheer Plant Based',
'Keens Curry Powder',
'Chocolate Easter Bunny',
'Selections Easter Chocolate',
'Dairy Milk Easter',
'Grand Rocher Milk',
'Hot Cross Buns',
'Surprise Chocolate Easter',
'Bubbly Mint Chocolate',
'Milk Chocolate Solid',
'Crunchie Hollow Chocolate',
'Cherry Ripe Hollow',
'Collection Boxed Egg',
'Mini Eggs Milk',
'Kitkat Chocolate Easter',
'Bueno Easter Eggs',
'Marvellous Creations Chocolate',
'Milkybar White Choc',
'Mint Slice Egg',
'Hazelnut Brittle White',
'Creme Egg 6-Pack',
'Peanut Chocolate Speckled',
'Bunny Ears Assorted',
'Mug',
'Honey Soy Chicken',
'Zero Sugar Soft',
'Classic Soft Drink',
'Sweet Gold Potatoes',
'Max No Sugar',
'Potato Chips Sea',

'Lamb Boneless Shoulder',
'Tennessee Whiskey 1L',
'Roses Boxed Chocolates',
'Natural Walnuts',
'Sea Salt &',
'Yellow Label Brut',
'Kabana From The',
'Virginian Ham From',
'Simply Thin BBQ',
'UP Rum 1L',
'Cordon Rouge NV',
'Beef Garlic &',
'12YO Single Malt',
'Ultra Low Carb',
'Thawed Basa Fillets',
'Draught Bottle 375mL',
'Whisky 1L',
'Kitchen Caesar Salad',
'Strawberries',
'Lindor Easter Assorted',
'Boneless Pork Leg',
'Humpty Dumpty Easter',
'Caramello Easter Chocolate',
'Oreo Chocolate Easter',
'Dream Chocolate Easter',
'Crispy Chocolate Speckled',
'Rocklea Road Bunny',
'Australia Puddles The',
'Hunt Basket',
'Dry Bottle 330mL',
'Perino Red Grape',
'Gold Block Can',
'No Added Hormone',
'Marlborough Sauvignon Blanc',
'Crinkle Cut Original',
'Super Dry Bottle',
'Irish Whiskey 700mL',
'Barossa Shiraz 750mL',
'Bottle 330mL',
'Rocher Chocolate Gift',
'Scotch Whisky 700mL',
'4 Leaf Salad',
'Vodka 700mL',
'Australian Roasted &',
'Can 375mL',
'Raffaello 26 Pack',
'Nastro Azzurro 5%',

'White Hungarian Salami',
'Prosecco NV 750mL',
'Pinot Noir Chardonnay',
'Gin 1L',
'Estate Shiraz 750mL',
'Italian Style Pork',
'Semi Dried Tomatoes',
'Clinkers Chocolate Giant',
'Gluten Free Hot',
'Caramilk Easter Chocolate',
'Nougat Easter Egg',
'Raspberry&coconut Slice Egg',
'Kit Kat Mini',
'Britecrawlers & Zappo',
'Deluxe Jam &',
'Blueberries Prepacked',
'Baby Cucumbers',
'Carrots Prepacked',
'Kent Pumpkin Cut',
'RSPCA Approved Free',
'Pistachios Roasted &',
'NV Sparkling Brut',
'Pork Belly Roast',
'Cola Can 375mL',
'Salada Original Crispbread',
'Triple Smoked Leg',
'Brewed Drink Ginger',
'Original Can 375mL',
'Roast Salted Macadamias',
'Extra Dry Bottle',
'Butcher Lamb Leg',
'Diet Coke Soft',
'Orange Soft Drink',
'Whole Chicken',
'Pub Style Extra',
'Gentleman Jack Tennessee',
'Black Label 12YO',
'Bourbon & Cola',
'Shiraz 750ml',
'Mixed Nuts Dry',
'Mid Strength Bottle',
'Pacific Ale Bottle',
'Hommus Dip',
'Traditional Hommus Dip',
'Jumbo Prawn Cutlets.',
'Blue Label Scotch',
'Cream Chocolate Bag',

```
'Aperitivo 700mL',
'Faith Shiraz 750mL',
'Japanese Gin 700mL',
'Rare Dry Gin',
'XPA Can 375mL',
'Made Easy Slow',
'Cabernet Sauvignon 750mL',
'Cruiser Mixed 275mL',
'Slow Cooked Lamb',
'Mild Hungarian Salami',
'Reserve Sparkling Prosecco',
'Semillon Sauvignon Blanc',
'Dip Hommus Traditional',
'Club Original Cheddar',
'Dollop Cream',
'Baked Pretzels Wheat',
'Selection Pack Sausages',
'Salad Baby Leaf',
'Deli Golden Ale',
'Chicken Mozzarella &',
'Turkish Delight Easter',
'Speckled Easter Eggs',
'100S & 1000S',
'Hollow Chocolate Bunny',
'Smarties Chocolate Easter',
'After Eight Chocolate',
'Egg Jumbo Chalks',
'Savoy Crackers Original',
'Favourites Boxed Chocolate',
'Green Beans Prepacked',
'Corn Chips Supreme',
'Graze Grassfed Beef',
'Celery Sticks Prepacked',
'Premium Bottle 375mL',
'Moscato 750mL',
'Sausages Beef Garlic',
'Tzatziki Dip',
'Lemonade Soft Drink',
'Energised Soft Drink',
'Ginger Beer Can',
'Cheese Brie Wrapped',
'Sugar Free Lemon',
'Sparkling Pinot Noir',
'Squeezy Sauce Tomato',
'700mL',
'Prosciutto',
'& No Sugar',
```

```
'Agrum Blood Orange',
'Lemon Lime Zero',
'Cerveza Bottle 355mL',
'Chardonnay Pinot Noir',
'AOC Luberon Rose',
"Max's Shiraz 750mL",
"'The Siding' Cabernet",
'Roasted And Salted',
'Pale Ale Bottle',
'Pink Gin RTD',
'Thick & Juicy',
'New York Style',
'Double Jack &',
'Cooked Prawns With',
'Mediterranean Olives',
'Liqueur 700mL',
'Pasito Passionfruit Soft',
'Antipasto Mix',
'Apple Cider Can',
'Ichiban Bottle 330mL',
'Snacking Carrots',
'Lager Can 375mL',
'Bubbles 750mL',
'Pepperoni Hot Salami',
'Spiced Rum 1L',
'Peanuts Unsalted',
'150 Lashes Pale',
'Limoncello Spritz 750mL',
'Giant Pitted Kalamata',
'Whiskey 700mL',
'Salt & Pepper',
'Macaron Selection',
'Mixers Tonic Water',
'French Style Cheesecake',
'Cruskits Original Crispbread',
'Brut Imperial NV',
'Riesling 750mL',
'Coffee Liqueur 700mL',
'Kentucky Straight Bourbon',
'Kalamata Olive Dip',
'Creaming Soda Soft',
'Hard Cider Can',
'Artisan Crackers Biscuits',
'Dark Chocolate Baking',
'Wafers Chocolate',
'Salted Mixed Nuts',
'Pralines Noir Dark',
```

'Snakes Lollies Party',
'Praline Master Collection',
'Non Vintage Sparkling',
'YB Prosecco 750mL',
'Crafted Guacamole Dip',
'Especial Tequila 700mL',
'Quiche Lorraine',
'Silver Tequila 700mL',
'London Dry Gin',
'Indian Tonic Water',
'Pretzel Twists',
'Sacred Hill Semillon',
'Eazy Hazy Can',
'Avocado Dip',
'Timber Spoons',
'BBQ Aluminium XL',
'Chinotto Drink Cans',
'Refreshing Ale Bottle',
'Bloody Shiraz Gin',
'Swiss Milk Chocolate',
'South Coast Pale',
'Soft Centres Gift',
'Cracked Black Pepper',
'10YO Single Malt',
'Yarra Valley Pinot',
'Rose Brut NV',
'Lime Soda Water',
'Cremant Blanc de',
'Blood Orange &',
'Rhubarb & Ginger',
'Party Mix Lollies',
'Pizza 11" Margherita',
'Smooth Classic Hommus',
'Midori Melon Liqueur',
'Blanc De Blancs',
'Grissini Rosemary &',
'Graphite Prosecco',
'Marinated Danish Fetta',
'Spritz Bottle 200mL',
'Taste Collection Buttery',
'15YO Scotch Whisky',
'Bulloak Rose 750mL',
'Tequila Blanco 700mL',
'Goats Cheese Polenta',
'Basque Cheesecake',
'Pumpkin & Sage',
'Bertie Cold Pressed',

```
'Limestone Coast Cabernet',
'Lunch Serviettes',
"Malt Master's Selection",
'Tradition Brut Vintage',
'Summer Ale Can',
'Low Sugar Grapefruit',
'Large Patterned Cups',
'Pancetta Crumb',
'Seafood Sauce Creamy',
'Patterned Dinner Plates',
'Pine Lime Mixed',
'Wholesome Banana Bread',
'Grapes Seedless White',
'Snack Crackers Mix',
'Fun Mix Variety',
'Bolognese with Chunky',
'Bananas',
'Spring Water 24X600mL',
'Water Multipack 24X350mL',
'Greek Style Natural',
'Weet-Bix Breakfast Cereal',
'Oranges Prepacked',
'Protein Energize Choc',
'Liquid Breakfast Vanilla',
'Grain Waves Chips',
'Tasty Cheese Slices',
'Tri Colour Capsicum',
'Strong & Bitey',
'Bread High Fibre',
'Rindless Pan Sized',
'Fruit Tangle Vodka',
'Dinner Plates',
'Roasting Pan',
"Jumbl'D Cinnamon Scroll",
'Signature Rose 750mL',
'Sauce & Dip',
'Tropical Punch Vodka',
'IPA Can 375mL',
'Portello Vodka Mixed',
'Cinnamon Crumble Muffin',
'Raspberries',
'Iceberg Lettuce',
'Chickpeas Lightly Salted',
'Favva Beans Red',
'Lightly Salted Crunchy',
'Fav-va Beans Lightly',
'All Butter Croissants',
```

'Peppermint Sugar Free',
'Sultana Bran High',
'Nutri-Grain Protein Breakfast',
'Butter Leaf Blend',
'Bacon Streaky',
'Pure Honey Squeeze',
'Short Cut Bacon',
'Cup A Soup',
'Valencia Oranges Medium',
'Leg Shaved Ham',
'Two Fold Whisky',
'Superior Rum 700mL',
'Little Yering Pinot',
'Salami Tasting Plate',
'Parmentier Potatoes',
'VSOP Cognac 700mL',
'Pear Cider Bottle',
'Frozen Tiramisu',
'Sauvignon Blanc 750mL',
'Carbon Neutral Beef',
'Bin 128 Shiraz',
'Ice Double Black',
'Trilogy Sparkling 750mL',
'Blanco Tequila 750mL',
'Garlic Crackers',
'Ready To Roast',
'Keyneton Euphonium 750mL',
'Dim Sims',
'Frosted Chocolate Brownie',
'Oval Plates',
'Chorizo Crumb',
'Rainbow Carrots With',
'Free Range Chorizo',
'Crisp Dry White',
'18YO Scotch Whisky',
'Church Block Red',
'Garden Spritz 750mL',
'16YO Single Malt',
'Newtowner Can 375mL',
'Crackers Checkers',
'Paper Dinner Plates',
'Dutch Gouda',
'Terre de Providence',
'Bitter Lemon Soft',
'Organic Pinot Grigio',
'Danish Salami',
'Belgian Chocolate Sea',

'Watermelon & Lemon',
'Eden Valley Chardonnay',
'Saffron Corn &',
'Cider Rose Bottle',
'Variety Pack Diamond',
'Hay Premium Molten',
'Refreshr Alcoholic Lemonade',
'Truffled Potato Croquettes',
'Tartare Sauce',
'Aranciata Rossa Drink',
'Fennel Sausage Roll',
'Small Patterned Cups',
'Tempranillo 750mL',
'Spanish Cava 750mL',
'Strawberry & Lime',
'Bowls',
'Crumpet Rounds Original',
'Travel Pack',
'Pot Set Greek',
'High Fibre Wholemeal',
'Yopro Vanilla Yoghurt',
'Kids Fav-va Beans',
'Mayonnaise',
'William Pears 1kg',
'Tempters Tuna Chunks',
'Multipacks Variety 10',
'Vegan Tasty Cheese',
'Finger Lime Soda',
'Sunset Tonic',
'Coloured Lunch Serviettes',
'Crushed Lemon Can',
'Cacio E Pepe',
'Passionfruit Vodka Mixed',
'Sarsaparilla Vodka Mixed',
'Greenhouse Truss Tomatoes',
'Medium Navel Oranges',
'Seedless Watermelon Cut',
'Blackberries Prepacked',
'Nutri Grain Protein',
'Dairylea Cheddar Block',
'Shredded Iceberg Lettuce',
'Tuna in Oil',
'Chewy Choc Chip',
'Apples Mini 1kg',
'Just Right Breakfast',
"I'M Perfect Apples",
'Le Snak Tasty',

```
'Microwave Jasmine Rice',
'Champagne Leg Ham',
'Light Rye Bread',
'Fit High Protein',
'Sesame Hamburger Rolls',
'Spearmint Sugar Free',
'Tiger Rolls',
'Froot Loops Breakfast',
'Salmon Tempters Springwater',
'Long Rindless Hickory',
'Hawaiian Pizza Roll',
'Durum Wheat Pasta',
'Tiny Teddy Multipack',
'Banana Bread Slices',
'Blueberry Muffins 4',
'Pancakes',
'Yoghurt Pouch Vanilla',
'B-Ready Biscuit Multipack',
'Spaghetti No.113 Italian',
'Oats Rolled',
'Streaky Rindless Bacon',
'Vanilla Yoghurt 6x160g',
'Multipack Best Ever',
'Stainless Steel Doubled',
'Raspberry Jam',
'Mango Slices',
'Pikelets 8 Pack',
'Pineapple Juice Sweetened',
'Trivelle Spirals No',
'Minis Sour Cream',
'A4 Copy Paper',
'Pesto Basilico',
'Fusilli No.114 Italian',
'Shitake Mushroom Crisps',
'Squeezie Yoghurt Pouch',
'Spice Mix For',
'Almond Hazelnut &',
'Multigrain Rolls',
'Angel Slice 6',
'Tronky Biscuit Bars',
'Creamy Tomato &',
'Convenience Meals Butter',
'Tetra Drink Golden',
'Beer Battered Steakhouse',
'Passata Cooking Sauce',
'Pulp Free No',
'Sunshine Punch Fruit',
```

```
'Wholemeal & Grain',
'Manuka Honey Leg',
'Margherita Pizza Rolls',
'Coco Pops Chocolatey',
'Damper Inspired Rolls',
'English Style Sliced',
'Imperial Mandarins',
'Lamington Fingers 18',
'Petit Miam Squeezie',
'Natures Kitchen Basil',
'Wrap Foil',
'Moo Flavoured Chocolate',
'Family Pizza Pepperoni',
'Resealable Sandwich Bags',
'Cheesy Nachos Tortilla',
'Tomato Passata',
'Sugo Passata Basil',
'Chips Deli Crisps',
'Seriously Good Garlic',
'Special Treats Chocolate',
'Oat Slices Belgian',
'Crisps Original Blend',
'Chilli Tuna',
'Chunk Style Tuna',
'Sliced Smoked Salmon',
'Dipped with White',
'Ristorante Mozzarella Pizza',
'Crunchy Peanut Butter',
'Grated Parmesan Cheese',
'Dried Apricot Nut',
'Rice Bubbles Puffed',
'Popcorn Sea Salt',
'Hearty Chicken Soup',
'Whiting Bites Battered',
'Spread',
'Coconut Yoghurt Vanilla',
'Stringers Original Cheese',
'Singles Tzatziki Classic',
'Facial Sunscreen Spray',
'Deo Roll On',
'Top 20 Variety',
'Burrito Tortilla Wraps',
'Lip Balm Original',
'Clickzip Sandwich Bags',
'95% Sugar Free',
'Clusters Berry Bliss',
'Buttermilk',
```

'Chokito Mueli Bars',
'Cling Wrap',
'Pitted Dates',
'UHT Skim Multipack',
'Cheesey Peelers 8x20g',
'Assorted Sensitive Strips',
'Really Good Crunchy',
'Silverside',
'Pimple Patch Party',
'Regular Butter Blend',
'Wafer Crackers',
'Madeira Cake',
'Meals Satay Chicken',
'Crusty Round Rolls',
'Cranberry, Blueberry &',
'Clix Crackers',
'Grass-Fed Beef Meatballs',
'Shapes Cracker Biscuits',
'BAGEL SESAME 4',
'Genovese Pesto',
'Superfood Mix',
'Lentil Soup',
'Girl-Teen Regular Pads',
'Colby Cheese Slices',
'Penne Rigati Pasta',
'Pasta Penne Rigati',
'Biscuit Crackers Multipack',
'Alfoil Traditional Strength',
'Reduced Sugar Banana',
'Margarine Spread Light',
'Vita-Weat Original Crispbread',
'Mexican Style Brown',
'Roll Ups Apple',
'Freeze Dried Strawberry',
'Nothing But Kids',
'Nut Bar Choc',
'Tempura Chicken Pops',
'To Go Ribbon',
'Veggie Straws Salt',
'2 Minute Instant',
'Plant Protein Bar',
'Muesli Bars Strawberry',
'Southern Blue Whiting',
'1.2M USB-A To',
'Pea Mix Multipack',
'Vintage Cheese Block',
'Bluey Chocolate Flavoured',

'Seaweed Sea Salt',
'LCMs Split Stix',
'Asia Hokkien Ready',
'Adventurer Stainless Steel',
'Taco Seasoning Mix',
'Brown Rice Chips',
'Ultimate Gluten Free',
'Plus Rectangle Containers',
'Caramelised Biscuit Spread',
'Cookies & Cream',
'Twinzip Resealable Large',
'Raisins',
'Truffle & Mushroom',
'Rye Wraps 8',
'Wall Charger Dual',
'On The Go',
'Probiotics Banana Kids',
'SPF 50+ Sunscreen',
'Stackers Spicy Chorizo',
'MSC Tuna Portions',
'Jam Lamingtons 6',
'Fruity Bites Wildberry',
'Cracker Squares',
'50% Less Sugar',
'Bakery Cake Madeira',
'Daily Moisturiser Sunscreen',
'Escape True Wireless',
'Date & Walnut',
'Blu Tack',
'Ricotta Spinach Agnolotti',
'Arrabbiata Pasta Sauce',
'Choc Chip Banana',
'Milo Energy Dairy',
'Monterey Jack Cheese',
'Healthy Touch Liquid',
'Panda Brand Oyster',
'Napoletana with Chunky',
'Carrot Cake',
'Tender Squid Chips',
'Unicorn Fancies 8',
'Shanghai Stir Fry',
'Share Topped Mini',
'25% Salt Reduced',
'Click Zip Sandwich',
'Food Flask',
'Wired Earphones With',
'Apricots',

'Petroleum Jelly Original',
'Vine Ripened Natural',
'Flip Sipper Bottle',
'Wild Caught New',
'Bites Burger',
'Stacker Chicken Cheese',
'Microsafe Easy Heat',
"Mum's Sause Bolognese",
'Sport Ultrathin Pads',
'Instant Noodles Cup',
'Dressings Creamy Caesar',
'Boostup 30w Usb',
'Porridge Sachets Super',
'K-Time Baked Twists',
'Combination Word 40mm',
'Microwavable Brown Rice',
'Insulated Double Wall',
'XL Double Wall',
'42w (Pd30w Pps',
'Infused Tuna Basil',
'Milky Max Dairy',
'Hand Balm',
'Boostcharge USB-A To',
'Fancy Large Dried',
'Leak Proof Snack',
'Car Mobile Phone',
'Diamond Padlock Long',
'Wireless Mouse',
'2M USB-C To',
'Hydrate Square Bottle',
'Popped Rice Cakes',
'Tritan Drink Bottle',
'Lite Tuna In',
'School Glue',
'Clinical Protection Antiperspirant',
'Snap & Seal',
'Sticky tape 18mmx25m',
'Napoli Style Pasta',
"Women's Roll On",
'Be Gentle Anti-Bacterial',
'Uptown Dual Sip',
'Protein+ Bowl Caps',
'Groovy Grape Roll',
'Microdart Pimple Patches',
'Lint Roller',
'Sweetened Dried Blueberries',
'Smoothie Bars Berrylicious',

'Meal Base Pouch',
'Crispbread Pumpkin',
'Jaw Clip Extra',
'V8 Breakfast Juice',
'Antiperspirant Original Deodorant',
'Illusions Body Spray',
'Caramel Cake Bars',
'Lightning To 3.5Mm',
'Fluo XL Assorted',
'Soothing After Sun',
'Boss Highlighter Assorted',
'Design Your Own',
'Goji Berries',
'Snackright Multipack Cracker',
'Comfy Natural Ponytailers',
'Elastics All Hair',
'Purple Case With',
'Organics Seaweed Multipack',
'Ham And Chicken',
'Marmite Yeast Extract',
'Sunscreen SPF50+ Ultra',
'Klip It Plus',
'Trios Twiggy Cheese',
'Indulge Wireless Bluetooth',
'Back To Basics',
'Puffs BBQ',
'Strips Tropicana',
'Tween Period Brief',
'Teen Roll On',
'Everyday Sunscreen SPF50',
'Rosemary Crackers',
'Magnetic Wireless Charging',
'Crumbed Chicken Kyiv',
'Spreadable Sweet Mustard',
'Cleansing Facial Wipes',
'Pad Thai Style',
'Papaw Ointment Lip',
'Hydro Chugger',
'Sandwich Bags Resealable',
'Diced Apricots',
'4mm No Damage',
'Westloop Auto Seal',
'Sun Protect &',
'Peckables Rice Snacks',
'Festival Summer Edition',
'Sanitiser Fig &',
'XXL Double Wall',

'Cabanossi Jatz &',
'Gel Ice',
'Metallic Markers',
'Active Sipper 1.7L',
'Pickle-licious Pickled Flavoured',
'Sip Straw Junior',
'USB 2.0 SM02',
'5 Formaggi With',
'570 Notebook A5',
'Cocoa Butter Lip',
'Bluetooth Portable Speaker',
'SM02 USB 2.0',
'Tellfresh Cake Storer',
'Aussie Made Diced',
'Fish Crispy Goujons',
'Aqua Case With',
'USB-C To 3.55MM',
'Licensed Riverside Bottle',
'Silicone Straight Straws',
'Mens Logo Low',
'Markers Fine Tip',
'97% Fat Free',
'Table Of Plenty',
'Antifog Lens Wipes',
'S-Gel Pens Assorted',
'Laundry Marker',
'Glu Stik 21g',
'Atlantis Pens Black',
'Womens Socks Quarter',
'Period Proof Girl-Teen',
'Letra Tag Labeller',
'Aloe Vera Pure',
'Berry Frugos',
'Trek Wireless Sport',
'Bendy Stainless Steel',
'Go Dressing Tubs',
'Keralan Coconut Curry',
'Bamboo Handle Lint',
'Correction Pen White',
'Picnic Reusable Cutlery',
'Face Day Wear',
'Men Advanced Antiperspirant',
'Fineliner Pens',
'Cristal Fashion Pens',
'Sultanas With Cheddar',
'Nail Grooming Set',
'PVC Free Eraser',

'A5 Spiral Notebook',
'Oaty & Crispy',
'Barrette Snap Clips',
'Intensity Colouring Felt',
'Woodcase Hb Pencils',
'Connector Pens',
'Microfibre Cloth 20x20cm',
'Plastic Clear Tape',
'Quad Bands Tritan',
'Kelso Licensed Tumbler',
'Deep Sleep',
'Chow Mein Mince',
'Vitamin C 500mg',
'Curried Sausages Recipe',
'Probiotic + Daily',
'Odourless Fish Oil',
'Napisan Oxi Action',
'Sensitivity & Enamel',
'Crunch Milkshake Flavour',
'Folate 500mg Folic',
'Lash Sensational Sky',
'Flat Iron Sharing',
'Complete Care Sensitive',
'By Laurent Wholemeal',
'Multivitamin For Women',
'Aluminium Foil Wrap',
'Fueld Creamy Chicken',
'Lutein Defence Eye',
'Repair & Protect',
'Preen Everyday Stain',
'Revitalift Filler Eye',
'Probiotics+ Daily Health',
'Age Perfect Gold',
'Aged Red Leicester',
'Bio C 1000mg',
'Island Rose &',
'Falsies Surreal Washable',
'Colossal Washable Mascara',
'B5 Perforated Notebook',
'Clear Zinc Kids',
'Colourhide Notebook A4',
'Silence Pro Noise',
'Ponytailers Girl Soft',
'Polyblands Girl Pastel',
'Snackbox Assorted',
'Bento Box With',
'Whiteboard Markers Fine',

'Drumstix Crackers Multipack',
'Spicy Capsicum Dip',
'Executive B Stress',
'Sensation Cheesy Broccoli',
'Non-Stick Baking Paper',
'Multivitamins 50+',
'Mega B Complex',
'Excellence 4 Dark',
'Mushroom & Cheese',
'Glucosamine Sulfate 1500mg',
'Elvive Glycolic Gloss',
'Multi Purpose Trigger',
'Lavender Disinfectant Liquid',
'Macu Vision Eye',
'Granola Grain Free',
'Delicate Korma Curry',
'Blackcurrant And Chia',
'Colouring Felt Pen',
'SDMI Card 32GB',
'Harmony Double Wall',
'Tape Plastic',
'Inkjoy 100 Assorted',
'Omega Triple Concentrated',
'Confectionery Very Berry',
'Splits Multi Flavoured',
'Juice Orange',
'200 Fineline Black',
'Replacement Straw',
'1st Cutlery Set',
'Gaming Headset',
'Magic Eraser Spong',
'Personalised Bottle Set',
'Thai Green Curry',
'Dubbin Neutral Leather',
'Gelocity Original Pens',
'Wreless Mouse Keyboard',
'Shepherds Pie Recipe',
'Nails Hair &',
'Rapid Relief Sensitive',
'CoQ10 150mg Heart',
'Middle Bacon',
'Sky High Mascara',
'Greener Baking Paper',
'Youjuice Blue Tropics',
'Drawtight Super Strong',
'Paris Age Perfect',
'Total Calcium Magnesium',

'Preference Brasilia Permanent',
'Clean Chicken With',
'Bake Non Stick',
'Evening Primrose Oil',
'Ashwagandha',
'Pregnancy + Breastfeeding',
'Caribbean Vetiver &',
'Eucalyptus & Mediterranean',
'Filled Gnocchi Mozz',
'Macaroni Cheese Pasta',
'Cocktail Frankfurts',
'Disinfectant Wipes Lemon',
'Eyeliner Kajal Black',
'UK Peppercorn Cheddar',
'Stock Pot Beef',
'Seasonal Edition Green',
'Whitening Denture Cleaner',
'Jatz & Extra',
'Pomegranate Floor Disinfectant',
'Eraser Concealer Ivory',
'Eye Shadow Palette',
'4-in-1 Instant Perfector',
'Master Shape Eye',
'Superstay Lumi Matte',
'Jasmine & Sri',
'Oil Italian Style',
'Color Sensational Made',
'Mozzarella Cheese Block',
'Adhesive Cream Flavour',
'Superb Chicken Stock',
'Limited Edition Triple',
'Wrinkle Expert Anti-Wrinkle',
'Alfredo Pasta &',
'Witch Hazel Unscented',
'Conceive Well Tab',
'Wensleydale With Cranberry',
'Stone Baked Garlic',
'Soy & Ginger',
'Oriental Fried Rice',
'Sensations Harvest Pumpkin',
'Essential Mist Peace',
'Build A Brow',
'Powder Natural Beige',
'True Match Serum',
'Duo Cabanossi &',
'Angus Beef Ragu',
'Dulche De Leche',

```
        'Sunkisser Blush Sunset',
        'Eyestudio Colortattoo Inspired',
        'Aerosol- Free Automatic',
        'Four Cheeses Pasta',
        …]
```

[26]:
```python
new_brand_names = ['Sensodyne','Tassal','Western Star', 'Harvey',
↪'Pauls','Tablelands', 'Bega', 'Mersey Valley Club','Happy Snack Company',
                   'La Espanola', 'Slice X Chupa',
↪"Abe's",'Vegemite','Gravox','MasterFoods', 'Carmans', 'Liberate','John West',
                   'Powerade', 'Coca-Cola', 'Gym Bod','Paddle Pop', 'Weis',
↪'Easter Hunt', 'Canesten', 'Oral B', 'Libra', 'Quick Screen',
                   'Lady Jayne', 'Gaviscon', 'Quick Flick', 'Swisspers', 'Nair',
↪'Babybel', 'Weet-Bix', "Campbell's", 'Swisspers', 'Aptamil',
                   'Health Lab', 'Kinder', 'The Natural Confectionery',
↪'Sakata', 'Uncle Tobys', 'Golden Circle', 'Twinings','Red Bull',
                   'Magnum', "Ben & Jerry's", 'Easter Bunny Pom', 'Swisse',
↪'Glam By Manicare', 'Tamar Valley', 'Sealord', 'Heaps', 'Windex',
                   "Bosisto's", 'Finish Ultimate', 'Earth Choice', 'Nilodor',
↪'Radiant', "Bosisto's", 'Epzen', 'Vaseline', 'Murine', 'Great Northern',
                   'Carlton', 'Gage Roads', 'Grant Burge', 'Oyster Bay Pinot',
↪'Heineken', 'XXXX Gold', 'Sprite', 'Blackmores', 'Don', 'Good Farmer',
                   'Olympian', 'Jarlsberg', 'Moro Primero', 'Lindt', 'Soda
↪Stream', 'Connoisseur', 'Destination Italy Tiramisu', 'McCain',
↪'Haagen-Dazs',
                   "Kb's", 'Morning Fresh', 'Bref', 'Quilton']
```

[26]: pandas.core.series.Series

[38]:
```python
def get_unique_by_first_word(brand_list):
    seen_first_words = set()
    unique_brands = []

    for brand in brand_list:
        first_word = brand.split()[0].lower()  # case-insensitive comparison
        if first_word not in seen_first_words:
            seen_first_words.add(first_word)
            unique_brands.append(brand)

    return unique_brands
unique_brand_names = get_unique_by_first_word(new_brands)
```

[45]:
```python
len(unique_brand_names)
```

[45]: 3860

[48]:
```python
merged_df['item_name'][77688:]
```

```
[48]: 77688              Veggie Tubes Sweetcorn & Carrot  12m+
      77689                          100% Pure Cotton Ripple
      77690      Pumpkin Potato & Beef Baby Food Pouch 6+ Months
      77691    Lamb Vegetables & Polenta Baby Food Pouch 6+ M…
      77692                      Beef With Vegetables 6+ Months
      Name: item_name, dtype: object
```

```
[47]: merged_df['item_name'].duplicated()
```

```
[47]: 0        False
      1        False
      2        False
      3        False
      4        False
                …
      77688     True
      77689     True
      77690     True
      77691     True
      77692     True
      Name: item_name, Length: 77693, dtype: bool
```

```
[49]: updated_brand_list = actual_brand_names + unique_brand_names
```

```
[50]: len(updated_brand_list)
```

```
[50]: 3912
```

```
[53]: merged_df['brand'] = merged_df['item_name'].apply(lambda x: match_brand(x,␣
      ↪updated_brand_list))
```

```
[55]: merged_df.head(25)
```

```
[55]:                          _id  product_code        category  \
      0    67edcf33cb2258cf4430592a       8371390          Easter
      1    67edcf33cb2258cf4430592b       7473849          Easter
      2    67edcf33cb2258cf4430592c       5726070          Easter
      3    67edcf33cb2258cf4430592d       7107738          Easter
      4    67edcf33cb2258cf4430592e       4885191          Easter
      5    67edcf33cb2258cf4430592f       3378370          Easter
      6    67edcf33cb2258cf44305930       9633786          Easter
      7    67edcf33cb2258cf44305931       2686693          Easter
      8    67edcf33cb2258cf44305932       9077853          Easter
      9    67edcf33cb2258cf44305933       9246065          Easter
      10   67edcf33cb2258cf44305934       5854739     Footy's Back
      11   67edcf33cb2258cf44305935       1786182     Footy's Back
      12   67edcf33cb2258cf44305936       5854717     Footy's Back
      13   67edcf33cb2258cf44305937       1786160     Footy's Back
```

```
14   67edcf33cb2258cf44305938      229828              Footy's Back
15   67edcf33cb2258cf44305939     1786171              Footy's Back
16   67edcf33cb2258cf4430593a     5072676              Footy's Back
17   67edcf33cb2258cf4430593b     6364382              Footy's Back
18   67edcf33cb2258cf4430593c     3468508              Footy's Back
19   67edcf33cb2258cf4430593d     5112318              Footy's Back
20   67edcf33cb2258cf4430593e      125050  Bonus Credit Products
21   67edcf33cb2258cf4430593f     3959190  Bonus Credit Products
22   67edcf33cb2258cf44305940     9156257  Bonus Credit Products
23   67edcf33cb2258cf44305941     1167759  Bonus Credit Products
24   67edcf33cb2258cf44305942     2855600  Bonus Credit Products

                                           item_name best_price  \
0      Coles Hot Cross Buns Traditional Fruit | 6 Pack       $3.0
1              Coles Hot Cross Buns Choc Chip | 6 Pack       $3.0
2    Coles Hot Cross Buns Traditional Fruit Mini | …       $3.0
3    Cadbury Dairy Milk Chocolate Easter Egg Hunt C…     $8.00
4    Cadbury Dairy Milk Easter Chocolate Eggs Bag |…     $6.70
5        Coles Hot Cross Buns Apple & Cinnamon | 6 Pack    $3.0
6    Cadbury Selections Large Easter Chocolate Eggs…    $13.00
7    Coles Deli Thawed Australian Cooked Black Tige…     $9.00
8    Greens Deluxe Jam & Cream Lamington Cake Roll …     $6.00
9          Coles Lamb Leg Roast Half | approx. 1.74kg    $26.10
10   Mcvities Digestives Biscuits Milk Chocolate | …     $2.75
11   Red Rock Deli Potato Chips Sweet Chilli And So…     $3.15
12   Mcvities Digestives Biscuits Dark Chocolate | …     $2.75
13   Red Rock Deli Honey Soy Chicken Potato Chips |…     $3.15
14    Ferrero Rocher Chocolate Gift Box 16 Pack | 200g    $9.00
15   Red Rock Deli Sea Salt & Bal Vinegar Potato Ch…     $3.15
16       Cadbury Dairy Milk Chocolate Roll Pack | 55g     $1.25
17   Cadbury Favourites Ultimate Share Boxed Chocol…    $16.00
18   Peckish Gluten Free Fancies Aged Cheddar & Chi…     $1.50
19       Four'N Twenty Frozen Meat Pies 4 Pack | 700g     $4.50
20               Sirena Tuna in Oil Italian Style | 95g    $2.95
21   Maybelline Lash Sensational Sky High Washable …     $17.5
22   Sensodyne Repair & Protect Whitening Sensitive…     $7.35
23       Sensodyne Repair & Protect Toothbrush | 1 pack    $3.00
24   Maybelline Colossal Big Shot Volume Express Bl…    $21.00


    best_unit_price item_price       unit_price price_was special_text  \
0     $0.73 per 1ea      $4.40    $0.73 per 1ea       NaN      SPECIAL
1     $0.73 per 1ea      $4.40    $0.73 per 1ea       NaN      SPECIAL
2     $0.49 per 1ea      $4.40    $0.49 per 1ea       NaN      SPECIAL
3    $2.61 per 100g      $8.00   $2.61 per 100g     13.50      SPECIAL
4    $5.88 per 100g      $6.70   $5.88 per 100g       NaN          NaN
5     $0.73 per 1ea      $4.40    $0.73 per 1ea       NaN      SPECIAL
6    $2.39 per 100g     $13.00   $2.39 per 100g     16.00      SPECIAL
```

```
7   $36.00 per 1kg    $9.00  $36.00 per 1kg      NaN       NaN
8   $2.00 per 100g    $6.00  $2.00 per 100g      NaN       NEW
9   $15.00 per 1kg   $26.10  $15.00 per 1kg      NaN       NaN
10  $1.03 per 100g    $2.75  $1.03 per 100g     5.50       50%
11  $1.91 per 100g    $3.15  $1.91 per 100g     6.30       50%
12  $1.03 per 100g    $2.75  $1.03 per 100g     5.50       50%
13  $1.91 per 100g    $3.15  $1.91 per 100g     6.30       50%
14  $4.50 per 100g    $9.00  $4.50 per 100g    18.00       50%
15  $1.91 per 100g    $3.15  $1.91 per 100g     6.30       50%
16  $2.27 per 100g    $1.25  $2.27 per 100g     2.50       50%
17  $1.95 per 100g   $16.00  $1.95 per 100g    32.00       50%
18  $1.67 per 100g    $1.50  $1.67 per 100g     3.00       50%
19  $0.64 per 100g    $4.50  $0.64 per 100g     9.00       50%
20  $31.05 per 1kg    $2.95  $31.05 per 1kg      NaN       NaN
21  $36.11 per 10ml  $26.00  $36.11 per 10ml     NaN   SPECIAL
22  $7.35 per 100g    $7.35  $7.35 per 100g    12.30       40%
23   $3.00 per 1ea    $3.00   $3.00 per 1ea     5.25       40%
24  $21.65 per 10ml  $21.00  $21.65 per 10ml     NaN       NaN

    promo_text                                          link  \
0    2 for $6  https://www.coles.com.au/product/coles-hot-cro…
1    2 for $6  https://www.coles.com.au/product/coles-hot-cro…
2    2 for $6  https://www.coles.com.au/product/coles-hot-cro…
3         NaN  https://www.coles.com.au/product/cadbury-dairy…
4         NaN  https://www.coles.com.au/product/cadbury-dairy…
5    2 for $6  https://www.coles.com.au/product/coles-hot-cro…
6         NaN  https://www.coles.com.au/product/cadbury-selec…
7         NaN  https://www.coles.com.au/product/coles-deli-th…
8         NaN  https://www.coles.com.au/product/greens-deluxe…
9         NaN  https://www.coles.com.au/product/coles-lamb-le…
10        NaN  https://www.coles.com.au/product/mcvities-dige…
11        NaN  https://www.coles.com.au/product/red-rock-deli…
12        NaN  https://www.coles.com.au/product/mcvities-dige…
13        NaN  https://www.coles.com.au/product/red-rock-deli…
14        NaN  https://www.coles.com.au/product/ferrero-roche…
15        NaN  https://www.coles.com.au/product/red-rock-deli…
16        NaN  https://www.coles.com.au/product/cadbury-dairy…
17        NaN  https://www.coles.com.au/product/cadbury-favou…
18        NaN  https://www.coles.com.au/product/peckish-glute…
19        NaN  https://www.coles.com.au/product/four'n-twenty…
20        NaN  https://www.coles.com.au/product/sirena-tuna-i…
21  2 for $35  https://www.coles.com.au/product/maybelline-la…
22        NaN  https://www.coles.com.au/product/sensodyne-rep…
23        NaN  https://www.coles.com.au/product/sensodyne-rep…
24        NaN  https://www.coles.com.au/product/maybelline-co…

                brand
```

```
0                 Coles
1                 Coles
2                 Coles
3               Cadbury
4               Cadbury
5                 Coles
6               Cadbury
7                 Coles
8                Greens
9                 Coles
10   Mcvities Digestives
11             Red Rock
12   Mcvities Digestives
13             Red Rock
14       Ferrero Rocher
15             Red Rock
16              Cadbury
17              Cadbury
18              Peckish
19        Four'N Twenty
20               Sirena
21           Maybelline
22    Sensodyne Repair &
23    Sensodyne Repair &
24           Maybelline
```

[57]: `merged_df['item_size'] = merged_df['item_name'].apply(extract_clean_item_size)`

[58]: `merged_df[['item_name','item_size']]`

[58]:

	item_name	item_size
0	Coles Hot Cross Buns Traditional Fruit | 6 Pack	6 Pack
1	Coles Hot Cross Buns Choc Chip | 6 Pack	6 Pack
2	Coles Hot Cross Buns Traditional Fruit Mini | …	9 Pack
3	Cadbury Dairy Milk Chocolate Easter Egg Hunt C…	306g
4	Cadbury Dairy Milk Easter Chocolate Eggs Bag |…	114g
…	…	…
77688	Veggie Tubes Sweetcorn & Carrot 12m+	None
77689	100% Pure Cotton Ripple	None
77690	Pumpkin Potato & Beef Baby Food Pouch 6+ Months	None
77691	Lamb Vegetables & Polenta Baby Food Pouch 6+ M…	None
77692	Beef With Vegetables 6+ Months	None

```
[77693 rows x 2 columns]
```

[ ]:

```
[59]: # Define the columns to clean
      price_columns = ['best_price', 'item_price']
      unit_price_columns = ['best_unit_price', 'unit_price']

      # Remove "$" and convert to float for simple price columns
      for col in price_columns:
          merged_df[col] = merged_df[col].replace('[\$,]', '', regex=True).
       ↪astype(float)

      # Extract just the numeric part from unit price columns (e.g., "$0.73 per 1ea"␣
       ↪→ 0.73)
      for col in unit_price_columns:
          price = merged_df[col].str.extract(r'\$?([\d\.]+)').astype(float)

          merged_df[f'unit_weight'] = merged_df[col].str.extract(r'per\s+(\w+)',␣
       ↪expand=False)
          merged_df[col] = price
```

```
[60]: merged_df.head(30)
```

```
[60]:                          _id  product_code               category  \
      0    67edcf33cb2258cf4430592a       8371390                 Easter
      1    67edcf33cb2258cf4430592b       7473849                 Easter
      2    67edcf33cb2258cf4430592c       5726070                 Easter
      3    67edcf33cb2258cf4430592d       7107738                 Easter
      4    67edcf33cb2258cf4430592e       4885191                 Easter
      5    67edcf33cb2258cf4430592f       3378370                 Easter
      6    67edcf33cb2258cf44305930       9633786                 Easter
      7    67edcf33cb2258cf44305931       2686693                 Easter
      8    67edcf33cb2258cf44305932       9077853                 Easter
      9    67edcf33cb2258cf44305933       9246065                 Easter
      10   67edcf33cb2258cf44305934       5854739           Footy's Back
      11   67edcf33cb2258cf44305935       1786182           Footy's Back
      12   67edcf33cb2258cf44305936       5854717           Footy's Back
      13   67edcf33cb2258cf44305937       1786160           Footy's Back
      14   67edcf33cb2258cf44305938        229828           Footy's Back
      15   67edcf33cb2258cf44305939       1786171           Footy's Back
      16   67edcf33cb2258cf4430593a       5072676           Footy's Back
      17   67edcf33cb2258cf4430593b       6364382           Footy's Back
      18   67edcf33cb2258cf4430593c       3468508           Footy's Back
      19   67edcf33cb2258cf4430593d       5112318           Footy's Back
      20   67edcf33cb2258cf4430593e        125050  Bonus Credit Products
      21   67edcf33cb2258cf4430593f       3959190  Bonus Credit Products
      22   67edcf33cb2258cf44305940       9156257  Bonus Credit Products
      23   67edcf33cb2258cf44305941       1167759  Bonus Credit Products
      24   67edcf33cb2258cf44305942       2855600  Bonus Credit Products
      25   67edcf33cb2258cf44305943       8100086  Bonus Credit Products
```

```
26  67edcf33cb2258cf44305944        3959204   Bonus Credit Products
27  67edcf33cb2258cf44305945        8086371   Bonus Credit Products
28  67edcf33cb2258cf44305946        8605537   Bonus Credit Products
29  67edcf33cb2258cf44305947        5919324   Bonus Credit Products


                                         item_name  best_price  \
0      Coles Hot Cross Buns Traditional Fruit | 6 Pack        3.00
1                Coles Hot Cross Buns Choc Chip | 6 Pack        3.00
2   Coles Hot Cross Buns Traditional Fruit Mini | …        3.00
3   Cadbury Dairy Milk Chocolate Easter Egg Hunt C…        8.00
4   Cadbury Dairy Milk Easter Chocolate Eggs Bag |…        6.70
5        Coles Hot Cross Buns Apple & Cinnamon | 6 Pack        3.00
6   Cadbury Selections Large Easter Chocolate Eggs…       13.00
7   Coles Deli Thawed Australian Cooked Black Tige…        9.00
8   Greens Deluxe Jam & Cream Lamington Cake Roll …        6.00
9           Coles Lamb Leg Roast Half | approx. 1.74kg       26.10
10  Mcvities Digestives Biscuits Milk Chocolate | …        2.75
11  Red Rock Deli Potato Chips Sweet Chilli And So…        3.15
12  Mcvities Digestives Biscuits Dark Chocolate | …        2.75
13  Red Rock Deli Honey Soy Chicken Potato Chips |…        3.15
14   Ferrero Rocher Chocolate Gift Box 16 Pack | 200g        9.00
15  Red Rock Deli Sea Salt & Bal Vinegar Potato Ch…        3.15
16       Cadbury Dairy Milk Chocolate Roll Pack | 55g        1.25
17  Cadbury Favourites Ultimate Share Boxed Chocol…       16.00
18  Peckish Gluten Free Fancies Aged Cheddar & Chi…        1.50
19        Four'N Twenty Frozen Meat Pies 4 Pack | 700g        4.50
20              Sirena Tuna in Oil Italian Style | 95g        2.95
21  Maybelline Lash Sensational Sky High Washable …       17.50
22  Sensodyne Repair & Protect Whitening Sensitive…        7.35
23       Sensodyne Repair & Protect Toothbrush | 1 pack        3.00
24  Maybelline Colossal Big Shot Volume Express Bl…       21.00
25  Maybelline Lash Sensational Firework Mascara |…       17.50
26  Maybelline Lash Sensational Sky High Waterproo…       26.00
27        Sensodyne Soft Daily Care Toothbrush | 3 pack        6.30
28  Sensodyne Repair & Protect Sensitive Toothpast…        7.35
29  Sensodyne Complete Care and Smart Clean Extra …        7.35


    best_unit_price  item_price  unit_price  price_was special_text promo_text  \
0              0.73        4.40        0.73        NaN      SPECIAL   2 for $6
1              0.73        4.40        0.73        NaN      SPECIAL   2 for $6
2              0.49        4.40        0.49        NaN      SPECIAL   2 for $6
3              2.61        8.00        2.61      13.50      SPECIAL        NaN
4              5.88        6.70        5.88        NaN          NaN        NaN
5              0.73        4.40        0.73        NaN      SPECIAL   2 for $6
6              2.39       13.00        2.39      16.00      SPECIAL        NaN
7             36.00        9.00       36.00        NaN          NaN        NaN
8              2.00        6.00        2.00        NaN          NEW        NaN
```

9	15.00	26.10	15.00	NaN	NaN	NaN
10	1.03	2.75	1.03	5.50	50%	NaN
11	1.91	3.15	1.91	6.30	50%	NaN
12	1.03	2.75	1.03	5.50	50%	NaN
13	1.91	3.15	1.91	6.30	50%	NaN
14	4.50	9.00	4.50	18.00	50%	NaN
15	1.91	3.15	1.91	6.30	50%	NaN
16	2.27	1.25	2.27	2.50	50%	NaN
17	1.95	16.00	1.95	32.00	50%	NaN
18	1.67	1.50	1.67	3.00	50%	NaN
19	0.64	4.50	0.64	9.00	50%	NaN
20	31.05	2.95	31.05	NaN	NaN	NaN
21	36.11	26.00	36.11	NaN	SPECIAL	2 for $35
22	7.35	7.35	7.35	12.30	40%	NaN
23	3.00	3.00	3.00	5.25	40%	NaN
24	21.65	21.00	21.65	NaN	NaN	NaN
25	10.36	29.00	10.36	NaN	SPECIAL	2 for $35
26	43.33	26.00	43.33	NaN	NaN	NaN
27	2.10	6.30	2.10	11.00	40%	NaN
28	7.35	7.35	7.35	12.30	40%	NaN
29	7.35	7.35	7.35	12.30	40%	NaN

	link	brand \
0	https://www.coles.com.au/product/coles-hot-cro…	Coles
1	https://www.coles.com.au/product/coles-hot-cro…	Coles
2	https://www.coles.com.au/product/coles-hot-cro…	Coles
3	https://www.coles.com.au/product/cadbury-dairy…	Cadbury
4	https://www.coles.com.au/product/cadbury-dairy…	Cadbury
5	https://www.coles.com.au/product/coles-hot-cro…	Coles
6	https://www.coles.com.au/product/cadbury-selec…	Cadbury
7	https://www.coles.com.au/product/coles-deli-th…	Coles
8	https://www.coles.com.au/product/greens-deluxe…	Greens
9	https://www.coles.com.au/product/coles-lamb-le…	Coles
10	https://www.coles.com.au/product/mcvities-dige…	Mcvities Digestives
11	https://www.coles.com.au/product/red-rock-deli…	Red Rock
12	https://www.coles.com.au/product/mcvities-dige…	Mcvities Digestives
13	https://www.coles.com.au/product/red-rock-deli…	Red Rock
14	https://www.coles.com.au/product/ferrero-roche…	Ferrero Rocher
15	https://www.coles.com.au/product/red-rock-deli…	Red Rock
16	https://www.coles.com.au/product/cadbury-dairy…	Cadbury
17	https://www.coles.com.au/product/cadbury-favou…	Cadbury
18	https://www.coles.com.au/product/peckish-glute…	Peckish
19	https://www.coles.com.au/product/four'n-twenty…	Four'N Twenty
20	https://www.coles.com.au/product/sirena-tuna-i…	Sirena
21	https://www.coles.com.au/product/maybelline-la…	Maybelline
22	https://www.coles.com.au/product/sensodyne-rep…	Sensodyne Repair &
23	https://www.coles.com.au/product/sensodyne-rep…	Sensodyne Repair &

```
24  https://www.coles.com.au/product/maybelline-co…              Maybelline
25  https://www.coles.com.au/product/maybelline-la…              Maybelline
26  https://www.coles.com.au/product/maybelline-la…              Maybelline
27  https://www.coles.com.au/product/sensodyne-sof…                 Unknown
28  https://www.coles.com.au/product/sensodyne-rep…  Sensodyne Repair &
29  https://www.coles.com.au/product/sensodyne-com…                 Unknown

    item_size unit_weight
0     6 Pack         1ea
1     6 Pack         1ea
2     9 Pack         1ea
3       306g        100g
4       114g        100g
5     6 Pack         1ea
6       545g        100g
7       250g         1kg
8       300g        100g
9     1.74kg         1kg
10      266g        100g
11      165g        100g
12      266g        100g
13      165g        100g
14      200g        100g
15      165g        100g
16       55g        100g
17      820g        100g
18       90g        100g
19      700g        100g
20       95g         1kg
21     7.2mL        10ml
22      100g        100g
23    1 pack         1ea
24     9.7mL        10ml
25       28g         10g
26       6mL        10ml
27    3 pack         1ea
28      100g        100g
29      100g        100g
```

[61]: `data.head(30)`

[61]:
```
                         _id  product_code              category  \
0   67ed14d6a8bdc17e3ddceed5       8371390                Easter
1   67ed14d6a8bdc17e3ddceed6       7473849                Easter
2   67ed14d6a8bdc17e3ddceed7       5726070                Easter
3   67ed14d6a8bdc17e3ddceed9       4885191                Easter
4   67ed14d6a8bdc17e3ddceedb       3378370                Easter
```

```
5   67ed14d6a8bdc17e3ddceedc    9633786              Easter
6   67ed14d6a8bdc17e3ddceede    9077853              Easter
7   67ed14d6a8bdc17e3ddceedf    5854739         Footy's Back
8   67ed14d6a8bdc17e3ddceee0    5854717         Footy's Back
9   67ed14d6a8bdc17e3ddceee1    1786182         Footy's Back
10  67ed14d6a8bdc17e3ddceee4    1786160         Footy's Back
11  67ed14d6a8bdc17e3ddceee6    5072676         Footy's Back
12  67ed14d6a8bdc17e3ddceeeb    2855600  Bonus Credit Products
13  67ed14d6a8bdc17e3ddceeec    8100086  Bonus Credit Products
14  67ed14d6a8bdc17e3ddceeee    5239330  Bonus Credit Products
15  67ed14d6a8bdc17e3ddceef1    2364438  Bonus Credit Products
16  67ed14d6a8bdc17e3ddceef4    9246065       Meat & Seafood
17  67ed14d6a8bdc17e3ddceefa    8112799       Meat & Seafood
18  67ed14d6a8bdc17e3ddcef04    5111654    Fruit & Vegetables
19  67ed14d6a8bdc17e3ddcef0d    5066120    Dairy, Eggs & Fridge
20  67ed14d6a8bdc17e3ddcef0f    3349870    Dairy, Eggs & Fridge
21  67ed14d6a8bdc17e3ddcef13    8956189               Bakery
22  67ed14d6a8bdc17e3ddcef18    4215337               Bakery
23  67ed14d6a8bdc17e3ddcef19    4150672               Bakery
24  67ed14d6a8bdc17e3ddcef1a    1647164               Bakery
25  67ed14d6a8bdc17e3ddcef1c    5400130                 Deli
26  67ed14d6a8bdc17e3ddcef20    3272845                 Deli
27  67ed14d6a8bdc17e3ddcef25     125050               Pantry
28  67ed14d6a8bdc17e3ddcef26    2683924               Pantry
29  67ed14d6a8bdc17e3ddcef2b    5311187               Pantry

                                        item_name  best_price  \
0      Coles Hot Cross Buns Traditional Fruit | 6 Pack        3.00
1                Coles Hot Cross Buns Choc Chip | 6 Pack        3.00
2   Coles Hot Cross Buns Traditional Fruit Mini | …        3.00
3   Cadbury Dairy Milk Easter Chocolate Eggs Bag |…        6.70
4       Coles Hot Cross Buns Apple & Cinnamon | 6 Pack        3.00
5   Cadbury Selections Large Easter Chocolate Eggs…       13.00
6   Greens Deluxe Jam & Cream Lamington Cake Roll …        6.00
7   Mcvities Digestives Biscuits Milk Chocolate | …        2.75
8   Mcvities Digestives Biscuits Dark Chocolate | …        2.75
9   Red Rock Deli Potato Chips Sweet Chilli And So…        3.15
10  Red Rock Deli Honey Soy Chicken Potato Chips |…        3.15
11       Cadbury Dairy Milk Chocolate Roll Pack | 55g        1.25
12  Maybelline Colossal Big Shot Volume Express Bl…       21.00
13  Maybelline Lash Sensational Firework Mascara |…       17.50
14   Maybelline Sky High Mascara Cosmic Black | 7.2mL       26.00
15  Maybelline Lash Sensational Very Black Waterpr…       25.00
16          Coles Lamb Leg Roast Half | approx. 1.74kg       26.10
17  Coles No Added Hormone Beef 5 Star Extra Trim …       10.00
18        Coles Pink Lady Apples Medium | approx. 200g        1.70
19               Brownes Dairy Milk Full Cream | 2L        4.45
```

```
20           Fresh Fodder Greek Style Tzatziki Dip | 200g          5.40
21               Tip Top The One White Toast Bread | 700g          4.50
22        Noshu 95% Sugar Free Iced Carrot Cakes | 180g          5.90
23           Mr Kipling Chocolate Slices 6 Pack | 155g          6.00
24               Mr Kipling Lemon Slices 6 Pack | 165g          6.00
25                         Castello Aged Havarti | 200g          6.25
26           Castello Double Cream Brie Chilli | 150g          6.25
27               Sirena Tuna in Oil Italian Style | 95g          2.95
28  Leggo's Bolognese with Chunky Tomato Garlic & …          2.75
29                             Wicked Kimchi | 460g          3.50


    best_unit_price  item_price  unit_price price_was special_text  \
0              0.73        4.40        0.73       NaN      SPECIAL
1              0.73        4.40        0.73       NaN      SPECIAL
2              0.49        4.40        0.49       NaN      SPECIAL
3              5.88        6.70        5.88       NaN          NaN
4              0.73        4.40        0.73       NaN      SPECIAL
5              2.39       13.00        2.39     16.00      SPECIAL
6              2.00        6.00        2.00       NaN          NEW
7              1.03        2.75        1.03      5.50          50%
8              1.03        2.75        1.03      5.50          50%
9              1.91        3.15        1.91      6.30          50%
10             1.91        3.15        1.91      6.30          50%
11             2.27        1.25        2.27      2.50          50%
12            21.65       21.00       21.65       NaN          NaN
13            10.36       29.00       10.36       NaN      SPECIAL
14            36.11       26.00       36.11       NaN          NaN
15            26.32       25.00       26.32       NaN          NaN
16            15.00       26.10       15.00       NaN          NaN
17            20.00       10.00       20.00       NaN          NaN
18             8.50        1.70        8.50       NaN          NaN
19             2.23        4.45        2.23       NaN    EVERY DAY
20             2.70        5.40        2.70       NaN          NaN
21             0.64        4.50        0.64       NaN    EVERY DAY
22             3.28        5.90        3.28      7.50      SPECIAL
23             3.87        6.00        3.87       NaN          NaN
24             3.64        6.00        3.64       NaN          NaN
25            42.50        8.50       42.50       NaN      SPECIAL
26            53.34        8.00       53.34       NaN      SPECIAL
27            31.05        2.95       31.05       NaN          NaN
28             0.55        2.75        0.55      4.60      SPECIAL
29             7.61        3.50        7.61       NaN    EVERY DAY


     promo_text                                            link  \
0      2 for $6  https://www.coles.com.au/product/coles-hot-cro…
1      2 for $6  https://www.coles.com.au/product/coles-hot-cro…
2      2 for $6  https://www.coles.com.au/product/coles-hot-cro…
```

```
3           NaN   https://www.coles.com.au/product/cadbury-dairy…
4       2 for $6   https://www.coles.com.au/product/coles-hot-cro…
5           NaN   https://www.coles.com.au/product/cadbury-selec…
6           NaN   https://www.coles.com.au/product/greens-deluxe…
7           NaN   https://www.coles.com.au/product/mcvities-dige…
8           NaN   https://www.coles.com.au/product/mcvities-dige…
9           NaN   https://www.coles.com.au/product/red-rock-deli…
10          NaN   https://www.coles.com.au/product/red-rock-deli…
11          NaN   https://www.coles.com.au/product/cadbury-dairy…
12          NaN   https://www.coles.com.au/product/maybelline-co…
13      2 for $35   https://www.coles.com.au/product/maybelline-la…
14          NaN   https://www.coles.com.au/product/maybelline-sk…
15          NaN   https://www.coles.com.au/product/maybelline-la…
16          NaN   https://www.coles.com.au/product/coles-lamb-le…
17          NaN   https://www.coles.com.au/product/coles-no-adde…
18          NaN   https://www.coles.com.au/product/coles-pink-la…
19          NaN   https://www.coles.com.au/product/brownes-dairy…
20          NaN   https://www.coles.com.au/product/fresh-fodder-…
21          NaN   https://www.coles.com.au/product/tip-top-the-o…
22          NaN   https://www.coles.com.au/product/noshu-95perce…
23          NaN   https://www.coles.com.au/product/mr-kipling-ch…
24          NaN   https://www.coles.com.au/product/mr-kipling-le…
25   2 for $12.50   https://www.coles.com.au/product/castello-aged…
26   2 for $12.50   https://www.coles.com.au/product/castello-doub…
27          NaN   https://www.coles.com.au/product/sirena-tuna-i…
28          NaN   https://www.coles.com.au/product/leggo's-bolog…
29          NaN   https://www.coles.com.au/product/wicked-kimchi…


                brand item_size unit_weight  similarity_score
0                Coles    6 Pack        1ea          0.190840
1                Coles    6 Pack        1ea          0.190840
2                Coles    9 Pack        1ea          0.190840
3              Cadbury      114g       100g          0.203252
4                Coles    6 Pack        1ea          0.190840
5              Cadbury      545g       100g          0.125313
6               Greens      300g       100g          0.203252
7   Mcvities Digestives      266g       100g          0.284091
8   Mcvities Digestives      266g       100g          0.284091
9             Red Rock      165g       100g          0.290698
10            Red Rock      165g       100g          0.290698
11             Cadbury       55g       100g          0.223714
12           Maybelline     9.7mL       10ml          0.153728
13           Maybelline       28g        10g          0.140746
14           Maybelline     7.2mL       10ml          0.203874
15           Maybelline     9.5mL       10ml          0.203874
16               Coles    1.74kg        1kg          0.067981
17               Coles      500g        1kg          0.111483
```

```
18          Coles          200g      1kg      0.267594
19        Brownes           2L        1l      0.311042
20    Fresh Fodder          200g     100g     0.404040
21         Tip Top          700g     100g     0.543478
22           Noshu          180g     100g     0.609756
23      Mr Kipling          155g     100g     0.595238
24      Mr Kipling          165g     100g     0.595238
25        Castello          200g      1kg      0.465116
26        Castello          150g      1kg      0.487805
27          Sirena           95g      1kg      0.361664
28         Leggo's          500g     100g     0.350263
29          Wicked          460g      1kg      0.355240
```

[ ]:
```python
merged_df = pd.concat([data_1, data_2, data_3, data_4], ignore_index=True,
     sort=False)
```

[62]:
```python
data.shape
```

[62]: (140, 16)

[63]:
```python
merged_df.shape
```

[63]: (77693, 15)

[65]:
```python
data.columns
```

[65]:
```
Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
       'best_unit_price', 'item_price', 'unit_price', 'price_was',
       'special_text', 'promo_text', 'link', 'brand', 'item_size',
       'unit_weight', 'similarity_score'],
      dtype='object')
```

[66]:
```python
data_new = data.drop('similarity_score', axis=1)
```

[67]:
```python
data_new.columns
```

[67]:
```
Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
       'best_unit_price', 'item_price', 'unit_price', 'price_was',
       'special_text', 'promo_text', 'link', 'brand', 'item_size',
       'unit_weight'],
      dtype='object')
```

[68]:
```python
data_new.shape
```

[68]: (140, 15)

[69]:
```python
merged_df = pd.concat([data_new, merged_df], ignore_index=True, sort=False)
```

```
[70]: merged_df.shape
```

```
[70]: (77833, 15)
```

```
[71]: merged_df = merged_df.groupby('category', group_keys=False).
      ↪apply(compute_similarity)
```

```
[72]: merged_df[['category', 'similarity_score']]
```

```
[72]:                 category  similarity_score
      0                 Easter          0.198373
      1                 Easter          0.198373
      2                 Easter          0.198373
      3                 Easter          0.214731
      4                 Easter          0.198373
      ...                  ...               ...
      77828        INFANT FOOD          0.439613
      77829   INFANT PERSONAL          0.267265
      77830        INFANT FOOD          0.439613
      77831        INFANT FOOD          0.439613
      77832        INFANT FOOD          0.439613

      [77833 rows x 2 columns]
```

1.5 Plot by category

```
[73]: plt.figure(figsize=(12, 6))
      ax = sns.countplot(data=merged_df, x='category', order=merged_df['category'].
        ↪value_counts().index)

      # Rotate x-axis labels vertically
      plt.xticks(rotation=90)

      for p in ax.patches:
          count = int(p.get_height())
          ax.text(p.get_x() + p.get_width() / 2, p.get_height() + 0.5,  # x, y␣
        ↪position
                  str(count), ha='center', va='bottom', fontsize=9)

      plt.title('Product Count by Category')
      plt.xlabel('Category')
      plt.ylabel('Count')
      plt.tight_layout()
      plt.show()
```

Product Count by Category

```
[78]: top_categories = merged_df['category'].value_counts().head(10)

      # Plot
      plt.figure(figsize=(10, 6))
      top_categories.plot(kind='bar')
      plt.title('Top 10 Categories by Count')
      plt.xlabel('Category')
      plt.ylabel('Count')
      plt.xticks(rotation=45)
      plt.tight_layout()
      plt.show()
```

Top 10 Categories by Count
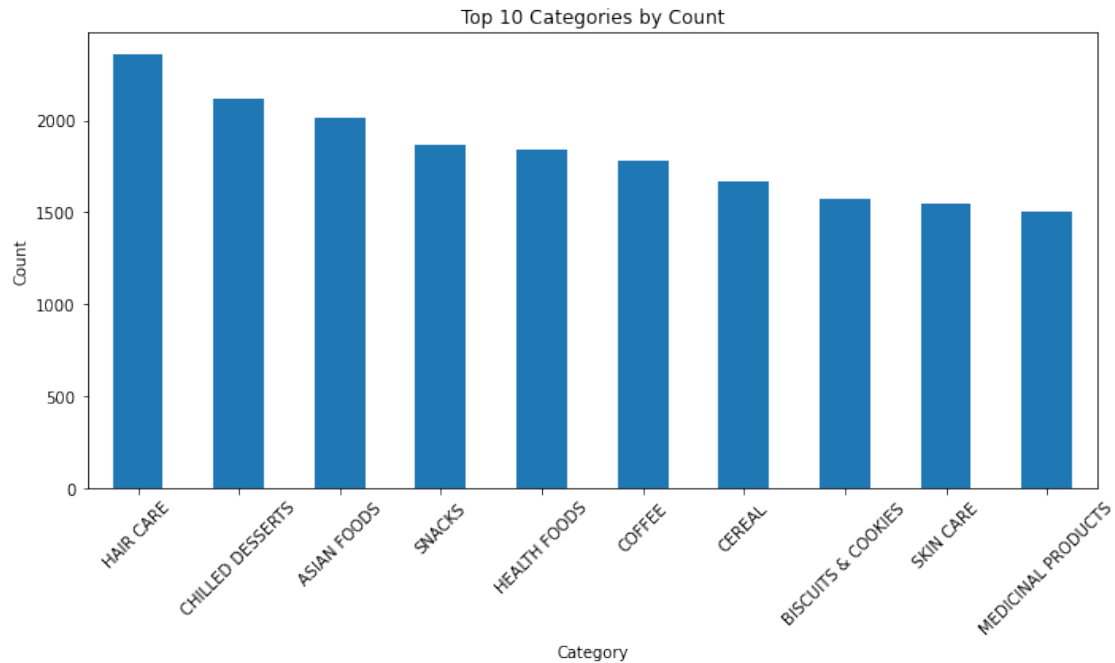
[79]: `merged_df.to_csv('merged_data.csv')`

[80]: `merged_df.isna().sum()`

[80]:
```
_id                     0
product_code            0
category                0
item_name               0
best_price              0
best_unit_price     77053
item_price              0
unit_price           2237
price_was           77641
special_text        69680
promo_text          59274
link                    0
brand                   0
item_size           66683
unit_weight          2238
similarity_score        0
dtype: int64
```

[30]:
```python
# load merged.csv
file_name = 'merged_data.csv'
data = pd.read_csv(file_name, index_col=False)
```

```
C:\ProgramData\Anaconda3\lib\site-
packages\IPython\core\interactiveshell.py:3165: DtypeWarning: Columns (9) have
mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

[31]: `data.columns`

[31]:
```
Index(['Unnamed: 0', '_id', 'product_code', 'category', 'item_name',
       'best_price', 'best_unit_price', 'item_price', 'unit_price',
       'price_was', 'special_text', 'promo_text', 'link', 'brand', 'item_size',
       'unit_weight', 'similarity_score'],
      dtype='object')
```

[32]: `df = data.drop(columns=['Unnamed: 0'], errors='ignore')`

[33]: `df.columns`

[33]:
```
Index(['_id', 'product_code', 'category', 'item_name', 'best_price',
       'best_unit_price', 'item_price', 'unit_price', 'price_was',
       'special_text', 'promo_text', 'link', 'brand', 'item_size',
       'unit_weight', 'similarity_score'],
      dtype='object')
```

Remove unnecessary columns: - '_id' - 'price_was' - 'special_text' - 'promo_text' - 'link' - 'best_unit_price' - 'best_price'

[34]:
```
df = df.drop(columns=['_id', 'price_was','special_text', 'promo_text', 'link',
 ↪'best_unit_price', 'best_price'])
```

[35]:
```python
#rearrange column order
order = ['product_code', 'category', 'item_name', 'brand', 'item_price',
 ↪'item_size', 'unit_price','unit_weight', 'similarity_score']
data = df[order]
```

[36]: `data.head()`

[36]:
```
   product_code category                                      item_name  \
0       8371390   Easter    Coles Hot Cross Buns Traditional Fruit | 6 Pack
1       7473849   Easter            Coles Hot Cross Buns Choc Chip | 6 Pack
2       5726070   Easter    Coles Hot Cross Buns Traditional Fruit Mini | …
3       4885191   Easter    Cadbury Dairy Milk Easter Chocolate Eggs Bag |…
4       3378370   Easter       Coles Hot Cross Buns Apple & Cinnamon | 6 Pack

      brand  item_price item_size  unit_price unit_weight  similarity_score
0     Coles         4.4    6 Pack        0.73         1ea          0.198373
1     Coles         4.4    6 Pack        0.73         1ea          0.198373
2     Coles         4.4    9 Pack        0.49         1ea          0.198373
3   Cadbury         6.7      114g        5.88        100g          0.214731
```

```
4      Coles          4.4      6 Pack          0.73          1ea          0.198373
```

[37]: `coles_new = pd.read_csv('coles_new.csv')`

[38]: `coles_new.shape`

[38]: (20604, 11)

[39]: `coles_new.columns`

[39]:
```
Index(['product_code', 'category', 'item_name', 'essential_flag ',
       'item_name ', 'Unnamed: 5', 'subcat', 'item_price', 'unit_price',
       'brand_name', ' weights'],
      dtype='object')
```

[40]: `coles_brand_names = list(coles_new['brand_name'].unique())`

[41]: `coles_new.head()`

[41]:
```
   product_code        category  \
0       4174794  Meat & Seafood
1       9224255  Meat & Seafood
2       9223332  Meat & Seafood
3       7382539  Meat & Seafood
4       3091126  Meat & Seafood


                                      item_name  essential_flag  \
0         Coles Boneless Pork Leg Roast | approx 2.1kg                 1
1  Cleaver's Organic Grass-Fed Premium Beef Mince…                 1
2  Cleaver's Organic Grass-Fed Extra Lean Beef Mi…                 1
3         The Bare Bird Chicken Thigh Fillets | 480g                 1
4         The Bare Bird Whole Chicken | approx 1.6kg                 1


                          item_name  Unnamed: 5   subcat  item_price  \
0         Boneless Pork Leg Roast          NaN     Pork        14.7
1      Grass-Fed Premium Beef Mince        NaN     Beef        14.0
2  Grass-Fed Extra Lean Beef Mince       NaN     Beef        15.0
3      Bird Chicken Thigh Fillets        NaN  Chicken        13.0
4              Bird Whole Chicken        NaN  Chicken        13.6


   unit_price        brand_name       weights
0       7.001              Coles  approx 2.1kg
1      28.001  Cleaver's Organic          500g
2      30.001  Cleaver's Organic          500g
3      27.081           The Bare          480g
4       8.501           The Bare  approx 1.6kg
```

[42]: `sub_categories = list(coles_new['subcat'].unique())`

```
[43]: len(sub_categories)
```

```
[43]: 129
```

```
[44]: def get_subcat(item_name):
          if not isinstance(item_name, str):
              return "Unknown"

          item_name_lower = item_name.lower()

          for subcat in sub_categories:
              if subcat.lower() in item_name_lower:
                  return subcat  # return the original-cased brand
          return "Unknown"
```

1.6 Use predefined Subcat from coles.csv to extract subcategories

```
[45]: data['subcat'] = data['item_name'].apply(get_subcat)
```

<ipython-input-45-0344954ad8fd>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data['subcat'] = data['item_name'].apply(get_subcat)

```
[46]: data.head()
```

```
[46]:    product_code category                                      item_name  \
       0       8371390   Easter    Coles Hot Cross Buns Traditional Fruit | 6 Pack
       1       7473849   Easter         Coles Hot Cross Buns Choc Chip | 6 Pack
       2       5726070   Easter    Coles Hot Cross Buns Traditional Fruit Mini | …
       3       4885191   Easter    Cadbury Dairy Milk Easter Chocolate Eggs Bag |…
       4       3378370   Easter      Coles Hot Cross Buns Apple & Cinnamon | 6 Pack

            brand  item_price item_size  unit_price unit_weight  similarity_score  \
       0    Coles         4.4    6 Pack        0.73         1ea          0.198373
       1    Coles         4.4    6 Pack        0.73         1ea          0.198373
       2    Coles         4.4    9 Pack        0.49         1ea          0.198373
       3  Cadbury         6.7      114g        5.88        100g          0.214731
       4    Coles         4.4    6 Pack        0.73         1ea          0.198373

            subcat
       0     Fruit
       1   Unknown
       2     Fruit
       3      Milk
```

4 Unknown

```
[47]: data.isna().sum()
```

```
[47]: product_code           0
      category               0
      item_name              0
      brand                  0
      item_price             0
      item_size          66683
      unit_price          2237
      unit_weight         2238
      similarity_score       0
      subcat                 0
      dtype: int64
```

```
[48]: data['unit_weight'].unique()
```

```
[48]: array(['1ea', '100g', '10ml', '10g', '1kg', '1l', '100ml', '10', '100',
             '1', '100ea', nan, '100gwas', '1kgwas', '100eawas', '1L', '100mL',
             '1M', '10mL', '1kgM'], dtype=object)
```

```
[24]: data['item_size'].unique()
```

```
[24]: array(['6 Pack', '9 Pack', '114g', '545g', '300g', '266g', '165g', '55g',
             '9.7mL', '28g', '7.2mL', '9.5mL', '1.74kg', '500g', '200g', '2L',
             '700g', '180g', '155g', '150g', '95g', '460g', '700mL', '330g',
             '400GRAM', '1kg', '350g', '40g', '306g', '375g', '395g', '24 pack',
             '95 GRAM', '30 pack', '900mL', '50 pack', '90g', '6mL', '2.8kg',
             '2.1kg', '112g', '1.2kg', '1 each', '125g', '650g', '240g', '750g',
             '288g', '4 pack', '600g', '1 Each', '4L', '16 pack', '9mL', '170g',
             '850g', '250g', '1.8kg', '4 Pack', '10 pack', '445 GRAM', '425g',
             '31 pack', '820g', '8mL', '1.4kg', '400g', '6g', '230g', '100g',
             '1 pack', '3 pack', '1L', '1KG', '235g', '160 GRAM', '360g',
             '220g', '120g', '65g', '210g', '190g', '340g', '600mL', '390mL',
             '400mL', '544mL', '475mL', '264mL', '1 Pack', '500mL', '14 pack',
             '0.4g', '120 pack', '312g', '243g', '187g', '80g', '60g', '63g',
             '40 pack', '900g', '160g', '466g', '185g', '6 pack', '3L', '428mL',
             '458mL', '60 Pack', '100 pack', '20 pack', '240 pack', '64 pack',
             '34 pack', '7.5mL', '28 pack', '750mL', '200mL', '15mL', '30 Pack',
             '24 Pack', '341g', '1.25L', '60PACK', '110g', '800g', '50g', '70g',
             '1pack', '457mL', '1 KG', '350mL', '60 pack', '420g', '145g',
             '30g', '3 Pack', '115g', '390g', '415g', '520g', '550g', '2 pack',
             '10L', '165 gram', '150 Pack', '8 pack', '560mL', '440g', '36mL',
             '680mL', '50 Pack', '310g', '32 Pack', '9 pack', '250 Pack',
             '450g', '357g', nan, '24x375mL', '330mL', '375mL', '30x375mL',
             '16 Pack', '15 Pack', '5 Pack', '4x375mL', '345mL', '750ml',
```

'275mL', '355mL', '10x375mL', '2 Pack', '10 Pack', '330ml',
'24x375ml', '10 x 375mL', '10X375Ml', '4x200mL', '400Ml',
'12 Pack', '20 Pack', '24X600mL', '24X350mL', '12 x 250ml',
'6x250mL', '4X300mL', '320mL', '5 pack', '8 Pack', '5L', '3x27g',
'6x22g', '2x125g', '6x160g', '650mL', '12 pack', '4x115g', '5X18g',
'15 pack', '18 pack', '5pack', '3x65g', '130g', '3X60G', '8x20g',
'1300mL', '2x120g', '4 PACK', '18 Pack', '250mL', '4x95g', '1.1L',
'1.3L', '5x15g', '2X125G', '2.7L', '15x18g', '1.2L', '1.4L',
'6X20g', '250Ml', '5x18g', '6 x 250ml', '6X10 Pack', '7 Pack',
'1.9L', '6Pack', '920mL', '24pack', '12x100g', '1.48L', '12X20g',
'6X21g', '1.5L', '410mL', '800mL', '8x2g', '180mL', '10X15g',
'1.6L', '460mL', '2.2L', '6x30g', '3x60g', '1.7L', '520ml',
'1.76L', '12X21g', '6L', '517mL', '21g', '35mL', '20x20cm',
'443mL', '1.15L', '946mL', '532mL', '2x40g', '56L', '2x19mL',
'4Pk', '2 X 19mL', '2pk', '76L', '1Kg', '2pack', '1 Kg', '12Pack',
'5Kg', '3X50G', '4X125G', '15kg', '15x250g', '4X140g', '6x200mL',
'84 Pack', '4x350mL', '12 x100g', '2Pk', '2x170g', '12x250g',
'4x150g', '2x150g', '2x80g', '2X62g', '452g', '12X50g', '26g',
'280g', '2X150G', '25g', '2X150Gram', '2Pack', '2 x 150g', '2l',
'2x200g', '2x110g', '6pk', '3pk', '7 pack', '9pk', '4pk', '8pack',
'5x85g', '4X40G', '4x90g', '137g', '92.3g', '80 pack', '3 x 250ml',
'12x250mL', '3x250mL', '145mL', '5pk', '5x120g', '9x3g', '5x140g',
'6X200mL', '5x80g', '25 pack', '4X89g', '5X40g', '12X375g',
'200 Pack', '52g', '6x330mL', '48g', '14 Pack', '530g', '5X110g',
'5x60g', '4X125g', '4x40g', '132g', '25 Pack', '140g', '200 pack',
'4 x 35g', '24x70mL', '105g', '5X30g', '4x220g', '5x40g', '5x34g',
'4X60G', '5X30G', '5x30g', '13.5L', '40 Pack', '16g', '32 pack',
'100 Pack', '9X27g', '2 X 125G', '5X90g', '8g', '2x250g', '4x120g',
'6X250mL', '5X128g', '5x90g', '4X35g', '4X250mL', '95gram',
'4x250mL', '4Pack', '4 X 500mL', '4x14g', '3Pack', '4X14G',
'8Pack', '1000mL', '6x14g', '8x250mL', '8pk', '8X2 pack',
'3x10 Multipack', '8x15g', '12X42G', '35 Pack', '12pk',
'4 x 500mL', '12x300mL', '6 x 250mL', '4x500mL', '12x600mL',
'8x17g', '12X500mL', '4x330mL', '6x275mL', '6x240mL', '15x250mL',
'8x330mL', '4x250ml', '300mL', '10X375mL', '4x473ml', '4x237mL',
'4X500Ml', '4x500ml', '540ml', '4x450mL', '10x375ml', '24x250mL',
'1l', '4X330mL', '4x300mL', '6x250ml', '500ml', '50mL', '5g',
'2x50g', '27L', '4G', '2X50G', '96 Pack', '3.5L', '2x75g',
'2X250Ml', '7.5L', '236mL', '8L', '950mL', '5G', '2.5L', '860mL',
'3x50g', '2.63L', '2x42g', '480mL', '2.4L', '2x36mL', '1.73L',
'3x125g', '870ml', '5.5L', '25L', '2X10Ml', '8.5L', '4.7L',
'250ml', '18pk', '450mL', '4g', '1.52L', '150mL', '710mL', '35g',
'1.65L', '4pack', '9L', '420mL', '473mL', '7Kg', '3.6L', '12.5L',
'850mL', '5.3L', '280mL', '220mL', '11Kg', '40ml', '380mL', '4.2L',
'96g', '6 X 90G', '4 X 90G', '3X125g', '2 X 90G', '2X90g',
'4X 90G', '7pk', '145pk', '100G', '75Gram', '62.5 mL', '8x50mL',
'100pk', '4X70G', '2.7mL', '6 X90G', '40pk', '0.28g', '1pk', '15g',

```
       '70pk', '44x69mm', '75x75mm', '6x10Cm', '10pk', '90pk', '150pk',
       '50pk', '20g', '900G', '57Kg', '210mL', '4x4g', '6x15g', '6.8g',
       '120mL', '300ml', '295mL', '240mL', '18Pack', '440mL', '490G',
       '14pk'], dtype=object)
```

[17]:
```python
# Keep rows where item_size contains 'pack' (case-insensitive)
df_packs = df[df['item_size'].str.contains('pack', case=False, na=False)]

# If you only want the item name and size columns
df_packs = df_packs[['item_name', 'item_size']]
```

[27]:
```python
names = df_packs['item_name'].unique().tolist()
```

[28]:
```python
names
```

[28]:
```
['Coles Hot Cross Buns Traditional Fruit | 6 Pack',
 'Coles Hot Cross Buns Choc Chip | 6 Pack',
 'Coles Hot Cross Buns Traditional Fruit Mini | 9 Pack',
 'Coles Hot Cross Buns Apple & Cinnamon | 6 Pack',
 'Coles Spring Water 24X600mL | 24 pack',
 'Pepsi Max No Sugar Cola Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack',
 'OMO Ultimate 3 in 1 Laundry Capsules 50 Washes | 50 pack',
 'Fever-Tree Premium Mixers Light Meditteranean T … 4 pack',
 'Tide Laundry 3 In 1 Pods Original | 16 pack',
 'Danone The Ultimate Black Cherry Yoghurt 4x115g | 4 pack',
 'Danone The Ultimate Golden Honey Yoghurt 4x115g | 4 Pack',
 'Illy Espresso Classico Capsules | 10 pack',
 'Solo Thirst Crusher Original Lemon Soft Drink Cans Multipack  … 30 pack',
 'Pepsi Cola Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack',
 'OMO Touch of Comfort 3 in 1 Laundry Capsules 50 Washes | 50 pack',
 'Tide Laundry 3 In 1 Pods Original | 31 pack',
 'Schweppes Lemonade Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack',
 'Sunkist Orange Soft Drink Cans Multipack 375mL x 30 Pack | 30 pack',
 'OMO Active Clean 3 in 1 Laundry Capsules 50 Washes | 50 pack',
 'Sensodyne Repair & Protect Toothbrush | 1 pack',
 'Sensodyne Soft Daily Care Toothbrush | 3 pack',
 'Tip Top English Muffins Pizza Flavoured | 6 Pack',
 'Coca-Cola Zero Sugar Soft Drink Multipack Cans 24x375mL | 24 pack',
 'Oral B IO2 Electric Toothbrush Black | 1 Pack',
 'Libra Goodnights Extra Long With Wings Pads | 10 pack',
 'Libra Pads Ultra Thin Regular Wings | 14 pack',
 'Swisspers Cotton Tips Paperstems | 120 pack',
 'Oral B IO2 Electric Toothbrush White | 1 Pack',
 'Coles Finest By Laurent White Sourdough Vienna | 1 pack',
 'Swisspers Baby Organic Cotton Pads | 40 pack',
 'Golden Circle Tropical Punch Fruit Drink Lunch Box Multipack Poppers 6x250mL |
6 pack',
```

'Coca-Cola Classic Soft Drink Multipack Cans 10x375mL | 10 pack',
'Red Bull Energy Drink 4X250mL | 4 pack',
'Swisse Ub Magnesium Glycinate | 60 Pack',
'Glam By Manicare Kendall Lash | 1 pack',
'Glam By Manicare Eyelash Adhesive | 1 pack',
'Glam By Manicare Pre Glued Kelsey Lash | 1 pack',
'Glam By Manicare Sophia Lash | 1 pack',
'Swisspers Cotton Tips Cosmetic Dual Tips Paperstem | 100 pack',
'Swisspers Cotton Tips Paperstems | 20 pack',
'Swisspers Cotton Tips Paper | 240 pack',
'Coles Finest By Laurent White Sourdough Vi … 1 pack',
'Golden Circle Tropical Punch Fruit Drink Lunch Box Mu … 6 pack',
'Heaps Normal Half Day Hazy Can 375mL | 4 Pack',
'Finish Ultimate Plus Baking Soda Dishwasher … 64 pack',
'Finish Ultimate 0% Dishwasher Tablets | 34 pack',
'Radiant Laundry Capsules Active Clean | 28 pack',
'Great Northern Super Crisp Lager Block Can 375mL | 30 Pack',
'Carlton Dry Bottle 330mL | 24 Pack',
'Gage Roads Single Fin Summer Ale Can 330mL | 24 Pack',
'Great Northern Super Crisp Lager Bottle 330mL | 24 Pack',
"Gage Roads Pinky's Sunset Cider Can 330mL | 6 Pack",
'Heaps Normal Third IPA Can 375mL | 4 Pack',
'Heineken Bottle 330mL | 24 Pack',
'XXXX Gold Block Can 375mL | 30 Pack',
'Blackmores Multivitamins 50+ | 60 Pack',
'Blackmores Valerian Forte | 60PACK',
'Blackmores Cranberry Forte 50000Mg | 30 Pack',
'Coles Hot Cross Buns Fruit Free | 6 Pack',
'Soda Stream Gas Cylinder | 1pack',
'Quilton 3 Ply White Toilet Paper | 20 pack',
'Quilton White Toilet Tissue | 6 pack',
"Swisse Men's 65+ Ultivite Tablets | 60 pack",
'Victoria Bitter Block Can 375mL | 30 Pack',
'Coles Finest Luxurious Spiced Fruit Hot Cross Buns | 4 Pack',
'Coles Arnotts Iced Vovo Hot Cross Buns | 4 Pack',
'Coles Gluten Free Hot Cross Bun | 3 Pack',
'Coca-Cola Zero Sugar Soft Drink Multipack Cans 30x375mL | 30 Pack',
'Sprite Zero Sugar Lemonade Soft Drink Multipack Cans 1 … 10 pack',
"Blackmores Probiotics+ Women's Flora Balance | 30 pack",
'Coles Baby Cos Lettuce | 2 pack',
'Golden Circle Tropical Punch Fruit Drink Lunch Box Multipack Pop … 6 pack',
'Blackmores Vitamin B12 | 150 Pack',
'Coles Gluten Free Chocolate Hot Cross Buns | 3 Pack',
'Coles Spring Water 8x1.5L | 8 pack',
'Coca-Cola Classic Soft Drink Multipack Cans 30x375mL | 30 Pack',
"Swisse Ultivite Women's 65+ Tablets | 60 Pack",
'Life Space Gummies Probitoics Blueberry | 50 Pack',

'Corona Bottle 355mL | 24 Pack',
'Kleenex Complete Clean Toilet Paper | 32 Pack',
'Quilton King Size Gold 4 Ply Toilet Tissue | 20 pack',
'Quilton 3 Ply Floral Print Soft Double Length Toilet Paper | 9 pack',
'Ostelin Vitamin D | 250 Pack',
'Great Northern Original Lager Block 375mL | 30 Pack',
'Great Northern Original Lager Bottle 330mL | 24 Pack',
'Jack Daniels Cola Can 250mL | 4 Pack',
'Finish Ultimate Dishwashing Tablets Baking Soda | 50 pack',
'Quilton 3 Ply Double Length White Toilet Paper | 4 pack',
'Handee Ultra Paper Towels White | 3 pack',
'Victoria Bitter Bottle 375mL | 24 Pack',
'Asahi Super Dry Bottle 330mL | 24 Pack',
'Coca-Cola Zero Sugar Soft Drink Multipack Cans 24x375 … 24 pack',
'Sprite Zero Sugar Lemonade Soft Drink Multipack Ca … 10 pack',
'Schweppes Lemonade Soft Drink Cans Multipack 375 … 30 pack',
'Solo Thirst Crusher Original Lemon Soft Drink Cans M … 30 pack',
'Coca-Cola Zero Sugar Soft Drink Multipack Cans 30x375 … 30 Pack',
'Pepsi Max No Sugar Cola Soft Drink Cans Multipack 375 … 30 pack',
'Sunkist Orange Soft Drink Cans Multipack 375mL x 30 Pa … 30 pack',
'Nair Precision Easiwax Wax Strips For Face & Bikini | 20 pack',
'Max No Sugar Cola Soft Drink Cans Multipack 375mL x 24 Pack',
'Simply Thin BBQ Sausages 24 pack',
'Chocolate Easter Bunnies 3 Pack',
'Rocher Chocolate Gift Box 16 Pack',
'Collection Rocher Raffaello Rondnoir Chocolate Gift Box 15 Pack',
'Easter Mini Gold Bunny Chocolate 5 Pack',
'Rocher Chocolate Gift Box 30 Pack',
'Collection Rocher Raffaello Rondnoir Chocolate Gift Box 24 Pack',
'Rocher Chocolate Gift Box 5 Pack',
'Easter Figurines 6 Pack',
'Energised Soft Drink Multipack Cans 375mL x 24 Pack',
'Thick & Juicy Beef Burger 2 Pack',
'Max No Sugar Cola Soft Drink Cans Multipack 375mL x 10 Pack',
'Zero Sugar Original Lemon Soft Drink Cans Multipack 375mL x 24 Pack',
'Mini Dessert Selection 20 pack',
'Soda Water Bottle Classic Mixers Glass Multipack 300mL x 4 Pack',
'Lemon Lime Bitters Soft Drink Classic Mixers Multipack Bottles 300mL x 4
Pack',
'Lamb Burgers With Rosemary And Garlic 4 Pack',
'Indian Tonic Water Classic Mixers Glass Bottle Multipack 300mL x 4 Pack',
'Zero Sugar Lemon Lime Bitters Soft Drink Classic Mixers Multipack Bottles
300mL x 4 Pack',
'Simply Beef BBQ Burgers 10 pack',
'Agrum Blood Orange Soft Drink Bottles Glass Multipack 300mL x 4 Pack',
'French Raspberry Tart 2 pack',
'French Petits Fours 16 pack',

'Snack Crackers Mix 2 Multipack 12 Pack',
'Fun Mix Variety Multipack 20 Pack',
'Australian Still Spring Water Bottles Multipack 600mL x 24 Pack',
'Beef Bacon & Stout Party Pies 12 Pack',
'Snack Mix Variety Multipack 20 Pack',
'Dairy Natural Tasty Cheese Slices 30 Pack',
'Dairy Stringers Original Cheese 12 Pack',
'Dark Choc & Raspberry Roasted Chickpeas 5 pack',
'Roasted Chickpeas Dark Choc 5 Pack',
'Chocolate Roasted Chickpeas 5 Pack',
'Favva Beans Red Pepper & Chilli 6 Pack',
'All Butter Croissants 4 Pack',
'Roasted Seasoned Seaweed 8 Pack',
'Kitchen Mozzarella Bites 10 pack',
'Kitchen Sweet Corn Fritters 10 Pack',
'Milk Chocolate 8 pack',
'Raffaello Coconut And Almond Gift Box 15 Pack',
'Lemonade Soft Drink Bottle Glass Multipack 300mL x 4 Pack',
'Zero Sugar Mixers Indian Tonic Water Bottles Multipack 300mL x 4 Pack',
'Dry Ginger Ale Soft Drink Bottle Mixers Multipack 300mL x 4 Pack',
'Zero Sugar Dry Ginger Ale Soft Drink Mixers Multipack 300mL x 4 Pack',
'Bitter Lemon Soft Drink Classic Mixers Multipack 300mL x 4 Pack',
'Chicken Mushroom & Truffle Party Pies 12 Pack',
'Favva Beans Sea Salt & Vinegar 6 Pack',
'Multipacks Variety 10 Pack',
'Free Range Eggs 12 Pack',
'Frozen Meat Pies 4 Pack',
'Cheese Slices Tasty 24 Pack',
'Frozen Pies Meat Beef & Cheese 4 Pack',
'Mini Chocolate Chip Muffins 9 pack',
'Happy Hippo Cocoa Biscuit Bar Multipack 5 Pack',
'Dairy Milk Chocolate Freddo Sharepack 12 Pack',
'Tiny Teddy Multipack Variety Biscuits 15 Pack',
'Banana Bread Slices 5 Pack',
'Original Mini Wraps 8 pack',
'Blueberry Muffins 4 pack',
'Multipack Best Ever Butter 5 Pack',
'Chicken Potato Chips Multipack School Snacks 5 Pack',
'Pikelets 8 Pack',
'Cheese Chips Multipack Lunch Box School Snacks 6 Pack',
'Classic Fruit & Nut Muesli Bars 12 pack',
'Minis Sour Cream & Onion Potato Chips 5 pack',
'Almond Hazelnut & Vanilla 10 Pack',
'Angel Slice 6 pack',
'Tronky Biscuit Bars Multipack 5 Pack',
'Tasty Light Cheese Slices 24 Pack',
'Soft White Wraps 8 pack',

'Frozen Fish Fillets With Original Lightly Battered Oven Bake 6 Pack',
'Potato Chips Crinkle Flavour Mix Variety 20 Pack',
'Wholemeal & Grain Soft Wraps 8 Pack',
'Favourites Crackers Multipack Variety 15 pack',
'Dairy Milk Chocolate Sharepack 12 Pack',
'Frozen Fish Fillets With Lemon Pepper Crumb Oven Bake 6 pack',
'Baked Beans in Tomato Sauce Beans 3 Pack',
'Lamington Fingers 18 pack',
'Minis Multi Pack Original 5 Pack',
'Potato Chips Crinkle Cut Original 20 pack',
'Dairy Milk Chocolate Freddo Strawberry Sharepack 12 Pack',
'Mini Rolls Chocolate 5pack',
'Oat Slices Belgian Chocolate 5 Pack',
'Nutri Grain Original Bars 12 Pack',
'Dairy Milk Chocolate Caramello Koala Sharepack 12 Pack',
'Mini Scotch Finger 8 Pack',
'Mini Blueberry Muffins 9 pack',
'English Muffins 6 Pack',
'Popcorn Sea Salt 5 Pack',
'Frozen Bake At Home Butter Croissants 6 Pack',
'Stringers Original Cheese 8 pack',
'Cheese & Bacon Pizza Slices 2 Pack',
'Fresh Lasagne Sheets 8 Pack',
'Top 20 Variety Chips Multipack Lunch Box Snacks 20 Pack',
'Burrito Tortilla Wraps 10 pack',
'Dark Choc & Cranberry Gourmet Protein Bar 5 pack',
'Tasty Cheese & Crackers 4 Pack',
'Natural 25% Red Fat Tasty Slices 30 pack',
'Milk Chocolate Wafer Stick Cookies 10 Pack',
'Cranberry, Blueberry & Goji Super Berry Muesli Bars 6 pack',
'Tiny Teddy Multipack Hundreds & Thousands 8 Pack',
'Apple Poppers Fruit Drink 30% Juice Poppers Multipack Lunch Box Bottles 250mL
X 6 Pack',
'Shapes Cracker Biscuits Variety Multipack 15 Pack',
'BAGEL SESAME 4 PACK',
'Original Oats Quick Sachets 10 pack',
'Chocolate Chip Muffins 4 Pack',
'Popcorn Multipack 10 Pack',
'Minis Variety Bags Cookies 18 Pack',
'Potato Chips Multipack Lunch Box School Snacks Gluten Free 6 Pack',
'Dairy Dessert Choc A Lot 12 pack',
'Tasty Mix Potato Chips Variety Multipack of 20 Pack',
'Chocolate Mud Cake Slices 5 pack',
'Veggie Straws Salt & Vinegar 5 Pack',
'2 Minute Instant Noodles Wholegrain Chicken Flavour 5 Pack',
'Plant Protein Bar Raspberry Pistachio 5 Pack',
'Muesli Bars Strawberry Yoghurt Top 8 Pack',

'Double Chocolate Muffins 4 pack',
'Chocolate Slices 6 Pack',
'Protein Nut Bars Salted Caramel with Superseeds 5 Pack',
'Chips Multi 12 Pack',
'Apple Blackcurrant Fruit Drink 30% Juice Poppers Multipack Lunch Box Bottles
250mL X 6 Pack',
'Dark Choc Cherry & Coconut Bars 6 pack',
'Original Grab & Go Cookies 9 Pack',
'LCMs Split Stix Yoghurty 12 Pack',
'Cherry & Almond Bakewells Tarts 6 pack',
'Natures Kitchen Banana Bread Slices 5 Pack',
'Nut Butter Bars Peanut Butter & Salted Caramel Snacks 5 Pack',
"I'M Free From Blueberry Muffins 2 pack",
'Asia Hokkien Ready to Eat Noodles 2 Pack',
'Oat Slices Berry 5 Pack',
'Nut Bars Trail Mix Multipack Snacks 6 Pack',
'Mini Bites Rollin Salt & Vinegar 6 Pack',
'Crunchy Granola Bar Golden Oat & Honey 5 Pack',
'Cookies & Cream Protein Bars 5 Pack',
'Banana Muffins 9 pack',
'Rye Wraps 8 pack',
'LCMs Treatz Original with Marshmallows 5 pack',
'Variety Chips Multipack Lunch Box School Snacks 18 Pack',
'On The Go Tasty Cheese & Crackers 4 Pack',
'Original Mix Variety Multipack 18 Pack',
'Snack Rainbow Sticks 24 pack',
'Tiny Teddy Multipack Chocolate 8 Pack',
'Sea Salt Potato Chips Multipack School Snacks 5 Pack',
'Dairy Cheese Dippers 4 Pack',
'Facial Tissue Everyday 6X10 Pack',
'Jam Lamingtons 6 Pack',
'LCMs Rice Bubbles Kaleidos 12 Pack',
'Mini Crackers Original Flavour Multipack 8 Pack',
'Kids Yoghurt 12 Pack',
'Crumpet Rounds 6 Pack',
'Date & Walnut Slices 5 Pack',
'2 Minute Instant Noodles Oriental Flavour 5 Pack',
'Frozen Sausage Roll Jumbo 6 Pack',
'Cup A Soup Creamy Mushroom with Croutons 2 pack',
'White Large Jumbo Wraps 8 Pack',
'Mini Choc Chip Cookies 7 Pack',
'Cheese Stringers Lactose Free 8 Pack',
'Minis Frozen Yoghurt Strawberry & Vanilla Bean and Chocolate & Vanilla Bean 4
Pack',
'Choc Chip Banana Bread Slices 5 Pack',
'Sweet & Salty 10 Pack',
'Two Fruits In Strawberry Jelly 4 Pack',

'Muesli Bars Chocolate Chips 8 Pack',
'Shapes Cracker Biscuits Original Barbecue Multipack 8 Pack',
'Multipack Assorted Chips 15 Pack',
'Unicorn Fancies 8 Pack',
'Mini Bites Surfin Sea Salt 6 Pack',
'Greek Style Yoghurt & Berry Gourmet Protein Bars 5 pack',
'Salted Caramel Nut Butter Gourmet Protein Bars 5 pack',
'LCMs Variety 18 Pack',
'Variety Potato Chips Multi-Pack School Snacks 10 Pack',
'Classic Fruit & Nut Muesli Bars 6 pack',
'Salted Dark Choc & Almond Gourmet Protein Bars 5 pack',
'LCMs Rice Bubbles Choc Chip 12 Pack',
'Mini Cheese Original Protein 5 Pack',
'Frozen Bake At Home Pain Au Chocolat 6Pack',
'Protein Nut Bars Dark Choc & Cacao 5 Pack',
'Shapes Cracker Biscuits Original Pizza Multipack 8 Pack',
'Lemon & Raspberry Tarts 6 Pack',
'Fruit Salad In Syrup 4 Pack',
'Dark Choc Macadamia & Coconut Nut Bar 5 pack',
'Shapes Cracker Biscuits Original Chicken Crimpy Multipack 8 Pack',
'Protein Bars Violet Crumble 5 pack',
'Porridge Sachets Super Berry & Coconut 8 Pack',
'K-Time Baked Twists Raspberry & Apple Flavour Filled Snack Bars 5 pack',
'Microwavable Brown Rice & Quinoa Pouch 5pack',
'Coconut Yoghurt & Roasted Nut Gourmet Protein Bars 5 pack',
'Fruit Strings Multipack School Snacks Gluten Free 8 Pack',
'Original Fruit Free Muesli Bars 6 pack',
'Multipack Popped Rice Cakes Salt & Vinegar 6 Pack',
'Mini Pikelets 24pack',
'LCMs Treatz Choctastic with Marshmallows 5 pack',
'Chocolate Brownies 4 Pack',
'Gluten Free Bagels 4 Pack',
'LCMs Split Stix Variety 18 Pack',
'Popped Rice Cakes Cheddar Cheese 6 Pack',
'Nut Bars Peanut & Almond with Milk Choc Multipack Snacks 6 Pack',
'Mini Wrap Wholegrain 8 pack',
'Snacking Roasted Garlic Hommus 3 Pack',
'Nutri-Grain Original Bars 5 Pack',
'Smoothie Bars Berrylicious Blast 5 Pack',
'Snacking Zesty Lemon Hommus 3 Pack',
'Mini Bites Mixin Cheese 6 Pack',
'Fruit Oddities Multipack School Snacks Gluten Free 8 Pack',
'Variety Potato Chips Multipack School Snacks Gluten Free 10 Pack',
'K-Time Baked Twists Strawberry & Blueberry Flavour Filled Snack Bars 9 pack',
'Whole Wheat Mini Wraps 8 pack',
'K-Time Baked Apple Blueberry Twist 9 pack',
'LCMs Rice Bubbles Choc Chip 5 Pack',

'LCMs Rice Bubbles Kaleidos 5 Pack',
'Nut Butter Bars Peanut Butter Milk Choc Multipack Snacks 5 Pack',
'Multipack Mini Choc Chip Cookies 8 Pack',
'K-Time Baked Twists Strawberry & Yoghurt Flavour Filled Snack Bars 5 pack',
'Porridge Sachets Honey Vanilla & Cinnamon 8 Pack',
'Nut Bars Apricot with Dark Choc Multipack Snacks 6 Pack',
'Chicken Chips Multipack Lunch Box School Snacks 6 Pack',
'LCMs Rice Bubbles Unicorn 5 Pack',
'Mini Bites Rockin Chicken 6 Pack',
'Shapes Cracker Biscuits Vegemite & Cheese Multipack 8 Pack',
'Nut Bars Salted Caramel with Dark Choc Multipack Snacks 6 Pack',
'Smoothie Bars Raspberry Smash 5 Pack',
'Nut Bars Peanut & Cashew with Milk Choc Multipack Snacks 6 Pack',
'LCMs Split Stix Yoghurty 5 Pack',
'Diced Apricot Snack 6 pack',
'K-Time Bakery Favourites Apple Custard Pie Flavour Snack Bars 5 Pack',
'Diced Mango Snack Pack X 6 Pack',
'Fruit Fun Mix Variety Multi Pack School Snacks Gluten Free 16 Pack',
'K-Time Bakery Favourites Strawberry Cheesecake Flavour Snack Bars 5 Pack',
'Multipack Popped Rice Cakes Pizza 6 Pack',
'Crunch Milkshake Flavour Iced Coffee Limited Edition 8 Pack',
'Creamy Classics Choc Honeycomb Cone 4 Pack',
'Creamy Classics Choc Caramel Peanut Cone 4 Pack',
'Drumstix Crackers Multipack 8 Pack',
'Creamy Classic Vanilla Choc Fudge Cone 4 Pack',
'Stackers Leg Ham Cheese & Crackers 3 Pack',
'Chips Mini Rice Crackers Multi 12 Pack',
'Assorted Fruit Flavoured Snacks 8 pack',
'Creamy Classics Neapolitan Cone 4 Pack',
'Splits Multi Flavoured Ice Cream Sticks 10 pack',
'All Butter Croissants 2 Pack',
'Classic Chorizo 2 Pack',
'Crunch Vanilla Ice Cream 8 pack',
'Crunch Frozen Double Chocolate Ice Cream 8 Pack',
'Stackers Salami 3 Pack',
'Crunch Multi Flavoured Ice Cream Sticks 8 pack',
'Mini Tart Selection 6 pack',
'Splits Summer Splits 10 Pack',
'Splits Ice Cream Raspberry 10 pack',
'French Lemon Meringue Tart 2 pack',
'French Chocolate Tart 2 Pack',
'No Added Hormone Beef Porterhouse Steak 2 Pack',
'Beef Scotch Steak Fillet 2 Pack',
'Splits Ice Cream Selection 10 pack',
'RSPCA Approved Satay Chicken Kebabs 10 Pack',
'Tasmanian Salmon Portions Skin Off 4 Pack',
'Tasmanian Salmon Portions Skin On 2 Pack',

'Premium Angus Beef & Caramelised Onion Sausages 6 pack',
'Lamb Rump Steak 2 Pack',
'Medium Lemons 5 Pack',
'Graze Lamb Mini Shoulder Roast 2 Pack',
'Graze Lamb Leg Steak 2pack',
'Basa Portions Skin Off 2 Pack',
'Lamb Rump Roast 2pack',
'Free Range Large Eggs 12 pack',
'Free Range Extra Large Eggs 12 pack',
'Free Range Big Brekkie Browns Eggs 12 pack',
'Free Range Eggs 12Pack',
'Simply Sandwich Slices 24 Pack',
'Cage Free Eggs 18 Pack',
'Fermented Light Probiotic Milk 5 Pack',
'Gluten Free Fermented Milk Drink 10 pack',
'Free Range X Large Eggs 6 pack',
'Free Range Natural Living Extra Large Eggs 12 pack',
'Organic Free Range Eggs 12 Pack',
'Tasty Natural Cheese Slices 15 pack',
'Cage Free Eggs 12 Pack',
'Free Range Large Eggs 18 pack',
'Cage Free Barn Laid Eggs 12 pack',
'Organic Eggs 12 pack',
'Simply Light Sandwich Slices 24 Pack',
'Custard Snack Pack Vanilla 6 Pack',
'Thick & Creamy Yoghurt Mango Passionfruit 4 pack',
'Jumbo Free Range Eggs 12 Pack',
'Free Range Eggs 6 pack',
'Free Range Eggs Regular 6 Pack',
'Simply Smokey Flavoured Sandwich Slices 12 Pack',
'Regular Greek Yoghurt 4 pack',
'Creme Caramel Dessert 2 pack',
'Liberty Cage Free Eggs 12 Pack',
'Simply Sandwich Slices 84 Pack',
'Bread Slices Garlic 9 Pack',
'Natural Living Eggs Free Range 12 pack',
'Free Range Eggs 18 Pack',
'Natural Living Free Range Large Eggs 12 pack',
'Cage Free Barn Laid Eggs 12 Pack',
'Fermented Probiotic Milk 5 pack',
'Lactose Free Plain Yoghurt 4 pack',
'Dairy Cream Cheese Original Block 2 Pack',
'Vintage Cheese Slices 12 Pack',
'Plant Based Chicken Schnitzel 2 Pack',
'Pie Beef Pepper 2 Pack',
'Natural Dairy Cheese Slices 12 Pack',
'Sweet Corn Fritters 6 Pack',

'Peninsula Cage Free Eggs 12 Pack',
'Organic Free Range Large Eggs 6 pack',
'Multipack Full Cream Milk 6 Pack',
'Farmers Tasty Cheese & Crackers Multi Pack 4 Pack',
'Dairy Colby Cheese Slices 12 Pack',
'Party Pies 12 pack',
'Thai Spiced Fritters 6 Pack',
'Pizza Base 12 pack',
'Pies Butter Chicken 2 pack',
'Plain Roti 4 Pack',
'Greek Style Yoghurt 6 Pack',
'Vintage Cheese & Crackers 4 Pack',
'Dairy Tasty Cheese & Crackers 10 Pack',
'Sausage Rolls 12 pack',
'Raspberry Cheesecake 2 Pack',
'Low Fat Passionfruit Yoghurt 4 Pack',
'Plant Based Chicken Tender Swt Chilli 5 Pack',
'Double Chocolate Dessert 2 Pack',
'Roasted Garlic Naan 2 pack',
'Multipack Skim Uht Milk 6 Pack',
'Pie Chunky Beef 2 Pack',
'Parsley Sage & Rosemary Sausages 6 Pack',
'Blue Cheese Mini 5 Pack',
'Fresh Spring Onion Pancake 2Pack',
'Chocolate Dairy Dessert 12 Pack',
'Biteables Loaded Potato Rostis 6 Pack',
'Cage Free Large Eggs 18 pack',
'Everything 4 Pack',
'Brioche Sliders 10 Pack',
'Chocolate Chip Brioche Rolls 8 Pack',
'Split Top Brioche Hot Dog Rolls 6 Pack',
'Pain Au Chocolate 4 Pack',
'Belgian Cheesecake Slice 2 pack',
'Low Carb High Protein + Fibre 6 pack',
'Mini Pain Au Choc 8 pack',
'All Butter Croissants 3 Pack',
'Turkish Rolls 4 pack',
'Croissants 12 Pack',
'Lemon Slices 6 Pack',
'Macarons Selection 10 pack',
'Original Wraps 8 pack',
'Vanilla Slice 2 Pack',
'White Dinner Rolls 6 pack',
'Mixed Grain Wraps 8 pack',
'Brioche Hot Dog Rolls 4 Pack',
'Traditional Butter Waffles 4 pack',
'Crumpet Square 6 pack',

'Gourmet Bun 4 Pack',
'Tofu Tots Ginger Soy 6 Pack',
'Barn Cage Free Large Eggs 18 pack',
'Neenish Tarts 4 Pack',
'High Protein Souvlaki 5 Pack',
'Wholemeal Wraps 8 pack',
'White Wraps 8 pack',
'Bagels Plain 4 pack',
'Preservative Free Premium White Large Wraps 8 Pack',
'Pita Khobz White 5 pack',
'Spinach & Herb Wraps 8 pack',
'Preservative Free Mixed Wholegrain Large Wraps 8 Pack',
'Natural 4 Pack',
'Salted Caramel Slice 6 Pack',
'New York Style Baked Cheesecake Slice 2 Pack',
'Brioche Hamburger Buns 4 Pack',
'Brioche Rolls 8 Pack',
'Brioche Burger Buns With Sesame Seeds 4 Pack',
'Chocolate Cake Slices 2 pack',
'Pita Khobz Wholemeal 5 pack',
'Spinach & Garden Herb Wraps 7 pack',
'Pita Pockets Wholemeal 8 pack',
'Low Carb 4 Pack',
'Cafe Style Chocolate Chunk Muffin 2 pack',
'Chocolate Mini Cupcakes 9 Pack',
'Soft & Fluffy Wrap 12 pack',
'White Gourmet Pizza Base 2 pack',
'Thins Original 6 Pack',
'High Protein Low Carb Wholemeal Wrap 5 Pack',
'Cafe Style Blueberry Oat Muffin 2 pack',
'Sesame 4 Pack',
'Premium Mini Pizza Bases 6 Pack',
'Perform High Protein Wholemeal & Quinoa Wraps 8 Pack',
'Low Carb Spinach And Herb 5 pack',
'Wholewheat Pita Pocket 4 Pack',
'Low Carb High Protein Pizza Base 2 pack',
'Thins Mixed Grains 6 Pack',
'Lemon Cheesecake Slices 2 Pack',
'Roasted Garlic Wraps 8 pack',
'Pizza Base Pane 2 Pack',
'Pita Pockets White 8 pack',
'Natural Wraps 8 Pack',
'Snow Drops 8 Pack',
'Low Carb Keto 5 pack',
'Gluten Free Super Soft Sliced White Rolls 4 Pack',
'Wraps Low Gi Wholemeal 6 Pack',
'High Protein Pizza Base 2 Pack',

'Tart Shells 6 pack',
'Natures Kitchen Vegan Chocolate Cupcake 2 Pack',
'Gluten Free Crumpets 4 Pack',
'Vanilla Mini Cupcakes 9 Pack',
'Mini Rolls Raspberry 5 Pack',
'Meringue Nests 8 Pack',
'Gluten Free White Rolls 2 Pack',
'Sourdough Wraps 8 Pack',
'Vegan Brioche Burger Bun 4 Pack',
'Carb Balance Wrap 6 pack',
'Trancetto Cacao Cakes 10 pack',
'Sponge Mini Jam Rolls 6 pack',
'Ginger Kisses Cream Filled Sponge Cake 8 pack',
'Caramel Bakewell 6 Pack',
'Low Carb Mini 6 pack',
'70% Lower Carb English Muffins 4 Pack',
'Carrot Cupcakes 4 Pack',
'Chocolate Cake Slices 4 pack',
'French Fancies Selection Cakes 8 Pack',
'Gluten Free Brioche Style Rolls 2 Pack',
'Soft & Fluffy Butter Wrap 12 pack',
'Low Carb High Protein 5 pack',
'Vegan Croissant 3 Pack',
'Nap Mini Wraps Value Pack 12 Pack',
'Mix Max Cocoa Cakes 10 pack',
'Thins Wholemeal 6 Pack',
'No Artificial Preservatives Wholemeal Mini Wraps 12 Pack',
'Pancakes 6 Pack',
'Sponge Roll Mini Choc 6 pack',
'Cinnamon Buns 12 pack',
'Wholegrain Wraps 8 pack',
'Cupcakes Red Velvet 4 Pack',
'Choco Chip Waffles 4 pack',
'Flourless Orange Slice 2 Pack',
'Fantasy Cupcakes 4 Pack',
'Vegan Chocolate Macaroons 6 Pack',
'Lite Wrap 8 pack',
'Gluten Free Soft Sliced Brioche Rolls 4 Pack',
'Mini Pizza Base 4 Pack',
'Multigrain Dinner Rolls 6 pack',
'Cupcakes Chocolate 4 Pack',
'Salt Reduced Wraps 8 pack',
'Vol Au Vent 4 Pack',
'Premium Stone Baked Garlic And Herb Pizza 2 Pack',
'Premium Mini Pizza Bases Garlic And Herb 6 Pack',
'Dairy Milk Freddo Cupcakes 6 Pack',
'Sticky Toffee Cake Slices 4 pack',

'Mini Brownies Assorted 9 pack',
'Choc Pikelets 8pack',
'Gourmet Wholemeal Pizza Bases 2 pack',
'Complete Protein Buns 2 Pack',
'2 Minute Instant Noodles Chicken Flavour 5 Pack',
'Mixed Pate Pods 4 pack',
'Shin Ramyun Noodle 5 Pack',
'Mi Goreng Hot & Spicy Instant 5 Pack',
'Buldak Hot Chicken Ramen Carbonara 5 Pack',
'Buldak Hot Chicken Ramen Original 5 Pack',
'Pure Ceylon Premium Tea Bags 100 pack',
'English Breakfast Tea Bags 100 pack',
'Neon Food Colours 4 Pack',
'2 Minute Instant Noodles Chicken Flavour 12 Pack',
'Noodle Shrimp Tomyum 5 Pack',
'Shin Ramyun Black 4 pack',
'Pure Camomile Infusions Tea Bags 80 pack',
'Kimchi Ramyun Noodle Soup 5 Pack',
'Pure Green Tea Bags 100 pack',
'Regular Tortillas Fajita 16 Pack',
'2 Minute Instant Noodles Beef Flavour 5 Pack',
'Pure Ceylon Extra Strength Tea Bags 100 pack',
'Green Jasmine & Pear Tea Bags18 pack',
'Ramen Tokyo Shoyu Noodles 5 Pack',
'2 Minute Instant Noodles Spicy Chicken Flavour 5 Pack',
'Ramen Spicy 5 Pack',
'Croissant With Chocolate Cream Filling 6 Pack',
'Canned Regular 4 Pack',
'Fold & Roll Tortilla Wraps 6 Pack',
'English Breakfast Extra Strong Tea Bags 80 pack',
'Earl Grey Tea Bags 100 pack',
'Regular Tomato Paste 2 Pack',
'Buldak Hot Chicken Ramen Spicy 5 Pack',
'Tomato Paste No Added Salt 4 pack',
'Instant Noodle Sesame Oil 5 Pack',
'Lemon & Ginger Infusions Tea Bags 80 pack',
'Baked Beans in Tomato Sauce Beans 3 pack',
'Taco Shells 12 Pack',
'Tortillas Mini Tacos 10 Pack',
'Diced Peaches in Syrup 4 Pack',
'Stand N Stuff Taco Shell 10 Pack',
'Indian Roti Plain 6 Pack',
'Tomato Paste No Added Salt 2 pack',
'Regular Tortillas Fajita 10 Pack',
'Fusian Mi Goreng Instant Noodles Hot & Spicy 5 Pack',
'Jumbo Tortillas Burrito 6 Pack',
'Hokkaido Miso Noodles 5 Pack',

'2 Minute Beef Flavour Noodles 12 Pack',
'Gluten Free Jumbo Tortillas Burrito 4 Pack',
'Ramen Mild 5 Pack',
'Tomato Paste Sachet 4 Pack',
'Matcha Traditional Green Tea Bags 20 Pack',
'Baked Beans in Tomato Sauce Beans 4 pack',
'Peach In Mango Flavoured Jelly 4 Pack',
'Yorkshire Black Tea Bags 100 pack',
'Street Tortilla Wraps 10 Pack',
'French Earl Grey Flavoured Black Tea Bags 25 pack',
'Corn Jumbo Taco Shells 10 pack',
'Noodle Pork 5 Pack',
'Pizza Tomato Paste 2 Pack',
'Extra Strength Premium Ceylon Tea Bags 50 Pack',
'Chamomile & Spearmint Herbal Infusions Tea Bags 10 pack',
'Assam Bold Tea Bags 80 pack',
'Ramen Seafood Party 5 Pack',
'Wakame Instant Miso Soup 12 pack',
'Roasted Seasoned Seaweed 2 Pack',
'Indian Naan Garlic 3 Pack',
'Extra Strength Pure Ceylon Bags 200 Pack',
'Seed Beans Four Bean Mix 4 Pack',
'Variety Pack assorted breakfast cereals sachets 8 pack',
'Sticky Date Steamy Pud 2 Pack',
'Regular Tortillas Fajita 6 Pack',
'Super Sweet Corn Kernels 4 Pack',
'Tea Bags 100 pack',
'Irish Breakfast Tea Bags 100 pack',
'Pure Green Tea Bags 50 pack',
'Premium Coconut Milk Powder 3 Pack',
'Pure Ceylon Premium Tea Bags 50 Pack',
'Meal Replacement Variety 14 Pack',
'Pure Green Tea Bags 10 pack',
'Protein Crisp Bars Choc Peanut 4 Pack',
'Noodle Chicken 5 Pack',
'Dolce Gusto Cafe Au Lait Coffee Capsules 16 Pack',
'Jelly Sticks 20 Pack',
'Canned Baked Beans Salt Reduced 4 Pack',
'Solid Light Tuna In Olive Oil 2 Pack',
'Porridge Sachets Protein Rich Almond Vanilla & Cinnamon 6 Pack',
'Pure Peppermint Herbal Infusions Tea Bags 40 pack',
'Chopped Tomatoes 4 Pack',
'Dairylea Cheestiks Cheese 12 pack',
'Microwavable Brown Rice 5 Pack',
'Melbourne Breakfast Flavoured Black Tea Bags 25 pack',
"Stand 'N Stuff Flour Soft Tortillas 8 Pack",
'Protein Crisp Bars Vanilla Caramel 4 Pack',

'Japanese Instant Miso Soup Fried Tofu 2 Pack',
'Slim 7 Day Variety 14 Pack',
'Sardines IN Water 4 Pack',
'Fusian Mi Goreng Instant Noodles Soy & Mild Spice Flavour 5 Pack',
'Hazelnut Chocolate Spread 12 pack',
'Tortilla Pockets 8 Pack',
'Decaffeinated Black Tea Bags 100 pack',
'English Breakfast Tea Bags 10 pack',
'Classic Chai Latte Tea Sachets 8 pack',
'Mandarin In Orange Flavoured Jelly 4 Pack',
'Lite Raspberry Jelly Crystals 2 pack',
'Night Time Tea Sachets 20 pack',
'Mixed Mini Pudding 25 Pack',
'Chai Latte 10 Pack',
'Mini Toasty Plain Naan 5 Pack',
'Artificial Food Colours 8 pack',
'Indian Naan Plain 3 Pack',
'English Breakfast Tea Bags 50 pack',
'Cocktail Frankfurts 25 pack',
'All Rounder Pure Black Tea Bags 200 pack',
'Gooey Caramel Steamy Puds 2 Pack',
'White Corn Tortillas Tacos 10 Pack',
'Low Carb Caramel Chocolate Nut Roll Bars 5 Pack',
'Porridge Sachets Natural 5 Grain & Super Seed 8 Pack',
'Espressotoria Espresso Coffee Capsules 12 pack',
'Mandarin Segments in Syrup 4 Pack',
'Raspberry Jelly Cups 4 Pack',
'Gluten Free Original Sesame Snaps 3 pack',
'Dolce Gusto Cappuccino Capsules 16 pack',
'Lite Strawberry Jelly Crystals 2 pack',
'Pears In Blackcurrant Jelly 4 Pack',
'Three Mint Herbal Tea Sachets 20 pack',
'Fruit Jelly Mini Cups 20 pack',
'All Natural Food Colours Rainbow 4 Pack',
'Chocolate Steamy Puds 2 Pack',
'Caramel Milk Chocolate Wafers Biscuits 8 Pack',
'Lite Port Wine Jelly Crystals 2 pack',
'Plum Pudding 2 Pack',
'2-minute Mug Instant Noodles Chicken 4 pack',
'Lemon Ginger & Manuka Honey Tea Sachets 20 pack',
'Mini Toasty Garlic And Herb Naan 5 Pack',
'Australian Afternoon Tea Bags 10 pack',
'Pineapple In Syrup 4 Pack',
'Spaghetti in Tomato Sauce Cheese Pasta 3 pack',
'Natural Green Tea Bags 50 pack',
'Snack Bar Chococalte Crunch 5 Pack',
'Sesame Snaps 3 Pack',

'Spice Chai Tea Sachets 20 pack',
'Spaghetti in Tomato Sauce Pasta 3 pack',
'Quelch Apple Berry Fruit Juice Icy Tubes 24 pack',
'Cup Or Pot Tea Bags 200 pack',
'Porridge Sachets Apple Sultana & Cinnamon 8 Pack',
'Low Carb Tortilla 5 Pack',
'Oat Cups Protein Porridge Creamy Honey 2 Pack',
'Caffeine Free Lemongrass Lime & Ginger Tisane Tea Bags 18 pack',
'Earl Grey Tea Bags 10 pack',
'Garlic Herb Naan 4 Pack',
'Vanilla Bean Pods 2 Pack',
'Quelch Tropical Fruit Juice Icy Tubes 24 pack',
'Organics Pure Chai Tea Bags 50 pack',
'Pure Ceylon Green Tea Bags 20 pack',
'Canned Rich Tomato 4 Pack',
'Lemongrass & Ginger Herbal Tea Bags 25 Pack',
'Lite Pineapple Jelly Crystals 2 pack',
'Premium Blend Tea Bags 100 pack',
'Organic Lemon & Ginger Tea Bags 40 pack',
'Jelly Cups Strawberry 4 Pack',
'Rainbow Food Colour 4 pack',
'Gooey Golden Syrup Steamy Puds 2 Pack',
'Organics Pure Green Tea Bags 50 pack',
'Quick Oats Honey Flavoured Sachets 12 Pack',
'Lite Orange Mango Jelly Crystals 2 pack',
'Writing Icing Multipack Colour Tubes 5 pack',
'Lite Lime Jelly Crystals 2 pack',
'Plain Naan 4 Pack',
'Earl Grey Tea Bags 50 pack',
'Diced Pears In Syrup 4 Pack',
'Pure Ceylon Jasmine Flavour Green Tea Bags 20 pack',
'Herbal Collection Tea Bags 20 pack',
'Organics Peppermint Tea Bags 40 Pack',
'Bengal Spice Tea Bags 20 pack',
'Snacks Nourish Balls Peppermint Choc Chip 4 pack',
'Espressotoria Italian Blend Coffee Capsules 12 pack',
'Dandy Chai Blend Tea Bags 32 pack',
'Pot Tea Bags 100 Pack',
'Premium Pure Ceylon Tea Bags 200 pack',
'Rooibos Tea Bags 40 Pack',
'Lite Vanilla Berry Jelly Crystals 2 pack',
'Dark Chocolate Almond High Protein Bar 4 Pack',
'Endulge Peanut Butter Cups 10 Pack',
'Lite Mango Passionfruit Jelly Crystals 2 pack',
'Creamy Rice Apple Cinnamon 2 Pack',
'Oat Cups Protein Porridge Berry 2 Pack',
'Chocolate Brownie Crunch High Protein Bar 4 Pack',

'Baked Beans In Tomato Sauce 3 pack',
'Green Tea Bags 50 pack',
'Canned With Cheese 4 Pack',
'Orange Jelly Cups 4 Pack',
'Spagasaurus 4 Pack',
'High Protein Bar Raspberry Truffle 4 Pack',
'Earl Grey Tea Bags 50 Pack',
'Berry Cashew & Cacao Smooshed Wholefood Balls 6 pack',
'Shelf Fresh Noodles Hokkien Style 4 pack',
'Camomile Herbal Infusion Tea Bags 40 pack',
'Creamy Rice Creme Brulee 2 Pack',
'Dolce Gusto Espresso Intenso Coffee Capsules 16 pack',
'Lemon & Ginger Infusions Tea Bags 40 pack',
'Decaffeinated Green Tea Bags 50 pack',
'Snack Bar Choc Caramel 5 Pack',
'Peanut Butter & Caramel Smooshed Wholefood Balls 6 pack',
'Soft Taco Kit 10 Pack',
'Mini Plain Naan 6 Pack',
'Spaghetti in Tomato Sauce Pasta 4 pack',
'Club Milk 5 Pack',
'Honey Soy Chicken 5 Pack',
'Dolce Gusto Americano Coffee Capsules 16 pack',
'Canned Ham Sauce 4 Pack',
'Snacks Nourish Balls Choc Hazelnut 4 pack',
'Malabar Paratha 4 Pack',
'2 Minn Curry 5 Pack',
'Protein Classic Beef Pie 2 Pack',
'Protein Beef & Cheese Pie 2 Pack',
'Choc Coated Vanilla Flavoured Sticks 12 Pack',
'White Tea Bags 25 pack',
'Sprinkles 5 In 1 Pack',
'Phulka Roti 30 pack',
'Gluten Free Fish Fillets Crumb 4 pack',
'Vegetable Cocktail Spring Rolls 60 Pack',
'Low Carb High Protein Salted Caramel Bars 4 Pack',
'Vegetable Spring Rolls 16 Pack',
'Peanut Butter Sandwichs Ice Cream 4 Pack',
'Twin Pack Peanut Butter & Chocolate 4 Pack',
'Collagen 50mLX10 Pack',
'Frozen Party Sausage Rolls 12 pack',
'Protein Bars Dark Choc 5 Pack',
'Frozen Chocolate Lava Cake 2 pack',
'Frozen Multi Flavoured Yoghurt Sticks 8 pack',
'Fruit Collection Ice Cream Mini Cups 4 Pack',
'Cookies & Creme Sticks 4Pack',
'Wholemeal Roti Bread 15 Pack',
'Frozen Variety Minis Yoghurt 14 pack',

'Sweet Mint Sugar Free Chewing Gum 3 Pack',
'Ice Cream Sticks 4 Pack',
'Irresistible Ice Cream Mini Almond 6 Pack',
'Low Carb High Protein White Choc Raspberry 4 Pack',
'Blueberry Sugar Free Chewing Gum 3 Pack',
'Protein Slices Pepperoni Pizza 2 Pack',
'Tandoori Naan 16 Pack',
'Frozen Shortcrust Pastry Sheets 5 pack',
'Low Carb High Protein Cookies & Cream Bars 4 Pack',
'Frozen Puff Pastry Sheets Extra Light 6 pack',
'Protein Bakes Salted Caramel 5 Pack',
'Irresistible Salted Caramel 4 Pack',
'Ice Cream Dairy Free Classic 3Pack',
'Frozen Steam Fresh Fish Fillets With Thai Red Curry Sauce 2 Pack',
'Plant Based Chicken Style Burgers 4 Pack',
'Sugar Free Pastilles Forest Berries 3 pack',
'Ice Cream Dairy Free Almond 3Pack',
'Tandoori Garlic Naan 16 Pack',
'Ice Cream Choc Bars Vanilla 8Pack',
'Fruit Apple Pie Snack 4 pack',
'Protein Vege Supreme Pizza Slices 2 Pack',
'Iced Coffee Protein Bars 5 Pack',
'Protein Slice Bar Choc Peanut 5 Pack',
'Sticks Peanut Butter Cups 4 Pack',
'Irresistable Boysenberry Sticks 4 Pack',
'Frozen Shortcrust Pastry 5 pack',
'Sultanas Snack 6 Pack',
'Sticks Blackforest 4 Pack',
'Irresistible Almond Ice Cream 4 Pack',
'Irresistible Vanilla Ice Cream 4 Pack',
'Frosty Fruits Watermelon Ice Blocks 8 Pack',
'Sticks Boysenberry Cheesecake 4 Pack',
'Sticks Salted Caramel 4 Pack',
'Plant Based Schnitzels 4 Pack',
'Roti Paratha Bread 6 Pack',
'Choc Fudge Ice Cream 8Pack',
'Irresistible Ice Cream Mini Classic Vanilla 6 Pack',
'Golden Gaytime Plant Based 4 pack',
'Active Protein Mixed Berry Multipack School Snacks 5 Pack',
'Drumstick Plant Based Classic Vanilla Cones 4 Pack',
'Mocha Protein Bars 5 Pack',
'Caramel Fudge Protein Bar 5 Pack',
'Active Protein Chocolate Chips Multipack School Snacks 5 Pack',
'Fudge Bars Caramel 8 Pack',
'Garlic Naan 4 Pack',
'Plant Based Vanilla Brownie Ice Cream 4 Pack',
'Frozen Spring Roll Pastry Sheets 20 pack',

'Creamy Classics Vanilla Sandwich 4 pack',
"I'M Free From Tempura Prawns 10 Pack",
'Lemonade Ice Poles 8 Pack',
'Samosas & Puffs 16 Pack',
'Creamy Classic Sandwich Neapolitan 4 pack',
'99% Sugar Free Caramelised White Choc 10 pack',
'Lemonade 10 Pack',
'Bake At Home Mini Cinnamon Swirl 6 Pack',
'Natures Kitchen Hearty Mushroom & Lentil Pies 2 pack',
'Natures Kitchen Cheesy Cauliflower Pies 2 pack',
'Animal Shaped Egg Custard Bao 6 Pack',
'No Sugar Mints Spearmint 3 Pack',
'Ice Potong Sweet Corn 6 pack',
'Tropical 10 Pack',
'Golden Blond Chocolate Mochi Bites 6 pack',
'Bake At Home Maple Pecan Plait 4 Pack',
'Bake At Home Chocolate Scrolls 4 Pack',
'Organic Salted Caramel Frozen Dessert 4 pack',
'No Sugar Mints Peppermints 3 Pack',
'Easter Bunnies 12 Pack',
'Chocolate Fudge Brownies 5 pack',
'Twirl Chocolate Sharepack 12 Pack',
'Picnic Chocolate Sharepack 12 Pack',
'Biscoff Biscuits 8X2 pack',
'Breakfast Biscuits Chocolate 6 Pack',
'Surprise Chocolate 3 pack',
'Crunchie Chocolate Sharepack 12 Pack',
'Flake Chocolate Sharepack 12 Pack',
'Healthier Choice Go Ahead Biscuits Forest Fruits 4 Pack',
'Cherry Ripe Chocolate Sharepack 12 Pack',
'Salted Caramel Squares 5 pack',
'Greek Style Yoghurt & Blueberry Nut Bars 5 pack',
'Peanut Butter Cups Milk Chocolate 2 pack',
'Oat Slices Golden Oat And Coconut 5 Pack',
'Healthier Choice Go Ahead Biscuits Apple 4 Pack',
'Multigrain Muesli Bars Triple Chocolate 5 Pack',
'Best Of Mini Tube 50 Pack',
'Breakfast Biscuits Strawberry Yoghurt Sandwich 5 Pack',
'Healthier Choice Go Ahead Biscuits Raspberry 4 Pack',
'Twirl Chocolate Bar 4 Pack',
'Grape And Strawberry 4Pack',
'Microwave Popcorn Triple Butter 6 Pack',
'Multigrain Muesli Bars Honey Yoghurt 5 Pack',
"Fry's Turkish Delight Chocolate Sharepack 12 Pack",
'Fruit Blocks Multipack Lunch Box School Snacks Gluten Free 6 Pack',
'Dairy Milk Chocolate Freddo Milky Top Sharepack 12 Pack',
'Breakfast Biscuits Honey & Nut With Chocolate Chips 6 Pack',

'Nut Bars Coconut with Milk Choc Multipack Snacks 6 Pack',
'Multigrain Muesli Bars Milk Chocolate Strawberry 5 Pack',
'Australian Sultanas 6 pack',
'Breakfast Biscuits Milk & Cereals 6 Pack',
'Multigrain Muesli Bars Cookies & Cream 5 Pack',
'Joyful Choc Brownie Baked Bars 5 Pack',
'Fruit 3 pack',
'Simply Cheese Spread Snack 12 Pack',
'Honey & Lemon Flavour Throat Lozenges 3x10 Multipack',
'Xtra Milk Chocolate Bar With Caramel Biscuit 2 Pack',
"Mix Of Mini's 35 Pack",
'Mini Berry Choc Top Cookies 10 Pack',
'Nut Bars Yoghurt Multipack Snacks 6 Pack',
'Joyful Berry & White Choc Baked Bars 5 Pack',
'Nut Bars Honeycomb with Milk Choc Multipack Snacks 6 Pack',
'Microwave Popcorn Butter 3 Pack',
'Caramilk Wallaby Chocolate Sharepack 12 Pack',
'Nut Bars Mixed Berry Multipack Snacks 6 Pack',
'Dark Choc Espresso Nut Bar 5 pack',
'Cream Candies 3 pack',
'Fruit Filled Mini Bars Stawberry 10 Pack',
'Nut Bars with Milk Choc Multipack Snacks 6 Pack',
'Peanut Butter Cup White 2 pack',
'Max No Sugar Cola Soft Drink Cans Multipack 375mL x 30 Pack',
'Zero Sugar Blood Orange Infused Mineral Water Cans Multipack 375mL X 10 Pack',
'Thirst Crusher Original Lemon Soft Drink Cans Multipack 375mL x 10 Pack',
'Zero Sugar Original Lemon Soft Drink Cans Multipack 375mL x 10 Pack',
'Thirst Crusher Original Lemon Soft Drink Cans Multipack 375mL x 30 Pack',
'Lemonade Soft Drink Cans Multipack 375mL x 30 Pack',
'Sultana & Apple 6 Pack',
'Cola Soft Drink Cans Multipack 375mL x 30 Pack',
'Orange Soft Drink Cans Multipack 375mL x 30 Pack',
'Max No Sugar Cola Soft Drink Mini Cans Multipack 275mL x 6 Pack',
'Instant Hot Chocolate 6 pack',
'99% Sugarfree Drinking Sachets Chocolate 10 Pack',
'Cola Soft Drink Cans Multipack 375mL x 10 Pack',
'Zero Sugar Lime & Lemon Infused Natural Mineral Sparkling Water 375mL x 10 Pack',
'Zero Sugar Raspberry & Blueberry Infused Natural Mineral Sparkling Water 375mL x 10 Pack',
'Lemonade Multipack Mini Cans 275mL x 6 Pack',
'Cola Soft Drink Mini Cans Multipack 275mL X 6 Pack',
'Lime Flavour Sparkling Water No Sugar Multipack Cans 375mL x 8 Pack',
'Hot Dark Chocolate 8 pack',
'Mixers Ginger Beer 4 Pack',
'Hot Choc Mint 10 pack',
'Apple Fruit Juice Tetra 6 Pack',

'Apple Mango Juice Tetra 6 Pack',
'Melon Ice Bar 4 Pack',
'Vanilla Choc 4 Pack',
'Golden Gaytime Ice Cream Sticks 4 pack',
'Tropical Juice Tetra 6 Pack',
'Red Rooster Chicken Pie 4 Pack',
'Frozen Party Classic Beef Pies 12 pack',
'Frozen Chicken & Vegetable Pies 4 Pack',
'Boysenberry Sticks 4 Pack',
'Dairy Milk Vanilla Mini Ice Cream 6 Pack',
'Caramel Choc 4 Pack',
'Ice Cream Mini Classic 6Pack',
'Dairy Milk Caramello Mini Ice Cream 6 Pack',
'Paddle Pop Twister Mini Tropical 6 Pack',
'Ice Cream Almond 4Pack',
'Ice Cream Classic 4Pack',
'Melona Mango Ice Bar 4 Pack',
'Strawberry & Chocolate Milkshake Pops 8 Pack',
'Ice Cream Mini Almond 6Pack',
'Drumstick Minis Classic Vanilla Cones 6 Pack',
'Frozen Beef Party Pies 24 Pack',
'Frozen Gourmet King Island Beef In Rich Gravy Pie 2 pack',
'Drumstick Boysenberry Swirl Cones 4 Pack',
'Frosty Fruits Fruit Stack Sorbet 6 Pack',
'Ice Cream Mini Bars Mango 6 Pack',
'Murray River Salted Caramel & Macadamia Ice Cream 4 Pack',
'Ice Cream Peppermint 4Pack',
'Frozen Mini Combo Party Pies & Sausage Rolls 40 pack',
'Ice Cream Honeycomb Crunch 4Pack',
'Frozen Chicken & Leek Pie 2 pack',
'Ice Cream White 4Pack',
'Party 30 Pack',
'Mango & Macadamia Ice Cream Bars 4 pack',
'Classic Vanilla Ice Cream 4 Pack',
'Kitchen Mini Chicken Kyiv 8 pack',
'Maxibon Vanilla Slice Ice Cream 4 Pack',
'AFL Party Sausage Rolls 24 Pack',
'Frosty Fruits Tropical Ice Blocks 8 Pack',
'Frozen Dessert Coco Crunch 4 Pack',
'Seasonal Garden Vegetable Pastie 4 pack',
'Frozen Gourmet Chunky Sous Vide Beef In Rich Gravy Pies 2 pack',
'Bueno White Frozen Dessert Cones 4 Pack',
'Sicilian Blood Orange & Chocolate Ice Cream 4 Pack',
'Frozen Deli New Zealand Whiting Cracked Pepper Blend & Sea Salt Fillets 5
Pack',
'Frozen Dessert Cookie Crunch 4 Pack',
'Ice Cream Cones 4 Pack',

```
    'Creamy Ice Bar 5 Pack',
    'Puffy Dogs 20 Pack',
    'Kitkat Ice Cream 4 Pack',
    'Calippo Minis Lemon Orange & Lime Ice Blocks 10 pack',
    'Drumstick Saucy Strawberry Cones 4 Pack',
    'Drumstick Super Choc Cones 4 Pack',
    'Calippo Mini Ice Raspberry Pineapple Ice Cream 10 pack',
    'Drumstick Ice Cream Mega Pack Vanilla 16 pack',
    'Mango Ice Cream Bars 4 Pack',
    'Drumstick Classic Vanilla Cones 4 Pack',
    'Icy Pole Lemonade Ice Blocks 8 Pack',
    'Angus Beef Pies 4 Pack',
    'Angus Beef & Pepper Pies 4 Pack',
    'Drumstick Choc Mint Cones 4 Pack',
    'Frozen Dessert Double Sunlover 3 Pack',
    'Maxibon Hedgehog Ice Cream 4 Pack',
    'Steam Fresh Carrot Cauliflower & Broccoli Mix 5 pack',
    'Frozen King Island Frozen Beef Sausage Rolls 12 pack',
    'Frozen Deli Barramundi Crisp Light Batter 5 pack',
    'Slow Cooked Steak & Mushroom Pies 2 Pack',
    'Frozen Cookie Sandwich With Oreo Biscuit Pieces 4 pack',
    'Frozen South Atlantic Flathead Crisp Light Batter Fillets 5 Pack',
    'Ice Cream Cafe Grande 4 Pack',
    'Drumstick Minis Super Choc & Caramel Nut Cones Variety 6 Pack',
    'Dairy Milk Vanilla Ice Cream 4 Pack',
    'Paddle Pop Cyclone Ice Cream Sticks 8 pack',
    'Frozen Deli New Zealand Whiting Crisp Light Batter Fillets 5 Pack',
    'Lifesavers Ice Blocks 8 Pack',
    'Steam Fresh Carrot Broccoli & Corn Mix 5 pack',
    'Frozen Dessert Euphoria Raspberry Lemon 3 Pack',
    'Maxibon Ice Cream Original Vanilla 4 pack',
    'Maxibon Ice Cream Monster Cookie 4 pack',
    'Frozen Mini Beef Dim Sims 15 Pack',
    'Rocher Frozen Classic Dessert Hazelnut & Choc 4 pack',
    'Ice Cream Strawberry Chocolate & Caramel 4 Pack',
    'Ham & Pineapple Pizza Slices 2 Pack',
    'Bulk Ice Cream 18 Pack',
    …]
```

[38]:
```python
import seaborn as sns
import matplotlib.pyplot as plt

# Sample check: make sure required columns exist
assert all(col in df.columns for col in ['item_price', 'similarity_score',
  ↪'category'])

plt.figure(figsize=(10, 6))
```

```python
sns.scatterplot(data=df, x='item_price', y='similarity_score', hue='category',
 ↪palette='tab10')

plt.title('Similarity Score vs. Item Price by Category')
plt.xlabel('Item Price ($)')
plt.ylabel('Similarity Score')
plt.legend(title='Category', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.grid(True)
plt.show()
```

<ipython-input-38-ca30711aa4cf>:14: UserWarning: Tight layout not applied. The
bottom and top margins cannot be made large enough to accommodate all axes
decorations.
  plt.tight_layout()

Similarity Score vs. Item Price by Category

```
[24]: data.columns
```

```
[24]: Index(['product_code', 'category', 'item_name', 'brand', 'item_price',
             'item_size', 'unit_price', 'unit_weight', 'similarity_score', 'subcat',
             'unit_weight_new', 'unit_weight_type'],
            dtype='object')
```

```
[25]: data.head()
```

```
[25]:    product_code category                                    item_name  \
      0       8371390   Easter   Coles Hot Cross Buns Traditional Fruit | 6 Pack
      1       7473849   Easter          Coles Hot Cross Buns Choc Chip | 6 Pack
      2       5726070   Easter   Coles Hot Cross Buns Traditional Fruit Mini | …
      3       4885191   Easter   Cadbury Dairy Milk Easter Chocolate Eggs Bag |…
      4       3378370   Easter      Coles Hot Cross Buns Apple & Cinnamon | 6 Pack

           brand  item_price item_size  unit_price unit_weight  similarity_score  \
      0     Coles         4.4    6 Pack        0.73         1ea          0.198373
      1     Coles         4.4    6 Pack        0.73         1ea          0.198373
      2     Coles         4.4    9 Pack        0.49         1ea          0.198373
      3   Cadbury         6.7      114g        5.88        100g          0.214731
      4     Coles         4.4    6 Pack        0.73         1ea          0.198373

           subcat  unit_weight_new unit_weight_type
      0     Fruit              1.0               ea
      1   Unknown              1.0               ea
      2     Fruit              1.0               ea
      3      Milk            100.0                g
      4   Unknown              1.0               ea
```

1.7 Extract weight unit

```python
[49]: def extract_weight_unit_unified(value):
          """
          Handles both single-value sizes (e.g. '16g', '25mL')
          and pack-style sizes (e.g. '24x375mL').
          Returns (total_weight, unit).
          """
          if pd.isna(value):
              return (np.nan, np.nan)

          text = str(value).strip()

          # Case 1: Pack-style e.g. '24x375mL'
          match_pack = re.search(r'(\d+)\s*[xX]\s*([\d\.]+)\s*([a-zA-Z]+)', text)
```

```
    if match_pack:
        count = int(match_pack.group(1))
        per_unit_qty = float(match_pack.group(2))
        unit = match_pack.group(3).lower()
        return count * per_unit_qty, unit

    # Case 2: Single-value e.g. '16g', '25mL'
    match_single = re.match(r'([\d\.]+)\s*([a-zA-Z]+)', text)
    if match_single:
        qty = float(match_single.group(1))
        unit = match_single.group(2).lower()
        return qty, unit

    return np.nan, np.nan
```

[50]:
```python
data[['unit_weight_new', 'unit_weight_type']] = data['unit_weight'].
 ↪apply(lambda x: pd.Series(extract_weight_unit_unified(x)))
data[['item_weight_new', 'item_weight_type']] = data['item_size'].apply(lambda␣
 ↪x: pd.Series(extract_weight_unit_unified(x)))
```

```
C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\frame.py:3191:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self[k1] = value[k2]
```

[51]:
```python
data[['unit_weight_new', 'unit_weight_type']]
```

[51]:

	unit_weight_new	unit_weight_type
0	1.0	ea
1	1.0	ea
2	1.0	ea
3	100.0	g
4	1.0	ea
...	...	...
77828	100.0	g
77829	100.0	g
77830	100.0	g
77831	100.0	g
77832	100.0	g

[77833 rows x 2 columns]

[52]:
```python
data[['item_weight_new', 'item_weight_type']]
```

```
[52]:        item_weight_new item_weight_type
    0                 6.0              pack
    1                 6.0              pack
    2                 9.0              pack
    3               114.0                 g
    4                 6.0              pack
    ...               ...               ...
    77828             NaN               NaN
    77829             NaN               NaN
    77830             NaN               NaN
    77831             NaN               NaN
    77832             NaN               NaN

    [77833 rows x 2 columns]
```

```
[54]: data.loc[data['item_size'] == '24x375mL', ['item_size', 'item_weight_new',
      ↪'item_weight_type']]
```

```
[54]:        item_size  item_weight_new item_weight_type
    1498    24x375mL           9000.0               ml
    1499    24x375mL           9000.0               ml
    1655    24x375mL           9000.0               ml
    1656    24x375mL           9000.0               ml
    13329   24x375mL           9000.0               ml
    13347   24x375mL           9000.0               ml
    13418   24x375mL           9000.0               ml
    13668   24x375mL           9000.0               ml
    35265   24x375mL           9000.0               ml
    35338   24x375mL           9000.0               ml
    35361   24x375mL           9000.0               ml
    35602   24x375mL           9000.0               ml
    48078   24x375mL           9000.0               ml
    48193   24x375mL           9000.0               ml
    48382   24x375mL           9000.0               ml
    48863   24x375mL           9000.0               ml
    66247   24x375mL           9000.0               ml
    66249   24x375mL           9000.0               ml
    66535   24x375mL           9000.0               ml
    66981   24x375mL           9000.0               ml
```

```
[55]: data['item_weight_type'].unique()
```

```
[55]: array(['pack', 'g', 'ml', 'kg', 'l', 'gram', 'each', nan, 'cm', 'pk',
             'multipack', 'mm'], dtype=object)
```

```
[56]: data[data['item_weight_type']=='mm']
```

```
[56]:        product_code              category  \
      45702       3653113  MEDICINAL PRODUCTS
      45758       8415269  MEDICINAL PRODUCTS
      74494       3653113  MEDICINAL PRODUCTS
      75560       8415269  MEDICINAL PRODUCTS

                                       item_name                       brand  \
      45702                Blister Plastics 44x69mm  Blister Plastics 44x69mm
      45758  Extra Care Hydrogel Medistrips 75x75mm                   Unknown
      74494                Blister Plastics 44x69mm  Blister Plastics 44x69mm
      75560  Extra Care Hydrogel Medistrips 75x75mm                   Unknown

             item_price item_size  unit_price unit_weight  similarity_score  \
      45702         0.0   44x69mm        1.80         1ea          0.217882
      45758         0.0   75x75mm        2.67         1ea          0.217882
      74494         0.0   44x69mm        1.80         1ea          0.217882
      75560         0.0   75x75mm        2.67         1ea          0.217882

              subcat  unit_weight_new unit_weight_type  item_weight_new  \
      45702  Unknown              1.0               ea           3036.0
      45758  Unknown              1.0               ea           5625.0
      74494  Unknown              1.0               ea           3036.0
      75560  Unknown              1.0               ea           5625.0

            item_weight_type
      45702               mm
      45758               mm
      74494               mm
      75560               mm
```

```
[57]: data['unit_weight_type'].unique()
```

```
[57]: array(['ea', 'g', 'ml', 'kg', 'l', nan, 'gwas', 'kgwas', 'eawas', 'm',
             'kgm'], dtype=object)
```

```
[60]: data.loc[data['unit_weight_type'] == 'm', ['item_name', 'unit_weight',
       ↪'unit_weight_new', 'unit_weight_type']]
```

```
[60]:                              item_name unit_weight  \
      2659                          Wrap Foil          1M
      2831                          Cling Wrap          1M
      2949        Alfoil Traditional Strength 30cm          1M
      3241               Cling Wrap 60m x 33cm          1M
      4062              Aluminium Foil Wrap          1M
      ...                                 ...         ...
      76855          Waterproof Kinesiology Tape          1M
      76882                Conforming Bandage          1M
```

```
76959  First Aid Non-Irritating Paper Tape 2.5Cm x 9.1m          1M
76996                     Cohesive Compression Wrap 75MM          1M
77018                        Waterproof Kinesiology Tape          1M


       unit_weight_new unit_weight_type
2659               1.0                m
2831               1.0                m
2949               1.0                m
3241               1.0                m
4062               1.0                m
...                ...              ...
76855              1.0                m
76882              1.0                m
76959              1.0                m
76996              1.0                m
77018              1.0                m


[241 rows x 4 columns]
```

'm' is meters

[61]:
```
data.loc[data['unit_weight_type'] == 'kgm', ['item_name', 'unit_weight',␣
 ↪'unit_weight_new', 'unit_weight_type']]
```

[61]:
```
            item_name unit_weight  unit_weight_new unit_weight_type
41840       Super Tuff        1kgM              1.0              kgm
42068  Party Feet Heel Shield   1kgM              1.0              kgm
```

[67]:
```
data.loc[data['unit_weight_type'] == 'kgm', ['item_name', 'item_size',␣
 ↪'item_weight_new', 'item_weight_type']]
```

[67]:
```
            item_name item_size  item_weight_new item_weight_type
41840       Super Tuff       NaN              NaN              NaN
42068  Party Feet Heel Shield   NaN              NaN              NaN
```

[68]:
```
data.loc[data['unit_weight_type'] == 'm', ['item_name', 'item_size',␣
 ↪'item_weight_new', 'item_weight_type']]
```

[68]:
```
                                             item_name item_size  \
2659                                         Wrap Foil       NaN
2831                                        Cling Wrap       NaN
2949                     Alfoil Traditional Strength 30cm   NaN
3241                             Cling Wrap 60m x 33cm       NaN
4062                            Aluminium Foil Wrap       NaN
...                                                ...       ...
76855                         Waterproof Kinesiology Tape   NaN
76882                             Conforming Bandage       NaN
76959  First Aid Non-Irritating Paper Tape 2.5Cm x 9.1m   NaN
```

```
76996                    Cohesive Compression Wrap 75MM        NaN
77018                     Waterproof Kinesiology Tape          NaN

        item_weight_new item_weight_type
2659                NaN              NaN
2831                NaN              NaN
2949                NaN              NaN
3241                NaN              NaN
4062                NaN              NaN
...                 ...              ...
76855               NaN              NaN
76882               NaN              NaN
76959               NaN              NaN
76996               NaN              NaN
77018               NaN              NaN

[241 rows x 4 columns]
```

1.8 Convert to standard units

```python
[64]: # get standard unit names
      units_to_map = {'ea': 'each', 'gram': 'g', 'pk': 'pack', 'gwas': 'g', 'kgwas':
       ↪'kg', 'eawas': 'each'}

      data['unit_weight_type'] = data['unit_weight_type'].replace(units_to_map)
      data['item_weight_type'] = data['item_weight_type'].replace(units_to_map)
```

```
<ipython-input-64-281179e86d05>:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data['unit_weight_type'] = data['unit_weight_type'].replace(units_to_map)
<ipython-input-64-281179e86d05>:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data['item_weight_type'] = data['item_weight_type'].replace(units_to_map)
```

```python
[65]: data['unit_weight_type'].unique()
```

```
[65]: array(['each', 'g', 'ml', 'kg', 'l', nan, 'm', 'kgm'], dtype=object)
```

```python
[66]: data['item_weight_type'].unique()
```

```
[66]: array(['pack', 'g', 'ml', 'kg', 'l', 'each', nan, 'cm', 'multipack', 'mm'],
            dtype=object)
```

Standardize weights: - kg -> g - l -> ml - m -> cm - mm -> cm

1.9 Standardize weights

```python
[69]: def standardize_units(qty, unit):
          """
          Convert qty+unit into a standardized value.
          - g for weight
          - mL for volume
          - m for length
          """
          if pd.isna(qty) or pd.isna(unit):
              return np.nan, np.nan

          unit = unit.lower()

          # Weight
          if unit in ["g", "gram", "grams"]:
              return qty, "g"
          elif unit in ["kg", "kilogram", "kilograms"]:
              return qty * 1000, "g"

          # Volume
          elif unit in ["ml", "millilitre", "milliliter"]:
              return qty, "ml"
          elif unit in ["l", "litre", "liter"]:
              return qty * 1000, "ml"

          # Length
          elif unit in ["m"]:
              return qty, "m"
          elif unit in ["cm"]:
              return qty / 100, "m"
          elif unit in ["mm"]:
              return qty / 1000, "m"

          else:
              # Unknown unit → leave as is
              return qty, unit
```

```python
[70]: data[["std_item_weight", "std_item_unit"]] = data.apply(
          lambda row: standardize_units(row["item_weight_new"],␣
      ↪row["item_weight_type"]),
          axis=1, result_type="expand"
```

114

```
)
```

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\frame.py:3191:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self[k1] = value[k2]

```
[71]: data[["item_name","item_size", "item_weight_new",␣
      ↪"item_weight_type","std_item_weight", "std_item_unit"]]
```

```
[71]:                                           item_name item_size  \
      0          Coles Hot Cross Buns Traditional Fruit | 6 Pack    6 Pack
      1                    Coles Hot Cross Buns Choc Chip | 6 Pack    6 Pack
      2          Coles Hot Cross Buns Traditional Fruit Mini | …    9 Pack
      3          Cadbury Dairy Milk Easter Chocolate Eggs Bag |…      114g
      4             Coles Hot Cross Buns Apple & Cinnamon | 6 Pack    6 Pack
      …                                                  …        …
      77828              Veggie Tubes Sweetcorn & Carrot   12m+       NaN
      77829                       100% Pure Cotton Ripple             NaN
      77830      Pumpkin Potato & Beef Baby Food Pouch 6+ Months       NaN
      77831  Lamb Vegetables & Polenta Baby Food Pouch 6+ M…       NaN
      77832                   Beef With Vegetables 6+ Months         NaN

             item_weight_new item_weight_type  std_item_weight std_item_unit
      0                  6.0             pack              6.0          pack
      1                  6.0             pack              6.0          pack
      2                  9.0             pack              9.0          pack
      3                114.0                g            114.0             g
      4                  6.0             pack              6.0          pack
      …                    …                …                …            …
      77828              NaN              NaN              NaN           NaN
      77829              NaN              NaN              NaN           NaN
      77830              NaN              NaN              NaN           NaN
      77831              NaN              NaN              NaN           NaN
      77832              NaN              NaN              NaN           NaN

      [77833 rows x 6 columns]
```

```
[72]: data.isna().sum()
```

```
[72]: product_code           0
      category               0
      item_name              0
      brand                  0
```

```
item_price           0
item_size        66683
unit_price        2237
unit_weight       2238
similarity_score     0
subcat               0
unit_weight_new   2244
unit_weight_type  2244
item_weight_new  66683
item_weight_type 66683
std_item_weight  66683
std_item_unit    66683
dtype: int64
```

1.10 Check for items that does not have a weight

```
[73]: data[data['item_size'].isna()]
```

[73]:

	product_code	category	\
1460	1113465	EASTER	
1461	4886751	EASTER	
1462	4898604	EASTER	
1463	5766234	EASTER	
1464	4885191	EASTER	
…	…	…	
77828	4163437	INFANT FOOD	
77829	3754198	INFANT PERSONAL	
77830	6295520	INFANT FOOD	
77831	6295585	INFANT FOOD	
77832	3962038	INFANT FOOD	

	item_name	\
1460	Chocolate Easter Bunny	
1461	Selections Easter Chocolate Eggs Bag	
1462	Easter Gold Bunny Milk Chocolate	
1463	Chocolate Easter Hunt Assortment Pack	
1464	Dairy Milk Easter Chocolate Eggs Bag	
…	…	
77828	Veggie Tubes Sweetcorn & Carrot 12m+	
77829	100% Pure Cotton Ripple	
77830	Pumpkin Potato & Beef Baby Food Pouch 6+ Months	
77831	Lamb Vegetables & Polenta Baby Food Pouch 6+ M…	
77832	Beef With Vegetables 6+ Months	

	brand	item_price	item_size	unit_price	\
1460	Chocolate Easter Bunny	8.0	NaN	4.00	
1461	Selections Easter Chocolate	11.5	NaN	3.48	

116
```

```
1462                Unknown      18.0      NaN        6.25
1463                Unknown      17.5      NaN        2.40
1464        Dairy Milk Easter     6.7      NaN        3.51
...                     ...       ...       ...         ...
77828               Unknown       0.0      NaN       16.67
77829               Unknown       0.0      NaN        1.95
77830               Unknown       0.0      NaN        1.33
77831               Polenta       0.0      NaN        1.33
77832               Unknown       0.0      NaN        0.83

      unit_weight  similarity_score    subcat  unit_weight_new  \
1460         100g          0.171870   Unknown            100.0
1461         100g          0.170491   Unknown            100.0
1462         100g          0.103036      Milk            100.0
1463         100g          0.106898   Unknown            100.0
1464         100g          0.160387      Milk            100.0
...           ...               ...       ...              ...
77828        100g          0.439613   Unknown            100.0
77829        100g          0.267265   Unknown            100.0
77830        100g          0.439613      Beef            100.0
77831        100g          0.439613      Lamb            100.0
77832        100g          0.439613      Beef            100.0

      unit_weight_type  item_weight_new  item_weight_type  std_item_weight  \
1460                 g              NaN               NaN              NaN
1461                 g              NaN               NaN              NaN
1462                 g              NaN               NaN              NaN
1463                 g              NaN               NaN              NaN
1464                 g              NaN               NaN              NaN
...                ...              ...               ...              ...
77828                g              NaN               NaN              NaN
77829                g              NaN               NaN              NaN
77830                g              NaN               NaN              NaN
77831                g              NaN               NaN              NaN
77832                g              NaN               NaN              NaN

      std_item_unit
1460            NaN
1461            NaN
1462            NaN
1463            NaN
1464            NaN
...             ...
77828           NaN
77829           NaN
77830           NaN
77831           NaN
```

```
77832                 NaN

[66683 rows x 16 columns]
```

```
[74]: data[data['item_size'].isna() & data['unit_weight'].isna()]  # both item size
      ↪and unit weight are missing
```

```
[74]:         product_code        category                          item_name  \
      1753         3655141          EASTER        Easter Belgian Milk Choc Pops
      2164         3121431     KITCHENWARE                         Roasting Pan
      2369         3121453     KITCHENWARE               Roasting Pan with Rack
      2556         7208484      STATIONERY                 A4 Copy Paper 500 Sheets
      3045         5880626  TECH ACCESSORIES  1.2M USB-A To Lightning Cable White
      ...              ...             ...                                ...
      72473        3782863      STATIONERY                      Blu Tack Colour
      72478        9053888  SEASONAL EVENTS                               Saver
      72479        8471678   ANCIL SERVICES                          $1 donation
      75354        9198202       BEST BUYS                        Massage Pillow
      76089        1048556       HAIR CARE        Family Anti Dandruf Shampoo Foam

                     brand item_price item_size  unit_price unit_weight  \
      1753         Unknown        0.0       NaN         NaN         NaN
      2164    Roasting Pan        0.0       NaN         NaN         NaN
      2369    Roasting Pan        0.0       NaN         NaN         NaN
      2556   A4 Copy Paper        0.0       NaN         NaN         NaN
      3045   1.2M USB-A To        0.0       NaN         NaN         NaN
      ...              ...        ...       ...         ...         ...
      72473       Blu Tack        4.0       NaN         NaN         NaN
      72478          Saver        0.0       NaN         NaN         NaN
      72479    $1 donation        0.0       NaN         NaN         NaN
      75354        Unknown        0.0       NaN         NaN         NaN
      76089        Unknown        0.0       NaN         NaN         NaN

             similarity_score   subcat  unit_weight_new unit_weight_type  \
      1753           0.089724     Milk              NaN              NaN
      2164           0.267393  Unknown              NaN              NaN
      2369           0.267393  Unknown              NaN              NaN
      2556           0.141170  Unknown              NaN              NaN
      3045           0.156810  Unknown              NaN              NaN
      ...                 ...      ...              ...              ...
      72473          0.157536  Unknown              NaN              NaN
      72478          0.182327  Unknown              NaN              NaN
      72479          1.000000  Unknown              NaN              NaN
      75354          1.000000  Unknown              NaN              NaN
      76089          0.111614      Ham              NaN              NaN

             item_weight_new item_weight_type  std_item_weight std_item_unit
```

1753	NaN	NaN	NaN	NaN
2164	NaN	NaN	NaN	NaN
2369	NaN	NaN	NaN	NaN
2556	NaN	NaN	NaN	NaN
3045	NaN	NaN	NaN	NaN
…	…	…	…	…
72473	NaN	NaN	NaN	NaN
72478	NaN	NaN	NaN	NaN
72479	NaN	NaN	NaN	NaN
75354	NaN	NaN	NaN	NaN
76089	NaN	NaN	NaN	NaN

[2142 rows x 16 columns]

Below are the categories of the instances where both the item size and unit weight is missing.

```
[75]: data.loc[data['item_size'].isna() & data['unit_weight'].isna(), 'category'].
      ↪unique()
```

```
[75]: array(['EASTER', 'KITCHENWARE', 'STATIONERY', 'TECH ACCESSORIES',
             'HARDWARE', 'SHOE CARE', 'CHEESE DAIRY', 'FAMILY SOCKS',
             'HERBS & CHILLIES', 'SHOPPING BAGS', 'MAGAZINES', 'LBI MERCH',
             'SEASONAL EVENTS', 'ELECTRICAL', 'AUTO', 'HOUSEHOLD APPLIANCES',
             'TELCO', 'BEST BUYS', 'TOYS & HOBBIES', 'GARDEN NON GREENS',
             'HEATING & COOLING', 'FAMILY UNDERWEAR', 'FITNESS/HEALTH',
             'BRUSHWARE', 'PICNIC POOL BBQ ACC', 'HOSIERY', 'LEISURE',
             'SEASONAL APPAREL', 'CARDS/WRAPS', 'PARTY GOODS', 'ANCIL SERVICES',
             'BABYWEAR', 'HAIR CARE'], dtype=object)
```

```
[80]: data.loc[data['item_size'].isna() & data['unit_weight'].isna(), 'item_name'].
      ↪unique()
```

```
[80]: array(['Easter Belgian Milk Choc Pops', 'Roasting Pan',
             'Roasting Pan with Rack', 'A4 Copy Paper 500 Sheets',
             '1.2M USB-A To Lightning Cable White', 'Wall Charger Dual USB-C',
             'Escape True Wireless Sport Ear Buds', 'Blu Tack',
             'Wired Earphones With Mic Lightning',
             'Boostup 30w Usb C Pd Wall Charger', 'Combination Word 40mm',
             '1.2M USB-A To USB-C Cable Black',
             '42w (Pd30w Pps Usb-C & 12w Usb-A) Universal Wall Charger',
             'Boostcharge USB-A To USB-C Cable- 1M Black', 'Blu Tack Colour',
             'Car Mobile Phone Holder', 'Diamond Padlock Long Shackle 40mm',
             'Wireless Mouse', '2M USB-C To USB C Cable - Black', 'School Glue',
             'Sticky tape 18mmx25m', 'Lint Roller',
             '1.2M USB-C To USB-C Cable Black Braided',
             'Lightning To 3.5Mm Audio Adapter',
             '2M USB-A To USB-C Cable Black', 'Fluo XL Assorted Highlighters',
             '4 Colour Pens', 'Boss Highlighter Assorted Fluoro',
```

'Car Charger Dual USB-C',
'Trios Twiggy Cheese & BBQ Rice Crackers',
'Indulge Wireless Bluetooth Headphones',
'Wired Earphones With Mic 3.5mm Audio',
'Liquid Paper Dryline Grip Correction Tape',
'Magnetic Wireless Charging Pad',
'2M USB C To Lightning - White Braided',
'Boostcharge Usb A Wall Charger (12w)',
'1.2M USB C To Lightning - White Braided', 'A4 Notebook 120 Page',
'Metallic Markers', 'Paper Clips In Pouch Assorted Colours',
'USB 2.0 SM02 16GB', 'A4 Display Folder', 'Wall Charger Usb-A & C',
'1.2M USB-C To USB-C Cable Black', 'Super Sticky Cube',
'Boostcharge Usb-C To Usb-C Cable- 1m - White',
'2M USB-C To C - Black Braided', '570 Notebook A5 200 Page',
'Bluetooth Portable Speaker', 'SM02 USB 2.0 Flash Drive 64GB',
'A4 Spiral Notebook 200 Page',
'1.2M USB-C To Lightning Cable White',
'USB-C To 3.55MM Audio Adapter', 'Kids Logo QTR Crew Size 9-12',
'Mens Logo Low Cut 6-10', 'To Do List 52 Sheets',
'Markers Fine Tip Colour Burst',
'Kids Evolution Colouring Pencils',
'Sticky Notes Assorted Colours', 'S-Gel Pens Assorted',
'Laundry Marker', 'Atlantis Pens Black',
'Womens Socks Quarter Logo Size 8+', 'Letra Tag Labeller Blue',
'2M USB-C To Lightning - White', 'Trek Wireless Sport Earphones',
'Radiant Wired Headphones', 'Bamboo Handle Lint Roller',
'Correction Pen White', 'Kids Logo Low Cut 13-3', 'Fineliner Pens',
'Cristal Fashion Pens',
'Boostcharge Lightning to Usb-A Cable 1m Black', 'PVC Free Eraser',
'A5 Spiral Notebook 200 pages',
'Boostcharge Usb-C To Usb-C Cable- 1m - Black',
'Intensity Colouring Felt Pens', 'Woodcase Hb Pencils',
'Connector Pens', 'Plastic Clear Tape', 'Oval Lace White 45',
'Traditional 2b Graphite Pencils',
'Kids Volume Limiting Wireless Headphone',
'B5 Perforated Notebook 160Pg', 'Wired Earphones With Mic USB-C',
'Atlantis Pens Blue', 'Colourhide Notebook A4 120pages',
'Correction Tape', 'Kids Plastidecor Crayons',
'Silence Pro Noise Isolating Headphones',
'2M USB-A To Lighting Cable White', 'Clear School Glue',
'Cristal Original Pens Assorted', 'Intensity Fineliners Assorted',
'Whiteboard Markers Fine Tip',
'Kids Logo Light Quarter Crew Size 13 - 3',
'Kids Logo Low Cut Socks Size 9-12',
'Kids Evolution Illusion Colouring Pencils',
'Wireless Charging Stand', '4 Colours Pens Limited Edition',
'Colouring Felt Pen Wallet', 'SDMI Card 32GB UHS 1 Class 10 BW',

'Boostcharge Lightning to Usb-A Cable 1m White',
'Car Charger Dual USB-A', 'Plastic 2 Hole Sharpener',
'Tape Plastic', 'Colourhide Notebook A5 Assorted 200 page',
'Flat Lace Black 45 inch', 'Inkjoy 100 Assorted Colours',
'Double Sided Palm Cards', '200 Fineline Black Pen',
'Gaming Headset', 'USB A Car Charger 24W', 'Woodcase 2B Pencil',
'Traditional Hb Graphite Pencils',
'Boostcharge Usb A Usb C Cable 1m White', 'USB 3.0 BE03 32GB',
'Unicorn Stationery Set 5 Piece', 'Classic Colouring Pencils',
'Colourhide Notebook 5 Subject A4', 'Kids Turn & Colour Crayons',
'Flat Lace White Kids 27 inch', 'Gelocity Original Pens Black',
'Wreless Mouse Keyboard Combo', 'Potted Mixed Herbs',
'Country Fresh Potted Basil', 'Chiller Bag', 'Stick Frypan',
'Celebration Cake Pan', 'Large Sieve', 'Non Stick Frypan',
'Air Fryer Deep Baking Pan 17.5Cm', 'Silicone Jumbo Muffin Cups',
'3 In 1 Avocado Tool', 'Swivel Peeler', 'Chopping Board Set',
'Cooking Twine', 'Measuring Cup Set',
'True Wireless:Earphone With Enc', 'Multigrip Mandoline',
'Scissors', 'Compact Knife Organiser', 'Pizza Cutter',
'Silicone Icy Pole Mould', 'Mixing Bowl', 'Spaghetti Spoon',
'Madeleine Pan 12 Cup', 'Meat Thermometer', 'Chiller Bag A',
'Large Tote Carrot Bag', 'Sink Tidy', 'Silicone Muffin Cups',
'Silicone Rolling Pin', 'Reuseable Bag Clips', 'Funnel',
'Stick Sautepan', 'Meat Tenderiser', 'Silicone Trivet Set',
'Folding Chopping Board', '10000 Mah Powerbank',
'12 Cup Mini Bundt Pan', 'Aluminium Stock Pot 24cm',
'Easy Clean Garlic Press', 'Bundt Cake Pan',
'Compact Hand Held Spiralizer', 'Pull Chopper', 'Store & Pour',
"Woman's Day", 'Piping Set',
'Over Ear Wireless Headphone With Anc', 'Sink Strainer',
'Triple Screen Sifter', '8 Cup Mini Loaf Pan', 'Juicer',
'Compact Juicer', 'Compact Cutlery Organiser',
'Expanding Dish Rack', 'Silicone Baking Mat',
'Silicone Air Fryer Liners Round', 'Milk Frother', 'Garlic Press',
'Dumpling Press', 'Apple Corer', 'Air Fryer Baking Pan 18.5 Cm',
'Zester', 'New Idea', 'Glass Cookware Set', 'AFL Ball Size 3',
'Power Wall Charger Usbc Pd20W', 'Lubricating Spray',
'Premium Shopping Bag', 'Compact Food Processor', 'V Smart 2',
'Vines Curtain Lights', 'Sieves', 'Vehicle Power Grab',
'Samsung Galaxy A15', 'Frypan 28cm', 'Tripple Pack',
'Large Glass Container', '2 Slice Snackmaker',
'Magican Can Opener', 'Small Glass Container', 'TV Week',
'Frypan 26cm', 'Mini Glass Containers', 'Samsung Galaxy A14',
'Springform Cake Pan',
'iPad 5 Mini 64GB WiFi + Cellular Premium Refurbished', 'X Total',
'Medium Glass Container', 'Oppo A60', '7 Cup Rice Cooker',
'Car Gift Pack', '10w - 30w', 'Led Strip 5m',

'1 Hour Action Weed Kill Rtu', 'Bamboo Cocktail Sticks',
'No More Mess Twin Pack', 'Zero Weed Kill Value Pack Spray Twin',
'Electric Salt & Pepper Grinder', 'Who', 'Lamington Tray',
'Magical Liquid', 'Tag Set', 'Carbon Steel Wok 30cm',
'Basic Car Assorted', '12 Cup Muffin Pan', 'Wooden Wok Spatula',
'Saucepan 18cm', 'Motorola G04 4GX', 'Squeezy Bottles',
"Fix 'N' Go Supa Glue", 'Long Hot Water Bottle With Soft Cover',
'Power Board 4 Socket', 'Collapsible Steamer',
'Australia New Zealand Travel Adaptor', 'Regular Weedkiller',
'Lead Extension 3m', 'AWW Puzzle Book', 'Lighter',
'4 Socket Surge Protected 1.2m Power Board', 'Chopping Board',
'Permanent Markers Fine Tip', '4 Outlet Powerboard',
'Organic+ Blood & Bone Fertiliser', 'Take 5', 'Bamboo Board',
'Clear Packaging Tape', 'Garment Steamer', 'Hello',
'Better Homes & Gardens', "That's Life", 'Blind Packasst',
'OPPO A58 Black', 'Large Oven Tray', 'Oppo A16S 64Gb',
'Samsung Galaxy A05s', 'Nokia G11 Plus 64Gb 4GX', "Women's Weekly",
'Stackable Measuring Cups', 'X Tap 3', 'V Pro 2',
'Digital Kitchen Scales', 'Silicone Spatula',
'6 Outlet Powerboard', 'Silicone Tongs',
'Induction Non Stick Grill Pan 26Cm', "Feed 'n' Weed",
'Stapler With Staples', 'Mini Figures',
'Mens Guyfront Trunk Size L', 'Oil Motor Sae15W-40',
'Fast Action Weed Killer', 'Measuring Jug', 'Oppo A16S',
'Zero Weed Kill Ut Ultra Tough Weed Spray', 'Peeler',
'Manual Vacuum Pump', 'Detachable Kitchen Scissors', 'Steam Iron',
'Lovatts A5 Puzzle Puzzle Pocket And Fill In',
'H11032 Ladies Cushion Foot Socks Black Size 5-8',
'Premium Potting Mix', '150 Lumen Aluminium Flashlight 2 AA',
'Square Cake Tin 24cm', 'Lint Roller Refill', 'Tongs 30cm',
'Aluminium Saute Pan 28cm', 'Can Opener', 'Fxt Yoga Socks',
'Finals Official AFL Game', 'Mixed Rubber Bands In Pouch',
'Windscreen Wash', 'Power Board 6 Socket 2 USB', 'Loaf Pan',
'Low Odour', '24 Cup Mini Muffin Pan', 'Slotted Turner',
'Dustpan Set', 'BHG Puzzles', 'X Sight 3',
'Mens X-Temp Quarter Crew 6+', 'Permanent Markers Bullet Tip',
'Aluminium Frypan 24cm', 'Gel Activ Insole Casual Large',
'Power Grip', 'Masking Tape 24mmX18m', 'Table Tennis Anywhere',
'Mens Logo Light No Show Size 6+ & 11+', '32GB USB2.0 Twin Pack',
'Portable Modem +50GB', 'X Start 4', 'Supa Glue Brush',
'Terra White Sparkling Water Maker', 'Shoe Repair',
'Tough Lawn Seed', 'Wheel Cleaner', '600 Series',
'Extreme Mounting Tape', 'Healthy Food Guide', 'Cookie Scoop',
'Round Cake Tin 24cm', 'Weed Kill', 'Envelopes C6', 'Nokia C30',
'Deep Fill Jaffle Maker', 'Sewing Kit', "Men's Action Brief M",
'Concentrate Weed Kill', 'Aluminium Deep Frypan',
'Hot Wheels Gift Pack', 'Mens Briefs Large',

'Lawn Seed Lawn Patch Repair Seed', "Woman's Day Puzzler",
'Mens 1/4 Crew Socks Size 11-14',
'BBQ Digital Flip Tip Thermometer', 'Everyday Toaster',
'Double Sided Tape', 'Nordic Frypan 26Cm', 'Flip Card Game',
'Aluminium Saucepan 20cm', 'Mens Guyfront Trunk Size Medium',
'Motorola Moto E13 4GX', 'Nylon Tongs', 'Non Stick Frypan 20cm',
'Easy Clean Toilet Set', 'Multi Grip',
'Womens Logo Low Cut Size 3-8', 'Stick Mixer',
'Fxt Suspension Trainer', 'Cast Aluminium All Day Pan 28Cm',
'Assorted 0.7mm', 'Brownie Tin 20cm', 'Travel Adaptor',
'Osmocote Plus Organics Plant Food', 'USB To C 1M Black',
'Mens Work Socks 11-13 Black', 'Inkjoy Black Ball Point',
'Silicone Ice Cube Tray', 'Handheld Grater',
'Staysharp Kitchen Scissors', 'Colander',
'Silicone Air Fryer Trays Round', 'Spoon',
'Hair Straightener And Waver Set E34', '24 Hour Timer 1 Socket',
'Action Brief XL', 'Stainless Steel Toilet Set',
'Curvy Sheer Kh 15D Natural One Size', 'Storm',
'Pyrethrum Garden Insect Spray',
'Hl1032 Ladies Cushion Foot Socks Black Size 9-11',
'Mens Cushion Foot Crew Size 7-10', 'Paque Trouser Sock 70D',
'Womens Logo Light No Show Size 3 - 8', 'Mini Tongs',
'Metal Skewers', 'DL Envelopes', 'Fxt Glute Band', 'Ballpen Blue',
'Complete Garden Health Treatment', 'Alive', 'Essential Pro 3',
'Antimicrobial 7" Scissors', 'Stainless Steel 4 Slice Toaster',
'Bits Set 21 piece', 'PVC Electrical Tape', 'Silicone Whisk',
'Life 24 Can Collapsible Cooler',
'Womens Logo Light No Show Size 8 - 11',
'Green Up Organic+ Lawn Food', 'Dart Board',
'Foam Latex Comfort Insole 1 Pair', 'Pre-Paid 4GX Wi-Fi Plus 2',
'Sport Crew Sock 3+ & 8+', 'Plastic Push Pins In Pouch',
'Tape Measure', 'Aluminium Saucepan 16cm',
'Silky Control Ph Natural X-Tall',
'Fxt Adjustable Strength Cruncher', 'Extension Lead 10m',
'Frixion Ball Pen Black/Blue/Red', 'Balls Assorted', 'Staples',
'Double Adaptor Flat Left Hand', 'Race Car', 'MF910Y Wifi 4GX',
'X Lite 4', 'Umbrella Purple', '32800 Pantyhose Black Tall',
'SHEER VALUE 15 DENIER TALL NATURAL', 'Rookie Official AFL Game',
'Bug Green Organic Slow Release Fertiliser',
'Sheer Nylon Knee Hi Pantyhose Reinforced Toe',
'Soft Touch Scissors 9 inch', 'Ladies Marshmallow Anklet 5-10',
'Easy Slide Silicone', 'Salt & Pepper Set', 'Mini Frypan',
'Basketball', 'Silver Safety Pins Assorted Sizes',
'Sticky Tape Dispenser', 'Multifunctional Mandoline Slicer',
'Go Sockless', 'Footcare Work Insoles',
'MENS QUARTER CREW LOGO SIZE 11+', 'Mini Silicone Spatula',
'Fast Action Path Weeder RTU', 'Whisk', 'Profile Pen Blue',

'Bug & Fungus Kill', 'Rose Spray', 'Garden Fert Tub',
'Playing Cards', '4 Colours Pens Grip',
'Sticky Tape Refill 18mm x 25m', 'Heart Shaped Baking Tin',
'House & Garden', 'Home Beautiful', 'Household Scissors',
'Travel Sockblack 3-8', 'Threaded Needle Kit',
'Sphere Silicone Cube Tray', 'Clicker Erasable Pens',
'Open Umbrella Black', 'Pastel Highlighters Chisel Tip in Pouch',
'Bottle Opener', 'Essential Smart 4 4GX', 'Ice Cream Scoop',
'All Purpose NPK Fertiliser', 'Full Brief Size 18',
'Bamboo Shaper Bra Black 14/16', 'Wrench 8 Inch',
'Lightening Cable 1.2M', 'Original Protectant', 'Super Rig',
'Logo Crew Sock 11+', 'Rear Trigger Hand Sprayer',
'Hand Held Tax Calculator 8 Digit', 'Sticky Tape 18mmx66mm',
'Leather Garden Gloves Large', 'Permanent Black Marker',
'Tea Strainer', 'Trigger Sprayer', 'Bypass Pruner',
'Diamond Brass Padlock 30mm', '80 Lumen Rubber Flashlight 2 AAA',
'Double Adaptor Flat Right Hand',
'Large Gift Bag Pink & White Dots', 'Masher',
'2 Stage Knife Sharpener', 'Peppa Pig Play Pack', 'Pva Craft Glue',
'S-Gel Pens Black', 'Super Sticky Notes Miami',
'32800 Pantyhose Beige Tall', 'Silicone Pastry Brush',
'Fxt Massage Ball', 'Mens Logo Low Cut 11-14',
'Ladder Resist Natural Average', 'Multiguard Snail & Slug Killer',
'Kids Logo Light Quarter Crew Socks Size 13-3', 'Measuing Spoons',
'"5" Plush Assortment"', '4 Sided Box Grater',
'Flashlight 3 Button Batts', 'Value Anklet Socks Tan', 'Pliers',
'Kids Qtr Crew Sport Socks White 13-3', 'Epoxy 5min',
'Control Bottle', 'Measuring Cups', 'Silicone Spoonula',
'Automatic Night Light', 'Life Garden Tie Wire', 'Gue Candy Blast',
'Beach Doll', 'Gel Heel Cup', 'Weed N Feed Granular', 'Tube',
'The Complete Garden Health Treatment',
'Reporters Note Pad 300 Page', '90 Permanent Markers Black',
'Assorted Binder Clips 24mm', '1.5M HDMI',
'Blitzem Snail Slug Killer Bait Pellets',
'70D Opaque /No Dig Tight Black Medium/Large',
'Polyester Thread Multi Colours', 'Disc', 'White Oil',
'Deep Square Cake Tin 24Cm', 'Fxt Light Muscle Band',
'Adaptor Double Standard', 'Life Freezer',
'Windscreen Wash Additive', 'Smooth Footlet Natural',
'Micro SD Card With SD Adapter UHS-1 Class 10 R80 64GB Small',
'Neon Squish Ball', 'Boot Brown Lace 60"', 'Smooth Footlets',
'Eco 40D Opque Tights Black Tall/X Tall',
'Pastel & Neon Inspiration Set', 'Garden Spade', 'Claw Hammer',
'Marie Clare', 'A4 Writing Pad', 'Diabetic Living',
'Gel Activ Insole Work & Boot Large',
'Gel Activ Insole Work & Boot Small', 'Magnified Craft Light',
'Fast Action Organic Weed Kill',

'Ultra Low Cut Footlet Natural One Size',
'Whiteboard Markers Bullet Tip', 'Sheer Anklet 15D',
'Sheer Sockettes', 'Life 18 Can Cooler Backpack',
'Premium Garden Hose 12mmx15m', 'Polyester Thread Black & White',
'Salt & Pepper Grinder', 'Nappy Pins',
'Performance Lace 38 Solid Black', 'Rodara Screwdriver Set',
'Wooden Spoon', 'Profile Pen Black',
'Womens Logo No Show Size 3-8', 'Ice Cube Tray',
"That's Life Mega Monthly", 'Organic Weed Killer Spray', 'Kittens',
'Cooling Rack', 'Cristal Original Pens Blue',
'Pour+Feed House Plants Fertiliser', 'White Hank Elastic',
'2.4G 4-Axis 480P Camera', 'Minis Mystery Pack',
'Rope Cord 10 Metre', 'Unicorn Magic Slime Kit',
'Loose Base Quiche Pan 25.5cm', 'Value Anklet Socks Black',
'Egg Rings', 'Fxt Therapy Roller', 'Pizza Tray',
'Kids Logo Quarter Crew Socks Size Size 2-8',
'Correction Tape I Mini', 'Sheer Pantyhose Slim Black Medium',
'Fxt Message Roller Ball', 'Sink Plugs 50mm', 'Sneaker Balls',
'Multi Purpose Cleaner', 'Path Weeder Value Pack',
'F1 1:16 Scale RC Kit Set', 'Life Hose Set', 'Store & Shake',
'Greenlife Nordic Frypan', 'Cable Ties', 'Gourmet Traveller',
'Pizza Wheel', 'Markers Bullet Assorted', 'Ant Kill', 'Bluey',
'Utility Scraper', 'White Glue Stick', 'Fashion Tape',
'Party Toy Mix', 'Australian Country Style',
'Mens Sports Quarter Crew Socks', 'Shock Reducer Everyday Insole',
'Take 5 Monthly', 'Medium Gift Bag Blue', 'Liquid Nails',
'Bamboo Shaper Bra Black 10/12', 'Ladder Resist Natural Tall',
'Flosser', '4 Colour Pen Pastel',
'Large Gift Bag Blue & White Chevron', 'Lint Shaver',
'Girl Sneaker Footlet Black One Size', 'Clear Mounting Tape',
'All Purpose Plant Food', 'Fxt Phone Armband',
'Silicone Air Fryers Trays', 'Round Cookie Cutters', 'Stickers',
'Ladder Resist Natural/ X Tall', 'Gel Activ Insole Sport Large',
'6 Cup Jumbo Muffin Pan', '100 Lumen Spotlight Floating 4 AA',
'Gue Strawberry Donut', 'PANTYHOSE 20 DENIER X-TALL',
'Relief Pantyhose Trouser Sock Black', 'Sock Crew 13-3 / 3-8',
'4 Colour Pen', '2M Ethernet Cable Cat 6',
'Eco 40D Opque Tights Black Small/Medium', 'Sink Plug 38mm',
'Launcher', 'Life 9 Can Zipperless Cooler', 'Hand Mixer',
'Plastic Handle Scissors', 'Ballpen Black', '96mm Single Assorted',
'Medium Oven Tray', 'Low Cut Sport Socks White Sizes 5-8 & 9-12',
'Inkjoy Pen Black', 'Compact Needles Assorted',
'Manual Umbrella Black/white', 'Light', 'Opaque Anklets',
'Stitch Ripper', 'Bottle Bag Silver Spots',
'4 Outlets 5M Cassette Reel', '5 Cup Rice Cooker',
'Sheer Pantyhose Slim Black Small', 'A4 Binder Book 64 Page',
'Life Garden Hose 15m', 'Concentrate Advance Weedkiller',

'Spandex Gel Metatarsal Cushions', '17 Piece Tool Kit',
'Kilometrico Pens Blue', 'All Over Shaper Bronze X-Tall',
'Furblets Assorted', 'Pocketbook Take 5 Series',
'Multi Purpose Repair Sewing Needles', 'Brass Combination 30mm',
'Wall Repair Kit',
'Comfy Tops Semi Opaque Knee Hi 40 Denier Black 1 size',
'Glue Stick', '40d Opaque Knee High 1size Nude', 'Starter Set',
'Mini Greeting Card Girl & Balloons',
'Dynamic Lifter Pellet Fertiliser Organic',
'Kilometrico Pens Black', 'Mens Cushion Foot Crew Size 11-13',
'Fxt Fit Ball 65cm', 'Deep Round Pie Dish', 'Flair Candy Pop',
'Active Rib Legging Black M', 'Active Rib Legging Black L',
'Easican Can Opener', 'Life Hose Nozzle', 'Spakfilla Rapid',
'Directional Night Light', 'Womans Seamless Midi Brief Size 10',
'5M HDMI', 'Weed N Feed',
'Opaque Tights 70 Denier Black Tall/X-Tall', 'Umbrella Pin Dots',
'4 X USB 3.0 Hub', 'Selfie Ring 10 Inch E34', 'Set 98 Piece E34',
"Oval Lace Black 45''", 'Hoisery Footlet Size 8-11 Nude',
'Eco Cotton Tights Black Tall/X Tall', 'Street Machine',
'Clear Platic Ruler 30cm', 'A7 Pocket Notebook', '150d 1pk Md-lg',
'Fxt Medium Muscle Band', 'Herb Scissors', 'Rachet Pruner',
'150D Flex Leggings SM-MD', 'Mens Guyfront Trunk Assorted Small',
'Opaque Tights 40 Denier Black Ave/Tall', 'Opaque Tights 40d',
'Umbrella Black', 'Life Ice Bricks', 'Paint Brushes',
'Flat Lace White 45 inch', 'Sheer Pantyhose Slim Nude Medium',
'Sheer Pantyhose Slim Nude Large', 'Footlet Size 3+',
'Light Slimming Brief Natural Average', 'Oil & Grease Remover',
'Caddy For Remote Control', '2025 Kids Beanie',
'Opaque Knee Highs 40D', 'All Over Shaper Bronze Ave',
'Drive And Discover Baby Keys', 'Whiteboard Marker Blue',
'Sticky Tape Refills', 'Icons Series', 'Grid Grill Brush',
'Horizontal Lint Roller', 'Barbie Plush Assortment',
'200 Lumen Head Lamp 3 AAA',
'Womens Fluffy Qtr Crew Home Socks Size 5-10',
'Kids Low Cut Sport Socks White Size 13-3',
'Level Up & Learn Controller', 'Air Fyer Glass Baking Cups',
'New Idea Puzzler', 'Single Tub', 'Manual Umbrella Blue/ White',
'Heel & Arch Support Women', "Men's Heel and Arch Support",
'A4 Sheet Protector', 'Cookie Cutters Set',
'All Over Shaper Bronze Tall', 'Flexible Muffin Pan 6 Cup',
'Silicone Mini Muffin Pan 12 Cup',
'Monster Trucks 1 24 Assortment', "Men's Briefs Medium",
'Opaque Tights 70d', 'Umbrella Navy', 'Golf Ball',
'Washable Paper Bag', '2025 Socks',
'Bayer Baythroid Advanced Garden Insect Spray',
'Marker Brush Assorted', 'Animal Party Variety Pack',
'Fxt Heavy Muscle Band', 'Eco Cotton Tights Black Small/Medium',

'Eco Bio 130D Opque Tights Black Small/Medium',
'Womens Seamless Full Brief Size 12', 'Digital Timer', 'Pink',
"Men's Guyfront Trunk Size XL", 'Lux Velvet Tights 200D A/T',
'Luxury Velvet Tights 200 Denier Black Tall/X-Tall',
'Womens Hipster Bikini Brief Size 12',
'Mens Work Socks 7-10 Black', 'Fxt Push Up Bars',
'Boot Black Lace 60"', 'Curvy Trouser Sock 80D Black One Size',
'Kilometrico Pen Assorted', 'Womens Sport Tech No Show 8+',
'G2 Gel Ink Pens', 'Kittycorn Surprise', 'Lazy Susan Pantry Tray',
'A5 Spiral Notebook 140 pages',
'Footlet Sock Size 4 - 7 Ladies Tan',
'Eco Bio 130D Opque Tights Black Tall/X Tall', 'Bucket Hat',
'Shaped Colouring Book Assorted', 'P2 Particulate Respirator',
'Rainbowitos', 'Light Slimming Brief Natural X Tall',
'Womens Seamless Midi Brief Size 12',
'Womens Seamless Full Brief Size 14',
'Womens Seamless Bikini Brief Size 14', 'Belle',
'Outdoor Life 2 Hour Tap Timer',
'Womens Hipster Bikini Brief Size 10',
'Skinny Legging Mix Colour Medium/Large', 'A4 Colouring Book',
'Deep Round Cake Pan', 'Grip', 'No More Gaps',
'Go Sports Earphones', 'Ecopest Oil Insect Spray', 'Ovile Rodrigo',
'Pen Fun Colours 0.7mm', 'Invisible Footlet Nude 9-11', 'Lite 3',
'Mens Original Wool Crew 6-10', 'Notebook Carlton',
'Notebook Geelong', 'Notebook Hawthorn', 'Notebook Richmond',
'Notebook Gws', 'Notebook Gold Coast',
'Large Gift Bag Colourful Confetti',
'Kids Logo Light Quarter Crew Size 2 - 8',
'Automatic Red Clothes Brush',
'Kids Non Slip Fluffy Home Socks 13-3',
'Kids Qtr Crew Sport Socks White 5-8', 'Infant Cotton Rib Girl',
'Baking Dish', 'Markers Fine Tip Assorted',
'Skinny Legging Mix Colour Small/Medium', 'Cooking With AWW',
'Little People Babies Figure And Gear', 'Whiteboard Marker Black',
'6 In 1 Rachet Screwdriver', 'Shaver Lint Remover E34',
'No Shine Opaque Tights 70 Denier Black Tall/X-Tall',
'Little People Figure', 'Pop It Game', 'Notebook Port Adelaide',
'Notebook Sydney', 'Refill Black', 'Anti Static Spray',
'Life Circle Sprinkler', 'Grater', 'Plastic', 'Notebook Broncos',
'Notebook Panthers', 'Notebook Raiders', 'Notebook Roosters',
'Notebook Storm', 'Portion Freezer Pods',
'Prep Storage Box With 6 Dividers And Lid', 'Notebook Dragons',
'Notebook Knights', 'Notebook West Tigers', 'Silk Caresse 15D',
'Invisible Band Knee High 30D Natural One Size',
'Eco 15D Anklets Black One Size', 'Gift Wrap Tape Dispenser',
'Eels', 'Pantyhose Beige Average',
'Bottle Bag Kraft & Gold Stripes', 'Roller With 2 Refills E34',

'Gold 10M Curling Ribbon',
'Baby Blue Fabric Self Adhesive Gift Bow',
'Gold Fabric Selfadhesive Gift Bow', 'Footlet Size 8+',
'Iron On Viledon Tape', 'Erasable Felt Pens',
'Medium Gift Bag Pink', 'Cristal Original Pens Black', 'Rasoplast',
'Ping Pong Balls', 'Mini Greeting Card For You', 'Saver',
'40cm Pedestal Fan', 'Pattern Knee High', 'Freezer Pod',
'Or Asparagus Keeper', 'Adults Asst Colours',
'Pantyhose Black Average', 'Wrapping Paper Holographic Hearts',
'Holograph White 10M Curling Ribbon',
'Baby Pink Fabric Self Adhesive Gift Bow',
'White Fabric Self Adhesive Gift Bow',
'Sheer Pantyhose Slim Nude Small',
'Sheer Pantyhose Slim Black Large', 'Washbag Floral',
'Silicone Trivet', 'White Rayon Elastic 3M', 'Jute Bag',
'Mini Greeting Card Hooray',
'Mini Greeting Card Pretty Pink Floral', '$2 donation',
'United for Sick Kids $2 Donation', '$1 donation',
'Drink Bottles With Crystals', 'Handheld Sewing Machine E34',
'Thrive All Purpose Soluble Fertiliser',
'Power Grip Plastics Glue', 'Diamond Brass Padlock 20mm',
'Eco Sneaker Footlet Black One Size', 'Saver Container',
'Management Box', 'Jumbo Chalk Assorted', 'Poke Ball Collection',
'Silicone Wet Area White', 'Invisible Footlet Nude 5-8',
'Crew Sports Socks Storm', 'Womens Logo Low Cut Size 8-11',
'Vehicle Assorted', 'Notebook Adelaide', 'Notebook Essendon',
'Notebook Western Bulldogs', 'Hook & Loop Dots Self Stick',
'Fxt Weighted Skipping Rope',
'Kids Qtr Crew Sport Socks White 9-12', 'Notebook Bulldogs',
'Notebook Sharks', 'Notebook Warriors',
'Medium Gift Bag Colourful Balloons', 'Ladle',
'Sheer Knee Highs 15 Denier Black Size 1',
'Party Feet Ball Of Foot Gel Cushions', 'Counting Money Jar E35',
'Collectible Dino Asst', '4 Colours Pen Shine',
'Textured Rib Legging Medium/ Long',
'Curvy Very Thick Tights 120D Black - 2', 'Soaker Hose 7.5m',
'No Shine Opaque Tights 70 Denier', 'Invisible Footlet Blk 9-11',
'Invisible Footlet 5-8', 'Felt Pads', 'Profile Gel Pen Assortment',
'2M TV Aerial Cable', 'Eco 15D Knee High Black One Size',
'Glue Gun E34', 'Notebook Brisbane', 'Notebook Fremantle',
'Notebook Melbourne', 'Notebook North Melbourne',
'Notebook St Kilda', 'Notebook West Coast',
'Kids Low Cut Sports Socks White Size 9 -12',
'Infant Cotton Rib Assorted', 'Dish', 'Notebook Cowboys',
'Notebook Sea Eagles', 'Notebook Eels', 'Notebook Titans',
'Notebook Dolphins', 'Trading Cards Blister 1',
'Green Book Of Colouring', 'Carlton',

```
'Compact Lint Roller 2 Refills', 'Travel Sewing Kit',
'Relief 40d Compression Tall/xtall', 'Original Yellow Notes',
'Invisible Band Anklets Natural One Size', 'Weed Trimmer Line',
'Wrapping Paper Colourful Dots', 'Jug', 'Blue 10M Curling Ribbon',
'Black Laces 69Cm', 'Classic Fishnet Black One Size',
'Medium Gift Bag Solid White', 'Mini Greeting Card Xoxo',
'Pantyhose Beige X-Tall', 'Pink 10M Curling Ribbon',
'Core Doll Assorted', 'OSFA Pantyhose Knee Hi Mid Beige',
'Permanent Markers Ultra Fine Tip', 'Velocity Mechanical Pencils',
'Plunger', 'Mini Greeting Card Thank You', '$1 Donation',
'Family Anti Dandruf Shampoo Foam', 'Massage Pillow', 'Bib',
'Popover Bibs', 'Plastic Lined Terry Bibs', 'Dribble Bandana Bibs',
'A4 Notebook 240pg', 'Soft Grip Scissors 19Cm',
'Whiteboard Markers Starter Set', 'Hb Pencils',
'Lined Sticky Notes 100 Sheets', 'Gelocity Original Pens Blue',
'A5 Spiral Notebook', '4 Colour 3+1 Mechanical Pencil',
'Correction Bond White', 'Jumbo Crayons',
'Unicorn Barrel Pencil Case', '10cm USB Fan',
'20cm Folding USB Desk Fan', '30cm Desk Fan',
'Activity Assortment', 'Rechargeable Hand Held Fan',
'Antimicrobial 5" Scissors', 'Whiteboard Markers Chisel Tip',
'A4 Pastel 2 Ring Binder', 'Cristal Renew Barrel Pen + Refills',
'Tri Stik Glue', 'Gue Cotton Candy',
'Magnetic Whiteboard With Marker 35cm X 25cm', 'Notes Mini',
'Highlighter', 'Slime Starter Kit', 'Frixion Erasable Fineliner',
'Shopping List 50 sheets', '3.5"Clip On Plush Assorted',
'Sticky Note Tab Pet', 'Safety Dispenser Small',
'Dinosaur Night Slime Kit', 'Clicker 1 Pack And 3 Refills',
'2L Hot Water Bottle',
'Packing Tape Clear Dispenser With Recycled Handles 48mm X 20m',
'A4 Refill Paper Pad 240 Page', 'Page Markers', 'Flash Cards',
'Retractable Ballpoint Pen Assorted', 'Sticky Dots',
'Puffy Stickers', 'Metal Pens'], dtype=object)
```

We can observe that, these items do not require specific weight conditions for price changes when purchasing these items. Therefore, we can use the unit weight as the item weight here.

Below are the instances where item size is not available but the unit weight is available.

```
[78]: data[data['item_size'].isna() & data['unit_weight'].notna()]
```

```
[78]:       product_code        category  \
     1460       1113465          EASTER
     1461       4886751          EASTER
     1462       4898604          EASTER
     1463       5766234          EASTER
     1464       4885191          EASTER
      …             …               …
```

```
77828       4163437      INFANT FOOD
77829       3754198   INFANT PERSONAL
77830       6295520      INFANT FOOD
77831       6295585      INFANT FOOD
77832       3962038      INFANT FOOD

                                              item_name  \
1460                          Chocolate Easter Bunny
1461              Selections Easter Chocolate Eggs Bag
1462                    Easter Gold Bunny Milk Chocolate
1463              Chocolate Easter Hunt Assortment Pack
1464                Dairy Milk Easter Chocolate Eggs Bag
…                                                    …
77828              Veggie Tubes Sweetcorn & Carrot  12m+
77829                           100% Pure Cotton Ripple
77830      Pumpkin Potato & Beef Baby Food Pouch 6+ Months
77831      Lamb Vegetables & Polenta Baby Food Pouch 6+ M…
77832                    Beef With Vegetables 6+ Months

                              brand  item_price item_size  unit_price  \
1460          Chocolate Easter Bunny         8.0       NaN        4.00
1461      Selections Easter Chocolate        11.5       NaN        3.48
1462                         Unknown        18.0       NaN        6.25
1463                         Unknown        17.5       NaN        2.40
1464                    Dairy Milk Easter      6.7       NaN        3.51
…                                 …           …         …           …
77828                        Unknown         0.0       NaN       16.67
77829                        Unknown         0.0       NaN        1.95
77830                        Unknown         0.0       NaN        1.33
77831                        Polenta         0.0       NaN        1.33
77832                        Unknown         0.0       NaN        0.83

       unit_weight  similarity_score    subcat  unit_weight_new  \
1460          100g          0.171870   Unknown            100.0
1461          100g          0.170491   Unknown            100.0
1462          100g          0.103036      Milk            100.0
1463          100g          0.106898   Unknown            100.0
1464          100g          0.160387      Milk            100.0
…              …                 …         …                …
77828         100g          0.439613   Unknown            100.0
77829         100g          0.267265   Unknown            100.0
77830         100g          0.439613      Beef            100.0
77831         100g          0.439613      Lamb            100.0
77832         100g          0.439613      Beef            100.0

       unit_weight_type  item_weight_new  item_weight_type  std_item_weight  \
1460                  g              NaN               NaN              NaN
```

1461	g	NaN	NaN	NaN
1462	g	NaN	NaN	NaN
1463	g	NaN	NaN	NaN
1464	g	NaN	NaN	NaN
...	...	...	...	...
77828	g	NaN	NaN	NaN
77829	g	NaN	NaN	NaN
77830	g	NaN	NaN	NaN
77831	g	NaN	NaN	NaN
77832	g	NaN	NaN	NaN

	std_item_unit
1460	NaN
1461	NaN
1462	NaN
1463	NaN
1464	NaN
...	...
77828	NaN
77829	NaN
77830	NaN
77831	NaN
77832	NaN

[64541 rows x 16 columns]

Below are the categories of instances where item size is not available but the unit weight is available.

```python
[79]: data.loc[data['item_size'].isna() & data['unit_weight'].notna(), 'category'].
      ↪unique()
```

```
[79]: array(['EASTER', 'BOXED CHOCOLATES', 'BAKERY EASTER', 'SEASONAL EVENTS',
             'SEAFOOD DELI', 'VINE VEGETABLES', 'SNACKS', 'POTATOES', 'LAMB',
             'NUTS/DRIED', 'CONTINENTAL', 'SLICED MEATS', 'SAUSAGES',
             'VALUE ADDED FP', 'BERRIES & CHERRIES', 'PORK', 'FROZEN DESSERTS',
             'BH BOUGHT IN EASTER', 'TOMATOES', 'BAKERY BOUGHT IN', 'BEEF',
             'ANTIP/OLIVE/DIP/PATE', 'BAKERY PACKAGED CAKE', 'HARD VEGETABLES',
             'POULTRY (MT)', 'BISCUITS & COOKIES', 'FROZEN VEGETABLES',
             'PREPACKAGED SEAFOOD', 'ENTERTAINMENT', 'GOURMET CHEESE',
             'VALUE ADDED', 'SMALLGOODS DY', 'SOFTDRINKS', 'CHILLED DESSERTS',
             'SOFT VEGETABLES', 'BUNCH LINES', 'DRIED FRUIT/NUTS', 'WINE',
             'SAUCES/RELISH', 'BAKING MIXES', 'INSTORE CAKE', 'PROTEIN & MEALS',
             'CONFECTIONERY', 'PIES AND QUICHES', 'DISPOSABLE TABLEWARE',
             'PICNIC POOL BBQ ACC', 'MEALS', 'FROZEN SNACKS', 'GRAVY/STOCK',
             'GRAPES', 'MEAL BASES', 'BANANAS', 'CHEESE DAIRY',
             'CHILLED SPREADS', 'CEREAL', 'MELONS & TROPICAL', 'CITRUS',
             'APPLES & PEARS', 'INSTORE BREAD', 'BARS GUM POCKET PACK',
             'SPREADS', 'SOUP', 'BAKERY SNACKS', 'BEAUTY/TRIAL TRAVEL',
```

```
    'ASIAN FOODS', 'CANNED FISH', 'CHILLED HEALTHFOODS',
    'SMOKED & CURED FISH', 'NUTRITIONAL SNACKS', 'SMALLGOODS', 'RICE',
    'PASTA', 'JUICES & CORDIALS', 'MEXICAN FOODS', 'FACIAL TISSUES',
    'JUICES/DRINKS', 'FROZEN WORLD FOODS', 'KITCHEN NEEDS/BAGS',
    'HOMEWARES', 'SALAD DRESSINGS', 'FROZEN SEAFOOD', 'SKIN CARE',
    'WOMENS DEOS & GROOM', 'MEDICINAL PRODUCTS', 'ORGANIC MEAT',
    'SANITARY PROTECTION', 'FRUIT-SHELF STABLE', 'DY MILK',
    'INDIAN FOODS', 'CANNED MEALS', 'KITCHENWARE', 'SOAPS & BODY WASH',
    'POULTRY (DL)', 'INSTANT NOODLES', 'PEST CONTROL',
    'MENS DEOS & GROOMING', 'INFANT PERSONAL', 'HAIR CARE', 'VITAMINS',
    'DESSERTS (GR)', 'LAUNDRY ACCESSORIES', 'DENTAL HEALTH',
    'ICE CREAM', 'COSMETICS/TOILETRIES', 'AIRCARE', 'CLEANING GOODS',
    'HEALTH FOODS', 'SHOE CARE', 'CONVENIENCE MEALS',
    'SPEC/FRESH CHEESE', 'VINEGAR', 'HAMS', 'GAME', 'STONE FRUIT',
    'MUSHROOMS', 'ORGANIC FRUIT & VEG', 'AVOCADO & MANGO',
    'HERBS & CHILLIES', 'OILS', 'GROCERY MILK', 'MILK ADDITIVES',
    'DAIRY WORLD FOODS', 'GRAB & GO', 'BREAD ROLLS & FBREAD', 'BBQ',
    'DL PREPARED FOODS', 'SPICES/HERBS', 'COFFEE', 'FLOUR',
    'SUGAR/SWEETENERS', 'CANNED VEG', 'INTERNATIONAL FOODS',
    'PICKLED VEGETABLES', 'TEA', 'INFANT FOOD', 'INFANT FORMULA',
    'BETTER FOR YOU', 'COLD DRINK', 'FRESH PET FOOD DY',
    'ENERGY/SPORT/ICEDTEA', 'LIQUOR RTD', 'WATER', 'DISHWASHING',
    'TOILET PAPER', 'PAPER TOWELS', 'LAUNDRY', 'ELECTRICAL',
    'HOUSEHOLD GLOVES', 'PARTY GOODS', 'BRUSHWARE', 'HARDWARE', 'AUTO',
    'INFANT NAPPIES'], dtype=object)
```

[82]: `data.loc[data['item_size'].isna() & data['unit_weight'].notna(), 'item_name'].`
      `↪unique()`

[82]: 17043

1.11   Make NaN item weights using unit weights

This fill the NaN item weights of items where the unit weight is available.

[83]:
```python
# Fill item_weight_new with unit_weight where item_weight_new is NaN
data.loc[data['item_weight_new'].isna() & data['unit_weight'].notna(),
↪'item_weight_new'] = \
    data['unit_weight_new']
```

```
C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexing.py:1676:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self._setitem_single_column(ilocs[0], value, pi)
```

```
[85]: data[['item_name','item_size','item_weight_new','unit_weight']]
```

```
[85]:                                        item_name item_size  \
      0          Coles Hot Cross Buns Traditional Fruit | 6 Pack    6 Pack
      1                   Coles Hot Cross Buns Choc Chip | 6 Pack    6 Pack
      2        Coles Hot Cross Buns Traditional Fruit Mini | …    9 Pack
      3        Cadbury Dairy Milk Easter Chocolate Eggs Bag |…      114g
      4          Coles Hot Cross Buns Apple & Cinnamon | 6 Pack    6 Pack
      ...                                              ...        ...
      77828             Veggie Tubes Sweetcorn & Carrot  12m+       NaN
      77829                        100% Pure Cotton Ripple         NaN
      77830     Pumpkin Potato & Beef Baby Food Pouch 6+ Months     NaN
      77831  Lamb Vegetables & Polenta Baby Food Pouch 6+ M…       NaN
      77832                     Beef With Vegetables 6+ Months      NaN

            item_weight_new unit_weight
      0                 6.0         1ea
      1                 6.0         1ea
      2                 9.0         1ea
      3               114.0        100g
      4                 6.0         1ea
      ...               ...         ...
      77828           100.0        100g
      77829           100.0        100g
      77830           100.0        100g
      77831           100.0        100g
      77832           100.0        100g

      [77833 rows x 4 columns]
```

```
[86]: data.loc[data['item_size'].isna() & data['unit_weight'].notna(),
      ↪'item_weight_type'] = \
          data['unit_weight_type']
```

```
C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexing.py:1676:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self._setitem_single_column(ilocs[0], value, pi)
```

```
[87]: data[['item_name','item_size','item_weight_new','unit_weight','item_weight_type']]
```

```
[87]:                                        item_name item_size  \
      0          Coles Hot Cross Buns Traditional Fruit | 6 Pack    6 Pack
      1                   Coles Hot Cross Buns Choc Chip | 6 Pack    6 Pack
```

```
2          Coles Hot Cross Buns Traditional Fruit Mini | …     9 Pack
3          Cadbury Dairy Milk Easter Chocolate Eggs Bag |…      114g
4              Coles Hot Cross Buns Apple & Cinnamon | 6 Pack    6 Pack
…                                                          …       …
77828                  Veggie Tubes Sweetcorn & Carrot  12m+      NaN
77829                          100% Pure Cotton Ripple            NaN
77830       Pumpkin Potato & Beef Baby Food Pouch 6+ Months      NaN
77831   Lamb Vegetables & Polenta Baby Food Pouch 6+ M…          NaN
77832                     Beef With Vegetables 6+ Months          NaN


       item_weight_new unit_weight item_weight_type
0                  6.0         1ea             pack
1                  6.0         1ea             pack
2                  9.0         1ea             pack
3                114.0        100g                g
4                  6.0         1ea             pack
…                  …           …                 …
77828            100.0        100g                g
77829            100.0        100g                g
77830            100.0        100g                g
77831            100.0        100g                g
77832            100.0        100g                g


[77833 rows x 5 columns]
```

```
[77]: data[data['subcat']=='Unknown']
```

```
[77]:      product_code        category  \
      1         7473849          Easter
      4         3378370          Easter
      5         9633786          Easter
      6         9077853          Easter
      8         5854717     Footy's Back
      …             …             …
      77824     8421115       INFANT FOOD
      77826     8988935       INFANT FOOD
      77827     3292911       INFANT FOOD
      77828     4163437       INFANT FOOD
      77829     3754198   INFANT PERSONAL


                                          item_name                 brand  \
      1                Coles Hot Cross Buns Choc Chip | 6 Pack                Coles
      4          Coles Hot Cross Buns Apple & Cinnamon | 6 Pack                Coles
      5      Cadbury Selections Large Easter Chocolate Eggs…              Cadbury
      6      Greens Deluxe Jam & Cream Lamington Cake Roll …               Greens
      8      Mcvities Digestives Biscuits Dark Chocolate | …  Mcvities Digestives
      …                                                …                    …
```

```
77824                    Organic Tangy Tomato Chickpea Waves              Unknown
77826  Carrots Red Lentils & Cheddar Baby Food Pouch …              Unknown
77827                             Little Rollies Hazelnut              Unknown
77828                 Veggie Tubes Sweetcorn & Carrot  12m+              Unknown
77829                          100% Pure Cotton Ripple              Unknown

       item_price item_size  unit_price unit_weight  similarity_score  \
1            4.40    6 Pack        0.73         1ea          0.198373
4            4.40    6 Pack        0.73         1ea          0.198373
5           13.00      545g        2.39        100g          0.135630
6            6.00      300g        2.00        100g          0.213812
8            2.75      266g        1.03        100g          0.271592
…             …         …          …           …                 …
77824        7.50       NaN        7.81        100g          0.138376
77826        0.00       NaN        1.92        100g          0.439613
77827        0.00       NaN        8.00        100g          0.439613
77828        0.00       NaN       16.67        100g          0.439613
77829        0.00       NaN        1.95        100g          0.267265

          subcat  unit_weight_new unit_weight_type  item_weight_new  \
1        Unknown              1.0             each              6.0
4        Unknown              1.0             each              6.0
5        Unknown            100.0                g            545.0
6        Unknown            100.0                g            300.0
8        Unknown            100.0                g            266.0
…           …                 …                …                …
77824    Unknown            100.0                g              NaN
77826    Unknown            100.0                g              NaN
77827    Unknown            100.0                g              NaN
77828    Unknown            100.0                g              NaN
77829    Unknown            100.0                g              NaN

       item_weight_type  std_item_weight std_item_unit
1                  pack              6.0          pack
4                  pack              6.0          pack
5                     g            545.0             g
6                     g            300.0             g
8                     g            266.0             g
…                    …                …             …
77824               NaN              NaN           NaN
77826               NaN              NaN           NaN
77827               NaN              NaN           NaN
77828               NaN              NaN           NaN
77829               NaN              NaN           NaN

[59230 rows x 16 columns]
```

Below plot shows the categories in which both the item size and unt weight size is not available.

```
[89]: category_count_of_non_weights = data.loc[data['item_size'].isna() &␣
      ↪data['unit_weight'].isna(), 'category'].value_counts()
      import seaborn as sns

      plt.figure(figsize=(12, 8))
      sns.barplot(x=category_count_of_non_weights.values,␣
      ↪y=category_count_of_non_weights.index, palette="Blues_r")
      plt.xlabel('Number of Items')
      plt.ylabel('Category')
      plt.title('Number of Items with Missing Size and Weight by Category')
      plt.show()
```
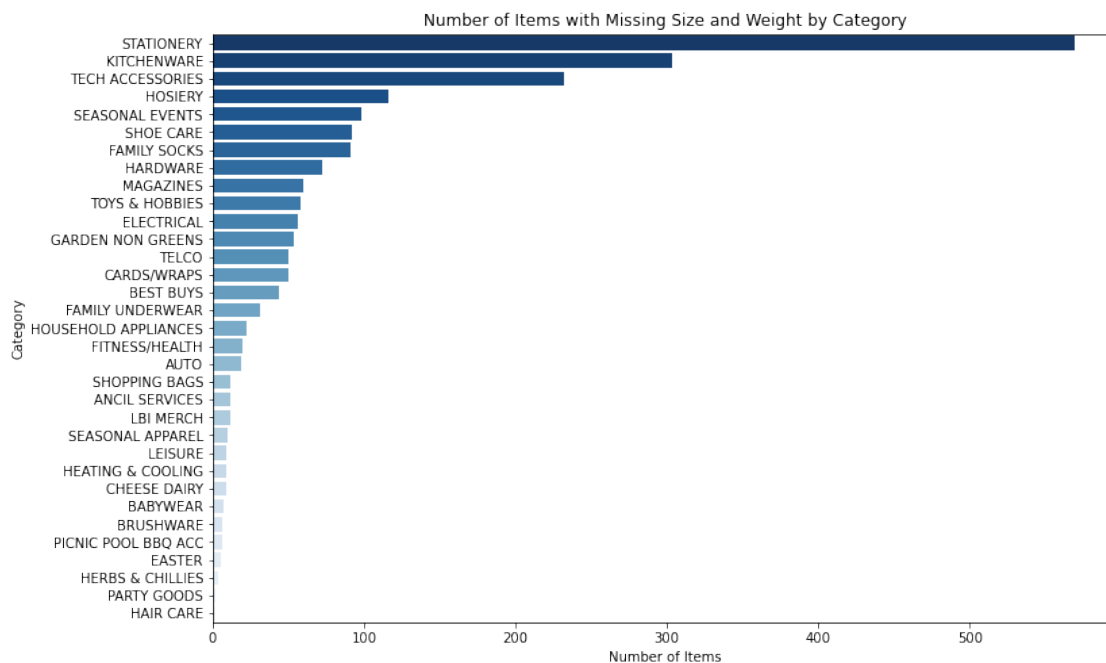


```
[101]: category_count_of_non_weights
```

```
[101]: STATIONERY            569
       KITCHENWARE           303
       TECH ACCESSORIES      232
       HOSIERY               116
       SEASONAL EVENTS        98
       SHOE CARE              92
       FAMILY SOCKS           91
       HARDWARE               72
       MAGAZINES              60
       TOYS & HOBBIES         58
       ELECTRICAL             56
```

```
GARDEN NON GREENS       54
TELCO                   50
CARDS/WRAPS             50
BEST BUYS               44
FAMILY UNDERWEAR        31
HOUSEHOLD APPLIANCES    22
FITNESS/HEALTH          20
AUTO                    19
SHOPPING BAGS           12
ANCIL SERVICES          12
LBI MERCH               12
SEASONAL APPAREL        10
LEISURE                  9
HEATING & COOLING        9
CHEESE DAIRY             9
BABYWEAR                 7
BRUSHWARE                6
PICNIC POOL BBQ ACC      6
EASTER                   5
HERBS & CHILLIES         4
PARTY GOODS              2
HAIR CARE                2
Name: category, dtype: int64
```

Most products are stationeries, kitchenware, tech accessories, and hosiery.

```python
[90]: # Fill subcategory with category where brand and subcategory are 'Unknown'
      data.loc[(data['brand'] == 'Unknown') & (data['subcat'] == 'Unknown'),␣
      ↪'subcat'] = \
          data.loc[(data['brand'] == 'Unknown') & (data['subcat'] == 'Unknown'),␣
      ↪'category']
```

```
C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexing.py:1676:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self._setitem_single_column(ilocs[0], value, pi)
```

```python
[91]: data[data['subcat']=='Unknown']
```

```
[91]:        product_code        category  \
      1            7473849          Easter
      4            3378370          Easter
      5            9633786          Easter
      6            9077853          Easter
```

```
8            5854717       Footy's Back
…                …                 …
77781        3784224   INFANT PERSONAL
77784        3841565   INFANT PERSONAL
77807        6027076      INFANT FOOD
77811        6252469      INFANT FOOD
77820        4151992   INFANT PERSONAL

                                          item_name  \
1                 Coles Hot Cross Buns Choc Chip | 6 Pack
4             Coles Hot Cross Buns Apple & Cinnamon | 6 Pack
5       Cadbury Selections Large Easter Chocolate Eggs…
6       Greens Deluxe Jam & Cream Lamington Cake Roll …
8       Mcvities Digestives Biscuits Dark Chocolate | …
…                                                     …
77781                            Training Cups Assorted
77784                                 Replacement Straw
77807   Spinach Apple Broccoli & Pea and Nothing Else …
77811                                         Melts 80g
77820                                  Weaning Spoon Set

                        brand  item_price item_size  unit_price unit_weight  \
1                       Coles        4.40    6 Pack        0.73         1ea
4                       Coles        4.40    6 Pack        0.73         1ea
5                     Cadbury       13.00      545g        2.39        100g
6                      Greens        6.00      300g        2.00        100g
8           Mcvities Digestives        2.75      266g        1.03        100g
…                           …         …         …           …           …
77781   Training Cups Assorted        0.00       NaN       11.00         1ea
77784        Replacement Straw        0.00       NaN       10.00         1ea
77807                 Broccoli        2.15       NaN        1.50        100g
77811                Melts 80g        0.00        80g        4.81        100g
77820                    Spoon        0.00       NaN        1.67         1ea

        similarity_score    subcat  unit_weight_new unit_weight_type  \
1               0.198373   Unknown              1.0             each
4               0.198373   Unknown              1.0             each
5               0.135630   Unknown            100.0                g
6               0.213812   Unknown            100.0                g
8               0.271592   Unknown            100.0                g
…                      …         …                …                …
77781           0.267265   Unknown              1.0             each
77784           0.267265   Unknown              1.0             each
77807           0.311051   Unknown            100.0                g
77811           0.439613   Unknown            100.0                g
77820           0.267265   Unknown              1.0             each
```

```
           item_weight_new item_weight_type  std_item_weight std_item_unit
1                      6.0             pack              6.0          pack
4                      6.0             pack              6.0          pack
5                    545.0                g            545.0             g
6                    300.0                g            300.0             g
8                    266.0                g            266.0             g
...                    ...              ...              ...           ...
77781                  1.0             each              NaN           NaN
77784                  1.0             each              NaN           NaN
77807                100.0                g              NaN           NaN
77811                 80.0                g             80.0             g
77820                  1.0             each              NaN           NaN

[23240 rows x 16 columns]
```

[92]: `data.isna().sum()`

```
[92]: product_code            0
      category                0
      item_name               0
      brand                   0
      item_price              0
      item_size           66683
      unit_price           2237
      unit_weight          2238
      similarity_score        0
      subcat                  0
      unit_weight_new      2244
      unit_weight_type     2244
      item_weight_new      2142
      item_weight_type     2142
      std_item_weight     66683
      std_item_unit       66683
      dtype: int64
```

[93]: `data[data['item_size'].isna() & data['unit_weight'].notna()]`

```
[93]:       product_code         category  \
      1460       1113465           EASTER
      1461       4886751           EASTER
      1462       4898604           EASTER
      1463       5766234           EASTER
      1464       4885191           EASTER
      ...            ...              ...
      77828      4163437      INFANT FOOD
      77829      3754198  INFANT PERSONAL
      77830      6295520      INFANT FOOD
```

```
77831        6295585        INFANT FOOD
77832        3962038        INFANT FOOD

                                          item_name  \
1460                            Chocolate Easter Bunny
1461                Selections Easter Chocolate Eggs Bag
1462                     Easter Gold Bunny Milk Chocolate
1463                Chocolate Easter Hunt Assortment Pack
1464                  Dairy Milk Easter Chocolate Eggs Bag
…                                               …
77828                  Veggie Tubes Sweetcorn & Carrot  12m+
77829                          100% Pure Cotton Ripple
77830      Pumpkin Potato & Beef Baby Food Pouch 6+ Months
77831      Lamb Vegetables & Polenta Baby Food Pouch 6+ M…
77832                        Beef With Vegetables 6+ Months

                               brand  item_price item_size  unit_price  \
1460            Chocolate Easter Bunny         8.0       NaN        4.00
1461      Selections Easter Chocolate        11.5       NaN        3.48
1462                          Unknown        18.0       NaN        6.25
1463                          Unknown        17.5       NaN        2.40
1464                    Dairy Milk Easter         6.7       NaN        3.51
…                                  …          …        …           …
77828                         Unknown         0.0       NaN       16.67
77829                         Unknown         0.0       NaN        1.95
77830                         Unknown         0.0       NaN        1.33
77831                         Polenta         0.0       NaN        1.33
77832                         Unknown         0.0       NaN        0.83


     unit_weight  similarity_score           subcat  unit_weight_new  \
1460        100g          0.171870          Unknown            100.0
1461        100g          0.170491          Unknown            100.0
1462        100g          0.103036             Milk            100.0
1463        100g          0.106898           EASTER            100.0
1464        100g          0.160387             Milk            100.0
…             …                 …                …                …
77828       100g          0.439613      INFANT FOOD            100.0
77829       100g          0.267265  INFANT PERSONAL            100.0
77830       100g          0.439613             Beef            100.0
77831       100g          0.439613             Lamb            100.0
77832       100g          0.439613             Beef            100.0


     unit_weight_type  item_weight_new item_weight_type  std_item_weight  \
1460                g            100.0                g              NaN
1461                g            100.0                g              NaN
1462                g            100.0                g              NaN
1463                g            100.0                g              NaN
```

```
      1464                   g         100.0              g              NaN
      ...                  ...           ...            ...            ...
      77828                  g         100.0              g              NaN
      77829                  g         100.0              g              NaN
      77830                  g         100.0              g              NaN
      77831                  g         100.0              g              NaN
      77832                  g         100.0              g              NaN

             std_item_unit
      1460             NaN
      1461             NaN
      1462             NaN
      1463             NaN
      1464             NaN
      ...              ...
      77828            NaN
      77829            NaN
      77830            NaN
      77831            NaN
      77832            NaN

      [64541 rows x 16 columns]
```

[97]: `data[data['category'] == 'STATIONERY']`

[97]:
```
             product_code    category                       item_name  \
      2556         7208484  STATIONERY     A4 Copy Paper 500 Sheets
      3266         5412934  STATIONERY                     Blu Tack
      3504         3782863  STATIONERY              Blu Tack Colour
      3543         3087242  STATIONERY                   School Glue
      3564         1842556  STATIONERY          Sticky tape 18mmx25m
      ...              ...         ...                           ...
      72404        3251142  STATIONERY    Permanent Markers Fine Tip
      72408        1840211  STATIONERY                  4 Colour Pen
      72410        8799579  STATIONERY        Clear Platic Ruler 30cm
      72412        3935013  STATIONERY                 A4 Writing Pad
      72473        3782863  STATIONERY              Blu Tack Colour

                         brand  item_price item_size  unit_price unit_weight  \
      2556       A4 Copy Paper         0.0       NaN         NaN         NaN
      3266            Blu Tack         3.1       NaN         NaN         NaN
      3504            Blu Tack         4.0       NaN         NaN         NaN
      3543         School Glue         7.0       NaN         NaN         NaN
      3564  Sticky tape 18mmx25m       0.0       NaN         NaN         NaN
      ...              ...          ...       ...         ...         ...
      72404     Markers Fine Tip       5.6       NaN         NaN         NaN
      72408             Unknown        3.2       NaN         NaN         NaN
```

141

```
72410        Unknown     0.0      NaN       NaN       NaN
72412        Unknown     0.0      NaN       NaN       NaN
72473        Blu Tack    4.0      NaN       NaN       NaN
```

```
        similarity_score      subcat  unit_weight_new unit_weight_type  \
2556            0.141170     Unknown             NaN             NaN
3266            0.155113     Unknown             NaN             NaN
3504            0.157536     Unknown             NaN             NaN
3543            0.144920     Unknown             NaN             NaN
3564            0.141170     Unknown             NaN             NaN
...                  ...         ...             ...             ...
72404           0.155627     Unknown             NaN             NaN
72408           0.155420  STATIONERY             NaN             NaN
72410           0.141170  STATIONERY             NaN             NaN
72412           0.141170  STATIONERY             NaN             NaN
72473           0.157536     Unknown             NaN             NaN
```

```
        item_weight_new item_weight_type  std_item_weight std_item_unit
2556                NaN              NaN              NaN           NaN
3266                NaN              NaN              NaN           NaN
3504                NaN              NaN              NaN           NaN
3543                NaN              NaN              NaN           NaN
3564                NaN              NaN              NaN           NaN
...                 ...              ...              ...           ...
72404               NaN              NaN              NaN           NaN
72408               NaN              NaN              NaN           NaN
72410               NaN              NaN              NaN           NaN
72412               NaN              NaN              NaN           NaN
72473               NaN              NaN              NaN           NaN

[582 rows x 16 columns]
```

1.11.1 Analyse categories where item size is not available and extract subcategory Stationary:

```python
[98]: data[data['category'] == 'STATIONERY']['item_name'].unique()
```

```
[98]: array(['A4 Copy Paper 500 Sheets', 'Blu Tack', 'Blu Tack Colour',
             'School Glue', 'Sticky tape 18mmx25m',
             'Fluo XL Assorted Highlighters', '4 Colour Pens',
             'Boss Highlighter Assorted Fluoro',
             'Liquid Paper Dryline Grip Correction Tape',
             'A4 Notebook 120 Page', 'Metallic Markers',
             'Paper Clips In Pouch Assorted Colours', 'USB 2.0 SM02 16GB',
             'A4 Display Folder', 'Super Sticky Cube',
             '570 Notebook A5 200 Page', 'SM02 USB 2.0 Flash Drive 64GB',
             'A4 Spiral Notebook 200 Page', 'To Do List 52 Sheets',
```

'Markers Fine Tip Colour Burst',
'Kids Evolution Colouring Pencils',
'Sticky Notes Assorted Colours', 'S-Gel Pens Assorted',
'Laundry Marker', 'Glu Stik 21g', 'Atlantis Pens Black',
'Letra Tag Labeller Blue', 'Correction Pen White',
'Fineliner Pens', 'Cristal Fashion Pens', 'PVC Free Eraser',
'A5 Spiral Notebook 200 pages', 'Intensity Colouring Felt Pens',
'Woodcase Hb Pencils', 'Connector Pens', 'Plastic Clear Tape',
'Traditional 2b Graphite Pencils', 'B5 Perforated Notebook 160Pg',
'Atlantis Pens Blue', 'Colourhide Notebook A4 120pages',
'Correction Tape', 'Kids Plastidecor Crayons', 'Clear School Glue',
'Cristal Original Pens Assorted', 'Intensity Fineliners Assorted',
'Whiteboard Markers Fine Tip',
'Kids Evolution Illusion Colouring Pencils',
'4 Colours Pens Limited Edition', 'Colouring Felt Pen Wallet',
'SDMI Card 32GB UHS 1 Class 10 BW', 'Plastic 2 Hole Sharpener',
'Tape Plastic', 'Colourhide Notebook A5 Assorted 200 page',
'Inkjoy 100 Assorted Colours', 'Double Sided Palm Cards',
'200 Fineline Black Pen', 'Woodcase 2B Pencil',
'Magic Blue Glue Stick 2x40g', 'Traditional Hb Graphite Pencils',
'USB 3.0 BE03 32GB', 'Unicorn Stationery Set 5 Piece',
'Classic Colouring Pencils', 'Colourhide Notebook 5 Subject A4',
'Kids Turn & Colour Crayons', 'Gelocity Original Pens Black',
'Magical Liquid', 'Permanent Markers Fine Tip',
'Clear Packaging Tape', 'Stapler With Staples',
'Mixed Rubber Bands In Pouch', 'Permanent Markers Bullet Tip',
'Masking Tape 24mmX18m', '32GB USB2.0 Twin Pack', 'Envelopes C6',
'Double Sided Tape', 'Assorted 0.7mm', 'Inkjoy Black Ball Point',
'Glu Stik Blister 35g', 'DL Envelopes', 'Ballpen Blue',
'Antimicrobial 7" Scissors', 'Plastic Push Pins In Pouch',
'Frixion Ball Pen Black/Blue/Red', 'Staples',
'Sticky Tape Dispenser', 'Profile Pen Blue', 'Playing Cards',
'4 Colours Pens Grip', 'Sticky Tape Refill 18mm x 25m',
'Clicker Erasable Pens', 'Pastel Highlighters Chisel Tip in Pouch',
'Hand Held Tax Calculator 8 Digit', 'Sticky Tape 18mmx66mm',
'Permanent Black Marker', 'Pva Craft Glue', 'S-Gel Pens Black',
'Super Sticky Notes Miami', 'Gue Candy Blast',
'Reporters Note Pad 300 Page', '90 Permanent Markers Black',
'Assorted Binder Clips 24mm',
'Micro SD Card With SD Adapter UHS-1 Class 10 R80 64GB Small',
'Pastel & Neon Inspiration Set', 'A4 Writing Pad',
'Whiteboard Markers Bullet Tip', 'Profile Pen Black',
'Cristal Original Pens Blue', 'Unicorn Magic Slime Kit',
'Correction Tape I Mini', 'Markers Bullet Assorted',
'White Glue Stick', '4 Colour Pen Pastel', 'Stickers',
'Gue Strawberry Donut', '4 Colour Pen', 'Ballpen Black',
'Inkjoy Pen Black', 'A4 Binder Book 64 Page',

```
        'Kilometrico Pens Blue', 'Glue Stick', 'Kilometrico Pens Black',
        'Flair Candy Pop', 'Clear Platic Ruler 30cm', 'A7 Pocket Notebook',
        'Whiteboard Marker Blue', 'Sticky Tape Refills',
        'A4 Sheet Protector', 'Marker Brush Assorted',
        'Animal Party Variety Pack', 'Kilometrico Pen Assorted',
        'G2 Gel Ink Pens', 'A5 Spiral Notebook 140 pages',
        'Shaped Colouring Book Assorted', 'A4 Colouring Book',
        'Pen Fun Colours 0.7mm', 'Markers Fine Tip Assorted',
        'Whiteboard Marker Black', 'Refill Black',
        'Gift Wrap Tape Dispenser', 'Erasable Felt Pens',
        'Cristal Original Pens Black', 'Rasoplast', 'Jumbo Chalk Assorted',
        '4 Colours Pen Shine', 'Profile Gel Pen Assortment',
        'Green Book Of Colouring', 'Original Yellow Notes',
        'Permanent Markers Ultra Fine Tip', 'Velocity Mechanical Pencils',
        'A4 Notebook 240pg', 'Soft Grip Scissors 19Cm',
        'Whiteboard Markers Starter Set', 'Hb Pencils',
        'Lined Sticky Notes 100 Sheets', 'Gelocity Original Pens Blue',
        'A5 Spiral Notebook', '4 Colour 3+1 Mechanical Pencil',
        'Correction Bond White', 'Jumbo Crayons',
        'Unicorn Barrel Pencil Case', 'Activity Assortment',
        'Antimicrobial 5" Scissors', 'Whiteboard Markers Chisel Tip',
        'A4 Pastel 2 Ring Binder', 'Cristal Renew Barrel Pen + Refills',
        'Tri Stik Glue', 'Gue Cotton Candy',
        'Magnetic Whiteboard With Marker 35cm X 25cm',
        'Erasable Highlighters 3pk', 'Notes Mini', 'Highlighter',
        'Slime Starter Kit', 'Frixion Erasable Fineliner',
        'Shopping List 50 sheets', 'Sticky Note Tab Pet',
        'Safety Dispenser Small', 'Dinosaur Night Slime Kit',
        'Clicker 1 Pack And 3 Refills',
        'Packing Tape Clear Dispenser With Recycled Handles 48mm X 20m',
        'A4 Refill Paper Pad 240 Page', 'Page Markers', 'Flash Cards',
        'Retractable Ballpoint Pen Assorted', 'Sticky Dots',
        'Puffy Stickers', 'Metal Pens'], dtype=object)
```

```python
[99]:  # define subcategories for Stationery
       subcategory_map = {
           "paper": "Paper",
           "notebook": "Notebook",
           "binder": "Notebook/Folder",
           "folder": "Notebook/Folder",
           "pad": "Notebook/Pad",
           "pen": "Pen",
           "pencil": "Pencil",
           "marker": "Marker",
           "highlighter": "Marker",
           "eraser": "Eraser",
           "correction": "Correction",
```

```python
        "glue": "Glue",
        "tack": "Adhesive",
        "tape": "Tape",
        "clip": "Clips & Pins",
        "rubber band": "Clips & Pins",
        "push pin": "Clips & Pins",
        "usb": "USB/Storage",
        "sd": "USB/Storage",
        "card": "USB/Storage",
        "scissor": "Scissors",
        "stapler": "Stapler",
        "staple": "Stapler",
        "label": "Labels",
        "envelope": "Envelope",
        "ruler": "Ruler",
        "whiteboard": "Board/Marker",
        "calculator": "Calculator",
        "sticker": "Stickers",
        "slime": "Fun/Creative Kits",
        "set": "Fun/Creative Kits",
        "kit": "Fun/Creative Kits"
}


def get_subcategory(item):
    item_lower = item.lower()
    for keyword, subcat in subcategory_map.items():
        if keyword in item_lower:
            return subcat
    return "Other"


data["subcat"] = data["item_name"].apply(get_subcategory)
```

```
<ipython-input-99-ecabb752d908>:44: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data["item_name"].apply(get_subcategory)
```

```python
[100]: data[data['category'] == 'STATIONERY']
```

```
[100]:       product_code    category                 item_name  \
       2556       7208484  STATIONERY    A4 Copy Paper 500 Sheets
       3266       5412934  STATIONERY                    Blu Tack
       3504       3782863  STATIONERY             Blu Tack Colour
       3543       3087242  STATIONERY                  School Glue
       3564       1842556  STATIONERY          Sticky tape 18mmx25m
```

```
...               ...         ...                                     ...
72404        3251142  STATIONERY        Permanent Markers Fine Tip
72408        1840211  STATIONERY                      4 Colour Pen
72410        8799579  STATIONERY          Clear Platic Ruler 30cm
72412        3935013  STATIONERY                     A4 Writing Pad
72473        3782863  STATIONERY                   Blu Tack Colour

                       brand  item_price item_size  unit_price unit_weight  \
2556            A4 Copy Paper         0.0       NaN         NaN         NaN
3266                 Blu Tack         3.1       NaN         NaN         NaN
3504                 Blu Tack         4.0       NaN         NaN         NaN
3543               School Glue         7.0       NaN         NaN         NaN
3564    Sticky tape 18mmx25m         0.0       NaN         NaN         NaN
...                      ...         ...       ...         ...         ...
72404      Markers Fine Tip         5.6       NaN         NaN         NaN
72408               Unknown         3.2       NaN         NaN         NaN
72410               Unknown         0.0       NaN         NaN         NaN
72412               Unknown         0.0       NaN         NaN         NaN
72473              Blu Tack         4.0       NaN         NaN         NaN

        similarity_score         subcat  unit_weight_new unit_weight_type  \
2556            0.141170          Paper              NaN              NaN
3266            0.155113       Adhesive              NaN              NaN
3504            0.157536       Adhesive              NaN              NaN
3543            0.144920           Glue              NaN              NaN
3564            0.141170           Tape              NaN              NaN
...                  ...            ...              ...              ...
72404           0.155627         Marker              NaN              NaN
72408           0.155420            Pen              NaN              NaN
72410           0.141170          Ruler              NaN              NaN
72412           0.141170   Notebook/Pad              NaN              NaN
72473           0.157536       Adhesive              NaN              NaN

        item_weight_new item_weight_type  std_item_weight std_item_unit
2556                NaN              NaN              NaN           NaN
3266                NaN              NaN              NaN           NaN
3504                NaN              NaN              NaN           NaN
3543                NaN              NaN              NaN           NaN
3564                NaN              NaN              NaN           NaN
...                 ...              ...              ...           ...
72404               NaN              NaN              NaN           NaN
72408               NaN              NaN              NaN           NaN
72410               NaN              NaN              NaN           NaN
72412               NaN              NaN              NaN           NaN
72473               NaN              NaN              NaN           NaN

[582 rows x 16 columns]
```

Kitchenware:

```
[102]: data[data['category'] == 'KITCHENWARE']['item_name'].unique()
```

```
[102]: array(['Roasting Pan', 'Roasting Pan with Rack', 'Ultra Container 1.3L',
              'Ultra Container 2.7L', 'Plus Rectangle Containers 1 Litre',
              'Ultra Container 700mL', 'Ultra Container 1.4L',
              'Ultra Container 1.9L', 'Microsafe Easy Heat Containers 900mL',
              'Ultra Container 920mL', 'Microsafe Vegetable Steamer 1.5L',
              'Microsafe Containers',
              'Microsafe Easy Heat Round Container 800mL',
              'Microsafe Easy Heat Value Pack',
              'Klip It Plus Rectangle Container 1L',
              'Klip It Plus Container 180mL',
              'Microsafe Easy Heat Container 1.6L', 'Ultra Container 460mL',
              'Klip It Plus Rectangle Container 2.2L',
              'Microsafe Easy Heat Container 900mL',
              'Plus Rectangle Containers 1.15 Litre',
              'Microsafe Easy Heat Containers 375mL',
              'Plus Square Containers 520ml', 'Tellfresh Cake Storer 6L',
              'Klip It Plus Rectangle Container 400mL',
              'Klip It Plus Square Container 1.15L', 'Klip It Plus Containers',
              'Stick Frypan', 'Celebration Cake Pan', 'Large Sieve',
              'Non Stick Frypan', 'Air Fryer Deep Baking Pan 17.5Cm',
              'Silicone Jumbo Muffin Cups', '3 In 1 Avocado Tool',
              'Food Containers 50mL', 'Swivel Peeler', 'Chopping Board Set',
              'Cooking Twine', 'Cooker', 'Measuring Cup Set',
              'Multigrip Mandoline', 'Scissors', 'Compact Knife Organiser',
              'Pizza Cutter', 'Silicone Icy Pole Mould', 'Mixing Bowl',
              'Spaghetti Spoon', 'Madeleine Pan 12 Cup', 'Meat Thermometer',
              'Food Containers 650mL', 'Sink Tidy', 'Silicone Muffin Cups',
              'Silicone Rolling Pin', 'Reuseable Bag Clips', 'Funnel',
              'Stick Sautepan', 'Meat Tenderiser', 'Silicone Trivet Set',
              'Folding Chopping Board', '12 Cup Mini Bundt Pan',
              'Aluminium Stock Pot 24cm', 'Easy Clean Garlic Press',
              'Bundt Cake Pan', 'Compact Hand Held Spiralizer', 'Pull Chopper',
              'Store & Pour', 'Piping Set', 'Sink Strainer',
              'Triple Screen Sifter', '8 Cup Mini Loaf Pan', 'Juicer',
              'Compact Juicer', 'Compact Cutlery Organiser',
              'Expanding Dish Rack', 'Egg Cooker', 'Silicone Baking Mat',
              'Silicone Air Fryer Liners Round', 'Milk Frother', 'Garlic Press',
              'Dumpling Press', 'Apple Corer', 'Air Fryer Baking Pan 18.5 Cm',
              'Zester', 'Food Containers 750mL', 'Food Containers 1L', 'Sieves',
              'Ultra Container 4L', 'Klip It Plus Rectangle Container 7.5L',
              'Frypan 28cm', 'Magican Can Opener', 'Frypan 26cm',
              'Springform Cake Pan', 'Food Containers 950mL',
              'Bake N Store Glass Container 2.5L', 'Bamboo Cocktail Sticks',
```

'Food Containers 300mL', 'Electric Salt & Pepper Grinder',
'Lamington Tray', 'Vent & Seal Leak Glass Container 1L',
"Let's Share Roaster 4L", 'Ultra Container 2.63L',
'Carbon Steel Wok 30cm', 'Container', '12 Cup Muffin Pan',
'Wooden Wok Spatula', 'Saucepan 18cm', 'Squeezy Bottles',
'Bamboo Chop Sticks', 'Collapsible Steamer', 'Saute Pan 3.5L',
'Bamboo Skewers', 'Food Containers 1.5L', 'Lighter',
'Chopping Board', 'Bamboo Board', 'Large Oven Tray',
'Oil Sprayer 300mL', 'Stackable Measuring Cups',
'Food Containers 2L', 'Ultra Container 1.73L',
'Digital Kitchen Scales', 'Silicone Spatula', 'Silicone Tongs',
'Induction Non Stick Grill Pan 26Cm', 'Dessert Spoons',
'Table Forks', 'Measuring Jug', 'Containers 870ml', 'Peeler',
'Food Containers 500mL', 'Detachable Kitchen Scissors',
'Klip It Plus Square Container 5.5L', 'Square Cake Tin 24cm',
'Tellfresh Super Storer 8.5L', 'Tongs 30cm',
'Aluminium Saute Pan 28cm', 'Can Opener',
'Stainless Steel Mixing Bowl 4L', 'Loaf Pan',
'Deeper Baking Dish 4.7L', '24 Cup Mini Muffin Pan',
'Slotted Turner', 'Containers 500ml',
'Simply Store Glass Rectangle Container 1.4L', 'Containers 250ml',
'Aluminium Frypan 24cm', 'Measuring Jug 500mL',
'Bamboo Cocktail Skewers', 'Cookie Scoop', 'Round Cake Tin 24cm',
'To Go Stax Container', 'Aluminium Deep Frypan',
'Klip It Plus Rectangle Container 3.35 Litre',
'Nordic Frypan 26Cm', 'Aluminium Saucepan 20cm',
'Vent & Seal Leak Proof Glass Container 750mL',
'Bake N Store Glass Container 1.52L', 'Deeper Baking Dish 3L',
'Nylon Tongs', 'Non Stick Frypan 20cm', 'Iconics Glass Bowl 3L',
'Cast Aluminium All Day Pan 28Cm', 'Teaspoons', 'Brownie Tin 20cm',
'1 Litre', 'Plus Rectangle Containers 2.2 Litre',
'Klip It Plus Square Container 3.5L', 'Silicone Ice Cube Tray',
'Tellfresh Cupcake Storer 4L',
'Simply Store Glass Rectangle Container 710mL', 'Handheld Grater',
'Staysharp Kitchen Scissors', 'Colander',
'Silicone Air Fryer Trays Round',
'Vent & Seal Leak Proof Glass Container 1.5L',
'Brilliance Food Storage 920mL', 'Spoon',
'Decor Tellfresh Container', 'Mini Tongs', 'Metal Skewers',
'Containers 3litre', 'Silicone Whisk', 'Aluminium Saucepan 16cm',
'Salt & Pepper Set', 'Mini Frypan',
'Multifunctional Mandoline Slicer', 'Mini Silicone Spatula',
'Whisk', 'Heart Shaped Baking Tin', 'Tellfresh Container 2L',
'Sphere Silicone Cube Tray', 'Bottle Opener', 'Ice Cream Scoop',
'Simply Store Glass Round Container 1.65L', 'Tea Strainer',
'Masher', 'Klip It Plus Cookie Container 2.2L',
'2 Stage Knife Sharpener', 'Food Containers 200mL',

```
       'Silicone Pastry Brush', 'Measuing Spoons', '4 Sided Box Grater',
       'Klip It Klip It Plus Cracker Container 900mL', 'Measuring Cups',
       'Silicone Spoonula', 'Deep Square Cake Tin 24Cm',
       'Salt & Pepper Grinder', 'Wooden Spoon', 'Ice Cube Tray',
       'Cooling Rack', 'Loose Base Quiche Pan 25.5cm', 'Egg Rings',
       'Pizza Tray', 'Store & Shake', 'Greenlife Nordic Frypan',
       'Pizza Wheel', 'Vent & Seal Leak Proof Glass Container 600mL',
       'VENT & SEAL LEAK PROOF GLASS CONTAINER',
       'Simply Store Glass Round Container 473mL',
       'Silicone Air Fryers Trays', 'Round Cookie Cutters',
       '6 Cup Jumbo Muffin Pan', 'Medium Oven Tray',
       'Cutlery Set 16 pieces', 'Starter Set', 'Deep Round Pie Dish',
       'Easican Can Opener', "Let's Share Roaster 2L", 'Herb Scissors',
       'Air Fyer Glass Baking Cups', 'Latte Spoons', 'Cookie Cutters Set',
       'Flexible Muffin Pan 6 Cup', 'Silicone Mini Muffin Pan 12 Cup',
       'Digital Timer', 'Deep Round Cake Pan', 'Baking Dish',
       'Iconics Glass Bowl 1L', 'Grater',
       'Brilliance Premium Food Storage 380mL', 'Silicone Trivet',
       'Measuring Jug 1L', 'Klip It Plus Cereal Container 4.2L', 'Ladle',
       'Simply Store Glass Round Container 946mL',
       'Food Containers 350mL'], dtype=object)
```

```
[103]: subcategory_map_kitchen = {
           # Cookware
           "frypan": "Cookware",
           "saute": "Cookware",
           "grill pan": "Cookware",
           "wok": "Cookware",
           "saucepan": "Cookware",
           "stock pot": "Cookware",
           "cooker": "Cookware",
           "roaster": "Cookware",
           "roasting pan": "Cookware",
           "baking dish": "Cookware",
           "oven tray": "Cookware",
           "all day pan": "Cookware",

           # Bakeware
           "cake pan": "Bakeware",
           "cake tin": "Bakeware",
           "brownie": "Bakeware",
           "bundt": "Bakeware",
           "quiche": "Bakeware",
           "pie dish": "Bakeware",
           "muffin": "Bakeware",
           "cupcake": "Bakeware",
           "loaf pan": "Bakeware",
```

```python
    "cookie cutter": "Bakeware",
    "cookie scoop": "Bakeware",
    "cooling rack": "Bakeware",
    "pizza": "Bakeware",
    "pastry": "Bakeware",
    "rolling pin": "Bakeware",
    "baking mat": "Bakeware",
    "air fryer": "Bakeware",
    "trivet": "Bakeware",

    # Food Storage
    "container": "Food Storage",
    "storer": "Food Storage",
    "storage": "Food Storage",
    "jar": "Food Storage",
    "cereal": "Food Storage",

    # Food Prep Tools
    "chopping board": "Food Prep Tools",
    "board": "Food Prep Tools",
    "peeler": "Food Prep Tools",
    "grater": "Food Prep Tools",
    "zester": "Food Prep Tools",
    "mandoline": "Food Prep Tools",
    "spiralizer": "Food Prep Tools",
    "masher": "Food Prep Tools",
    "tenderiser": "Food Prep Tools",
    "garlic press": "Food Prep Tools",
    "dumpling press": "Food Prep Tools",
    "apple corer": "Food Prep Tools",
    "juicer": "Food Prep Tools",
    "funnel": "Food Prep Tools",
    "scoop": "Food Prep Tools",
    "ladle": "Food Prep Tools",
    "measuring": "Food Prep Tools",
    "scales": "Food Prep Tools",
    "timer": "Food Prep Tools",

    # Utensils & Gadgets
    "spoon": "Utensils & Gadgets",
    "whisk": "Utensils & Gadgets",
    "spatula": "Utensils & Gadgets",
    "turner": "Utensils & Gadgets",
    "tongs": "Utensils & Gadgets",
    "colander": "Utensils & Gadgets",
    "sieve": "Utensils & Gadgets",
    "strainer": "Utensils & Gadgets",
```

```python
        "sifter": "Utensils & Gadgets",
        "egg": "Utensils & Gadgets",
        "frother": "Utensils & Gadgets",
        "mat": "Utensils & Gadgets",

        # Cutlery & Tableware
        "cutlery": "Cutlery & Tableware",
        "fork": "Cutlery & Tableware",
        "knife": "Cutlery & Tableware",
        "spoons": "Cutlery & Tableware",
        "chopstick": "Cutlery & Tableware",
        "skewer": "Cutlery & Tableware",
        "stick": "Cutlery & Tableware",
        "salt": "Cutlery & Tableware",
        "pepper": "Cutlery & Tableware",

        # Knives & Accessories
        "scissor": "Knives & Accessories",
        "sharpener": "Knives & Accessories",
        "knife organiser": "Knives & Accessories",

        # Miscellaneous
        "lighter": "Miscellaneous",
        "oil sprayer": "Miscellaneous",
        "store & shake": "Miscellaneous",
        "store & pour": "Miscellaneous",
        "bag clip": "Miscellaneous",
        "bottle opener": "Miscellaneous",
        "sink": "Miscellaneous",
        "dish rack": "Miscellaneous",
}

def get_kitchen_subcat(row):
    if row["category"].lower() == "kitchenware":
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_kitchen.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware

# apply row-wise
data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

```
<ipython-input-103-4d3a7cd354d7>:116: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

[104]: ```
data[data['category'] == 'KITCHENWARE']
```

[104]:
```
       product_code     category                             item_name  \
2164        3121431  KITCHENWARE                          Roasting Pan
2369        3121453  KITCHENWARE                Roasting Pan with Rack
3015        3633160  KITCHENWARE                    Ultra Container 1.3L
3046        4163460  KITCHENWARE                    Ultra Container 2.7L
3128        3794181  KITCHENWARE      Plus Rectangle Containers 1 Litre
...             ...          ...                                   ...
72362       5015457  KITCHENWARE                   Silicone Pastry Brush
72377       8419237  KITCHENWARE                       Ice Cream Scoop
72390       3116216  KITCHENWARE                         Bottle Opener
72399       4145990  KITCHENWARE                       Silicone Trivet
72405       3996379  KITCHENWARE  Microsafe Easy Heat Containers 900mL

                          brand  item_price item_size  unit_price  \
2164               Roasting Pan         0.0       NaN         NaN
2369               Roasting Pan         0.0       NaN         NaN
3015                  Container         0.0      1.3L       15.15
3046                  Container         0.0      2.7L       20.50
3128     Plus Rectangle Containers     15.0       NaN        2.50
...                         ...         ...       ...         ...
72362                   Unknown         0.0       NaN         NaN
72377                   Unknown         0.0       NaN         NaN
72390                   Unknown         0.0       NaN         NaN
72399                   Unknown         0.0       NaN         NaN
72405         Microsafe Easy Heat       0.0    900mL        5.67

      unit_weight  similarity_score            subcat  unit_weight_new  \
2164          NaN          0.267393          Cookware              NaN
2369          NaN          0.267393          Cookware              NaN
3015          1ea          0.267393      Food Storage              1.0
3046          1ea          0.267393      Food Storage              1.0
3128          1ea          0.069898      Food Storage              1.0
...           ...               ...               ...              ...
72362         NaN          0.267393          Bakeware              NaN
72377         NaN          0.267393   Food Prep Tools              NaN
72390         NaN          0.267393     Miscellaneous              NaN
72399         NaN          0.267393          Bakeware              NaN
72405         1ea          0.267393      Food Storage              1.0

       unit_weight_type  item_weight_new item_weight_type  std_item_weight  \
```

	std_item_unit			
2164	NaN	NaN	NaN	NaN
2369	NaN	NaN	NaN	NaN
3015	each	1.3	l	1300.0
3046	each	2.7	l	2700.0
3128	each	1.0	each	NaN
…	…	…	…	…
72362	NaN	NaN	NaN	NaN
72377	NaN	NaN	NaN	NaN
72390	NaN	NaN	NaN	NaN
72399	NaN	NaN	NaN	NaN
72405	each	900.0	ml	900.0

	std_item_unit
2164	NaN
2369	NaN
3015	ml
3046	ml
3128	NaN
…	…
72362	NaN
72377	NaN
72390	NaN
72399	NaN
72405	ml

[594 rows x 16 columns]

Tech Accessories:

```
[105]: data[data['category'] == 'TECH ACCESSORIES']['item_name'].unique()
```

```
[105]: array(['1.2M USB-A To Lightning Cable White', 'Wall Charger Dual USB-C',
        'Escape True Wireless Sport Ear Buds',
        'Wired Earphones With Mic Lightning',
        'Boostup 30w Usb C Pd Wall Charger',
        '1.2M USB-A To USB-C Cable Black',
        '42w (Pd30w Pps Usb-C & 12w Usb-A) Universal Wall Charger',
        'Boostcharge USB-A To USB-C Cable- 1M Black',
        'Car Mobile Phone Holder', 'Wireless Mouse',
        '2M USB-C To USB C Cable - Black',
        '1.2M USB-C To USB-C Cable Black Braided',
        'Lightning To 3.5Mm Audio Adapter',
        '2M USB-A To USB-C Cable Black', 'Car Charger Dual USB-C',
        'Indulge Wireless Bluetooth Headphones',
        'Wired Earphones With Mic 3.5mm Audio',
        'Magnetic Wireless Charging Pad',
        '2M USB C To Lightning - White Braided',
        'Boostcharge Usb A Wall Charger (12w)',
```

```
        '1.2M USB C To Lightning  - White Braided',
        'Wall Charger Usb-A & C', '1.2M USB-C To USB-C Cable Black',
        'Boostcharge Usb-C To Usb-C Cable- 1m - White',
        '2M USB-C To C - Black Braided', 'Bluetooth Portable Speaker',
        '1.2M USB-C To Lightning Cable White',
        'USB-C To 3.55MM Audio Adapter', '2M USB-C To Lightning - White',
        'Trek Wireless Sport Earphones', 'Radiant Wired Headphones',
        'Boostcharge Lightning to Usb-A Cable 1m Black',
        'Boostcharge Usb-C To Usb-C Cable- 1m - Black',
        'Kids Volume Limiting Wireless Headphone',
        'Wired Earphones With Mic USB-C',
        'Silence Pro Noise Isolating Headphones',
        '2M USB-A To Lighting Cable White', 'Wireless Charging Stand',
        'Boostcharge Lightning to Usb-A Cable 1m White',
        'Car Charger Dual USB-A', 'Gaming Headset',
        'USB A Car Charger 24W', 'Boostcharge Usb A Usb C Cable 1m White',
        'Wreless Mouse Keyboard Combo', 'True Wireless:Earphone With Enc',
        '10000 Mah Powerbank', 'Over Ear Wireless Headphone With Anc',
        'USB To C 1M Black', 'Lightening Cable 1.2M',
        'Go Sports Earphones'], dtype=object)
```

```python
[106]: subcategory_map_tech = {
           #Audio
           "earphone": "Earphones",
           "ear bud": "Earphones",
           "earbuds": "Earphones",
           "headphone": "Headphones",
           "headset": "Headphones",
           "speaker": "Speakers",

           # Charging
           "charger": "Chargers",
           "charging pad": "Chargers",
           "charging stand": "Chargers",
           "powerbank": "Power Banks",
           "power bank": "Power Banks",

           # ables & Adapters
           "usb-c": "Cables",
           "usb a": "Cables",
           "usb-a": "Cables",
           "lightning": "Cables",
           "cable": "Cables",
           "adapter": "Adapters",

           # Computer Accessories
           "mouse": "Computer Accessories",
```

```python
    "keyboard": "Computer Accessories",
    "keyboard combo": "Computer Accessories",

    # Car Accessories
    "car charger": "Car Accessories",
    "car mobile phone holder": "Car Accessories"
}

def get_tech_subcat(row):
    if row["category"].lower() == "tech accessories":
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_tech.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware


# apply row-wise
data["subcat"] = data.apply(get_tech_subcat, axis=1)
```

```
<ipython-input-106-7a4a082bb518>:46: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_tech_subcat, axis=1)
```

[120]:
```python
data[data['category'] == 'TECH ACCESSORIES']
```

[120]:
```
         product_code            category  \
3045          5880626   TECH ACCESSORIES
3176          5884048   TECH ACCESSORIES
3256          3969740   TECH ACCESSORIES
3363          5884300   TECH ACCESSORIES
3428          5159220   TECH ACCESSORIES
...               ...                ...
71789         5882417   TECH ACCESSORIES
71816         5159480   TECH ACCESSORIES
71857         5883996   TECH ACCESSORIES
71870         5159479   TECH ACCESSORIES
71942         5643347   TECH ACCESSORIES


                                      item_name                    brand  \
3045           1.2M USB-A To Lightning Cable White         1.2M USB-A To
3176                       Wall Charger Dual USB-C      Wall Charger Dual
3256        Escape True Wireless Sport Ear Buds   Escape True Wireless
```

```
3363              Wired Earphones With Mic Lightning  Wired Earphones With
3428              Boostup 30w Usb C Pd Wall Charger        Boostup 30w Usb
...                                              ...                     ...
71789               2M USB-C To USB C Cable - Black          2M USB-C To
71816  Boostcharge Usb-C To Usb-C Cable- 1m - White              Unknown
71857                     Car Charger Dual USB-A                 Unknown
71870  Boostcharge Usb-C To Usb-C Cable- 1m - Black             Unknown
71942                       Go Sports Earphones                  Unknown

       item_price item_size  unit_price  unit_weight  similarity_score  \
3045         0.00       NaN         NaN          NaN          0.156810
3176         0.00       NaN         NaN          NaN          0.156810
3256         0.00       NaN         NaN          NaN          0.156810
3363         0.00       NaN         NaN          NaN          0.156810
3428        34.95       NaN         NaN          NaN          0.032480
...           ...       ...         ...          ...               ...
71789        0.00       NaN         NaN          NaN          0.156810
71816       19.95       NaN         NaN          NaN          0.056408
71857        0.00       NaN         NaN          NaN          0.156810
71870       19.95       NaN         NaN          NaN          0.056408
71942        0.00       NaN         NaN          NaN          0.156810

          subcat  unit_weight_new unit_weight_type  item_weight_new  \
3045       Cables              NaN              NaN              NaN
3176     Chargers              NaN              NaN              NaN
3256    Earphones              NaN              NaN              NaN
3363    Earphones              NaN              NaN              NaN
3428     Chargers              NaN              NaN              NaN
...           ...              ...              ...              ...
71789      Cables              NaN              NaN              NaN
71816      Cables              NaN              NaN              NaN
71857    Chargers              NaN              NaN              NaN
71870      Cables              NaN              NaN              NaN
71942   Earphones              NaN              NaN              NaN

       item_weight_type  std_item_weight std_item_unit
3045                NaN              NaN           NaN
3176                NaN              NaN           NaN
3256                NaN              NaN           NaN
3363                NaN              NaN           NaN
3428                NaN              NaN           NaN
...                 ...              ...           ...
71789               NaN              NaN           NaN
71816               NaN              NaN           NaN
71857               NaN              NaN           NaN
71870               NaN              NaN           NaN
71942               NaN              NaN           NaN
```

```
    [232 rows x 16 columns]
```

Hosiery:

[108]: `data[data['category'] == 'HOSIERY']['item_name'].unique()`

[108]: array(['Curvy Sheer Kh 15D Natural One Size', 'Paque Trouser Sock 70D',
       'Regular Ph Natural Average/Tall 4 pack',
       'Silky Control Ph Natural X-Tall', '32800 Pantyhose Black Tall',
       'SHEER VALUE 15 DENIER TALL NATURAL',
       'Sheer Nylon Knee Hi Pantyhose Reinforced Toe',
       'Travel Sockblack 3-8', 'Bamboo Shaper Bra Black 14/16',
       '32800 Pantyhose Beige Tall', 'Ladder Resist Natural Average',
       'Value Anklet Socks Tan',
       '70D Opaque /No Dig Tight Black Medium/Large',
       'Smooth Footlet Natural', 'Smooth Footlets',
       'Curve Ph Black Size 2 4pack',
       'Eco 40D Opque Tights Black Tall/X Tall',
       'Ultra Low Cut Footlet Natural One Size', 'Sheer Anklet 15D',
       'Sheer Sockettes', 'Curve Ph Natural Size 2 4pack',
       'Value Anklet Socks Black', 'Sheer Pantyhose Slim Black Medium',
       'Bamboo Shaper Bra Black 10/12', 'Ladder Resist Natural Tall',
       'Girl Sneaker Footlet Black One Size',
       'Ladder Resist Natural/ X Tall', 'PANTYHOSE 20 DENIER X-TALL',
       'Relief Pantyhose Trouser Sock Black',
       'Regular Ph Natural Tall/X-Tall 4 pack',
       'Cotton Sole Kh 30D Natural One Size 2 pack',
       'Eco 40D Opque Tights Black Small/Medium', 'Opaque Anklets',
       'Sheer Pantyhose Slim Black Small',
       'All Over Shaper Bronze X-Tall',
       'Comfy Tops Semi Opaque Knee Hi 40 Denier Black 1 size',
       '40d Opaque Knee High 1size Nude', 'Active Rib Legging Black M',
       'Active Rib Legging Black L',
       'Opaque Tights 70 Denier Black Tall/X-Tall',
       'Hoisery Footlet Size 8-11 Nude',
       'Eco Cotton Tights Black Tall/X Tall', '150d 1pk Md-lg',
       '150D Flex Leggings SM-MD',
       'Opaque Tights 40 Denier Black Ave/Tall', 'Opaque Tights 40d',
       'Sheer Pantyhose Slim Nude Medium',
       'Sheer Pantyhose Slim Nude Large', 'Footlet Size 3+',
       'Light Slimming Brief Natural Average',
       'Cotton Sole Kh 30D Black One Size 2 pack',
       'Opaque Knee Highs 40D', 'All Over Shaper Bronze Ave',
       'All Over Shaper Bronze Tall', 'Opaque Tights 70d',
       'Eco Cotton Tights Black Small/Medium',
       'Eco Bio 130D Opque Tights Black Small/Medium',
       'Lux Velvet Tights 200D A/T',

```
         'Luxury Velvet Tights 200 Denier Black Tall/X-Tall',
         'Curvy Trouser Sock 80D Black One Size',
         'Footlet Sock Size 4 - 7 Ladies Tan',
         'Eco Bio 130D Opque Tights Black Tall/X Tall',
         'Light Slimming Brief Natural X Tall',
         'Skinny Legging Mix Colour Medium/Large',
         'Invisible Footlet Nude 9-11',
         'Skinny Legging Mix Colour Small/Medium',
         'No Shine Opaque Tights 70 Denier Black Tall/X-Tall',
         'Silk Caresse 15D',
         'Invisible Band Knee High 30D Natural One Size',
         'Eco 15D Anklets Black One Size', 'Pantyhose Beige Average',
         'Footlet Size 8+', 'Pattern Knee High', 'Pantyhose Black Average',
         'Sheer Pantyhose Slim Nude Small',
         'Sheer Pantyhose Slim Black Large', 'Washbag Floral',
         'Eco Sneaker Footlet Black One Size', 'Invisible Footlet Nude 5-8',
         'Sheer Knee Highs 15 Denier Black Size 1',
         'Textured Rib Legging Medium/ Long',
         'Curvy Very Thick Tights 120D Black - 2',
         'No Shine Opaque Tights 70 Denier', 'Invisible Footlet Blk 9-11',
         'Invisible Footlet 5-8', 'Eco 15D Knee High Black One Size',
         'Relief 40d Compression Tall/xtall',
         'Invisible Band Anklets Natural One Size',
         'Classic Fishnet Black One Size', 'Pantyhose Beige X-Tall',
         'OSFA Pantyhose Knee Hi Mid Beige'], dtype=object)
```

```python
[109]:  subcategory_map_hosiery = {
            # Socks & Footlets
            "sock": "Socks",
            "anklet": "Socks",
            "anklets": "Socks",
            "footlet": "Footlets",
            "sockette": "Footlets",
            "sneaker footlet": "Footlets",
            "knee hi": "Knee Highs",
            "knee high": "Knee Highs",

            # Pantyhose / Stockings
            "pantyhose": "Pantyhose",
            "ph": "Pantyhose",    # many items abbreviated as 'ph'
            "ladder resist": "Pantyhose",
            "sheer": "Pantyhose",
            "trouser sock": "Pantyhose",

            # Tights & Leggings
            "tight": "Tights",
            "tights": "Tights",
```

```python
        "legging": "Leggings",
        "leggings": "Leggings",

        # Shapewear / Undergarments
        "shaper": "Shapewear",
        "slimming brief": "Shapewear",
        "brief": "Shapewear",
        "bra": "Shapewear",

        # Special Hosiery
        "compression": "Compression Hosiery",
        "fishnet": "Fashion Hosiery",
        "pattern": "Fashion Hosiery",
        "velvet": "Fashion Hosiery",
        "washbag": "Accessories"
}

def get_hosiery_subcat(row):
    if row["category"].lower() == "hosiery":
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_hosiery.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware
# apply row-wise
data["subcat"] = data.apply(get_hosiery_subcat, axis=1)
```

```
<ipython-input-109-e235e67d2507>:49: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_hosiery_subcat, axis=1)
```

[110]: ```python
data[data['category'] == 'HOSIERY']
```

[110]:
```
       product_code category   item_name \
40342   4182883 HOSIERY   Curvy Sheer Kh 15D Natural One Size
40355   3122310 HOSIERY   Paque Trouser Sock 70D
40419   4150446 HOSIERY Regular Ph Natural Average/Tall 4 pack
40444   4144363 HOSIERY   Silky Control Ph Natural X-Tall
40479   4163630 HOSIERY   32800 Pantyhose Black Tall
...   .
72324   4163834 HOSIERY   Sheer Pantyhose Slim Black Large
72361   4150300 HOSIERY   All Over Shaper Bronze X-Tall
```

```
72367          7017227  HOSIERY        Relief 40d Compression Tall/xtall
72370          3121293  HOSIERY                    Smooth Footlet Natural
72371          3123095  HOSIERY                         Sheer Anklet 15D

                         brand  item_price item_size  unit_price unit_weight  \
40342          Curvy Sheer Kh         0.0       NaN         NaN         NaN
40355       Paque Trouser Sock        0.0       NaN         NaN         NaN
40419                  Unknown        0.0    4 pack         NaN         NaN
40444                  Unknown        0.0       NaN         NaN         NaN
40479   32800 Pantyhose Black        10.0       NaN         NaN         NaN
...                        ...         ...       ...         ...         ...
72324                  Unknown        9.0       NaN         NaN         NaN
72361                  Unknown        0.0       NaN         NaN         NaN
72367                  Unknown        11.0       NaN         NaN         NaN
72370                  Unknown        0.0       NaN         NaN         NaN
72371                  Unknown        0.0       NaN         NaN         NaN

        similarity_score               subcat  unit_weight_new  \
40342           0.173975            Pantyhose              NaN
40355           0.173975                Socks              NaN
40419           0.173975            Pantyhose              NaN
40444           0.173975            Pantyhose              NaN
40479           0.118156            Pantyhose              NaN
...                  ...                  ...              ...
72324           0.127593            Pantyhose              NaN
72361           0.173975            Shapewear              NaN
72367           0.108275  Compression Hosiery              NaN
72370           0.173975             Footlets              NaN
72371           0.173975                Socks              NaN

        unit_weight_type  item_weight_new item_weight_type  std_item_weight  \
40342                NaN              NaN              NaN              NaN
40355                NaN              NaN              NaN              NaN
40419                NaN              4.0             pack              4.0
40444                NaN              NaN              NaN              NaN
40479                NaN              NaN              NaN              NaN
...                  ...              ...              ...              ...
72324                NaN              NaN              NaN              NaN
72361                NaN              NaN              NaN              NaN
72367                NaN              NaN              NaN              NaN
72370                NaN              NaN              NaN              NaN
72371                NaN              NaN              NaN              NaN

        std_item_unit
40342            NaN
40355            NaN
40419            pack
```

```
40444          NaN
40479          NaN
…               …
72324          NaN
72361          NaN
72367          NaN
72370          NaN
72371          NaN

[123 rows x 16 columns]
```

Easter:

```
[119]: data[data['category'] == 'EASTER']['item_name'].unique()
```

```
[119]: array(['Chocolate Easter Bunny', 'Selections Easter Chocolate Eggs Bag',
          'Easter Gold Bunny Milk Chocolate',
          'Chocolate Easter Hunt Assortment Pack',
          'Dairy Milk Easter Chocolate Eggs Bag',
          'Easter Chocolate Selection Gift Box',
          'Chocolate Easter Bunny Tin',
          'Easter Mini Gold Bunny Chocolate Pouch Bag',
          'Surprise Chocolate Easter Bunny Blue',
          'Bubbly Mint Chocolate Easter Share Bag',
          'Milk Chocolate Solid Eggs',
          'Crunchie Hollow Chocolate Easter Egg',
          'Cherry Ripe Hollow Chocolate Easter Egg',
          'Easter Gold Bunny White Chocolate',
          'Easter Gourmet Milk Chocolate Egg Bag', 'Collection Boxed Egg',
          'Mini Eggs Milk Chocolate', 'Kitkat Chocolate Easter Bunny',
          'Easter Bunny Chocolate House', 'Bueno Easter Eggs',
          'Marvellous Creations Chocolate Easter Bunny Gift Box',
          'Surprise Chocolate Giant Easter Bunny',
          'Dairy Milk Coconut Rough Easter Bunny',
          'Milkybar White Choc Easter Bunny', 'Mint Slice Egg Gift Box',
          'Hazelnut Brittle White Chocolate Eggs',
          'Creme Egg 6-Pack Easter Chocolate Eggs Bag',
          'Peanut Chocolate Speckled Easter Egg Bag',
          'Life Savers Fruit Tingles Easer Eggs',
          'Milk Chocolate Mini Eggs Lemon Meringue',
          'Lindor Easter Assorted Chocolate Eggs Bag',
          'Dairy Milk Chocolate Easter Egg Hunt Carton',
          'Dairy Milk Chocolate Easter Bunny',
          'Easter Flower Gold Bunny Milk Chocolate',
          'Humpty Dumpty Easter Egg Gift Box',
          'Easter Gold Bunny Dark Chocolate',
          'Collection Hazelnut Easter Eggs',
          'Crunchie Chocolate Easter Bunny',
```

'Dairy Milk Hollow Chocolate Easter Egg',
'Caramello Easter Chocolate Eggs Bag',
'Chocolate Easter Bunnies 3 Pack',
'Oreo Chocolate Easter Gift Box', 'Dream Chocolate Easter Bunny',
'Lindor Easter Assorted Eggs & Milk Chocolate Egg Gift Box',
'Easter Hazelnut Dark Chocolate Eggs Pouch Bag',
'Milk Chocolate Bunnies & Chicks Nets', 'Mini Easter Eggs',
'Crispy Chocolate Speckled Easter Egg Bag',
'Kitkat Mini Pouch Chocolate Easter Bunny Share Pack',
'Easter Fried Eggs And Bunnies Gummies', 'Rocklea Road Bunny',
'Surprise Gift Pack Maxi Pink', 'Australia Puddles The Platypus',
'Chocolate Easter Rabbit',
'Clinkers Chocolate Giant Easter Egg Tube',
'Dairy Milk Deluxe Egg',
'Humpty Dumpty Easter Eggs With Beanies Carton',
'Clinkers Chocolate Easter Bunny Sharepack',
'Easter Gold Bunny & Milk Chocolate Egg',
'Creme Egg Minis Chocolate Easter Bag',
'Easter Mini Gold Bunny Chocolate 5 Pack', 'Grand Rocher Dark',
'Caramilk Easter Chocolate Eggs Bag',
'Easter Choc Coated Marshmallow', 'Nougat Easter Egg',
'Creme Egg Chocolate Easter Gift Box',
'Easter Sour Carrots Gummies', 'Milk Chocolate Easter Egg',
'Easter Milk Chocolate Coins',
'Raspberry&coconut Slice Egg Gift Box',
'Kit Kat Mini Easter Eggs Bag',
'Britecrawlers & Zappo Milk Chocolate Egg Boxes',
'Selections Large Easter Chocolate Eggs Bag',
'Oreo Easter Chocolate Eggs Bag', 'Rocher Easter Egg T3',
'Collection Dark Chocolate Easter Eggs',
'Turkish Delight Easter Chocolate Eggs Bag',
'Easter Figurines 6 Pack', 'Milk Chocolate Bilby',
'Crunchie Easter Chocolate Eggs Bag',
'Mint Easter Chocolate Eggs Bag',
'Lindor Easter Mini Eggs & Dark Chocolate Egg Gift Box',
'Easter Gold Bunny Salted Caramel Milk Chocolate',
'Easter Chocolate Carrot Pouch Bag',
'Dairy Milk Chocolate Easter Gift Box',
'Mini Eggs Chocolate Easter Gift Box', 'Rocher Easter Egg',
'Speckled Easter Eggs',
'Rocher Milk Chocolate And Hazelnut Hollow Squirrel',
'Mini Eggs Chocolate Easter Bunny',
'100S & 1000S Milk Chocolate Egg Casket',
'Caramilk Chocolate Easter Gift Box',
'Easter Bugs & Bees Chocolate Pouch Bag',
'Surprise Chocolate Gift Pack Maxi Egg Blue',
'Milk Choc Foiled Eggs', 'Hollow Chocolate Bunny',

```
        'Smarties Chocolate Easter Incredible Egg',
        'Kit Kat Chocolate Honeycomb Incredible Egg',
        'Milk Chocolate Bunny', 'After Eight Chocolate Easter Bunny',
        'Chocolate Hazelnut Mini Easter Eggs', 'Surprise Maxi Egg Pink',
        'Milk Chocolate Mini Eggs Popping Candy',
        'Easter Belgian Milk Choc Pops', 'Smarties Mini Eggs Bag',
        'Smarties Chocolate Easter Bunny', 'Easter Bunnies 12 Pack',
        'Low Sugar Half Eggs With Gooey Caramel', 'Koala Milk Chocolate',
        'Surprise Chocolate Easter Bunny Pink', 'Surprise Maxi Egg Blue',
        'Kit Kat Mini Bunny Burrow',
        'Rocklea Road With Speckled Eggs Gift Box',
        'Sea Salt & Caramel Almond Nougat Milk Chocolate Eggs',
        'Dairy Milk Extra Large Chocolate Easter Egg Bag',
        'Strawberry Easter Chocolate Eggs Bag'], dtype=object)
```

[122]:
```python
subcategory_map_easter = {
    # Easter Bunnies & Rabbits
    "bunny": "Chocolate Bunnies",
    "rabbit": "Chocolate Bunnies",
    "bilby": "Chocolate Bunnies",

    # Easter Eggs
    "egg": "Easter Eggs",
    "eggs": "Easter Eggs",
    "casket": "Easter Eggs",
    "carton": "Easter Eggs",
    "casket": "Easter Eggs",
    "chocolate eggs": "Easter Eggs",

    # Easter Gift Packs & Boxes
    "gift box": "Gift Packs & Boxes",
    "share pack": "Gift Packs & Boxes",
    "share bag": "Gift Packs & Boxes",
    "pouch": "Gift Packs & Boxes",
    "carton": "Gift Packs & Boxes",
    "surprise": "Gift Packs & Boxes",
    "collection": "Gift Packs & Boxes",
    "figurines": "Gift Packs & Boxes",

    # Novelty & Characters
    "house": "Novelty & Figures",
    "squirrel": "Novelty & Figures",
    "platypus": "Novelty & Figures",
    "bugs": "Novelty & Figures",
    "bees": "Novelty & Figures",
    "coins": "Novelty & Figures",
    "pops": "Novelty & Figures",
```

```python
    # Marshmallow / Gummies / Other
    "marshmallow": "Marshmallow Treats",
    "gummies": "Gummies & Candy",
    "carrot": "Gummies & Candy",
    "sour": "Gummies & Candy",


    "choc": "Chocolate",
    "chocolate": "Chocolate"
}

def get_easter_subcat(row):
    if row["category"].lower() == "easter":
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_easter.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware
# apply row-wise
data["subcat"] = data.apply(get_easter_subcat, axis=1)
```

```
<ipython-input-122-f094ffc0a6c4>:55: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_easter_subcat, axis=1)
```

Household Appliances:

```python
[123]: data[data['category'] == 'HOUSEHOLD APPLIANCES']['item_name'].unique()
```

```
[123]: array(['Stainless Steel Kettle 1.7L', 'Compact Food Processor',
       '2 Slice Snackmaker', '7 Cup Rice Cooker', 'Garment Steamer',
       'Everyday Kettle 1.7L', 'Steam Iron',
       'Terra White Sparkling Water Maker', '600 Series',
       'Deep Fill Jaffle Maker', 'Everyday Toaster', 'Stick Mixer',
       'Stainless Steel 4 Slice Toaster', 'Hand Mixer',
       '5 Cup Rice Cooker'], dtype=object)
```

```python
[124]: subcategory_map_kitchen = {
    "kettle": "Cookware",
    "rice cooker": "Cookware",
    "jaffle maker": "Bakeware",
```

```python
        "snackmaker": "Bakeware",
        "toaster": "Bakeware",
        "food processor": "Food Prep Tools",
        "mixer": "Food Prep Tools",    # covers "stick mixer", "hand mixer"
        "water maker": "Food Prep Tools",
        "garment steamer": "Miscellaneous",
        "steam iron": "Miscellaneous",
}
def get_kitchen_subcat(row):
    if row["category"].lower() == "household appliances":
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_kitchen.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware
# apply row-wise
data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

```
<ipython-input-124-ed90af54e594>:23: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

[125]:
```python
data[data['category'] == 'HOUSEHOLD APPLIANCES']
```

[125]:
```
       product_code          category                           item_name  \
38818       3790828  HOUSEHOLD APPLIANCES          Stainless Steel Kettle 1.7L
38902       4799668  HOUSEHOLD APPLIANCES              Compact Food Processor
39000       3790919  HOUSEHOLD APPLIANCES                   2 Slice Snackmaker
39184       3790894  HOUSEHOLD APPLIANCES                    7 Cup Rice Cooker
39535       7639442  HOUSEHOLD APPLIANCES                      Garment Steamer
39663       4799533  HOUSEHOLD APPLIANCES                Everyday Kettle 1.7L
39741       7500919  HOUSEHOLD APPLIANCES                           Steam Iron
40011       7532482  HOUSEHOLD APPLIANCES     Terra White Sparkling Water Maker
40031       2748269  HOUSEHOLD APPLIANCES                           600 Series
40071       7500817  HOUSEHOLD APPLIANCES                  Deep Fill Jaffle Maker
40114       4799511  HOUSEHOLD APPLIANCES                     Everyday Toaster
40234       4799657  HOUSEHOLD APPLIANCES                          Stick Mixer
40398       3790883  HOUSEHOLD APPLIANCES         Stainless Steel 4 Slice Toaster
41250       4799602  HOUSEHOLD APPLIANCES                           Hand Mixer
41306       7500839  HOUSEHOLD APPLIANCES                    5 Cup Rice Cooker
69205       4799602  HOUSEHOLD APPLIANCES                           Hand Mixer
69423       7532482  HOUSEHOLD APPLIANCES     Terra White Sparkling Water Maker
```

```
69439      4799668   HOUSEHOLD APPLIANCES              Compact Food Processor
69469      7500919   HOUSEHOLD APPLIANCES                          Steam Iron
69715      4799533   HOUSEHOLD APPLIANCES                  Everyday Kettle 1.7L
69902      4799511   HOUSEHOLD APPLIANCES                     Everyday Toaster
69903      4799657   HOUSEHOLD APPLIANCES                          Stick Mixer
69904      7500839   HOUSEHOLD APPLIANCES                   5 Cup Rice Cooker
69947      2748269   HOUSEHOLD APPLIANCES                           600 Series
70249      7500817   HOUSEHOLD APPLIANCES               Deep Fill Jaffle Maker


                        brand  item_price item_size  unit_price unit_weight  \
38818                 Unknown         0.0      1.7L         NaN         NaN
38902                 Unknown         0.0       NaN         NaN         NaN
39000                 Unknown         0.0       NaN         NaN         NaN
39184                  Cooker         0.0       NaN         NaN         NaN
39535         Garment Steamer         0.0       NaN         NaN         NaN
39663                 Unknown         0.0      1.7L         NaN         NaN
39741                 Unknown         0.0       NaN         NaN         NaN
40011   Terra White Sparkling         0.0       NaN         NaN         NaN
40031              600 Series         0.0       NaN         NaN         NaN
40071                 Unknown         0.0       NaN         NaN         NaN
40114                 Unknown         0.0       NaN         NaN         NaN
40234                 Unknown         0.0       NaN         NaN         NaN
40398                 Unknown         0.0       NaN         NaN         NaN
41250                 Unknown         0.0       NaN         NaN         NaN
41306                  Cooker         0.0       NaN         NaN         NaN
69205                 Unknown         0.0       NaN         NaN         NaN
69423   Terra White Sparkling         0.0       NaN         NaN         NaN
69439                 Unknown         0.0       NaN         NaN         NaN
69469                 Unknown         0.0       NaN         NaN         NaN
69715                 Unknown         0.0      1.7L         NaN         NaN
69902                 Unknown         0.0       NaN         NaN         NaN
69903                 Unknown         0.0       NaN         NaN         NaN
69904                  Cooker         0.0       NaN         NaN         NaN
69947              600 Series         0.0       NaN         NaN         NaN
70249                 Unknown         0.0       NaN         NaN         NaN


        similarity_score          subcat  unit_weight_new unit_weight_type  \
38818                1.0        Cookware              NaN              NaN
38902                1.0  Food Prep Tools              NaN              NaN
39000                1.0        Bakeware              NaN              NaN
39184                1.0        Cookware              NaN              NaN
39535                1.0   Miscellaneous              NaN              NaN
39663                1.0        Cookware              NaN              NaN
39741                1.0   Miscellaneous              NaN              NaN
40011                1.0  Food Prep Tools              NaN              NaN
40031                1.0           Other              NaN              NaN
40071                1.0        Bakeware              NaN              NaN
```

40114	1.0	Bakeware	NaN	NaN
40234	1.0	Food Prep Tools	NaN	NaN
40398	1.0	Bakeware	NaN	NaN
41250	1.0	Food Prep Tools	NaN	NaN
41306	1.0	Cookware	NaN	NaN
69205	1.0	Food Prep Tools	NaN	NaN
69423	1.0	Food Prep Tools	NaN	NaN
69439	1.0	Food Prep Tools	NaN	NaN
69469	1.0	Miscellaneous	NaN	NaN
69715	1.0	Cookware	NaN	NaN
69902	1.0	Bakeware	NaN	NaN
69903	1.0	Food Prep Tools	NaN	NaN
69904	1.0	Cookware	NaN	NaN
69947	1.0	Other	NaN	NaN
70249	1.0	Bakeware	NaN	NaN

	item_weight_new	item_weight_type	std_item_weight	std_item_unit
38818	1.7	l	1700.0	ml
38902	NaN	NaN	NaN	NaN
39000	NaN	NaN	NaN	NaN
39184	NaN	NaN	NaN	NaN
39535	NaN	NaN	NaN	NaN
39663	1.7	l	1700.0	ml
39741	NaN	NaN	NaN	NaN
40011	NaN	NaN	NaN	NaN
40031	NaN	NaN	NaN	NaN
40071	NaN	NaN	NaN	NaN
40114	NaN	NaN	NaN	NaN
40234	NaN	NaN	NaN	NaN
40398	NaN	NaN	NaN	NaN
41250	NaN	NaN	NaN	NaN
41306	NaN	NaN	NaN	NaN
69205	NaN	NaN	NaN	NaN
69423	NaN	NaN	NaN	NaN
69439	NaN	NaN	NaN	NaN
69469	NaN	NaN	NaN	NaN
69715	1.7	l	1700.0	ml
69902	NaN	NaN	NaN	NaN
69903	NaN	NaN	NaN	NaN
69904	NaN	NaN	NaN	NaN
69947	NaN	NaN	NaN	NaN
70249	NaN	NaN	NaN	NaN

HairCare:

```
[127]: data[data['category'] == 'HAIR CARE']['item_name'].unique()
```

```
[127]: array(['Jaw Clip Extra Large', 'Comfy Natural Ponytailers 12 pieces',
              'Elastics All Hair', '4mm No Damage Black Elastics',
              'Mixed Inner Polybands 300 pieces',
              'Active No Damage Reflective Elastics', 'Nail Grooming Set',
              'Barrette Snap Clips 10 pieces', 'No Slip Grip Octopus Clips',
              'Ponytailers Girl Soft 21 Piece',
              'Polyblands Girl Pastel 500 Piece', 'Excellence 4 Dark Brown',
              'Elvive Glycolic Gloss Shampoo',
              'Elvive Treatment Extraordinary Coconut Oil',
              'Polyblands & Elastics With Case Girl 212 Piece',
              'No Damage Elastics Value Pack',
              'Elastics Girl Snap Clips & Bracelet',
              'Elvive Glycolic Gloss Conditioner',
              'Preference Brasilia Permanent Hair Colour',
              'Excellence 7.1 Dark Ash Blonde', 'Excellence 3 Darkest Brown',
              'Excellence 4.15 Dark Frosted Brown',
              'Excellence 6.54 Light Mahogany Copper Brown',
              'Elvive Extraordinary Oil Nourishing Masque Dry Hair Mask',
              'Preference 4.15 Caracas Hair Colour', 'Excellence 1 Black',
              'Paris Elvive Bond Repair Serum',
              'Magic Retouch Permanent 4 Dark Brown',
              'Paris Elvive Bond Repair Conditioner',
              'Excellence 5.3 Golden Brown', 'Excellence 6.1 Light Ash Brown',
              'Magic Retouch Dark Brown', 'Elvive Colour Protect Purple Shampoo',
              'Magic Retouch Permanent 5 Brown', 'Preference 9L Ultra Platinum',
              'Preference Tahiti 4 Permanent Hair Colour',
              'Elvive Colour Protect Purple Conditioner',
              'Preference Vivids 4.261 Dark Purple', 'Excellence 6 Light Brown',
              'Paris Elvive Hyaluron Plump Mask',
              'Casting Creme Gloss Chocolate Chesnut Hair Colour',
              'Magic Retouch Permanent 3 Darkest Brown',
              'Excellence 5.15 Natural Frosted Brown',
              'Excellence 9 Light Blonde', 'Elvive Dream Lengths Mask',
              'Casting Creme Gloss Darkest Brown Hair Colour',
              'Preference 74 Dublin Hair Colour',
              'Excellence 7.3 Dark Golden Blonde',
              'Preference Vivids 1.102 Blue Black',
              'Excellence 5.5 Mahogany Brown', 'Excellence 5 Brown',
              'Preference 5 Bruges Hair Colour',
              'Elvive Treatment Dream Lengths No Hair Cut Leave-In Cream',
              'Casting Creme Gloss Plum Hair Colour',
              'Preference Acidic Toner Platinum Ice',
              'Casting Creme Gloss Dark Brown Hair Colour',
              'Casting Creme Gloss Amber Hair Colour',
              'Preference 11.21 Hair Colour Moscow', 'Magic Retouch Brown',
              'Excellence 8 Blonde', 'Excellence 9.1 Light Ash Blonde',
              'Excellence 01 Very Light Natural Blonde', '4.54 Dark Copper',
```

'Elnett Flexible Hold', 'Excellence 6.30 Light Golden Brown',
'Elnett Normal Strength Hair Spray',
'Preference Acidic Toner Pearly Boost',
'Casting Creme Gloss Medium Brown Hair Colour',
'Preference 1.0 Napoli Hair Colour', 'Magic Retouch Black',
'Paris Magic Retouch Auburn',
'Elvive Extraordinary Extra Rich Oil',
'Paris Elvive Treatment Dream Lengths Frizz Killer',
'Preference St Honore 3.12',
'Preference 7.23 Bali Dark Rose Gold Hair Colour',
'Magic Retouch Blonde', 'Magic Retouch Spray Cool Brown',
'Casting Creme Gloss Light Brown Hair Colour',
'Preference Vivids 3.161 Magnetic Plum',
'Preference Vivids 8.624 Bright Red', 'Excellence 7 Dark Blonde',
'Excellence 2 Black Brown',
'Excellence 10.21 Very Lgt Pearl Blonde',
'Elvive Treatment Extraodinary Oil',
'Magic Retouch Permanent 2 Black', 'Elnett Hairspray Supreme Hold',
'Preference Vienna Permanent Hair Colour',
'Light Brown Magic Retouch Instant Root Concealer Spray',
'Excellence 8.1 Ash Blonde',
'Magic Retouch Permanent 6 Light Brown',
'Paris Elvive Bond Repair Shampoo',
'Paris Elnett Satin Supreme Hold Hairspray',
'Castings Creme Gloss Hair Colour Ebony Black',
'Casting Creme Gloss Chocolate Hair Colour',
'Casting Creme Gloss Dark Chocolate Hair Colour',
'Preference Vivids 10.112 Silver Grey',
'Preference 9.1 Oslo Hair Colour',
'Preference 10.21 Stockholm Hair Colour',
'Preference Helsinki Permanent Hair Colour',
'Magic Retouch Spray Cool Dark Brown',
'Excellence 5.6 Rich Auburn', '8.13 Light Beige Blonde',
'Multipurpose Eyeliner Pen 4 In 1', 'Shampoo Gentle',
'Hair + Scalp Extra Hydration Shampoo', 'Shampoo Nourishing',
'Aom Weightless Oil Spray', 'Hair Strengthening Mask Tub',
'Detangling Brush Large',
'Essences Bio Renew Coco Milk & Aloe Leave IN',
'Bio Renew Coconut Milk & Aloe Deep Hydrate Mask',
'Shampoo Coconut Hydration', 'Flexi-Glide Brush',
'Bio Renew Argan Oil & Aloe Deep Repair Mask',
'Conditioner Coconut Hydration',
'Hair + Scalp Extra Hydration Conditioner',
'Frizz Ease Hydrating Jelly', 'Conditioner Nourishing',
'Bond Protein Shampoo', 'Argan Oil Repairing Hair Oil Pump',
'Conditioner Gentle', 'Profiller+ Super Fluid',
'All In One Treatment Conditioner', 'Bond Protein Conditioner',

'Pro-V Ultimate 10 Repair & Protect Shampoo',
'Classics Colour Shampoo', 'Clean & Classic Original Dry Shampoo',
'Coconut & Exotic Tropical Dry Shampoo',
'Medium Brunette Dry Shampoo',
'Summer Fruits & Berries Cherry Dry Shampoo', 'Dry Shampoo Bare',
'Fresh Dry Shampoo', 'Classics Normal Shampoo',
'Ultramen Shampoo 2 In 1 Old Spice', 'Shampoo Intensive Repair',
'Essences Potent Aloe Eucalyptus Shampoo', 'Original Dry Shampoo',
'Ultramen Shampoo 2 In 1 Deep Clean', 'Pure Grow Shampoo',
'Melbourne Sea Salt Texture Hair Spray', 'Blush Dry Shampoo',
'Conditioner Moisture Rich',
'Classics Lavender Anti Frizz Shampoo', 'Anti-Flyaway Hair Wand',
'Bio Renew Argan Oil Of Morocco Repair Shampoo',
'Miracles Collagen Repair & Protect Shampoo',
'Pro-V Classic Clean Shampoo', 'Wax Stick',
'Salon Professional Detangle Paddle & Finish Brush',
'By Manicare Pre Glued Kelsey Lash',
'Anti Dandruff Shampoo Apple Fresh', 'Bond Protein Sealing Serum',
'All In One Treatment Shampoo', 'Moisture Rich Shampoo',
'Pro-V Daily Moisture Renewal Shampoo',
'Dry Shampoo & Volume With Plumping Collagen',
'Floral & Flirty Blush Dry Shampoo', 'Eden Dry Shampoo',
'Active Dry Shampoo', 'Overnight Deep Cleanse Dry Shampoo',
'Daily Moisture Renewal Conditioner',
'Pretty & Delicate Rose Gold Dry Shampoo',
'Miracles Collagen Repair & Protect Conditioner',
'Refreshing Rosemary Mint Conditioner',
'Divine Dark Brown Dry Shampoo Plus',
'Blue Anti Dandruff Shampoo Replenishing',
'Argan Oil Repairing Shampoo Pump', 'Shampoo Smooth & Silky',
'Shampoo Daily Care', 'Pro-V Ultimate 10 Conditioner',
'Conditioner Smooth & Silky',
'Thicker Fuller Daily Densifying Serum',
'Styling Range Cruel Clay', 'Pro-V Always Smooth & Sleek Shampoo',
'Elvive Hyaluron Plump Shampoo', 'Shampoo Nourishing Moisture',
'Salon Professional Radial Porcupine Small Brush', 'T-Gel Shampoo',
'Shampoo Restoring Ritual', 'Classics Colour Conditioner',
'2 In 1 Anti Dandruff Shampoo Oil Control',
'Argan Oil Repairing Conditioner Pump', 'Clean & Balanced Shampoo',
'Pedi Pro Callus File',
'Extra Care Marrakesh Oil & Coconut Replenishing Shampoo',
'Styling Range Matte Shaper', 'Extra Care Restoring Argan Oil',
'Conditioner Intensive Repair', 'Keratin Smoothing Shampoo',
'Headlice All In One Treatment Spray',
'Pro-V Classic Clean Conditioner', 'Soft Black Ponytailers',
'Anti Dandruff Shampoo Itch Care', 'Hydrating Shampoo',
'5 In 1 Leave-In Spray Curl Care', 'Fashion Nail Shaper',

'Classics Normal Conditioner', 'Fierce Fiber Styler',
'Cushion Brush', 'Classic Colour Conditioner',
'Classic Colour Shampoo', 'Fructis Normal Shampoo',
'Colorsilk Medium Rich Brown 47 Hair Colour',
'Elvive Hyaluron Plump Conditioner', 'Complete Lice Kit',
'Anti Dandruff Shampoo Bare Soothing Hydration',
'Nature Fusion Shampoo',
'Extra Care Push Up Volume Lifting Shampoo',
'Addictive Brilliant Shine Shampoo', 'Tropical Dry Shampoo',
'Thick Ponytailers Snagless Black',
'Anti Dandruff Shampoo Cool Menthol', 'Thicker Fuller',
'Barbers Wax', 'Oatmeal Shampoo', 'By Manicare Eyelash Adhesive',
'Cleanse Replenish Shampoo', 'Olia Hair Colour 4.0 Dark Brown',
'Shampoo Moisture', 'Gold Eyelash Curler',
'Hair Colour Remover Max Effect',
'Dry Shampoo With Oat & Ceramide',
'Heavenly Hydration + Shine Cherry Blossom Shampoo For Thin And Fine
Hair',
'Bio Renew Argan Oil Of Morocco Repair Conditioner',
'Pro-V Always Smooth & Sleek Conditioner',
'Biotin Boost Thickening Shampoo', 'Bio Renew Shampoo Strawberry',
'Renewing + Repairing & Shine Argan Oil of Morocco Shampoo For Dry &
Damaged Hair',
'Live Colour Permanent 1.0 Black', 'Caffeine Shampoo',
'Miracles Biotin Moisture Renewal Shampoo',
'Shampoo Clean Balanced', 'Micellar Rose Water Conditioner',
'Olia Hair Colour 4.15 Iced Chocolate',
'Elvive Dream Long Conditioner',
'Miracles Keratin Smooth & Sleek Shampoo',
'Coconut Curls Styling Milk',
'Bio Renew White Grapefruit & Mosa Mint Shampoo',
'Divine Dark Dry Shampoo', 'Longer & Stronger Shampoo',
'Keratin Smoothing Conditioner',
'Frizz Ease Dream Curls Conditioner',
'Renewing + Hydrating & Shine Argan Oil of Morocco Penetrating Oil For
Dry & Heat Styled Hair',
'Thicker Fuller Conditioner', 'Normal Balance Shampoo',
'4 Way Buffer Black', 'Conditioner Daily Care',
'Biotin Boost Thickening Conditioner',
'Moisture Smoothie Coconut Oil Curl',
'Elvive Colour Protect Shampoo',
'Extra Care Ultimate Repair Strengthening Shampoo',
'Miracles Keratin Smooth & Sleek Conditioner',
'Nettle Dry Shampoo',
"Nice 'N Easy 6A Natural Light Ash Brown Permanent Hair Colour",
'Grooming Brush', 'Anti Dandruff Shampoo Bare Pure Clean',
'Classic Normal Shampoo', 'Lash & Brow Dye Kit Dark Brown',

'Taft Super Hold Styling Gel', 'Hydrating Conditioner',
'Clipper Cuticle', 'Daily Moisture Renewal Shampoo',
'Colorsilk Dark Brown 30 Hair Color',
'Elvive Extraordinary Oils Shampoo', 'Dry Shampoo Volume',
'5in1 Leave-in Spray Keratin Smoothing', 'Volumising Shampoo',
'Classics Lavender Anti Frizz Conditioner',
'Conditioner Nourishing Moisture',
'Extra Care Daily Oil Elixir Treatment',
'Soft & Smooth Conditioner', 'Fructis Coconut Water Shampoo',
'Potent Aloe Eucalyptus Conditioner', 'Keratin Shampoo',
'Shampoo Bar Manuka Honey & Almond', 'Nail Shaper 175mm',
'Scissors Nail Curve',
'Miracles Biotin Moisture Renewal Conditioner',
'Pro-V Colour Therapy Shampoo',
'Fructis Hydrating Aloevera Hair Food',
'Nutrisse 3 Espresso Permanent Hair Colour', 'Purple Shampoo',
'Conditioner Cleanse Replenish',
'Keratin Protein Smoothing Conditioner',
'Extra Care Ultimate Repair Strengthening Conditioner 400m',
'Soft & Smooth Shampoo', 'Shampoo Apple Fresh',
'Professional Advanced Oil Control Anti-Dandruff Shampoo',
'Fructis Nourishing Banana Hair Food',
'Olia Hair Colour 5.0 Brown', 'Bio Renew Coconut Milk Conditioner',
'Clay',
'Nourishing + Hydrating Coconut Milk Conditioner For Dry Hair',
'Pure Grow Conditioner',
'Bluey Kids 3In1 Shampoo- Conditioner & Bodywash',
'Corn Plane With Blade', 'Aloe & Eucalyptus Shampoo',
'Anti Dandruff Conditioner Dry Scalp Care',
'Addictive Brilliant Shine Conditioner', 'Combs Family Pack',
'Conditioner Curl Care Detangling',
'Olia Hair Colour 6.0 Light Brown', 'Long & Strong Shampoo',
'Moisture Shea Butter Shampoo',
'Bio Renew Hydrate Coconut Milk Shampoo',
'Extra Care Ultimate Style Hair Spray',
'Colour Protection Conditioner', 'Micellar Rose Water Shampoo',
'Toe Nail Clipper With File', 'Kids 3 In 1 Funny Honey',
'Tools Bioblender #3175', 'Lash & Brow Dye Kit Black',
'Extra Care Curl Flex Mousse', 'Taft Max Hold Styling Mousse',
'Ever Straightening + Smoothing & Shine Brazilian Keratin Therapy
Conditioner For Dull Hair',
'Extra Care Super Hold Hair Lacquer', 'Dry Scalp Care Conditioner',
'The Core Five Set', 'Duo Adhesive Clear', 'Dry Nourish Shampoo',
'Effortless Beauty Knit Pony Wraps',
'Expertise Volumising Growth Shampoo', 'Shampoo Tea Tree',
"Nice 'N Easy Frost & Tip Hair Colour Highlighting Kit",
'Smooth & Silky Refill Shampoo', 'Clean & Balanced Refill Shampoo',

'Dandruff Shampoo', 'Beautiful Brunette Dry Shampoo',
'Effortless Beauty Zigzag Headband',
'Indful Dandruff Repair Shampoo',
'Naturals 2 in 1 Antidandruff Shampoo & Conditioner',
'Ultra Hold Nail Glue 2.7mL',
'Professional Go Curly Enhance Define Gel',
'Family Anti Dandruf Shampoo Foam', 'Shampoo Strengthplex',
'Permanent Live Colour Blue Mercury', 'Thin Black Ponytailers',
'Classic Normal Conditioner', 'Brilliance 43 Red Passion',
'Fructis Damage Eraser Shampoo', 'Nail Clippers',
'Root Retoucher Spray Black', '4.62 Intense Deep Red',
'Extra Firm Hairspray',
'Professional Clean Blonde Violet Purple Toning Shampoo',
'Conditioner Strengthplex', 'Conditioner Restoring Ritual',
'Nourishing + Hydrating Coconut Milk Shampoo For Dry Hair',
'Bio Renew Conditioner Strawberry', 'Styling Powder',
'Got2b Glued Hairspray', 'Bio Renew Coconut Milk Shampoo',
'Conditioner Moisture', 'Pure Health Shampoo',
'Potent Aloe + Avocado Oil Nourishing Shampoo',
'Fructis Coconut Water Conditioner', 'Texture Clay',
'Extra Care Strong Hold Hair Spray',
'Styling Gel Tub Maximum Hold', 'Wet Look Gel',
'Dry Nourish Conditioner',
'Extra Care Normal Balance Leave In Conditioner',
'Thick Ponytailers Snagless Brown',
'Extra Care Colour Perfector Protecting Shampoo',
'Empower Volume & Shine Shampoo',
'Extra Care Ultimate Repair Biphase Repair & Gloss Treatment',
'Miracles Hyaluron Hydrate & Glow Daily Intensive Conditioner',
'Brilliance 46 Coffee Brown', 'Instant Brow Clear',
'Bio Renew Potent Aloe Honey Shampoo',
'Melbourne Hair Clay Crafting', 'Melbourne Hair Shaping Fibre',
'Paddle Brush', 'Colorsilk Black 10 Hair Color',
'Thicker Fuller Thickening', 'Styling Fiber',
'Longer & Stronger Conditioner',
'Extra Care Ultimate Repair Strengthening Conditioner',
'Touch Of Silver Brighten & Repair Shampoo',
'Nutrisse1 Liquorice Permanent Hair Colour',
"Nice 'N Easy 4 Natural Dark Brown Permanent Hair Colour",
'Elvive Dream Long Shampoo', 'Toe Nail Clipper',
'Extra Care Aqua Revive Moisturising Shampoo',
'Extra Strength Hydrate & Repair + Argan Oil of Morocco Shampoo For
Damaged Hair',
'Moisture Shea Butter Conditioner',
'Head Lice Removal Kit Comb Twin Pack',
'Extra Care Volume Styling Hair Spray',
'Touch Of Silver Colour Care Shampoo',

'Quenching + Coconut Curls Conditioner For Curly Hair',
'Normal Balance Conditioner', 'Extra Hold Hairspray',
'Colour Protection Shampoo', 'Shampoo Curl Care Moisturizing',
'Nature Fusion Conditioner',
'Violet Crush Purple Shampoo Intensive Toning',
'Fructis Hair Drink Banana', 'Taft Curl Defining Mousse',
'Aloe & Eucalyptus Conditioner', 'Multi-Tuft Brush',
'Brilliance 88 Dark Brown Allure', 'Pimple & Blackhead Remover',
'Hair Styling Spray Thermal Protections',
'Nutrisse 6 Acorn Permanent Hair Colour', 'Firm Hold Hairspray',
'Got2B Modelling Wax', 'Extra Care Curl Flex Creme',
'Expertise Volumising Growth Conditioner',
'Extra Care Super Styling Hold Hair Lacquer',
'Elvive Total Repair Conditioner',
'Nutrisse 5 Chocolate Brown Permanent Hair Colour',
"Nice 'N Easy 2 Natural Black Permanent Hair Colour",
'Got2b Farewell Flyaway',
'Thick & Full + Volumising Biotin & Collagen Shampoo For Fine Hair',
'Extra Care Strong Styling Hair Spray',
'Renewing + Repairing & Shine Argan Oil of Morocco Conditioner For Dry &
Damaged Hair',
'Nutrisse 8.13 Medium Ash Blonde Permanent Hair Colour',
'Extra Care 6 Miracles Oil Treatment',
'Permanent Live Colour Med Ash Blonde 7.1',
'Monoi Coconut Shampoo',
'3-65 Chocolate Brown Permanent Hair Colour',
'Frizz Ease Curls Reviver Mousse',
'Extra Care Liquid Silk Smoothing Shampoo',
'Kids Snap Clips Assortment',
'Brilliance 89 Bitter Sweet Chocolate',
'Dual Bristle Up Do Brush Natural & Nylon',
'Classic Coconut Conditioner', 'Classic Coconut Shampoo',
'Olia Hair Colour 5.3 Golden Brown', 'Instant Refresh Dry Shampoo',
'Extra Strength Hydrate & Repair + Argan Oil of Morocco Conditioner For
Damaged Hair',
'Conditioner Clean Balanced',
'Heavenly Hydration + Shine Cherry Blossom Conditioner For Thin & Fine
Hair',
'Extra Care Keratin Strength Hairspray',
'Quenching + Coconut Curls Shampoo For Curly Hair',
'Micellar Water Aloe Vera Shampoo', 'Purple Toning Conditioner',
"Nice 'N Easy 5 Natural Medium Brown Permanent Hair Colour",
'Dry Shampoo With Oat & Ceramide Dark Hair',
'Blonde It Up Toner Luminous Pearl',
'Elvive Extraordinary Oils Conditioner', 'Shockwaves Extreme',
'Fructis Normal Conditioner', 'Brilliance 03 Dark Amethyst',
'Fructis Nutri-Repair 3 Conditioner', 'Shampoo Colour Revitalise',

'Extra Care Push Up Volume Lifting Conditioner', 'Pomade',
        'Coconut Miracle Oil Mask', 'Elvive Total Repair Shampoo',
        'Colorsilk Light Ash Brown 50 Hair Colour',
        'Sensitive Scalp Care Conditioner',
        'Collagen Intense Repair & Shine Hair Mask',
        'Natural Balance Shampoo',
        'Renewing + Hydrating & Shine Argan Oil Of Morocco Extra Penetrating Oil
For Damaged & Heat Styled Hair',
        'Gold Tip Flat Tweezer', 'Shampoo Moisture Rich',
        'Extra Care Keratin Sea Salt Spray',
        'High Definition Curl Up Mousse',
        'Bio Renew Morocco Argan Oil Shampoo', 'Extra Strength Serum',
        'Brilliance 90 Black', '2 In 1 Ani-Dandruff Conditioning Shampoo',
        'Melbourne Hair Matt Balm', 'Shampoo Smooth And Silky',
        'Extra Care Keratin Heat Protect Spray',
        'Brilliance 59 Violet Wild Silk',
        'Fructis Coconut Water Nutri-Repair Shampoo',
        'Instant Brows Medium Brown', 'Repair & Protect Conditioner',
        'Super Charge Conditioner', 'Dandruff Conditioner',
        'Styling Hair Gel',
        'Extra Care Marrakesh Oil & Coconut Replenishing Conditioner',
        'Fructis Hair Food Pineapple', 'Black Bobby Pins',
        'Elvive Colour Protect Conditioner',
        'Fructis Hair Food Watermelon', 'Beach Curl Cream',
        'Argan Oil Vial Treatment', 'Soft Pomade Styler',
        'face Precision Tweezer Pink', 'Good 3.0 Dark Chocolate Brown',
        'Mess Grooming Wax', 'Refresh Me Quick Dry Shampoo',
        'Aspire Repair & Hydrate Shampoo', 'Good 2.0 Truffle Soft Black',
        'Violet Crush Purple Conditioner Tone Refresh',
        'Keratin Smooth Conditioner', 'Longer Stronger Shampoo',
        'Clipper Deluxe Nail With File',
        'Extra Care Liquid Silk Smoothing Conditioner',
        'Nutrisse Hair Colour 100 Truly Blonde Extra Light',
        'Clean & Balanced Refill 2 In 1', 'Classic Foundation',
        'Taft Maximum Hold Styling Lacquer',
        'Extra Care Ultimate Repair Mask Treatment',
        'Fructis Damage Eraser Conditioner',
        'Naturals Active Nourishment Shampoo',
        'Professional Advanced Soothing Care Anti-Dandruff Treatment',
        "Nice 'N Easy 3.5 Natural Darkest Brown Permanent Hair Colour",
        'Ever Straightening + Smoothing & Shine Brazilian Keratin Therapy Shampoo
For Dull Hair',
        'Thick & Full + Volumising Biotin & Collagen Conditioner For Fine Hair',
        'Mess Pliable Putty', 'Square Paddle Brush',
        'Good 5.0 Coffee Roast Brown', 'Dry Shampoo Volume Flax',
        'Pro-V Colour Therapy Conditioner', 'Long & Strong Conditioner',
        'Frizz Ease Secret Weapon Finishing Creme',

'Miracles Hyaluron Clarify & Hydrate Silicone Free Shampoo',
'Extra Strength Damage Remedy + Hydrating & Repairing Coconut Miracle Oil
Penetrating Oil For Dry & Damaged Hair',
'Volume Shampoo', 'Volumising Conditioner',
'Root Touch-up Colour Blending Gel 4 Dark Brown',
'Classic Clean Shampoo', 'Bio Renew Strawberry Shampoo',
'Taft Strong Hold Hair Spray', 'Pillar Box Red',
'Purple Conditioner', '1-0 Black Permanent Hair Colour',
'Kids 3 in 1 Merry Strawberry Shampoo Conditioner & Bodywash',
'Volume Blow Dry Spray', 'Tools Bioblender Duo #3163',
'Good 6.60 Pomegranate Red', 'Oatmeal Conditioner', 'Hairspray',
'Rock Solid Wax', '2 In 1 Anti Dandruff Shampoo Cool Menthol',
'Tweezer Fashion Slant', 'Pure Health Conditioner',
'Fashion Tape 36 Strip In Tin', 'Emery Boards',
'Hairspray Salon Finish Extra Hold',
'Live Colour Ultra Brights Magnetic Purple', 'Shampoo Dry Scalp',
'Repair & Protect Shampoo', 'Live Colour Permanent Silver Storm',
'Extreme Styling Matt Clay', 'Shampoo Cleanse & Replenish',
'Monoi Coconut Conditioner', 'Silky Smooth Conditioner',
'5 In 1 Leave In Spray Argan Oil', 'Classic Lavender Shampoo',
'Thin Brown Ponytailers', 'Decolour Hair Colour Stripper',
'Texturising Beach Spray', 'Empower Volume & Shine Conditioner',
'Colorsilk Light Gold Brown 54 Hair Colour',
'Frizz Ease Leave In Conditioning Spray',
'Manipro 2 In 1 Cleaner & Pusher', 'Profiller+ Thickening Shampoo',
'Frizz Ease Dream Curls Curl Defining Creme',
'Frizz Ease 3 Day Straight Spray', 'Full Powder Brush',
'Heatless Velcro Rollers', 'Fruits Apple Fresh Shampoo',
'Moisture Coconut Oil Shampoo',
'Bio Renew Hydrate Coconut Milk Conditioner',
'Purple Toning Shampoo', 'Leave In Conditioner',
'Bio Renew Strawberry Conditioner', 'Lamellar Shine Shampoo',
'Fructis Macadamia Hair Food', 'Creaseless Clips',
'Olia Hair Colour 1.0 Night Black',
'Olia Hair Colour 6.66 Very Light Intense Red',
'Colorsilk Medium Brown 41 Hair Color',
'Live Colour Permanent 3.0 Dark Brown',
'Aspire Intense Hydration Repairing 3 In 1 Mask',
'Emery Board 10 Pack', 'Hairspray Salon Finish Freeze Hold',
'Extra Care Aqua Revive Moisturising Conditioner',
'Fructis Watermelon Shampoo',
'Fashion Girl Rose Petals Detangler Spray',
'Silicone Scalp Massager', 'Brilliance 05 Dark Ruby',
'Fibre Styler', 'Volume Conditioner',
'Curl Care Intensive Deep Conditioner',
'Olia Hair Colour 10.1 Very Very Light Blonde',
'Dry Shampoo Rose Gold', 'Pure Grow Spray Treatment',

'Nutrisse 4 Tamarind Permanent Hair Colour',
'Blonde Care Purple Conditioner', 'Soft Headbands',
'Polish Nail Square Nude Matte 24 Piece',
'Extra Care Colour Perfector Protecting Conditioner',
'Blonde Shampoo', 'Got2B Glued Gel', 'Longer Stronger Conditioner',
'Extra Care Volume Styling Mousse',
'Heal & Hydrate + Shea Butter Hair Mask For Dry & Damaged Hair',
'Got2B Beachmatt Matt Paste', 'Conditioner Nourish & Shine',
'Wide Tooth Comb', 'Live Colour Permanent 1.1 Blue Black',
'Luminous Oils Moroccan Argan Oil & Camellia Conditioner',
'Root Touch Up Permanent Hair Colour - 4 Dark Brown',
'Hydrate Conditioner', 'Heat Protection Spray',
'Coconut Repair Conditioner',
'Protein Enriched Regular Hair Preparation For Men',
'Argan Oil Of Morocco Lightweight Shampoo',
'Fructis Hair Drink Pineapple', 'Conditioner Colour Revitalise',
'Nutrisse 5.25 Frosted Chestnut Permanent Hair Colour',
'Shampoo & Conditioner Bar Ocean Cruz 2 In 1',
'Keratin Sachet Conditioner', 'Kids 3 In 1 Fruit Blast',
'Bio Renew Conditioner Morocco Argan Oil',
'Bio Renew White Grapefruit & Mosa Mint Conditioner',
'Essentials Beauty Kit', 'Lamellar Serum Cream',
'Professional Clean Blonde Damage Rewind Violet Purple Toning
Conditioner',
'Argan Oil Shampoo', 'Argan Oil Conditioner',
'Blonde Care Purple Shampoo', 'Melbourne Volume Powder',
'Root Touch-up Colour Blending Gel 5 Medium Brown',
'Urban Matte Hair Wax Gel Pomade Styling Putty',
'Gold Tip Slant Tweezer',
'Flick Lash 2 In 1 Instant Lash Lift Comb & Curler',
'Organix Vegan Bleach Kit', 'Hydrate Shampoo',
"Nice 'N Easy 5G Natural Medium Golden Brown Permanent Hair Colour",
"Nice 'N Easy 9B Natural Light Beige Blonde Permanent Hair Colour",
"Nice 'N Easy 9A Natural Light Ash Blonde Permanent Hair Colour",
'Nail Shaper', 'Mega Hold Gel',
"Nice 'N Easy 8A Natural Medium Ash Blonde Permanent Hair Colour",
'No Crease Hair Clips', 'Curl Conditioner',
'Smooth & Sleek Shampoo',
'5-99 Rosewood Violet Permanent Hair Colour', 'Paste Styler',
'Professional Clean Blonde Violet Purple Toning Conditioner',
'Palette Hair Colour 6-0 Natural Light Brown',
'Fructis Colour Last Shampoo', 'Frizz Ease Air Dry Waves Foam',
'Blonde It Up Toner Iridescent Emerald',
'Blonde It Up Toner Radiant Opal',
'Root Touch Up Permanent Hair Colour - 5 Medium Brown',
'Basic Assorted Headwraps', 'Conditioner Purple',
'Shampoo Nourish & Shine', 'Conditioner Volume & Fullness',

'Miracles Hyaluron Clarify & Hydrate Silicon Free Shampoo',
'Luminous Oils Argan Oil & Shampoo',
'Nutrisse 9.13 Light Ash Beige Blonde Permanent Hair Colour',
'Semi Permanent Colour Purple',
'Nutrisse Hair Colour 6.6 Fiery Red', 'One Touch Clips Black',
'Silver Toner Live Colour', 'Aha Exfoliant', 'Scissor Cut Cvd',
'2 in 1 Normal Shampoo & Conditioner',
'Strengthplex Bond Repair Hair Mask',
'Live Colour Pastels Cotton Candy Pink', 'Skin Perfecting Brush',
'Smooth & Sleek Conditioner', 'Brilliance 95 Smokey Grey',
'Square Claw Clip', 'Hard Pomade', 'Tweezer Slant Tip',
'Treatment Catwalk Style Curls Rock Amplifier',
'Frizz Ease Miraculous Recovery Conditioner',
'Curved Nail Scissors', 'Naturals Active Nourishment Conditioner',
'French Nail Short Beige', 'By Manicare Kendall Lash',
'Moisture Coconut Oil Conditioner', 'Brown Bobby Pins',
'Smooth Conditioner', 'Brilliance 68 Chocolate Rush',
'Brilliance L43 Smouldering Red', 'Firm Hairspray',
'Monoi Coconut Oil 5 In 1 Leave In Spray', 'Sculpting Paste',
'Naturals Smooth & Shine Shampoo',
'Olia Hair Colour 7.0 Dark Blonde', 'Kids 3 In 1 Berry Bliss',
'Coal Tar Shampoo', 'Conditioner Cleanse & Replenish',
"Nice 'N Easy 8WR Natural Golden Auburn Permanent Hair Colour",
'Curl Shampoo', 'Curl Enhancer', 'Shampoo Itchy Scalp',
'Semi Permanent Colour Blue',
'Biotin Intense Moisture Lock Hair Mask',
'Professional Advanced Itch Care Anti-Dandruff Shampoo',
'Terry Towel Scrunchie Assortment', 'Extra Care Volume Powder',
"Nice 'N Easy 7CB Natural Dark Champagne Blonde Permanent Hair Colour",
'Melbourne Hair Texture Paste', 'Potent Aloe & Honey Conditioner',
'Slick Styler', 'Volumising Foam', 'Fruits Wild Berry Shampoo',
'Set Enhancing Eye Set', 'Powderful Volume Powder',
'Rapid Dry Hair Turban', 'Mens 2 In 1 Deep Clean & Anti- Dandruff',
'Live Colour Natrual Brown', 'Blonde+ Bond Building Shampoo',
'Marble Assortment Claw Clip', 'Tweezer Square Tip',
'Pro-V Miracle Oil Treatment', 'Naturals Moisture Intense Shampoo',
'Megahold Mousse', 'Good 7.0 Almond Creme Dark Blonde',
'Strength Plex Bond Repair Shampoo', 'Satin Headband',
'Thickening Paste', 'Chocolate Live Colour',
'Frizz Ease Dream Curls Shampoo',
'Brilliant Brunette Moisturising Conditioner',
'Blonde+ Bond Pre-Shampoo Treatment',
'Blonde+ Bond Building Conditioner', 'Kids Styling Gel',
'Frizz Ease Miraculous Recovery Shampoo',
'Ultra Men Sports Fresh 2 In 1 Anti-Dandruff Shampoo & Conditioner',
'Good 6.0 Mochaccino Brown',
'Elvive Extraordinary Oils Coconut Shampoo', 'Purple Toning Mask',

'Strengthplex Bond Repair Conditioner', 'Heat Protection Mist',
'Scrunchies Velvet',
'Micellar Aloe Vera Gentle Hydrating Conditioner',
'Plump It Up Dry Shampoo', 'Classic Clean Conditioner',
'Extra Care Marrakesh & Coco Oil Serum',
'Thickening Treatment Spray', 'Fruits Wild Berry Conditioner',
'Olia 7.40 Intense Copper',
'Root Touch Up Permanent Extra Lift Blonde',
'Natural Balance Conditioner',
"Nice 'N Easy 6 Natural Light Brown Permanent Hair Colour",
"Nice 'N Easy 5RB Natural Medium Chestnut Brown Permanent Hair Colour",
'Sensitive Scalp Care Shampoo', 'Smooth Shampoo',
'Silky Smooth Shampoo', 'Got2B Ultra Glued Gel',
'Instant Smooth Detangling Mist', 'Out Of Bed Clay',
'Blonde It Up Platinum Blonde',
'Potent Aloe & Avocado Conditioner', 'Fluffy Headband',
'Melbourne Hair Slick Pomade', 'Kids Soft Hair Ties Multipack',
'Lamellar Shine Spray', 'Mindful Dandruff Repair Conditioner',
'Long Handled Section Comb', 'Fructis Colour Last Conditioner',
'Lamellar Shine Mask', 'Luminous Oil Argan Oil & Shampoo',
'Live Colour Permanent Purple Punk', 'Nordic Lightening Spray G1',
'Sectioning Hair Clips', 'Nutrisse Naturals Hair',
'Flexible Control Shine + Hydrate Argan Oil Of Morocco Tame & Shine Cream
For Frizzy Hair',
'Extra Strength Damage Remedy + Hydrating & Repairing Coconut Miracle Oil
Conditioner For Damaged & Dry Hair',
'Argan Oil Mask', 'Mess Defining Paste', 'Texturising Spray',
'Detox For Men Shampoo', 'Shampoo Volume & Fullness',
'Extra Care Colour Perfector Express Leave In Conditioner',
'Soft Multi Ponytailers', 'Fruits Apple Fresh Conditioner',
'Fructis Sleek & Shine Flat Iron Spray', 'Targeted Treatment',
'Coconut Repair Shampoo', 'Semi Permanent Colour Pink',
'Crystal Nail Filer', 'Violet Crush Purple Shampoo Tone Refresh',
'Sheer Blonde Colour Renew Shampoo Go Blonder',
'Profiller+ Thickening Conditioner', 'Profiller+ Thickening Spray',
'Classic Lavender Conditioner', 'Original Wax', 'Moulding Cream',
'Decolour Hair Colour Remover',
'Naturals Intensive Moisture Conditioner',
'Luminous Moroccan Argan & Camellia oil Conditioner',
'Fructis Watermelon Conditioner',
'Sheer Blonde Conditioner Plat-Champ', 'Frizz Ease Original Serum',
'Frizz Ease Dream Curls Hair Spray', 'Glamour Dry Shampoo Volume',
'Conditioner Bar Manuka Honey & Vanilla',
'Aspire Repair & Hydrate Conditioner', 'Tame And Shine Spray',
'Fructis Full & Luscious Conditioner', 'Keratin Smooth Shampoo',
'Got2B Dry Shampoo Fresh It Up Volume', 'Oval Paddle Brush',
'Elvive Extraordinary Oils Coconut Conditioner',

'Ultimate Hold Hairspray Lacquer',
'Nutrisse 4.15 Mahogany Ash Brown Permanent Hair Colour',
'Professional Clean Blonde Damage Rewind Violet Purple Toning Shampoo',
'Naturals Intensive Moisture Shampoo', 'Black Satin Scrunchie',
'Dream Cream', 'Anti Flyaway Gel',
'Nutrisse 7.3 Honey Dip Permanent Hair Colour',
'Got2b Extra Fresh It Up Dry Shampoo',
'Brighter Blonder Purple Toning Shampoo',
'Brighter Blonder Purple Toning Conditioner',
'Sheer Blonde Colour Renew Conditioner Go Blonder',
'Super Charge Spray', 'Sponge Foundation Compact',
'Argan Oil Deep Sachet Conditioner',
'Extra Care Normal Balance Leave In Treatment',
'Live Colour Ultra Brights Shocking Pink',
'Extra Strength Damage Remedy + Hydrating & Repairing Coconut Miracle Oil
Shampoo For Damaged & Dry Hair',
'Argan Oil Of Morocco Lightweight Conditioner', 'Matte Cream',
'Colorsilk Medium Ash Blonde 70 Hair Colour',
'Touch Of Silver Colour Care Conditioner',
'Sheer Blonde Shampoo Moist Infusion',
'Fructis Hair Food Shampoo Pineapple',
'Brilliant Brunette Moisturising Shampoo',
'Bed Head Hair Cream After Party Treatment', 'Super Charge Serum',
'Frizz Free Cream', 'Casual Sea Salt Texturising Spray',
'Vent Brush', 'Precision Brow Grooming Kit',
'Nordic Blonde M1 Streaking Kit',
'Winged Eyeliner Stamp Intense Black Modest 10mm',
'Satin Hair Ties', 'Kids Mini Butterfly Clips',
"Nice 'N Easy 6RB Natural Light Chestnut Brown Permanent Hair Colour",
'Shampoo Purple', 'Expertise Daily Hydration Shampoo',
'Nordic L101 Silver Blonde', 'Nordic L1++ Ultimate Lightener',
'Got2b 4 Brows & Edges', 'Brow Lamination Gel Styling Kit',
'By Manicare Sophia Lash', 'Beauty Blender Sponge',
"Nice 'N Easy 5W Natural Medium Caramel Brown Permanent Hair Coluor",
"Nice 'N Easy 6R Natural Light Auburn Permanent Hair Colour",
"Nice 'N Easy 7 Natural Dark Blonde Permanent Hair Colour",
'Blonde Conditioner', 'Frizz Free Oil',
'Conditioner Colour Revital', 'Face Roller',
'Coconut & Lime Shampoo Bar',
'Blonde Care Purple Deep Conditioner',
'Brilliance 23 Burgundy Chic', 'Brilliance L68 Cinnamon Brown',
'Brilliance Timeless Blonde 25', 'Brilliance Magical Mahogany 76',
'Ultra Men Moisture Recharge 2 In 1 Anti-Dandruff Shampoo & Conditioner',
'Pure Health Mask', 'Lamellar Shine Conditioner',
'3-0 Dark Brown Permanent Hair Colour', 'Heat Protect Spray',
'Nutrisse 4.3 Cappuccino Permanent Hair Colour',
'Fructis Banana Conditioner', 'Nordic L1 Intensive Lightener',

'Keratin Smooth Serum', 'Coconut & Lime Conditioner Bar',
'Small Thin Elastic Hair Ties', 'Brilliance 85 Violet Vision',
'Brilliance 91 Blue Black', 'Brilliance 50 Venetian Sunset',
'Brilliance Cashmere Red 42',
'Touch Of Silver Brighten & Repair Conditioner',
'Nail Filer Scissor Straight 01-313',
'10-1 Ultra Light Ash Blonde Permanent Hair Colour',
'Root Retoucher Spray Dark Brown',
'Fructis Damage Eraser Treatment',
'Fructis Full & Luscious Shampoo',
'Thick Ponytailers Snagless Blonde', 'Spray Treatment',
'Thin Blonde Ponytailers', 'Blonde Bobby Pins',
"Nice 'N Easy 9.5 Natural Extra Light Blonde Permanent Hair Colour",
"Nice 'N Easy 3.5BG Natural Dark Burgundy Permanent Hair Colour",
'White Toner', 'Sapphire Nail File',
'Love Your Smooth Leave In Crme Treatment', 'Gua Sha',
'Miracles Hyaluron Hydration Lock Milk Serum',
'Fructis Hair Food Conditioner Pineapple',
'Naturals Smooth & Shine Conditioner',
'Curl Care Curl Defining Cream',
"Nice 'N Easy 8 Medium Blonde Permanent Hair Colour",
"Nice 'N Easy 9.5A Natural Baby Blonde Permanent Hair Colour",
'All In One Treatment', 'Shampoo Colour Revital',
'Expertise Daily Hydration Conditioner',
'Brilliance L75 Brown Sugar', 'Lengthening Starter Kit Lashes',
'Brilliance 73 Cool Frosty Browns',
'Touch Of Silver Toning Treatment Mask',
'Artificial Nail System Clear - Almond Shape',
'Good 4.0 Cacao Brown', 'Live Colour Permanent Pink Passion',
'Longer Stronger Anti Breakage Hair Mask',
'Olia 9.11 Silver Smoke', 'Hydrating Beach Waves Spray',
'Define & Protect Curl Spray',
'Colorsilk Dark Ash Blonde 60 Hair Colour',
'Professional Advanced Itch Care Anti-Dandruff Conditioner',
'Root Retoucher Spray Brown', 'Extra Care Strong Hold Mousse',
'Extra Care Ultimate Styling Hair Mousse',
'Luminous Oils Coconut And Frangipani Shampoo',
'Luminous Oils Coconut & Frangipani Conditioner',
'By Manicare Pre Glued Elsa Lash',
'C3 Nordic Blonde Anti-Yellow Toner',
'Natural Demi Lash 120 Black', 'Olia B+++ Max Bleach',
'Olia Hair Colour 8.0 Blonde',
'Got2b Guardian Angel Heat Protection Spray',
'Semi Permanent Rose Hair Colour',
'Semi Permanent Turquoise Hair Colour',
'Live Colour Pastels Cool Grey', 'Live Colour Pastels Lilac Blush',
'Live Colour Ultra Brights Electric Blue',

```
              'Palette Permanent Hair Colour 7-0 Light Brown',
              'Nutrisse 7 Almond Cream Permanent Hair Colour',
              'Fruits Shampoo Coconut & Lime', 'Fructis Banana Shampoo',
              'Colour Hairspray Glitter', 'Colour Hairspray Pink',
              'Elastomer Ponytailer', 'Coconut & Lime Conditioner',
              'Colour Hairspray Purple', 'Colour Hairspray Red',
              'Brazilian Keratin Mask', 'Anti Dandruff Shampoo Dry Scalp Care'],
            dtype=object)
```

[129]: ```
data[data['category'] == 'HAIR CARE'][['item_name','subcat']]
```

[129]:

	item_name	subcat
3616	Jaw Clip Extra Large	Clips & Pins
3664	Comfy Natural Ponytailers 12 pieces	Other
3666	Elastics All Hair	Other
3755	4mm No Damage Black Elastics	Other
3891	Mixed Inner Polybands 300 pieces	Other
...	...	...
77356	2 In 1 Ani-Dandruff Conditioning Shampoo	Other
77357	Shampoo Nourish & Shine	Other
77539	Kids 3 In 1 Funny Honey	Other
77747	Kids 3 In 1 Berry Bliss	Other
77800	Kids 3 in 1 Merry Strawberry Shampoo Condition…	Other

```
[2359 rows x 2 columns]
```

[147]: ```
hair_categories = {'Shampoo': [
        'shampoo', 'shampo','dry shampoo' # includes misspellings
    ],

    # Conditioner products
    'Conditioner': [
        'conditioner', 'conditioning'
    ],

    # Hair Colour/Dye products
    'Hair Colour': [
        'hair colour', 'hair color', 'permanent hair colour', 'permanent hair␣
 ↪color',
        'excellence', 'preference', 'casting creme gloss', 'magic retouch',
        'colorsilk', 'nice \'n easy', 'olia', 'live colour', 'nutrisse',
        'brilliance', 'good', 'palette', 'toner', 'lightener', 'bleach',
        'blonde it up', 'root touch', 'root retoucher', 'decolour',
        'colour remover', 'colour stripper', 'semi permanent', 'dye kit',
        'purple shampoo', 'purple conditioner', 'toning', 'violet',
        'colour protect purple', 'silver', 'blonde care purple'
    ],
```

```python
# Styling Products
'Styling Products': [
    'hairspray', 'hair spray', 'mousse', 'gel', 'wax', 'pomade',
    'paste', 'clay', 'fiber', 'styling', 'lacquer', 'serum',
    'cream', 'balm', 'putty', 'texturising', 'volumising',
    'curl defining', 'curl enhancer', 'wave', 'fibre styler',
    'moulding cream', 'sculpting paste', 'styling powder'
],

# Hair Treatments & Masks
'Hair Treatments': [
    'treatment', 'mask', 'oil', 'leave-in', 'leave in', 'sealing serum',
    'repair', 'nourishing', 'deep', 'intensive', 'bond', 'protein',
    'keratin', 'argan', 'coconut oil', 'extraordinary oil',
    'masque', 'elixir', 'penetrating oil'
],

# Hair Accessories
'Hair Accessories': [
    'ponytailers', 'elastics', 'polybands', 'hair ties', 'clips',
    'barrette', 'snap clips', 'octopus clips', 'bobby pins',
    'headband', 'scrunchie', 'claw clip', 'jaw clip', 'rollers',
    'headwraps', 'pony wraps', 'butterfly clips'
],

# Hair Tools/Brushes
'Hair Tools': [
    'brush', 'comb', 'detangling', 'paddle brush', 'vent brush',
    'radial', 'porcupine', 'section comb', 'wide tooth comb',
    'turban', 'hair wand'
],

# Specialty Shampoos
'Anti-Dandruff': [
    'anti dandruff', 'anti-dandruff', 'dandruff', 't-gel', 'coal tar',
    'scalp care', 'itchy scalp'
],


# Head Lice Products
'Lice Treatment': [
    'head lice', 'headlice', 'lice removal', 'lice kit'
],

# Beauty/Nail Care (mixed in with hair products)
'Beauty & Nail Care': [
```

```
        'nail', 'eyelash', 'lash', 'brow', 'tweezer', 'clipper',
        'cuticle', 'emery', 'file', 'scissors', 'grooming set',
        'adhesive', 'curler', 'eyeliner', 'foundation', 'blender',
        'beauty', 'face roller', 'gua sha'
    ]
}

def get_kitchen_subcat(row):
    if row["category"].lower() == "hair care":
        item_lower = row["item_name"].lower()
        for subcat, keywords in hair_categories.items():
            for keyword in keywords:
                if keyword in item_lower:
                    return subcat
        return "Other"
    else:
        return row["subcat"]  # keep the existing subcat if not kitchenware
# apply row-wise
data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

```
<ipython-input-147-923b7462d7c9>:86: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data["subcat"] = data.apply(get_kitchen_subcat, axis=1)
```

[148]: `data[data['category'] == 'HAIR CARE'][['item_name','subcat']]`

[148]:

	item_name	subcat
3616	Jaw Clip Extra Large	Hair Accessories
3664	Comfy Natural Ponytailers 12 pieces	Hair Accessories
3666	Elastics All Hair	Hair Accessories
3755	4mm No Damage Black Elastics	Hair Accessories
3891	Mixed Inner Polybands 300 pieces	Hair Accessories
…	…	…
77356	2 In 1 Ani-Dandruff Conditioning Shampoo	Shampoo
77357	Shampoo Nourish & Shine	Shampoo
77539	Kids 3 In 1 Funny Honey	Other
77747	Kids 3 In 1 Berry Bliss	Other
77800	Kids 3 in 1 Merry Strawberry Shampoo Condition…	Shampoo

```
[2359 rows x 2 columns]
```

[139]: `data.to_csv('new_data.csv',index=False)`

[140]: `data.columns`

```
[140]: Index(['product_code', 'category', 'item_name', 'brand', 'item_price',
              'item_size', 'unit_price', 'unit_weight', 'similarity_score', 'subcat',
              'unit_weight_new', 'unit_weight_type', 'item_weight_new',
              'item_weight_type', 'std_item_weight', 'std_item_unit'],
             dtype='object')
```

```
[142]: data['category'].unique()
```

```
[142]: array(['Easter', "Footy's Back", 'Bonus Credit Products',
              'Meat & Seafood', 'Fruit & Vegetables', 'Dairy, Eggs & Fridge',
              'Bakery', 'Deli', 'Pantry', 'Dietary & World Foods',
              'Chips, Chocolates & Snacks', 'Drinks', 'Frozen', 'Household',
              'Health & Beauty', 'Liquorland', 'Back to School', 'EASTER',
              'BOXED CHOCOLATES', 'BAKERY EASTER', 'SEASONAL EVENTS',
              'SEAFOOD DELI', 'VINE VEGETABLES', 'SNACKS', 'SOFTDRINKS',
              'POTATOES', 'LAMB', 'SPIRITS', 'NUTS/DRIED', 'WINE', 'CONTINENTAL',
              'SLICED MEATS', 'SAUSAGES', 'BEER', 'VALUE ADDED FP',
              'BERRIES & CHERRIES', 'PORK', 'FROZEN DESSERTS',
              'BH BOUGHT IN EASTER', 'TOMATOES', 'BAKERY BOUGHT IN', 'BEEF',
              'LIQUOR RTD', 'ANTIP/OLIVE/DIP/PATE', 'BAKERY PACKAGED CAKE',
              'HARD VEGETABLES', 'POULTRY (MT)', 'BISCUITS & COOKIES',
              'FROZEN VEGETABLES', 'PREPACKAGED SEAFOOD', 'ENTERTAINMENT',
              'GOURMET CHEESE', 'VALUE ADDED', 'SMALLGOODS DY',
              'CHILLED DESSERTS', 'SOFT VEGETABLES', 'BUNCH LINES',
              'DRIED FRUIT/NUTS', 'SAUCES/RELISH', 'BAKING MIXES',
              'INSTORE CAKE', 'PROTEIN & MEALS', 'CONFECTIONERY',
              'PIES AND QUICHES', 'DISPOSABLE TABLEWARE', 'PICNIC POOL BBQ ACC',
              'MEALS', 'FROZEN SNACKS', 'GRAVY/STOCK', 'GRAPES', 'MEAL BASES',
              'BANANAS', 'WATER', 'CHEESE DAIRY', 'CHILLED SPREADS', 'CEREAL',
              'MELONS & TROPICAL', 'CITRUS', 'APPLES & PEARS',
              'JUICES & CORDIALS', 'INSTORE BREAD', 'KITCHENWARE',
              'BARS GUM POCKET PACK', 'ASIAN FOODS', 'SPREADS', 'SOUP',
              'BAKERY SNACKS', 'BEAUTY/TRIAL TRAVEL', 'CANNED FISH',
              'CHILLED HEALTHFOODS', 'EGGS', 'SMOKED & CURED FISH',
              'NUTRITIONAL SNACKS', 'SMALLGOODS', 'RICE', 'PASTA',
              'BREAD ROLLS & FBREAD', 'HOMEWARES', 'STATIONERY', 'MEXICAN FOODS',
              'FACIAL TISSUES', 'JUICES/DRINKS', 'CANNED MEALS',
              'FROZEN WORLD FOODS', 'KITCHEN NEEDS/BAGS', 'GROCERY MILK',
              'SALAD DRESSINGS', 'FROZEN SEAFOOD', 'SKIN CARE',
              'WOMENS DEOS & GROOM', 'MEDICINAL PRODUCTS', 'ORGANIC MEAT',
              'SANITARY PROTECTION', 'FRUIT-SHELF STABLE', 'INSTANT NOODLES',
              'TECH ACCESSORIES', 'DY MILK', 'INDIAN FOODS', 'ICE CREAM',
              'SOAPS & BODY WASH', 'POULTRY (DL)', 'HARDWARE', 'PEST CONTROL',
              'MENS DEOS & GROOMING', 'INFANT PERSONAL', 'SHOE CARE',
              'HAIR CARE', 'FAMILY SOCKS', 'VITAMINS', 'DESSERTS (GR)',
              'LAUNDRY ACCESSORIES', 'DENTAL HEALTH', 'COSMETICS/TOILETRIES',
              'AIRCARE', 'CLEANING GOODS', 'HEALTH FOODS', 'CONVENIENCE MEALS',
```

'SPEC/FRESH CHEESE', 'VINEGAR', 'HAMS', 'GAME', 'STONE FRUIT',
'MUSHROOMS', 'ORGANIC FRUIT & VEG', 'AVOCADO & MANGO',
'HERBS & CHILLIES', 'OILS', 'MILK ADDITIVES', 'DAIRY WORLD FOODS',
'GRAB & GO', 'BBQ', 'DL PREPARED FOODS', 'SPICES/HERBS', 'COFFEE',
'FLOUR', 'SUGAR/SWEETENERS', 'CANNED VEG', 'INTERNATIONAL FOODS',
'PICKLED VEGETABLES', 'TEA', 'INFANT FOOD', 'INFANT FORMULA',
'BETTER FOR YOU', 'ENERGY/SPORT/ICEDTEA', 'COLD DRINK',
'FRESH PET FOOD DY', 'DISHWASHING', 'TOILET PAPER', 'PAPER TOWELS',
'SHOPPING BAGS', 'LAUNDRY', 'ELECTRICAL', 'TELCO', 'MAGAZINES',
'HOUSEHOLD GLOVES', 'LBI MERCH', 'HOUSEHOLD APPLIANCES', 'AUTO',
'PARTY GOODS', 'BRUSHWARE', 'BEST BUYS', 'TOYS & HOBBIES',
'GARDEN NON GREENS', 'HEATING & COOLING', 'FAMILY UNDERWEAR',
'FITNESS/HEALTH', 'HOSIERY', 'LEISURE', 'SEASONAL APPAREL',
'CARDS/WRAPS', 'ANCIL SERVICES', 'BABYWEAR', 'INFANT NAPPIES'],
      dtype=object)

Meat & Seafood:

```
[143]: data[data['category'] == 'Meat & Seafood']['item_name'].unique()
```

```
[143]: array(['Coles Lamb Leg Roast Half | approx. 1.74kg',
          'Coles No Added Hormone Beef 5 Star Extra Trim Mince | 500g',
          'Coles Lamb Whole Leg Roast | approx. 2.8kg',
          'Coles Boneless Pork Leg Roast | approx. 2.1kg',
          'Continental Stock Pot Beef | 112g',
          'Coles RSPCA Approved Chicken Thigh Fillets Large Pack | approx. 1.2kg',
          'Mount Barker Free Range Chicken Breast Fillets Skin Off | approx. 500g',
          'Coles Deli Thawed Australian Cooked Black Tiger Pra … approx. 250g',
          'Coles RSPCA Approved Chicken Breast Fillets Large Pack | approx. 1.4kg',
          'Continental Chicken Stock Pot | 112g',
          'Coles Deli Thawed Australian Cooked Black Tiger Prawns Extra Large |
      approx. 250g',
          'Tassal Tasmanian Smoked Salmon | 250g',
          'Tassal Hot Smoked Salmon Natural | 150g',
          'Tassal Tasmanian Smoked Salmon | 150g',
          'Tassal Tasmanian Smoked Salmon Premium Selection | 180g',
          'Tassal Hot Smoked Salmon Cracked Peppercorn | 150g',
          'Coles Deli Thawed Australian Cooked Black Tiger Praw … approx. 250g',
          'Tassal Tasmanian Smoked Salmon Premium Selec … 180g',
          'Coles Deli Thawed Australian Cooked Black T … approx. 250g',
          'Coles RSPCA Approved Chicken Breast Fillets Large  … approx. 1.4kg',
          'Coles No Added Hormone Beef 4 Star Lean Mince | 500g',
          'Harvey Beef Premium Mince | 500g',
          'Coles Deli Fresh Tasmanian Salmon Portions Skin On | approx. 200g',
          'Mount Barker Free Range Chicken Thigh Fillets Skin Off | approx. 550g',
          'Coles Beef Eye Fillet Steak Small Pack | 300g',
          'Coles RSPCA Approved Chicken Breast Fillets Small Pack | approx. 600g',
          'Coles Deli Thawed Australian Cooked Black Tiger Prawn … approx. 250g',
```

```
          'Coles RSPCA Approved Chicken Breast Fillets Large P … approx. 1.4kg',
          'Mount Barker Free Range Chicken Breast Fillets S … approx. 500g',
          'Coles RSPCA Approved Chicken Thigh Fillets Large P … approx. 1.2kg',
          'Tassal Tasmanian Smoked Salmon Premium Selecti … 180g',
          'Coles Deli Fresh Tasmanian Salmon Portions  … approx. 200g',
          'Coles Deli Thawed Australian Cooked Black Tiger Pr … approx. 250g',
          'Mount Barker Free Range Chicken Thigh Fillets S … approx. 550g',
          'Coles RSPCA Approved Chicken Breast Fillets Small Pa … approx. 600g'],
      dtype=object)
```

[8]:
```python
subcategory_map_meat_seafood = {
    "beef": "Beef",
    "eye fillet": "Beef",
    "lamb": "Lamb",
    "pork": "Pork",
    "chicken": "Chicken",
    "thigh fillet": "Chicken",
    "breast fillet": "Chicken",
    "prawn": "Seafood - Prawns",
    "tiger prawn": "Seafood - Prawns",
    "salmon": "Seafood - Salmon",
    "smoked salmon": "Seafood - Salmon",
    "stock pot": "Stock & Broth",
    "stock": "Stock & Broth",
}


def get_meat_seafood_subcat(row):
    if row["category"].lower() == 'meat & seafood':
        item_lower = row["item_name"].lower()
        for keyword, subcat in subcategory_map_meat_seafood.items():
            if keyword in item_lower:
                return subcat
        return "Other"
    else:
        return row["subcat"]


data["subcat"] = data.apply(get_meat_seafood_subcat, axis=1)
```

[9]:
```python
data[data['category'] == 'Meat & Seafood']
```

[9]:
```
     product_code         category  \
16        9246065  Meat & Seafood
17        8112799  Meat & Seafood
53        9246021  Meat & Seafood
54        4174794  Meat & Seafood
55        8392913  Meat & Seafood
…             …                …
```

```
1276        8106800   Meat & Seafood
1332        3477360   Meat & Seafood
1333        2686693   Meat & Seafood
1334        6915384   Meat & Seafood
1335        2263168   Meat & Seafood


                                      item_name           brand  \
16            Coles Lamb Leg Roast Half | approx. 1.74kg        Coles
17      Coles No Added Hormone Beef 5 Star Extra Trim …        Coles
53            Coles Lamb Whole Leg Roast | approx. 2.8kg       Coles
54        Coles Boneless Pork Leg Roast | approx. 2.1kg       Coles
55                   Continental Stock Pot Beef | 112g   Continental
…                                              …              …
1276  Coles Deli Fresh Tasmanian Salmon Portions  …        Coles
1332  Tassal Hot Smoked Salmon Cracked Peppercorn | …  Smoked Salmon
1333  Coles Deli Thawed Australian Cooked Black Tige…        Coles
1334  Mount Barker Free Range Chicken Thigh Fillets …  Mount Barker
1335  Coles RSPCA Approved Chicken Breast Fillets Sm…        Coles


      item_price item_size  unit_price unit_weight  similarity_score  \
16         26.10    1.74kg       15.00         1kg          0.077257
17         10.00      500g       20.00         1kg          0.123729
53         42.00     2.8kg       15.00         1kg          0.037386
54         16.80     2.1kg        8.00         1kg          0.140976
55          4.50      112g        4.02        100g          0.077868
…              …        …           …           …                 …
1276        8.80      200g       44.00         1kg          0.112483
1332       13.00      150g       86.67         1kg          0.144647
1333        9.00      250g       36.00         1kg          0.114650
1334       10.72      550g       19.50         1kg          0.129591
1335        8.70      600g       14.50         1kg          0.111381


               subcat  unit_weight_new unit_weight_type  item_weight_new  \
16               Lamb              1.0               kg             1.74
17               Beef              1.0               kg           500.00
53               Lamb              1.0               kg             2.80
54               Pork              1.0               kg             2.10
55               Beef            100.0                g           112.00
…                   …                …                …                …
1276  Seafood – Salmon             1.0               kg           200.00
1332  Seafood – Salmon             1.0               kg           150.00
1333            Other              1.0               kg           250.00
1334          Chicken              1.0               kg           550.00
1335          Chicken              1.0               kg           600.00


      item_weight_type  std_item_weight std_item_unit
16                  kg            1740.0             g
```

```
17                       g           500.0          g
53                      kg          2800.0          g
54                      kg          2100.0          g
55                       g           112.0          g
...                    ...             ...        ...
1276                     g           200.0          g
1332                     g           150.0          g
1333                     g           250.0          g
1334                     g           550.0          g
1335                     g           600.0          g

[100 rows x 16 columns]
```

```
[2]: import pandas as pd
     data = pd.read_csv('new_data.csv')
     data.head()
```

```
[2]:    product_code category                                       item_name  \
     0       8371390   Easter    Coles Hot Cross Buns Traditional Fruit | 6 Pack
     1       7473849   Easter            Coles Hot Cross Buns Choc Chip | 6 Pack
     2       5726070   Easter   Coles Hot Cross Buns Traditional Fruit Mini | …
     3       4885191   Easter   Cadbury Dairy Milk Easter Chocolate Eggs Bag |…
     4       3378370   Easter      Coles Hot Cross Buns Apple & Cinnamon | 6 Pack


          brand  item_price item_size  unit_price unit_weight  similarity_score  \
     0    Coles         4.4    6 Pack        0.73         1ea          0.198373
     1    Coles         4.4    6 Pack        0.73         1ea          0.198373
     2    Coles         4.4    9 Pack        0.49         1ea          0.198373
     3  Cadbury         6.7      114g        5.88        100g          0.214731
     4    Coles         4.4    6 Pack        0.73         1ea          0.198373


               subcat  unit_weight_new unit_weight_type  item_weight_new  \
     0          Other              1.0             each              6.0
     1      Chocolate              1.0             each              6.0
     2          Other              1.0             each              9.0
     3   Easter Eggs            100.0                g            114.0
     4          Other              1.0             each              6.0


       item_weight_type  std_item_weight std_item_unit
     0             pack              6.0          pack
     1             pack              6.0          pack
     2             pack              9.0          pack
     3                g            114.0             g
     4             pack              6.0          pack
```

```
[10]: category_count_of_non_weights = data.loc[data['item_size'].isna() &␣
      ↪data['unit_weight'].isna(), 'category'].value_counts()
```

```
[11]: category_count_of_non_weights
```

```
[11]: STATIONERY              569
      KITCHENWARE             303
      TECH ACCESSORIES        232
      HOSIERY                 116
      SEASONAL EVENTS          98
      SHOE CARE                92
      FAMILY SOCKS             91
      HARDWARE                 72
      MAGAZINES                60
      TOYS & HOBBIES           58
      ELECTRICAL               56
      GARDEN NON GREENS        54
      CARDS/WRAPS              50
      TELCO                    50
      BEST BUYS                44
      FAMILY UNDERWEAR         31
      HOUSEHOLD APPLIANCES     22
      FITNESS/HEALTH           20
      AUTO                     19
      ANCIL SERVICES           12
      SHOPPING BAGS            12
      LBI MERCH                12
      SEASONAL APPAREL         10
      HEATING & COOLING         9
      CHEESE DAIRY              9
      LEISURE                   9
      BABYWEAR                  7
      PICNIC POOL BBQ ACC       6
      BRUSHWARE                 6
      EASTER                    5
      HERBS & CHILLIES          4
      PARTY GOODS               2
      HAIR CARE                 2
      Name: category, dtype: int64
```

Seasonal Events:

```
[16]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'seasonal events')
      ]['item_name'].unique()
```

```
[16]: array(['AFL Ball Size 3', 'Finals Official AFL Game',
             'Rookie Official AFL Game', 'Lazy Susan Pantry Tray',
             'Notebook Carlton', 'Notebook Geelong', 'Notebook Hawthorn',
```

```
          'Notebook Richmond', 'Notebook Gws', 'Notebook Gold Coast',
          'Notebook Port Adelaide', 'Notebook Sydney', 'Notebook Broncos',
          'Notebook Panthers', 'Notebook Raiders', 'Notebook Roosters',
          'Notebook Storm', 'Portion Freezer Pods',
          'Prep Storage Box With 6 Dividers And Lid', 'Notebook Dragons',
          'Notebook Knights', 'Notebook West Tigers', 'Eels', 'Saver',
          'Freezer Pod', 'Or Asparagus Keeper', 'Saver Container',
          'Crew Sports Socks Storm', 'Notebook Adelaide',
          'Notebook Essendon', 'Notebook Western Bulldogs',
          'Notebook Bulldogs', 'Notebook Sharks', 'Notebook Warriors',
          'Notebook Brisbane', 'Notebook Fremantle', 'Notebook Melbourne',
          'Notebook North Melbourne', 'Notebook St Kilda',
          'Notebook West Coast', 'Dish', 'Notebook Cowboys',
          'Notebook Sea Eagles', 'Notebook Eels', 'Notebook Titans',
          'Notebook Dolphins', 'Carlton', 'Jug'], dtype=object)
```

[21]:
```python
# Define subcategory mapping
seasonal_events_subcats = {
    "afl ball": "Card Game",
    "afl game": "Card Game",
    "lazy susan pantry tray": "Food Storage",
    "portion freezer pods": "Food Storage",
    "prep storage box with 6 dividers and lid": "Food Storage",
    "freezer pod": "Food Storage",
    "asparagus keeper": "FoodStorage",
    "saver container": "Food Storage",
    "saver": "Food Storage",
    "notebook": "Stationery",
    "dish": "Kitchenware",
    "jug": "Kitchenware",
    "crew sports socks": "Apparel",
    "carlton": "Team Merchandise",
    "eels": "Team Merchandise",
}

# Function to map only Season Events rows with NaN size & unit_weight
def get_seasonal_event_subcat(item):
    item_lower = item.lower()
    for keyword, subcat in seasonal_events_subcats.items():
        if keyword in item_lower:
            return subcat
    return "Other"

# Apply only to filtered rows
mask = (
    (data['item_size'].isna()) &
    (data['unit_weight'].isna()) &
```

```
        (data['category'].str.lower() == 'seasonal events')
)

data.loc[mask, 'subcat'] = data.loc[mask, 'item_name'].
  ↪apply(get_seasonal_event_subcat)
```

```
[22]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'seasonal events')
      ]
```

```
[22]:       product_code         category                   item_name  \
      38846      9995335  SEASONAL EVENTS            AFL Ball Size 3
      39856      9361610  SEASONAL EVENTS  Finals Official AFL Game
      40499      9368051  SEASONAL EVENTS  Rookie Official AFL Game
      41709      9055001  SEASONAL EVENTS     Lazy Susan Pantry Tray
      41779      9949430  SEASONAL EVENTS           Notebook Carlton
      ...            ...              ...                        ...
      72351      9950144  SEASONAL EVENTS              Notebook Eels
      72352      9950279  SEASONAL EVENTS       Notebook West Tigers
      72353      9950290  SEASONAL EVENTS            Notebook Titans
      72379      9052604  SEASONAL EVENTS                        Jug
      72478      9053888  SEASONAL EVENTS                      Saver

                        brand item_price item_size  unit_price unit_weight  \
      38846            Unknown        0.0       NaN         NaN         NaN
      39856  Finals Official AFL        0.0       NaN         NaN         NaN
      40499  Rookie Official AFL        0.0       NaN         NaN         NaN
      41709   Lazy Susan Pantry        0.0       NaN         NaN         NaN
      41779    Notebook Carlton       10.0       NaN         NaN         NaN
      ...                  ...        ...       ...         ...         ...
      72351               Eels       10.0       NaN         NaN         NaN
      72352            Unknown       10.0       NaN         NaN         NaN
      72353            Unknown       10.0       NaN         NaN         NaN
      72379                Jug        0.0       NaN         NaN         NaN
      72478              Saver        0.0       NaN         NaN         NaN

             similarity_score       subcat unit_weight_new unit_weight_type  \
      38846          0.182327    Card Game             NaN              NaN
      39856          0.182327    Card Game             NaN              NaN
      40499          0.182327    Card Game             NaN              NaN
      41709          0.182327  Food Storage             NaN              NaN
      41779          0.153484   Stationery             NaN              NaN
      ...                 ...          ...             ...              ...
      72351          0.153484   Stationery             NaN              NaN
      72352          0.153484   Stationery             NaN              NaN
```

```
72353        0.153484    Stationery          NaN              NaN
72379        0.182327    Kitchenware         NaN              NaN
72478        0.182327    Food Storage        NaN              NaN


       item_weight_new  item_weight_type  std_item_weight  std_item_unit
38846             NaN               NaN              NaN            NaN
39856             NaN               NaN              NaN            NaN
40499             NaN               NaN              NaN            NaN
41709             NaN               NaN              NaN            NaN
41779             NaN               NaN              NaN            NaN
...               ...               ...              ...            ...
72351             NaN               NaN              NaN            NaN
72352             NaN               NaN              NaN            NaN
72353             NaN               NaN              NaN            NaN
72379             NaN               NaN              NaN            NaN
72478             NaN               NaN              NaN            NaN

[98 rows x 16 columns]
```

Shoe Care:

```python
[23]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'shoe care')
      ]['item_name'].unique()
```

```
[23]: array(['Lint Roller', 'Bamboo Handle Lint Roller', 'Oval Lace White 45',
             'Flat Lace Black 45 inch', 'Flat Lace White Kids 27 inch',
             'Lint Roller Refill', 'Gel Activ Insole Casual Large',
             'Sewing Kit', 'Foam Latex Comfort Insole 1 Pair',
             'Soft Touch Scissors 9 inch', 'Silver Safety Pins Assorted Sizes',
             'Go Sockless', 'Footcare Work Insoles', 'Household Scissors',
             'Threaded Needle Kit', 'Gel Heel Cup',
             'Polyester Thread Multi Colours', 'Boot Brown Lace 60"',
             'Tape Measure', 'Gel Activ Insole Work & Boot Large',
             'Gel Activ Insole Work & Boot Small',
             'Polyester Thread Black & White', 'Nappy Pins',
             'Performance Lace 38 Solid Black', 'White Hank Elastic',
             'Sneaker Balls', 'Fashion Tape', 'Shock Reducer Everyday Insole',
             'Lint Shaver', 'Gel Activ Insole Sport Large',
             'Plastic Handle Scissors', 'Compact Needles Assorted',
             'Stitch Ripper', 'Spandex Gel Metatarsal Cushions',
             'Multi Purpose Repair Sewing Needles', "Oval Lace Black 45''",
             'Flat Lace White 45 inch', 'Horizontal Lint Roller',
             'Heel & Arch Support Women', "Men's Heel and Arch Support",
             'Boot Black Lace 60"', 'Automatic Red Clothes Brush',
             'Anti Static Spray', 'Iron On Viledon Tape',
```

```
                  'White Rayon Elastic 3M', 'Hook & Loop Dots Self Stick',
                  'Party Feet Ball Of Foot Gel Cushions',
                  'Compact Lint Roller 2 Refills', 'Travel Sewing Kit',
                  'Black Laces 69Cm'], dtype=object)
```

[24]:
```python
shoe_subcategory_map = {
    "lace": "Laces",
    "laces": "Laces",
    "insole": "Insoles",
    "scissors": "Scissors",
    "needle": "Sewing Accessories",
    "thread": "Sewing Accessories",
    "sewing": "Sewing Accessories",
    "pins": "Pins",
    "tape": "Tapes",
    "roller": "Lint Rollers",
    "shaver": "Lint Rollers",
    "lint": "Lint Rollers",
    "brush": "Brushes",
    "elastic": "Elastic",
    "hook & loop": "Fasteners",
    "arch support": "Footcare",
    "heel": "Footcare",
    "cushion": "Footcare",
    "sockless": "Footcare",
    "spray": "Sprays",
    "balls": "Foot Fresheners",
}

def get_shoe_subcat(item):
    item_lower = item.lower()
    for keyword, subcat in shoe_subcategory_map.items():
        if keyword in item_lower:
            return subcat
    return "Other"

# Apply only to filtered rows
mask = (
    (data['item_size'].isna()) &
    (data['unit_weight'].isna()) &
    (data['category'].str.lower() == 'shoe care')
)

data.loc[mask, 'subcat'] = data.loc[mask, 'item_name'].apply(get_shoe_subcat)
```

[25]:
```python
data.loc[
    (data['item_size'].isna()) &
```

```
    (data['unit_weight'].isna()) &
    (data['category'].str.lower() == 'shoe care')
]
```

[25]:   product_code   category   item_name \
      3594   9382358 SHOE CARE   Lint Roller
      3968   3973792 SHOE CARE   Bamboo Handle Lint Roller
      4020   6524514 SHOE CARE   Oval Lace White 45
      4259   4309135 SHOE CARE   Flat Lace Black 45 inch
      4328   4314769 SHOE CARE   Flat Lace White Kids 27 inch
      ...   .
      71961   6417605 SHOE CARE   Performance Lace 38 Solid Black
      72069   4309135 SHOE CARE   Flat Lace Black 45 inch
      72157   3973894 SHOE CARE   Threaded Needle Kit
      72328   4309180 SHOE CARE   Flat Lace White 45 inch
      72341   3973919 SHOE CARE   Polyester Thread Black & White

                          brand item_price item_size unit_price unit_weight \
      3594   Lint Roller   0.0   NaN   NaN   NaN
      3968   Lint Roller   0.0   NaN   NaN   NaN
      4020   Unknown   0.0   NaN   NaN   NaN
      4259   Unknown   0.0   NaN   NaN   NaN
      4328   Unknown   0.0   NaN   NaN   NaN
      ...   .
      71961 Performance Lace 38   0.0   NaN   NaN   NaN
      72069   Unknown   0.0   NaN   NaN   NaN
      72157 Threaded Needle Kit   0.0   NaN   NaN   NaN
      72328   Unknown   0.0   NaN   NaN   NaN
      72341   Unknown   0.0   NaN   NaN   NaN

             similarity_score   subcat unit_weight_new unit_weight_type \
      3594   1.0   Lint Rollers   NaN   NaN
      3968   1.0   Lint Rollers   NaN   NaN
      4020   1.0   Laces   NaN   NaN
      4259   1.0   Laces   NaN   NaN
      4328   1.0   Laces   NaN   NaN
      ...   .
      71961   1.0   Laces   NaN   NaN
      72069   1.0   Laces   NaN   NaN
      72157   1.0 Sewing Accessories   NaN   NaN
      72328   1.0   Laces   NaN   NaN
      72341   1.0 Sewing Accessories   NaN   NaN

             item_weight_new item_weight_type std_item_weight std_item_unit
      3594   NaN   NaN   NaN   NaN
      3968   NaN   NaN   NaN   NaN
      4020   NaN   NaN   NaN   NaN
```

```
4259            NaN          NaN          NaN          NaN
4328            NaN          NaN          NaN          NaN
 …               …            …            …            …
71961           NaN          NaN          NaN          NaN
72069           NaN          NaN          NaN          NaN
72157           NaN          NaN          NaN          NaN
72328           NaN          NaN          NaN          NaN
72341           NaN          NaN          NaN          NaN

[92 rows x 16 columns]
```

Family Socks:

```python
[26]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'family socks')
      ]['item_name'].unique()
```

```
[26]: array(['Kids Logo QTR Crew Size 9-12', 'Mens Logo Low Cut 6-10',
             'Womens Socks Quarter Logo Size 8+', 'Kids Logo Low Cut 13-3',
             'Kids Logo Light Quarter Crew Size 13 - 3',
             'Kids Logo Low Cut  Socks Size 9-12',
             'Hl1032 Ladies Cushion Foot Socks Black Size 5-8',
             'Mens X-Temp Quarter Crew 6+',
             'Mens Logo Light No Show Size 6+ & 11+',
             'Mens 1/4 Crew Socks Size 11-14', 'Womens Logo Low Cut Size 3-8',
             'Mens Work Socks 11-13 Black',
             'Hl1032 Ladies Cushion Foot Socks Black Size 9-11',
             'Mens Cushion Foot Crew Size 7-10',
             'Womens Logo Light No Show Size 3 - 8',
             'Womens Logo Light No Show Size 8 - 11', 'Sport Crew Sock 3+ & 8+',
             'Ladies Marshmallow Anklet 5-10',
             'MENS QUARTER CREW LOGO SIZE 11+', 'Logo Crew Sock 11+',
             'Mens Logo Low Cut 11-14',
             'Kids Logo Light Quarter Crew Socks Size 13-3',
             'Kids Qtr Crew Sport Socks White 13-3',
             'Womens Logo No Show Size 3-8',
             'Kids Logo Quarter Crew Socks Size Size 2-8',
             'Mens Sports Quarter Crew Socks', 'Sock Crew 13-3 / 3-8',
             'Low Cut Sport Socks White Sizes 5-8 & 9-12',
             'Mens Cushion Foot Crew Size 11-13',
             'Womens Fluffy Qtr Crew Home Socks Size 5-10',
             'Kids Low Cut Sport Socks White Size 13-3',
             'Mens Work Socks 7-10 Black', 'Womens Sport Tech No Show 8+',
             'Mens Original Wool Crew 6-10',
             'Kids Logo Light Quarter Crew Size 2 - 8',
             'Kids Non Slip Fluffy Home Socks 13-3',
```

```
                'Kids Qtr Crew Sport Socks White 5-8',
                'Womens Logo Low Cut Size 8-11',
                'Kids Qtr Crew Sport Socks White 9-12',
                'Kids Low Cut Sports Socks White Size 9 -12'], dtype=object)
```

```python
[27]: subcategory_map_family_socks = {
          "crew": "Crew Socks",
          "quarter": "Crew Socks",
          "qtr": "Crew Socks",
          "low cut": "Low Cut Socks",
          "no show": "No Show Socks",
          "anklet": "Anklet Socks",
          "sport": "Sport Socks",
          "work": "Work Socks",
          "wool": "Wool Socks",
          "fluffy": "Fluffy/Home Socks",
          "home": "Fluffy/Home Socks",
          "non slip": "Fluffy/Home Socks",
          "cushion": "Cushion Socks",
      }

      def get_socks_subcat(item):
          item_lower = item.lower()
          for keyword, subcat in subcategory_map_family_socks.items():
              if keyword in item_lower:
                  return subcat
          return "Other"

      # Apply only to filtered rows
      mask = (
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'family socks')
      )

      data.loc[mask, 'subcat'] = data.loc[mask, 'item_name'].apply(get_socks_subcat)
```

```python
[28]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'family socks')
      ]
```

```
[28]:       product_code      category                              item_name  \
      3906       6845085  FAMILY SOCKS              Kids Logo QTR Crew Size 9-12
      3915       7665420  FAMILY SOCKS                    Mens Logo Low Cut 6-10
      3947       6847080  FAMILY SOCKS         Womens Socks Quarter Logo Size 8+
```

197

```
3975          7665500  FAMILY SOCKS                      Kids Logo Low Cut 13-3
4148          3629753  FAMILY SOCKS    Kids Logo Light Quarter Crew Size 13 - 3
...              ...         ...                                              ...
72343         3629753  FAMILY SOCKS    Kids Logo Light Quarter Crew Size 13 - 3
72346         6845085  FAMILY SOCKS             Kids Logo QTR Crew Size 9-12
72347         7016053  FAMILY SOCKS  Low Cut Sport Socks White Sizes 5-8 & 9-12
72348         7172616  FAMILY SOCKS    Kids Low Cut Sport Socks White Size 13-3
72349         7665500  FAMILY SOCKS                      Kids Logo Low Cut 13-3

                     brand  item_price item_size  unit_price unit_weight  \
3906               Unknown        14.0       NaN         NaN         NaN
3915         Mens Logo Low        16.0       NaN         NaN         NaN
3947  Womens Socks Quarter        16.0       NaN         NaN         NaN
3975               Unknown        14.0       NaN         NaN         NaN
4148               Unknown        14.0       NaN         NaN         NaN
...                    ...         ...       ...         ...         ...
72343              Unknown        14.0       NaN         NaN         NaN
72346              Unknown        14.0       NaN         NaN         NaN
72347              Unknown         0.0       NaN         NaN         NaN
72348              Unknown         0.0       NaN         NaN         NaN
72349              Unknown        14.0       NaN         NaN         NaN

      similarity_score         subcat  unit_weight_new unit_weight_type  \
3906          0.156089     Crew Socks              NaN              NaN
3915          0.146538  Low Cut Socks              NaN              NaN
3947          0.146538     Crew Socks              NaN              NaN
3975          0.156089  Low Cut Socks              NaN              NaN
4148          0.156089     Crew Socks              NaN              NaN
...                ...            ...              ...              ...
72343         0.156089     Crew Socks              NaN              NaN
72346         0.156089     Crew Socks              NaN              NaN
72347         0.089479  Low Cut Socks              NaN              NaN
72348         0.089479  Low Cut Socks              NaN              NaN
72349         0.156089  Low Cut Socks              NaN              NaN

      item_weight_new item_weight_type  std_item_weight std_item_unit
3906              NaN              NaN              NaN           NaN
3915              NaN              NaN              NaN           NaN
3947              NaN              NaN              NaN           NaN
3975              NaN              NaN              NaN           NaN
4148              NaN              NaN              NaN           NaN
...               ...              ...              ...           ...
72343             NaN              NaN              NaN           NaN
72346             NaN              NaN              NaN           NaN
72347             NaN              NaN              NaN           NaN
72348             NaN              NaN              NaN           NaN
72349             NaN              NaN              NaN           NaN
```

```
[91 rows x 16 columns]
```

Hardware:

```
[29]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'hardware')
      ]['item_name'].unique()
```

```
[29]: array(['Combination Word 40mm', 'Diamond Padlock Long Shackle 40mm',
             'Tripple Pack', 'No More Mess Twin Pack', "Fix 'N' Go Supa Glue",
             'Power Grip', 'Supa Glue Brush', 'Shoe Repair',
             'Extreme Mounting Tape', 'Multi Grip', 'Storm',
             'Bits Set 21 piece', 'PVC Electrical Tape', 'Tape Measure',
             'Wrench 8 Inch', 'Diamond Brass Padlock 30mm', 'Pliers',
             'Epoxy 5min', 'Control Bottle', 'Tube', 'Claw Hammer',
             'Rodara Screwdriver Set', 'Rope Cord 10 Metre', 'Sink Plugs 50mm',
             'Cable Ties', 'Utility Scraper', 'Liquid Nails',
             'Clear Mounting Tape', 'Sink Plug 38mm', '17 Piece Tool Kit',
             'Brass Combination 30mm', 'Wall Repair Kit', 'Spakfilla Rapid',
             'Paint Brushes', 'P2 Particulate Respirator', 'Grip',
             'No More Gaps', '6 In 1 Rachet Screwdriver',
             'Power Grip Plastics Glue', 'Diamond Brass Padlock 20mm',
             'Silicone Wet Area White', 'Felt Pads', 'Plunger'], dtype=object)
```

```
[30]: subcategory_map_hardware = {
          "padlock": "Locks & Security",
          "combination": "Locks & Security",
          "glue": "Adhesives",
          "epoxy": "Adhesives",
          "liquid nails": "Adhesives",
          "supa glue": "Adhesives",
          "tape": "Tapes & Mounting",
          "mounting tape": "Tapes & Mounting",
          "electrical tape": "Tapes & Mounting",
          "wrench": "Hand Tools",
          "pliers": "Hand Tools",
          "screwdriver": "Hand Tools",
          "hammer": "Hand Tools",
          "tool kit": "Hand Tools",
          "spakfilla": "Fillers & Sealants",
          "no more gaps": "Fillers & Sealants",
          "wall repair": "Fillers & Sealants",
          "sink plug": "Plumbing",
          "plunger": "Plumbing",
          "rope": "Hardware Accessories",
```

```
    "cord": "Hardware Accessories",
    "bits set": "Hand Tools",
    "paint brushes": "Painting Tools",
    "control bottle": "Hardware Accessories",
    "tube": "Hardware Accessories",
    "grip": "Hardware Accessories",
    "felt pads": "Hardware Accessories",
    "respirator": "Safety Equipment"
}
def get_hardware_subcat(item):
    item_lower = item.lower()
    for keyword, subcat in subcategory_map_hardware.items():
        if keyword in item_lower:
            return subcat
    return "Other"

# Apply only to filtered rows
mask = (
    (data['item_size'].isna()) &
    (data['unit_weight'].isna()) &
    (data['category'].str.lower() == 'hardware')
)

data.loc[mask, 'subcat'] = data.loc[mask, 'item_name'].
  ↪apply(get_hardware_subcat)
```

```
[31]: data.loc[
          (data['item_size'].isna()) &
          (data['unit_weight'].isna()) &
          (data['category'].str.lower() == 'hardware')
      ]
```

[31]:   product_code category   item_name \
      3443   4868137 HARDWARE   Combination Word 40mm
      3521   4868046 HARDWARE Diamond Padlock Long Shackle 40mm
      15214   4868137 HARDWARE   Combination Word 40mm
      15361   4868046 HARDWARE Diamond Padlock Long Shackle 40mm
      38982   3796267 HARDWARE   Tripple Pack
      ...   .
      71747   7026410 HARDWARE   Claw Hammer
      71821   6943539 HARDWARE   PVC Electrical Tape
      72007   6943517 HARDWARE   Wrench 8 Inch
      72013   7024302 HARDWARE   Paint Brushes
      72318   7021610 HARDWARE   Rope Cord 10 Metre

                            brand item_price item_size unit_price unit_weight \
      3443   Combination Word 40mm   0.0   NaN   NaN   NaN

```
3521   Diamond Padlock Long      0.0    NaN    NaN    NaN
15214  Combination Word 40mm     0.0    NaN    NaN    NaN
15361  Diamond Padlock Long      0.0    NaN    NaN    NaN
38982            Tripple Pack    0.0    NaN    NaN    NaN
...                       ...    ...    ...    ...    ...
71747            Claw Hammer     0.0    NaN    NaN    NaN
71821                Unknown     0.0    NaN    NaN    NaN
72007          Wrench 8 Inch     0.0    NaN    NaN    NaN
72013          Paint Brushes     0.0    NaN    NaN    NaN
72318            Rope Cord 10    0.0    NaN    NaN    NaN


       similarity_score               subcat  unit_weight_new  \
3443           0.311365      Locks & Security              NaN
3521           0.311365      Locks & Security              NaN
15214          0.311365      Locks & Security              NaN
15361          0.311365      Locks & Security              NaN
38982          0.311365                 Other              NaN
...                 ...                   ...              ...
71747          0.311365            Hand Tools              NaN
71821          0.311365      Tapes & Mounting              NaN
72007          0.311365            Hand Tools              NaN
72013          0.311365         Painting Tools             NaN
72318          0.311365  Hardware Accessories              NaN


       unit_weight_type  item_weight_new  item_weight_type  std_item_weight  \
3443                NaN              NaN               NaN              NaN
3521                NaN              NaN               NaN              NaN
15214               NaN              NaN               NaN              NaN
15361               NaN              NaN               NaN              NaN
38982               NaN              NaN               NaN              NaN
...                 ...              ...               ...              ...
71747               NaN              NaN               NaN              NaN
71821               NaN              NaN               NaN              NaN
72007               NaN              NaN               NaN              NaN
72013               NaN              NaN               NaN              NaN
72318               NaN              NaN               NaN              NaN


       std_item_unit
3443             NaN
3521             NaN
15214            NaN
15361            NaN
38982            NaN
...              ...
71747            NaN
71821            NaN
72007            NaN
```

```
72013          NaN
72318          NaN

[72 rows x 16 columns]
```

1.12 Drop unnecessary NaNs

- Items missing item_size but having unit_size
  - Can often compute or approximate item_size from unit_size if you know the quantity or packaging.
- Items missing both item_size and unit_weight
  - These are trickier because we don't know the scale of the product. They could still be used if we have a price per unit, but for models that need quantities to normalize prices, this could be problematic.
- Items missing item_price but having unit_price
  - We can compute item_price = unit_price × item_size (or unit_weight) if item_size or unit_weight is known.
  - This means we don't necessarily need item_price upfront.
- Items missing both item_price and unit_price
  - These are truly unusable for price-based substitution models, because we cannot infer the price from anything in the dataset.

Recommendation:

- Keep items where at least unit_price is available. We can generate item_price from it.
- Drop items only when both item_price and unit_price are missing, because we cannot recover the price.
- Consider filling missing item_size from unit_size where possible, instead of dropping the row.

[34]: `data.shape`

[34]: (77833, 16)

[35]: `data.isna().sum()`

[35]:
```
product_code          0
category              0
item_name             0
brand                 0
item_price            0
item_size         66683
unit_price         2237
unit_weight        2238
similarity_score      0
subcat                0
unit_weight_new    2244
unit_weight_type   2244
item_weight_new    2142
item_weight_type   2142
```

```
std_item_weight      66683
std_item_unit        66683
dtype: int64
```

[38]:
```python
# Drop rows where item_price == 0.0 and unit_price is NaN
data = data[~((data['item_price'] == 0.0) & (data['unit_price'].isna()))]
data = data.reset_index(drop=True)
```

[41]:
```python
# Fill item_weight with unit_weight if item_weight is NaN
data['item_weight_new'] = data['item_weight_new'].
 ↪fillna(data['unit_weight_new'])
data['item_weight_type'] = data['item_weight_type'].
 ↪fillna(data['unit_weight_type'])
data['std_item_weight'] = data['std_item_weight'].
 ↪fillna(data['unit_weight_new'])
data['std_item_unit'] = data['std_item_unit'].fillna(data['unit_weight_type'])
```

1.13 Save the data

[44]:
```python
data.to_csv('preprocessed_data.csv',index=False)
```

[ ]: