Nama: Dina Cahyo Wibowo NIM: G.231.22.0004 Matkul: Data Mining

Kode dia atas ketika dijalankan akan mengimport beberapa library dan juga mengimport file excel dari link yang sudah ada di atas

```
[3] 1 df['Description'] = df['Description'].str.strip()
2 df.dropna(axis=0, subset=['InvoiceNo'], inplace=True)
3 df['InvoiceNo'] = df['InvoiceNo'].astype('str')
4 df = df[~df['InvoiceNo'].str.contains('C')]

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will n and should_run_async(code)
```

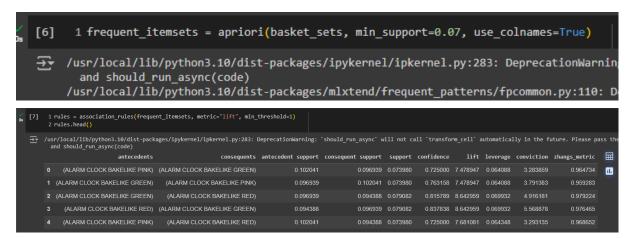
kode ini berguna untuk membersihkan data mulai dari Menghapus spasi di awal dan akhir teks pada kolom Description dan menghapus baris yang memiliki nilai NaN pada kolom Invoicedo

sedangkan untuk kode ini berfungsi untuk menyiapkan data mulai dari Menyaring transaksi untuk negara Prancis. Mengelompokkan data berdasarkan nomor faktur dan deskripsi produk, kemudian menjumlahkan kuantitas produk. Mengubah bentuk data dengan unstack sehingga setiap produk menjadi kolom. Mengatur ulang indeks dan mengisi nilai NaN dengan 0.

```
[5] 1 def encode_units(x):
    2    if x <= 0:
    3        return 0
    4    if x >= 1:
    5        return 1
    6
    7 basket_sets = basket.applymap(encode_units)
    8 basket_sets.drop('POSTAGE', inplace=True, axis=1)

// usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning and should_run_async(code)
```

Kode ini berfungsi untuk Mendefinisikan fungsi encode_units untuk mengubah kuantitas produk menjadi 0 atau 1.Menerapkan fungsi encode_units ke semua data pada basket. Menghapus kolom 'POSTAGE' dari basket_sets



Kedua kolom code di atas berfungsi untk Mengimpor fungsi apriori dan association_rules dari pustaka mlxtend. Menemukan kumpulan item yang sering muncul dengan apriori dan minimum support 0.07.Menghasilkan aturan asosiasi dengan association_rules berdasarkan metrik "lift" dan ambang batas minimal 1. Dan juga menampilkan lima aturan asosiasi pertama.



Sedangkan kdoe di atas ini berfungsi untuk Menyaring aturan asosiasi yang memiliki nilai lift minimal 6 dan confidence minimal 0.8. Serta menampilkan aturan yang memenuhi kriteria tersebut

```
[9] 1 basket['ALARM CLOCK BAKELIKE GREEN'].sum()
2
3 340.0
4
5 basket['ALARM CLOCK BAKELIKE RED'].sum()
6
7 316.0

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: Degrand should_run_async(code)
```

Mengakses kolom ALARM CLOCK BAKELIKE GREEN dan ALARM CLOCK BAKELIKE GREEN dalam dataframe basket. Serta Menghitung total kuantitas untuk produk dengan menggunakan fungsi sum(). dan menunjukkan total unit yang terjual untuk produk tersebut. Hasilnya adalah 340.0. dan 316.0

Sedangkan untuk kode di atas berguna untuk menspesifikkan transaksi yang hanya berasal dari Jerman. Mengelompokkan data berdasarkan InvoiceNo dan Description, lalu menjumlahkan kuantitas produk. Mengubah bentuk data dengan unstack sehingga setiap produk menjadi kolom. Serta mengatur ulang indeks dan mengisi nilai NaN dengan 0. dan juga menerapkan fungsi encode_units untuk mengubah kuantitas produk menjadi 0 atau 1 (biner). Sedangkan sisanya kurang lebih sama seperti kode kode sebelumnyha