Penjelasan Basket

Masukan library Pandas(pd), Apriori, dan Asosiasi

Membaca data excel melalui URL(df)

Tampilkan sebagian dari data tersebut

Menghapus spasi di kolom Description

Menghapus baris yang bernilai Nan di kolom InvoiceNo

Mengubah nilai di kolom InvoiceNo menjadi string

Menghapus nilai InvoiceNo yang mengandung 'C'

Memfilter data untuk negara France pada kolom Country

* Mengelompokan data berdasarkan InvoiceNo dan Description
* Membentuk pivot table, mengubah data sehingga setiap 'Description' menjadi kolom, dengan 'InvoiceNo' sebagai indeks, dan nilai 'Quantity' sebagai isi tabel
* Mengatur ulang indeks sehingga 'InvoiceNo' kembali menjadi kolom biasa, bukan indeks
* Mengganti nilai NaN dengan 0, karena jika tidak ada nilai, artinya item tersebut tidak dibeli dalam transaksi tersebut
* Mengatur 'InvoiceNo' sebagai indeks dari DataFrame

Dengan fungsi ‘encode\_units’ dapat mengubah nilai kuantitas menjadi biner, jika kurang/sama dari 0 akan bernilai salah dan sebaliknya.

Menerapkan nilai ‘encode\_units’ ke setiap elemen.

Menghapus kolom POSTAGE

Menggunakan fungsi apriori untuk menemukan itemset yang sering muncul dalam dataset transaksional.

Itemset harus muncul di setidaknya 7%

Pastikan nama kolom digunakan sebagai label item, bukan indeks numerik

Terapkan fungsi asosiasi

Menghasilkan aturan asosiasi menggunakan fungsi association\_rules

Menentukan metrik yang digunakan untuk mengevaluasi aturan asosiasi(lift)

Menentukan nilai ambang batas minimum untuk metrik yang dipilih

Menggunakan fungsi association\_rules dengan parameter metric="lift" dan min\_threshold=1

Memilih aturan dengan lift setidaknya 6 dan confidence setidaknya 0.8

Melakukan agregasi pada data untuk menemukan jumlah total dari kuantitas item ALARM CLOCK BAKELIKE GREEN & ALARM CLOCK BAKELIKE RED

340.0 dan 316.0

Memfilter data untuk transaksi dari Jerman

Mengelompokkan data berdasarkan 'InvoiceNo' dan 'Description', dan menjumlahkan kuantitas

Membentuk pivot table dengan 'InvoiceNo' sebagai indeks dan 'Description' sebagai kolom

Menggunakan fungsi encode\_units untuk mengubah nilai kuantitas menjadi biner

Menghapus kolom 'POSTAGE' karena tidak relevan dalam analisis

Menggunakan algoritma Apriori untuk menemukan itemset frekuen dengan support minimal 0.05

Menggunakan fungsi association\_rules untuk menghasilkan aturan asosiasi berdasarkan metrik lift

Memfilter aturan berdasarkan lift dan confidence sesuai dengan kriteria yang ditentukan