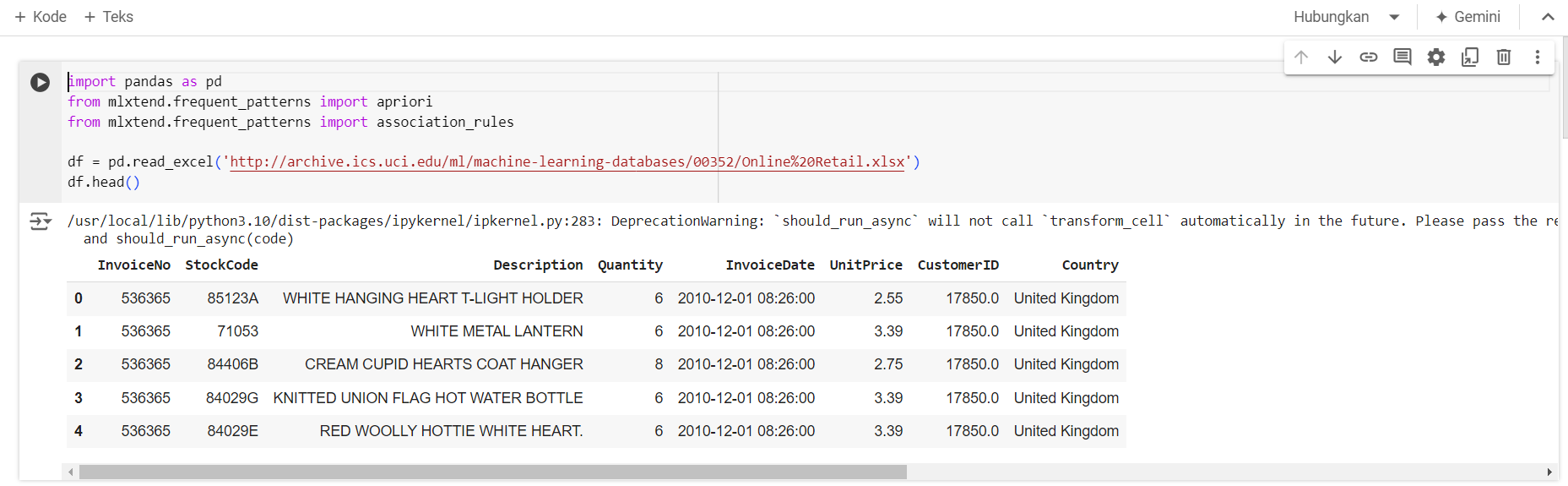
Nama : Jovita Wayan R

Nim : G.231.22.0103

Kelas : Teknik Informatika B (Sore)

Matkul : Data Mining

PRAKTIKUM 4 MARKET BASKET ANALISIS



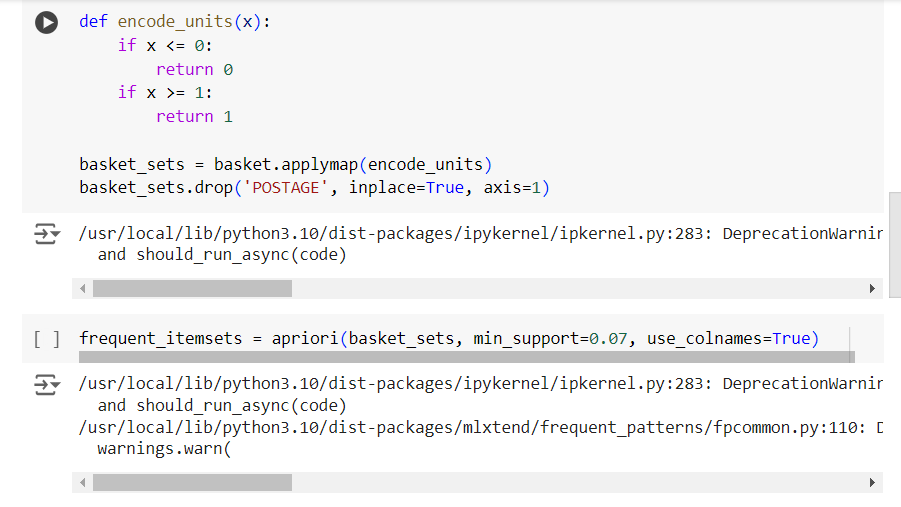
Penjelasan :

Memasukkan data dengan menggunakan link yang akan langsung terbaca tanpa memasukkan file terlebih dahulu, sehingga ada sedikit pembersihan, yang perlu kita lakukan. Pertama, beberapa deskripsi memiliki spasi yang perlu dihilangkan. Dan juga akan menghapus baris yang tidak memiliki nomor faktur dan menghapus transaksi kredit (yang memiliki nomor faktur berisi C).



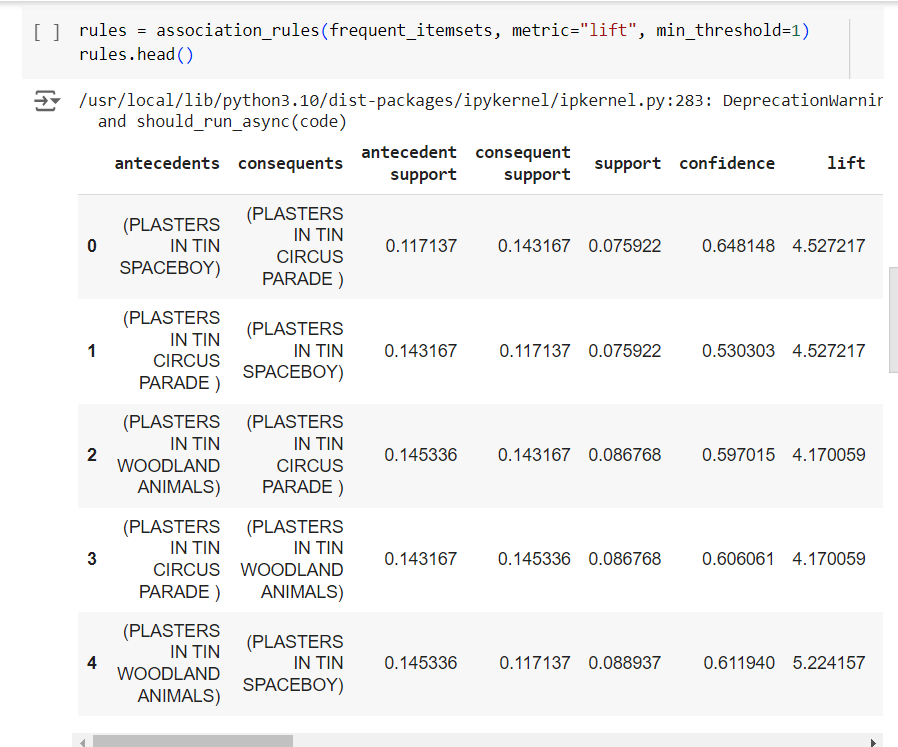
Penjelasan :

Setelah pembersihan, sehingga perlu menggabungkan item menjadi 1 transaksi per baris dengan setiap produk 1 hot encoded. Demi menjaga kumpulan data tetap kecil, cara yang dilakukan hanya melihat penjualan untuk Perancis. Namun pada kode tambahan di bawah ini, saya akan membandingkan hasil tersebut dengan penjualan dari Jerman. Perbandingan negara lebih lanjut akan menarik untuk diselidiki.



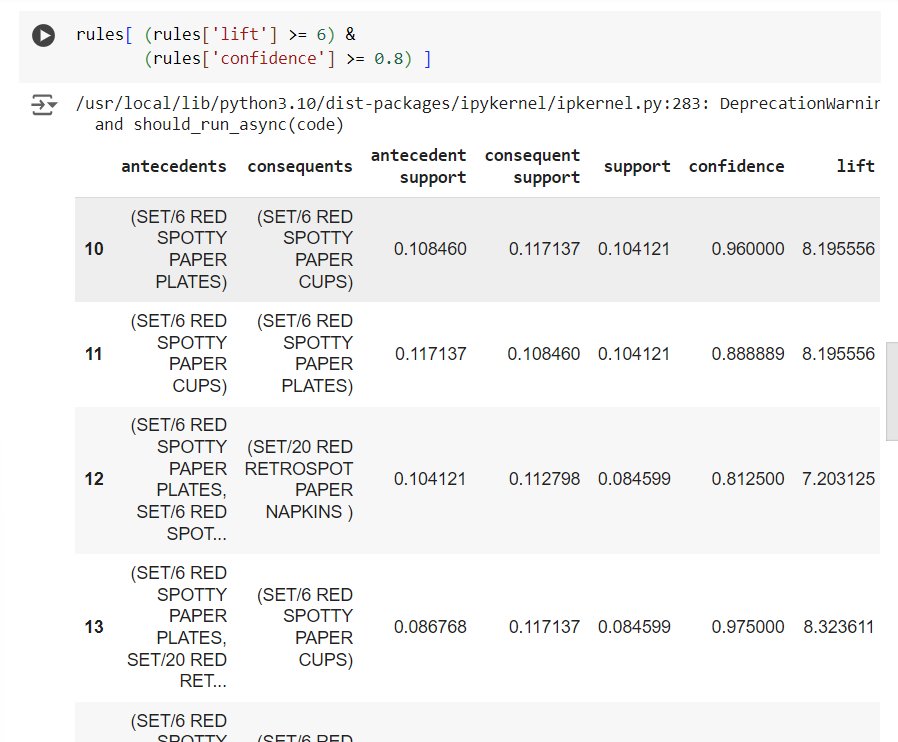
Penjelasan :

Ada banyak angka nol dalam data tetapi kita juga perlu memastikan setiap nilai positif diubah menjadi 1 dan nilai apa pun yang kurang dari 0 disetel ke 0. Langkah ini akan menyelesaikan pengkodean data yang paling menarik dan menghapus kolom ongkos kirim (karena biaya tersebut bukanlah biaya yang ingin kami jelajahi). Dan rumus yang kedua yaitu setelah data terstruktur dengan benar, kita dapat menghasilkan kumpulan item frekuensi yang memiliki dukungan minimal 7 % (angka ini dipilih sehingga saya bisa mendapatkan cukup banyak contoh berguna).



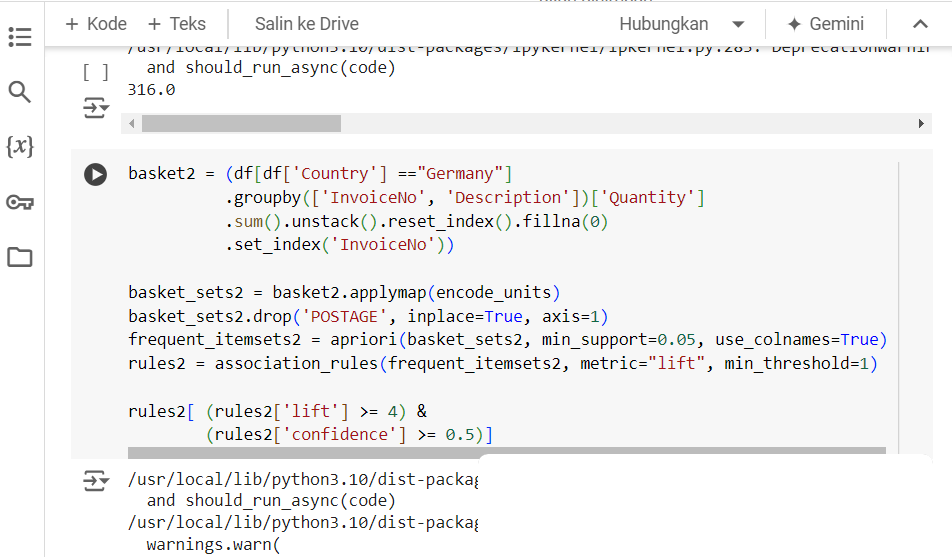
Penjelasan :

Kode di atas adalah langkah terakhir untuk menghasilkan aturan dengan dukungan, keyakinan, dan dorongan yang sesuai. Bangun item yang sering digunakan apriori lalu buat aturannya dengan association\_rules.



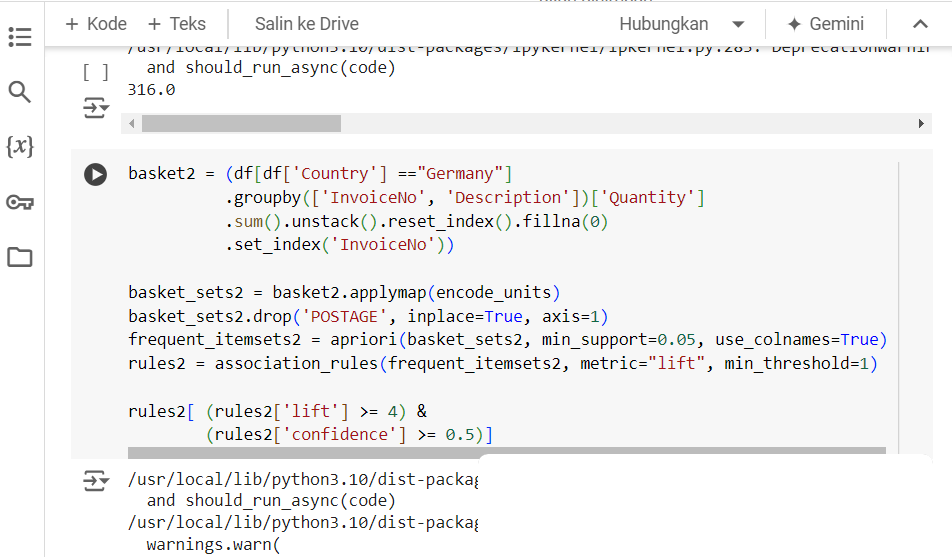
Penjelasan :

Dengan menggunakan kode di atas kita dapat memfilter kerangka data menggunakan kode pandas standar. Dalam hal ini carilah peningkatan yang besar (6) dan keyakinan yang tinggi (0,8).



Penjelasan :

Pada kode di atas, dapat melihat seberapa besar peluang yang ada untuk menggunakan popularitas satu produk untuk mendorong penjualan produk lain. Misalnya, kita dapat melihat bahwa kita menjual 340 jam alarm hijau tetapi hanya 316 jam alarm merah, jadi mungkin kita dapat mendorong lebih banyak penjualan jam alarm merah melalui rekomendasi



Penjelasan :

Berikut dapat dilihat hasil dari MLxtend yang menggunakan analisis dasar excel sehingga tidak perlu repot memasukkan file ke dalam colab. Jika Anda tidak memiliki akses ke MLxtend dan analisis asosiasi ini, akan sangat sulit menemukan pola ini menggunakan analisis dasar Excel. Dengan python dan MLxtend, proses analisisnya relatif mudah dan karena Anda menggunakan python, Anda memiliki akses ke semua teknik visualisasi tambahan dan alat analisis data di ekosistem pyrhon.