**ANALISIS MARKET BASKET UNTUK MENENTUKAN *ASSOCIATION RULES* MENGGUNAKAN ALGORITMA APRIORI**

**Fadila Shely Amalia**

*Universitas Gadjah Mada*

*Email: fadilashelyamalia@mail.ugm.ac.id*

**Abstrak**

Peranan teknologi informasi saat ini sangatlah penting terutama pada bidang bisnis. Pemanfaatan teknologi informasi dalam dunia bisnis telah menjadi sebuah kebutuhan yang mutlak untuk dimiliki oleh setiap perusahaan agar dapat bersaing dengan kompetitor lain. Lypetshop merupakan salah satu toko pada marketplace shopee yang menjual makanan dan perlengkapan hewan peliharaan khususnya kucing. Banyaknya data transaksi dan belum adanya analisis pola penjualan untuk menentukan produk mana yang paling sering dibeli secara bersamaan dibutuhkan sebuah metode atau teknik yang dapat merubah banyaknya data agar dapat dimanfaatkan menjadi sebuah informasi berharga atau pengetahuan yang bermanfaat. PAda penelitian ini menggunakan algoritma apriori dalam penyelesaiannya. Algoritma Apriori merupakan salah satu jenis algoritma yang ada pada data mining yang memakai aturan asosiasi. Kegunaan algoritma apriori itu sendiri untuk mencari frekuensi dan keterkaitan itemset dengan itemset lainnya dari kumpulan data yang diolah yang dimana telah di tentukan syarat *Minimum*  nilai *support* dan syarat *Minimum*  nilai *confidence* terlebih dahulu. Berdasarkan hasil analisis yang telah dilakukkan maka dalam penelitian ini dapat ditarik kesimpulan yaitu dengan menggunakan algoritma apriori membantu toko mengetahui pola belanja konsumen dan item yang sering dibeli oleh konsumen. Dengan minimal *support* dan *confidance* yang ditentukan menghasilkan lima rules dengan nilai *support* 0.001276 dan nilai *confidence* 1 (100%) pada masing-masing *rules*.

**Kata Kunci:** *Apriori, Association Rules, Confidance, Support*

1. **Pendahuluan**
2. **Latar Belakang**

Peranan teknologi informasi saat ini sangatlah penting terutama pada bidang bisnis. Pemanfaatan teknologi informasi dalam dunia bisnis telah menjadi sebuah kebutuhan yang mutlak untuk dimiliki oleh setiap perusahaan agar dapat bersaing dengan kompetitor lain. Dibutuhkan sebuah metode atau teknik yang dapat merubah tumpukan data tersebut menjadi sebuah informasi berharga atau pengetahuan (*knowledge*) yang bermanfaat untuk mendukung pengambilan keputusan bisnis. Suatu teknologi yang dapat digunakan untuk mewujudkannya adalah data mining. Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan mengunakan teknik atau metode tertentu. Data mining berkembang menjadi alat bantu untuk mencari pola-pola yang berharga dalam suatu basis data yang sangat besar jumlahnya, sehingga tidak memungkinkan dicari secara manual.

Lypetshop merupakan salah satu toko pada marketplace shopee yang menjual makanan dan perlengkapan hewan peliharaan khususnya kucing. Banyaknya data transaksi dan belum adanya analisis pola penjualan untuk menentukan produk mana yang paling sering dibeli secara bersamaan dibutuhkan sebuah metode atau teknik yang dapat merubah banyaknya data agar dapat dimanfaatkan menjadi sebuah informasi berharga atau pengetahuan yang bermanfaat.

1. **Tinjuan Pustaka**
2. ***Data Mining***

Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database dan visualisasi untuk pengenalan permasalahan pengambilan informasi dari database yang besar. Sementara dalam pengertian lain, Data mining merupakan suatu proses pencarian pola dari data-data dengan jumlah yang sangat banyak yang tersimpan dalam suatu tempat penyimpanan dengan menggunakan teknologi pengenal pola, teknik statistik, dan matematika [1].

1. ***Association Rules***

*Association* *Rule* atau Aturan Asosiasi adalah teknik data *mining* untuk menemukan aturan asosiatif atau pola kombinasi dari suatu item. Bila kita mengambil contoh aturan asosiatif dalam suatu transaksi pembelian barang disuatu minimarket adalah kita dapat mengetahui berapa besar kemungkinan seorang konsumen membeli suatu item bersamaan dengan item lainnya [2].

1. **Algoritma Apriori**

Algoritma Apriori merupakan salah satu jenis algoritma yang ada pada data mining yang memakai aturan asosiasi. Kegunaan algoritma apriori itu sendiri untuk mencari frekuensi dan keterkaitan itemset dengan itemset lainnya dari kumpulan data yang diolah yang dimana telah di tentukan syarat *Minimum*  nilai *support* dan syarat *Minimum*  nilai *confidence* terlebih dahulu [3].

1. ***Market Basket Analyst***

*Market Basket Analyst* berhubungan dengan penemuan pola pembelian konsumen dari database transaksiona. MBA berfokus peda penemuan pola pembelian pada ribuan data transaksi. Dalam hal ini aturan asosiasi berperan penting untuk menemukan item set yang sering dan berhubungan antara produk yang berbeda, dimana deklarasi eksplisit dari jenisnya ditetapkan, seperti jika item X dibeli maka item Y juga dibeli [4].

1. **Metodologi Penelitian**
2. **Metode Pengumpulan Data**
3. Pengamatan Langsung

Pengumpulan data yaitu dilakukan pengamatan langsung pada transaksi penjualan yang terjadi di Toko Lypetshop pada *marketplace* Shopee.

1. Dokumentasi

Dokumentasi dilakukan untuk mendapatkan data yang valid dan sesuai untuk kebutuhan penelitian, yaitu dengan mendokumentasikan data transaksi penjualan dari bulan Juli 2022 sampai September 2022.

1. **Desain Penelitian**

Pada penelitian ini peneliti menggunakan *Knowledge Discovery in Databases* (KDD). *Knowledge Discovery in Databases* (KDD) mempunyai lima langkah [5].

1. ***Data collection***

Pada tahap ini peneliti melakukan pengumpulan (*data colllection*) yaitu menyeleksi data berupa data transaksi penjualan di Toko Lypetshop pada *marketplace* Shopee.

1. ***Data Preprocessing***

Tahap ini peneliti melakukan proses pembersihan pada data, pembersihan yang dilakukkan seperti memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data.

1. ***Data Reduction***

Tahap ini peneliti melakukan transformasi yaitu mengubah data transaksi penjualan yang sudah dibersihkan (*cleaning*) yang awal data *microsoft* *excel* diubah dalam format *tabular*.

1. **Data Mining**

Pada tahapan ini data diolah dengan menerapkan algoritma Apriori dengan menggunakan Bahasa pemrogrman *python* dan tools *Google Colaboratory* untuk menemukan hubungan berdasarkan nilai minimun *support* dan *Minimum*  *confidence* dari data yang diolah.

1. **Evaluasi & Analisis Hasil**

Pada tahap ini peneliti menganalisis hasil dari perhitungan metode algoritma Apriori dengan Bahasa pemrograman *python* dan tools *Google Colaboratory*.

1. **Hasil**
2. **Analisis Data**

Data yang ada akan diesktrasi menjadi data yang lebih terstruktur dengan menghilangkan atribut-atribut yang tidak tiperlukan dalam proses analisis pola belanja konsumen. Pada data transaksi asli terdapat 46 atribut dan pada analisis untuk menemukan pola belanja konsumen menggunakan metode apriori ini hanya memerlukan tiga atribut yakni nomor invoice nama dan jumlah barang yang dibeli. Berikut sampel data transaksi yang dapat digunakan untuk analisis ini.

**Tabel 1.** Keterangan Atribut Data Transaksi

|  |  |  |
| --- | --- | --- |
| Nomor *Invoice* | Nama Produk | Qty |
| 220702DQE97W1V | SUSU TOP GROWTH 1 Murah Sachet 30 gr susu anak kucing / susu kucing | 1 |
| 220702DQE97W1V | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220702DQE97W1V | Makanan kucing ORI CAT Murah repack 1kg bentuk ikan | 1 |
| 220701AXHG8DH2 | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220704HCHHD9SR | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220704HCHHD9SR | Bolt kitten murah kemasan repack 1kg | 1 |
| 220705K46E4DQF | Cat Choize kitten tuna&salmon MURAH | 1 |
| 220705K46E4DQF | Cat Choize kitten tuna&salmon MURAH | 1 |
| 220701B8TX70HB | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220702CD4AW53Y | Cat Choize kitten tuna&salmon MURAH | 1 |
| 220702CD4AW53Y | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705KQ39UTMB | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705KQ39UTMB | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705KQ39UTMB | LIFE CAT POUCH murah Baim Wong 85gr Wet Food makanan kucing | 1 |
| 220702BNA1T6UX | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1HK7JHC | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1HK7JHC | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220701BBC9PGXU | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1AXMXD0 | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1AXMXD0 | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1AXMXD0 | LIFE CAT POUCH murah Baim Wong 85gr Wet Food makanan kucing | 1 |

Berdasarkan Tabel diatas maka mendapatkan data yang sudah di sortir berdasarkan quantity barang yang terjual sesuai nama barang sebagai berikut.

**Tabel 2.** Keterangan Atribut Data Persediaan

|  |  |  |
| --- | --- | --- |
| Nomor Invoice | Nama Produk | Qty |
| 220702DQE97W1V | SUSU TOP GROWTH 1 Murah Sachet 30 gr susu anak kucing / susu kucing | 1 |
| NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| Makanan kucing ORI CAT Murah repack 1kg bentuk ikan | 1 |
| 220701AXHG8DH2 | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220704HCHHD9SR | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| Bolt kitten murah kemasan repack 1kg | 1 |
| 220705K46E4DQF | Cat Choize kitten tuna&salmon MURAH | 2 |
| 220701B8TX70HB | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220702CD4AW53Y | Cat Choize kitten tuna&salmon MURAH | 1 |
| NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705KQ39UTMB | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 2 |
| LIFE CAT POUCH murah Baim Wong 85gr Wet Food makanan kucing | 1 |
| 220702BNA1T6UX | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1HK7JHC | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 2 |
| 220701BBC9PGXU | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 1 |
| 220705M1AXMXD0 | NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah | 2 |
| LIFE CAT POUCH murah Baim Wong 85gr Wet Food makanan kucing | 1 |

Maka berdasarkan data yang tertera pada tabel dapat dilihat pada transaksi dengan invoice 220702DQE97W1V terjual dengan tiga kombinasi barang yaitu SUSU TOP GROWTH 1 Murah Sachet 30 gr susu anak kucing / susu kucing, NEW LIFE CAT KALENG Murah Adult Kitten 400gr lifecat wet food makanan kucing kaleng basah, dan Makanan kucing ORI CAT Murah repack 1kg bentuk ikan.

1. ***Data Design***

Setelah melakukan analisis maka dari data yang akan digunakan dirubah dalam bentuk tabular dan merubah nama barang dalam bentuk angka 0 dan angka 1. Dimana jika barang kurang dari sama dengan 0 maka sistem perhitungan keranjang tersebut akan bernilai 0 dan jika nilai lebih dari 1 maka sistem perhitungan kerjang akan memberi output (nilai) 1, sehingga jika sebuah nota membeli barang A sebanyak 10 buah maka hanya akan dihitung 1.

Untuk mempermudah pembuatan tabel maka saya akan menyingkat masing-masing nama barang yang dijadikan contoh. Berikut merupakan contoh data tabular yang siap digunakan untuk langkah selanjutnya dalam menentukan nilai terjualnya barang supaya sistem dapat memahami informasi dataset.

**Tabel 3** Contoh Data Transaksi

|  |  |  |
| --- | --- | --- |
| Nomor Invoice | Nama Produk | Qty |
| 220702DQE97W1V | Susu Kucing | 1 |
| Wet Food 400gr | 1 |
| Ori Cat 1KG | 1 |
| 220701AXHG8DH2 | Wet Food 400gr | 1 |
| 220704HCHHD9SR | Wet Food 400gr | 1 |
| Bolt 1KG | 1 |
| 220705K46E4DQF | Kitten Food | 2 |
| 220701B8TX70HB | Wet Food 400gr | 1 |
| 220702CD4AW53Y | Kitten Food | 1 |
| Wet Food 400gr | 1 |
| 220705KQ39UTMB | Wet Food 400gr | 2 |
| Wet Food 800gr | 1 |
| 220702BNA1T6UX | Wet Food 400gr | 1 |
| 220705M1HK7JHC | Wet Food 400gr | 2 |
| 220701BBC9PGXU | Wet Food 400gr | 1 |
| 220705M1AXMXD0 | Wet Food 400gr | 2 |
| Wet Food 800gr | 1 |

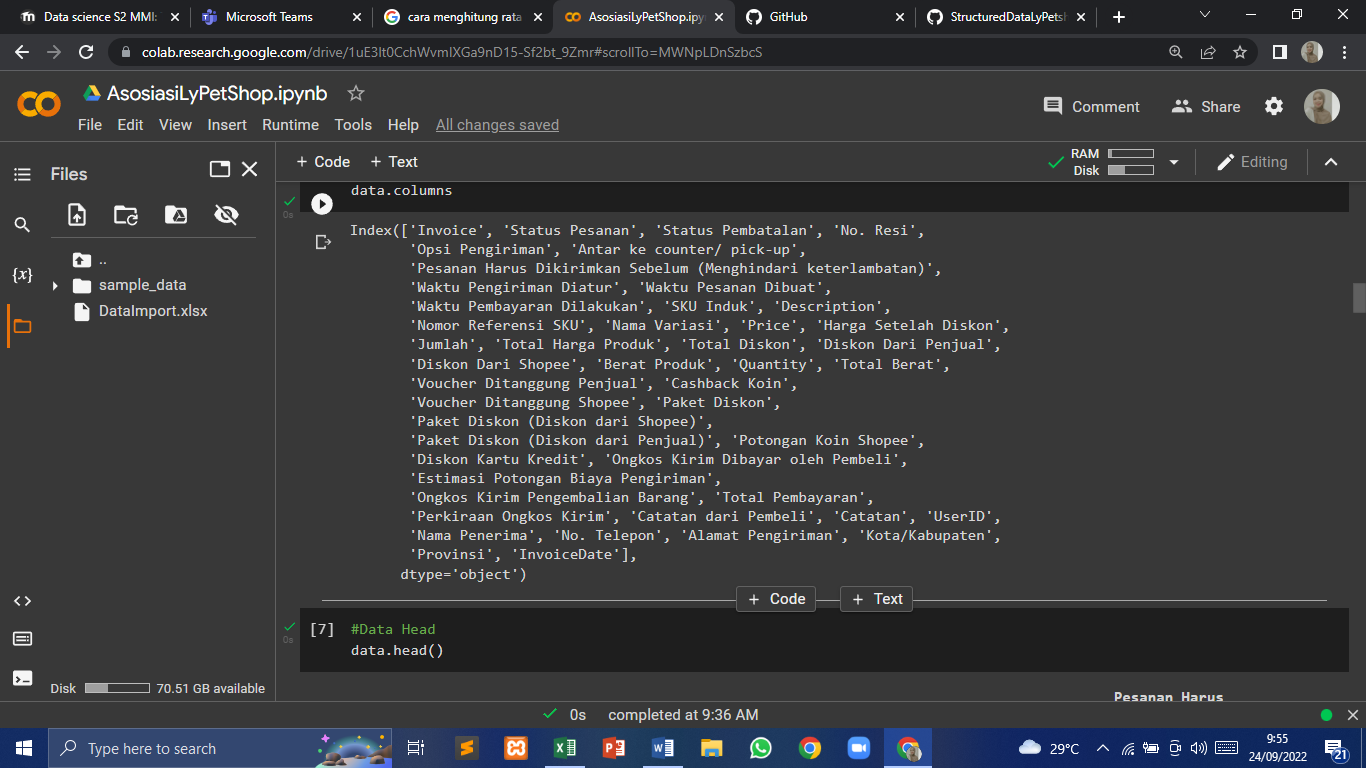
Setelah itu nama barang di rubah posisinya menjadi bentuk baris dan keterangan terjual dengan quantity >= 1 maka ditulis 1 dan tidak terjual adalah 0 terlihat di baris selanjutnya berdasarkan nomor invoice seperti berikut.

**Tabel 4** Contoh Data Transaksi Format *Tabular*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Susu Kucing | Wet Food 400gr | Ori Cat 1KG | Wet Food 400gr | Bolt 1KG | Kitten Food | Kitten Food | Wet Food 800gr |
| T1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| T2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| T3 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| T4 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| T5 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| T6 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| T7 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| T8 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| T9 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| T10 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| T11 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| ∑ | 1 | 1 | 1 | 9 | 1 | 1 | 2 | 2 |

1. **Penyelesaian Masalah dengan Algoritma Apriori dan *Tools Google Colaboratory***

Untuk mempermudah proses selanjutnya maka saya mengubah nama masing-masing field yang akan digunakan. Seperti No. Pesanan dirubah menjadi Invoice, Nama Produk menjadi Description, Harga awal menjadi *Price, Username* (Pembeli) menjadi *UserID* dan Waktu Pesanan Selesai menjadi *InvoiceDate*.



Dikarenakan hanya beberapa kolom saja yang diperlukan untuk proses selanjutnya, disini saya mengambil kolom invoice, description, p*rice, quantity , customerID*, dan *invoice date*. Maka, selain daripada kolom tersebut dihapus menggunakan fungsi drop.

Selanjutnya mengecek jumlah record yang kosong pada masing-masing kolom. Berikut merupakan total record yang kosong pada masing-masing kolom yang sudah final. Dimana dapat dilihat bahwa ada enam kolom yang saya gunakan dan masing-masing sudah bernilai 0. Maka, sudah tidak ada lagi data dengan record yang kosong atau NULL.

Invoice 0

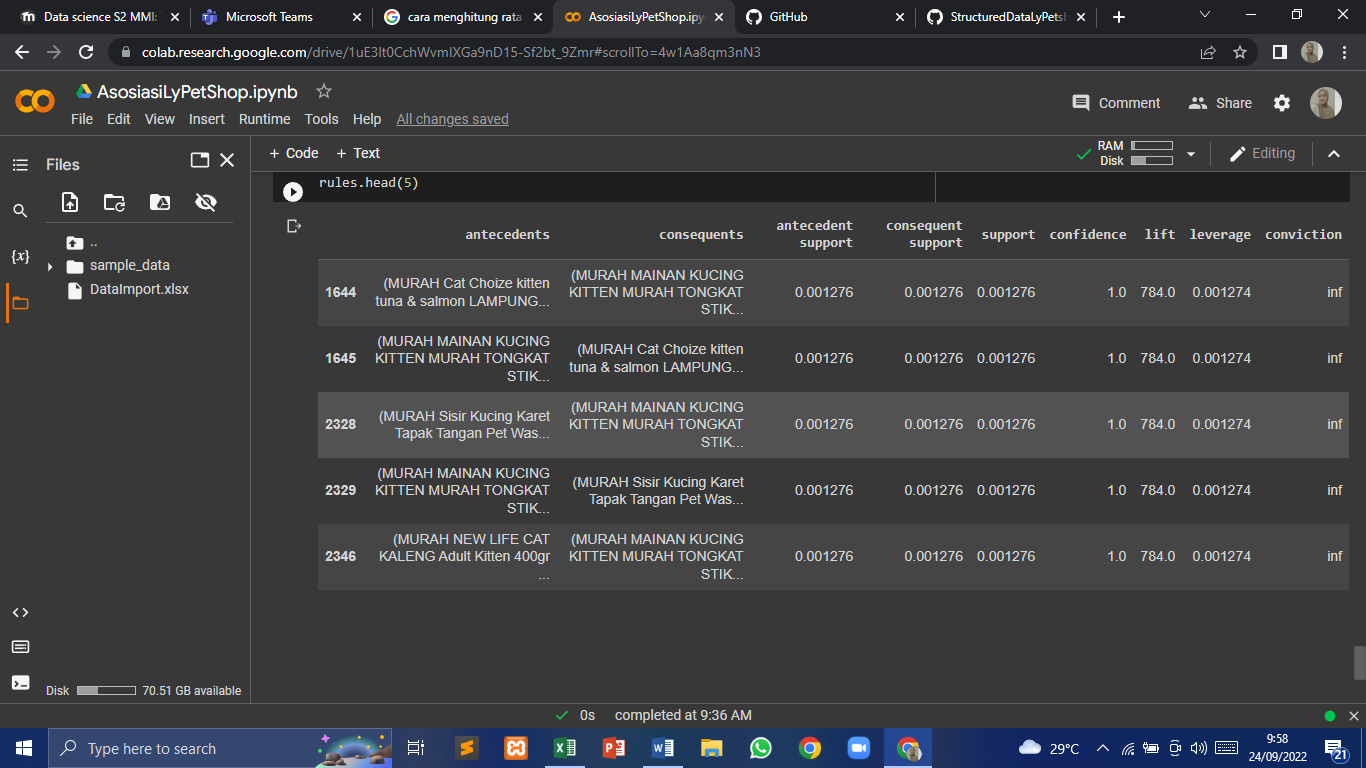
Description 0

Price 0

Quantity 0

CustomerID 0

InvoiceDate 0



Berdasarkan hasil dari penerapan algoritma apriori diatas dapat dilihat bahwa saya menggunakan minimal support 2 0.001 untuk menentukan 5 barang teratas yang paling banyak dibeli secara bersamaan dengan masing masing hasil sebagia berikut :

1) Apabila customer membeli Catchoize kitten maka akan membeli juga tongkat mainan kucing dengan nilai support 0.001276 dan nilai confidence 1 (100%).

2) Apabila customer membeli tongkat mainan kucing maka akan membeli juga cat choize kitten nilai support 0.001276 dan nilai confidence 1 (100%).

3) Apabila customer membeli sisir kucing maka akan membeli juga tongkat mainan kucing dengan nilai support 0.001276 dan nilai confidence 1 (100%).

4) Apabila customer membeli tongkat mainan kucing maka akan membeli juga sisir kucing dengan nilai support 0.001276 dan nilai confidence 1 (100%).

5) Apabila customer membeli Life cat kaleng adult maka akan membeli juga tongkat mainan kucing dengan nilai support 0.001276 dan nilai confidence 1 (100

1. **KESIMPULAN**

Berdasarkan hasil analisis yang telah dilakukkan maka dalam penelitian ini dapat ditarik kesimpulan yaitu dengan menggunakan algoritma apriori membantu toko mengetahui pola belanja konsumen dan item yang sering dibeli oleh konsumen. Dengan minimal support dan confidance yang ditentukan menghasilkan nilai support 0.001276 dan nilai confidence 1 (100%) pada masing-masing rules.

**Daftar Pustaka**

[1] M. M. Rahman, Y. Watanobe, T. Matsumoto, R. U. Kiran, and K. Nakamura, “Educational Data Mining to Support Programming Learning Using Problem-Solving Data,” *IEEE Access*, vol. 10, pp. 26186–26202, 2022, doi: 10.1109/ACCESS.2022.3157288.

[2] A. M. Khedr, Z. Al Aghbari, A. Al Ali, and M. Eljamil, “An efficient association rule mining from distributed medical databases for predicting heart diseases,” *IEEE Access*, vol. 9, pp. 15320–15333, 2021, doi: 10.1109/ACCESS.2021.3052799.

[3] M. Tian *et al.*, “Data Dependence Analysis for Defects Data of Relay Protection Devices Based on Apriori Algorithm,” *IEEE Access*, vol. 8, pp. 120647–120653, 2020, doi: 10.1109/ACCESS.2020.3006345.

[4] M. A. Valle and G. A. Ruz, “Finding Hierarchical Structures of Disordered Systems: An Application for Market Basket Analysis,” *IEEE Access*, vol. 9, pp. 1626–1641, 2021, doi: 10.1109/ACCESS.2020.3047537.

[5] B. Molina-Coronado, U. Mori, A. Mendiburu, and J. Miguel-Alonso, “Survey of Network Intrusion Detection Methods from the Perspective of the Knowledge Discovery in Databases Process,” *IEEE Trans. Netw. Serv. Manag.*, vol. 17, no. 4, pp. 2451–2479, 2020, doi: 10.1109/TNSM.2020.3016246.