

Analisa Keranjang Belanja Konsumen Pada Data Penjualan Bulan Ramadhan Menggunakan Algoritma Apriori (Studi Kasus : Distro Coffepark Clothes Pekanbaru)

¹M Iqbal, ²Muatin

^{1,2}Prodi Sistem Informasi Fakultas Sains dan Teknologi UIN Suska Riau
Jl.HR. Soebrantas KM.15 Simpang Baru Panam Pekanbaru-Riau
Email :¹M.Iqbal1@students.uin-suska.ac.id,²muatin@students.uin-suska.ac.id

Abstrak

Distro coffeepark clothes adalah salah satu usaha yang menjual pakaian jadi yang langsung di supply dari Bandung, yang mana omset penjualan tertinggi distro adalah pada bulan ramadhan. Namun pada bulan ramadhan 2016 penjualan distro mengalami penurunan yang sangat signifikan yaitu dari 318 juta pada ramadhan 2015 menjadi 186 juta pada tahun 2016, penurunan berkisar 41,51%. Jika dibiarkan berkemungkinan berdampak terhadap penurunan penjualan ramadhan mendatang. Untuk itu perlu dilakukan evaluasi, salah satunya dengan menganalisa keranjang belanja konsumen atau market basket analysis untuk melihat pola belanja konsumen sebagai acuan strategi penjualan dibulan puasa mendatang. Penganalisaan dilakukan dengan menggunakan algoritma apriori dengan metode asociation rule. Dari hasil penelitian penganalisaan keranjang belanja konsumen yang dilakukan pada data bulan ramadhan 2016 didapatkan hasil pembelian kombinasi item set yang mendominasi adalah kombinasi item G dan D dengan nilai support tertinggi yaitu 0,502. Dari penelitian di dapatkan kombinasi item yang memiliki tingkat hubungan paling tinggi adalah kombinasi item E dan D yaitu kombinasi item kemeja panjang dan celana pendek yang dapat dilihat dari nilai confidence tertinggi yaitu 0,925. Dari hasil tersebut dapat ditarik kesimpulan bahwasanya setiap konsumen membeli kemeja panjang ada indikasi konsumen juga akan membeli kemeja pendek.

Kata kunci : Algoritma Apriori, association rule, market basket analysis, data mining

Abstract

Distro coffeepark clothes is one business that sells apparel directly in the supply from Bandung, which is the highest sales turnover distribution is in the month of ramadan. However in the month of ramadhan 2016 sales distro has a very significant decline of 318 million in 2015 To 186 million by 2016, the decline ranges from 41.51%. If allowed to be likely to impact the decline in sales of upcoming ramadan. For that need to be evaluated, one of them by analyzing the consumer shopping cart or market basket analysis to see the pattern of consumer spending as a reference sales strategy in the coming month of fasting. Analyzing is done by using apriori algorithm with asociation rule method. From result of research of consumer shopping cart analyzers conducted on data of ramadhan month 2016 got result of purchasing combination of item set which dominate is combination of item G and D with highest support value that is 0,502. From the research in getting the combination of items that have the highest level of relationship is a combination of items E and D is a combination of long shirt items and shorts that can be seen from the highest confidence value of 0.925. From these results can be concluded that every consumer to buy a long shirt there are indications consumers will also buy a short shirt.

Keywords : Algorithm Apriori, Association rule, market basket analysis, data mining.

1. Pendahuluan

Distro coffeepark clothes adalah salah satu usaha yang bergerak di bidang penjualan pakaian jadi, khususnya pakaian distro, yang langsung di *supply* dari Bandung. Coffeepark clothes ini telah berdiri sejak tahun 2009 di daerah Panam Pekanbaru. Saat ini coffeepark clothes telah melayani sekian banyak transaksi penjualan pakaian dan setiap data transaksi tersebut di simpan dalam suatu sistem dengan nama sistem *point of sales*. Dari sekian banyak transaksi, bulan ramadhan adalah bulan yang banyak memberikan kontribusi terhadap penjualan distro coffepark, yang dapat dilihat dari data penjualan pertahun distro[1].

Namun pada bulan ramadhan tahun 2016 distro coffeepark clothes mengalami penurunan penjualan yang cukup signifikan yakni dari 318 juta pada ramadhan tahun 2015 menjadi 186 juta pada ramadhan tahun 2016 penurunan berkisar 41,51%[1]. dari hasil wawancara dengan pemilik, penurunan ini diakibatkan karna lemahnya strategi pemasaran, promosi yang tidak inovatif dan kurang mengetahui *trend* atau minat konsumen pada bulan

ramadhan. Jika hal ini terus dibiarkan maka kemungkinan penjualan ramadhan tahun 2017 atau berikutnya akan mengalami penurunan dan akan berpengaruh pada omset penjualan pertahun. Selain berdampak kepada omset penjualan pertahun hal ini juga akan berdampak larinya pelanggan tetap atau konsumen karna tidak adanya inovasi dalam promosi dan pemasaran. Berangkat dari permasalahan tersebut peneliti akan menganalisa keranjang belanja konsumen distro coffeepark clothes melalui data penjualan yang ada di distro pada bulan ramadhan tahun 2016. Dengan menganalisa keranjang belanja konsumen, penulis akan melihat pola belanja dari konsumen distro coffeepark clothes untuk rekomendasi promosi dan strategi penjualan yang tepat. Untuk proses analisa peneliti menggunakan data penjualan ramadhan tahun 2016 karna merupakan data penjualan ramadhan terakhir dan terbaru yang paling tepat untuk mencari rekomendasi.

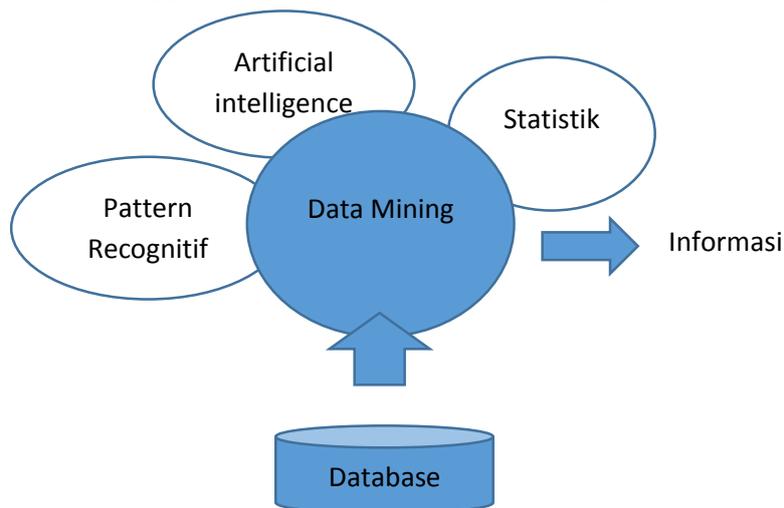
Untuk menganalisa keranjang belanja konsumen distro coffeepark, penulis menggunakan metode *association rule* dengan algoritma apriori. *Association rule* adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item[2]. Sedangkan algoritma apriori adalah algoritma dasar yang di usulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *boolean*[3]. Sebelumnya algoritma ini telah di gunakan pada penelitian Goldie Gunadi dan Dana Indra Sensua pada tahun 2012[2]. Merujuk dari penelitian tersebut peneliti juga akan menggunakan metode dan algoritma yang sama yaitu algoritma apriori dengan studi kasus yang berbeda. Untuk mendapat hasil yang maksimal dalam mencari kombinasi *item* penulis menggunakan bantuan aplikasi *RapitMinner studio 5.3.0*. dengan aplikasi ini di harapkan mendapatkan kombinasi item yang tepat.

Dengan di lakukan penelitian ini diharapkan dapat menghasilkan informasi-informasi yang bermanfaat bagi pemilik distro coffeepark clothes terutama dalam pengambilan keputusan yang berkaitan dengan perumusan atau pembuatan strategi untuk pemasaran dan penjualan barang distro coffeepark clothes khususnya pada bulan ramadhan, dengan melihat dari analisa pola pembelian konsumen yang dilakukan pada penelitian.

2. Metode penelitian dan bahan

2.1 Data mining

Daryl Pregibon menyatakan bahwa data mining adalah campuran dari statistik, kecerdasan buatan, dan riset basis data yang masih berkembang (Gononescu,2011). Munculnya data mining di dasarkan pada jumlah data yang tersimpan dalam basis data semakin besar. Ada istilah lain yang mempunyai makna yang sama dengan data mining yaitu *knowledge-discovery in database* (KDD). Data mining atau KDD bertujuan untu memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna[4].



Gambar 1 Akar ilmu data mining[4]

2.2. Association Rule

Analisa asosiasi di gunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang di temukan biasanya mempresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien[4]. *Interestingness measure* yang dapat di gunakan dalam data mining adalah :

- Support*, adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi.
- Confidence*, adalah suatu ukuran yang menunjukkan hubungan antar dua item secara *conditional* (berdasarkan suatu kondisi tertentu).

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap, yakni melakukan analisa pola frekuensi tinggi (*frequent pattern*) dan berikutnya adalah proses pembentukan aturan asosiasi [5].

2.3 Market Basket Analysis

Market basket analysis adalah suatu metodologi untuk melakukan analisis *buying habit* konsumen dengan menemukan asosiasi antara beberapa *item* yang berbeda, yang di letakkan konsumen dalam *shopping basket* (keranjang belanja) yang di beli pada suatu transaksi tertentu. Tujuan dari *market basket analysis* adalah untuk mengetahui produk-produk mana yang mungkin akan dibeli secara bersamaan [2].

2.4 Algoritma Apriori

Algoritma Apriori adalah suatu algoritma dasar yang di usulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan frequent itemset untuk aturan asosiasi boolean [2]. Bentuk algoritma dari metode apriori dapat dituliskan sebagai berikut [2]:

```

L1 = {frequent itemset with one element} for (k = 2; Lk-1 ≠ ∅; k++)
{
Ck = Apriori - gen(Lk-1); //pembuatan kandidat
//Baru for all transaction t
{
C't = Subset(Ct); //Kandidat yang
//tampil pada t for all candidates c ∈ C't do c.count++;
}
Lk = {c ∈ C't | c.count ≥ minsup}
}
Return ∪k Lk;
    
```

Dimana :

- L* : himpunan frequent itemset minsup : minimum support
- C* : himpunan kandidat itemset
- c* : kandidat itemset
- t* : transaksi

2.5. Analisis Pola Frekuensi Tinggi dengan Algoritma Apriori

Mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam basis data.

Nilai *support* sebuah *item* diperoleh dengan menggunakan rumus berikut:

$$Support(A) = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total Transaksi}}$$

Nilai support dari 2 item diperoleh dengan menggunakan rumus :

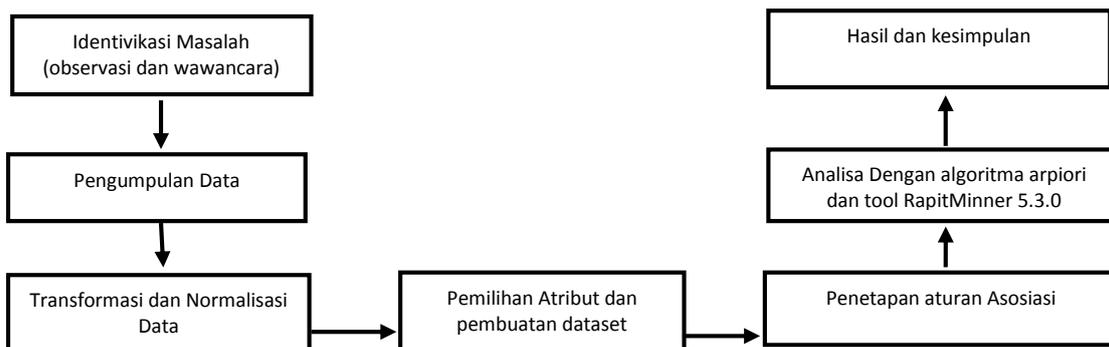
$$Support(A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}}$$

Frequent itemset menunjukkan itemset yang memiliki frekuensi kemunculan lebih dari nilai minimum yang ditentukan (\emptyset). Misalkan $\emptyset = 2$, maka semua yang itemset yang frekuensi kemunculan lebih dari atau sama dengan 2 kali disebut *frequent*. Himpunan dari frequent k-itemset dilambangkan dengan F_k .

3. Metodologi

Metode penelitian di mulai dari identifikasi masalah, pengumpulan data, tranformasi dan normalisasi, pemilihan atribut dan pembuatan dataset, penetapan aturan asosiasi, analisa menggunakan algoritma apriori, hingga hasil dan kesimpulan. yang di tunjukkan pada gambar2.



Gambar 2 Tahapan metodologi penelitian

4. Analisa dan Hasil

4.1 Analisa

Dalam penelitian ini data yang akan digunakan adalah data transaksi barang distro coffepark clothes pada bulan ramadhan 2016 yang di peroleh dari sistem yang ada di distro yaitu sistem penjualan atau *point of sales*. Data tersebut memiliki sembilan atribut diantaranya adalah tanggal, nomor struk, kode barang, nama barang, jenis barang, merek, transaksi, harga, stok, dan memiliki 249 record. Adapapun data yang di gunakan sesuai dengan gambar tabel berikut:

Gambar3. Tabel data Transaksi bulan ramadhan distro 2016

Tanggal	No Struk	Kode	Nama Barang	Jenis	Merek	Transaksi	Harga	Stok
04/06/2016	56432	22810	kmj pndk lucaman	kemeja pendek	lucaman	penjualan	135.000	12
04/06/2016	56432	50112	dln pj street soft typo	celana panjang	typo	penjualan	235000	30
04/06/2016	56432	10301	kaos Ap surf	kaos pendek	surf	penjualan	85000	3
04/06/2016	56433	90916	topi built up cp	topi	coffepark	penjualan	85000	9
04/06/2016	56434	23010	kmj pnjng plnl rockland	kemeja panjang	rockland	penjualan	165000	9
.....
.....
07/04/2016	57200	50314	dlna pj jean destory maggt	celana panjang	maggot	penjualan	265000	82
07/04/2016	57201	21206	kmja pngkek jnr magma	kemeja pendek	magma	penjualan	185000	12
07/04/2016	57201	10912	kaos plstsl typo	kaos pendek	typo	penjualan	100000	69
07/04/2016	57201	10109	kaos slur new bronze	kaos pendek	bronze	penjualan	100000	19

Gambar4. Tabel data transaksi transformasi bulan ramadhan distro 2016

Tanggal	No Struk	Kode	Nama Barang	Jenis	Merek	Transaksi	Harga	Stok
04/06/2016	1	22810	kmj pndk lucaman	D	lucaman	penjualan	135.000	12
04/06/2016	1	50112	cln pj street soft typo	G	typo	penjualan	235000	30
04/06/2016	1	10301	kaos Ap surf	A	surf	penjualan	85000	3
04/06/2016	2	90916	topi built up cp	M	coffepark	penjualan	85000	9
04/06/2016	3	23010	kmj pnjng plnl rockland	E	rockland	penjualan	165000	9
.....
.....
07/04/2016	248	50314	clna pj jean destory maggt	E	maggot	penjualan	265000	82
07/04/2016	249	21206	kmja pngk jnr magma	D	magma	penjualan	185000	12
07/04/2016	249	10912	kaos plstsl typo	A	typo	penjualan	100000	69
07/04/2016	249	10109	kaos slur new bronze	A	bronze	penjualan	100000	19

Gambar tabel 4 adalah data yang telah di tranformasi ke bentuk angka dan huruf. Adapun data yang ditransformasi adalah nomor struk dan jenis barang yang akan di pakai dalam penganalisaan association rule.

4.2 Pemilihan atribut

Dari sembilan atribut pada tabel diatas untuk melihat pola pembelian konsumen akan diambil dua atribut yaitu nomor struk dan jenis barang yang akan dianalisa menggunakan algoritma apriori dengan bantuan *tools rapitminer studio 5.3.0.* adapun atribut yang akan digunakan sesuai dengan tabel berikut.

Tabel 1. Tabel transaksi barang bulan ramadhan distro 2016

No Struck	Jenis
1	kemeja pendek, celana panjang, kaos pendek
2	Topi
3	kemeja panjang, celana panjang
4	kemeja pendek, celana panjang
.....
246	kaos pendek, jaket
247	kaos pendek, kaos pendek, celana panjang
248	kaos pendek, kaos pendek, kemeja panjang, celana panjang
249	kemeja pendek, kaos pendek, kaos pendek

Tabel 2. Tabel transaksi

No Struck	Jenis
1	D,G,A
2	M
3	E,G
4	D,G
.....
246	A,H
247	A,A,G
248	A,A,E,E
249	D,A,A

Gambar 5. Keterangan tabel transaksi

Jenis	Keterangan
A	kaos Pendek
B	Kaos Panjang
C	Kaos Kerah Pendek
D	Kemeja Pendek
E	Kemeja Panjang
F	Celana Pendek
G	Celana Panjang
H	Jaket
I	Sweater
J	Gesper
K	Kalung
L	Tas
M	Topi
N	Dompot
O	Sebo
P	Sandal
Q	Gelang

4.3 Hasil

Sebelum dilakukan penghitungan menggunakan *tools rapitminer* 5.3.0 data jenis barang yang di pakai diubah menjadi bilangan biner. Adapun data yang digunakan sesuai gambar tabel berikut.

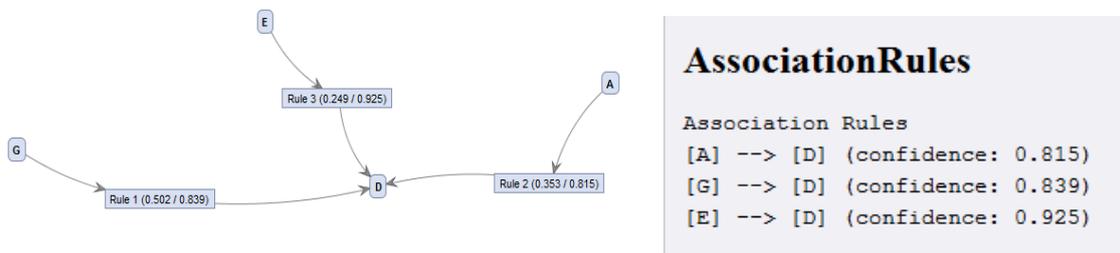
Gambar 6. Data yang telah di ubah dalam bentuk biner

no struck	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
.....
.....
247	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
248	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
249	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Setelah di ubah menjadi bilangan biner, selanjutnya dilakukan perhitungan menggunakan *tools rapitminer* dengan $\Phi=3$ dan *confidence* 40%, 60% dan 80% yang datanya diambil melalui microsoft excel. Berikut hasil tampilan analisa menggunakan *tools rapitminer*:

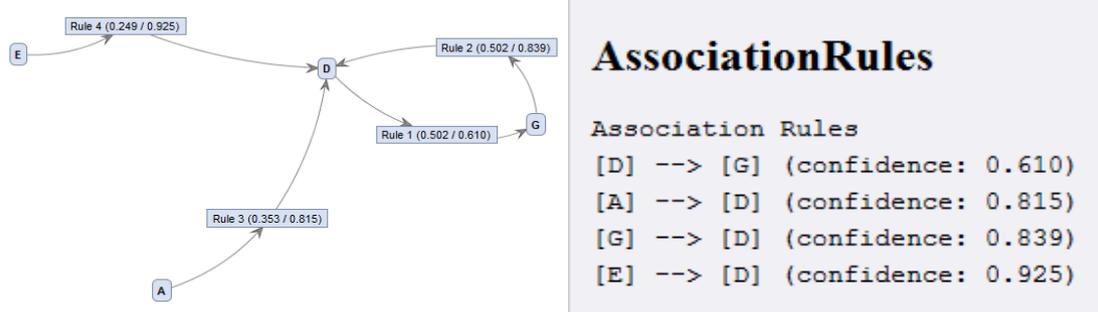
Gambar 7. Tabel hasil dan grafik perhitungan pada *rapitminer* 5.3.4 dengan *confidence* 80%

No.	Premises	Conclusion	Support	Confid...	LaPla...	Gain	p-s	Lift	Convi...
1	A	D	0.353	0.815	0.944	-0.514	-0.004	0.990	0.954
2	G	D	0.502	0.839	0.940	-0.695	0.009	1.019	1.097
3	E	D	0.249	0.925	0.984	-0.289	0.027	1.124	2.368



Percobaan pertama dilakukan dengan memasukkan *support* 0,95 dan *confidence* 0,8. Dari percobaan yang dilakukan diperoleh hasil tingkat dominasi item set yang paling tinggi adalah pada kombinasi item G dan D yaitu pada kombinasi item celana panjang dan kemeja pendek yang di lihat dari nilai *support* tertinggi yaitu 0,502. Dari hasil ini dapat ditarik kesimpulan bahwasanya item yang paling banyak di beli konsumen atau yang mendominasi adalah kombinasi item celana panjang dan kemeja pendek. Pada percobaan ini juga di dapatkan hasil barang yang memiliki hubungan keterkaitan atau kesamaan paling tinggi adalah item E dan D yaitu kombinasi item kemeja pendek dan kemeja panjang yang dilihat dari nilai *confidence* yang tinggi yaitu 0,925. Dapat ditarik kesimpulan ketika konsumen membeli kemeja panjang ada indikasi konsumen juga akan membeli kemeja pendek.

No.	Premises	Conclusion	Support	Confid...	LaPla...	Gain	p-s	Lift	Convi...
1	D	G	0.502	0.610	0.824	-1.145	0.009	1.019	1.029
2	A	D	0.353	0.815	0.944	-0.514	-0.004	0.990	0.954
3	G	D	0.502	0.839	0.940	-0.695	0.009	1.019	1.097
4	E	D	0.249	0.925	0.984	-0.289	0.027	1.124	2.368

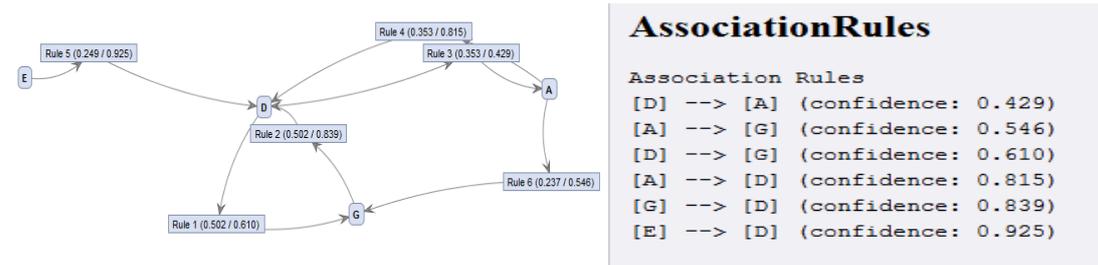


Gambar 8. Tabel hasil dan grafik perhitungan pada *rapitminner* 5.3.4 dengan confidence 60%

Percobaan kedua dilakukan dengan memasukkan *support* 0,95 dan *confidence* 0,6. Dari percobaan ini hasil yang di dapat serupa dengan percobaan pertama yaitu tingkat dominasi item set yang paling tinggi adalah pada kombinasi item G dan D yaitu pada kombinasi item celana panjang dan kemeja pendek dan kebalikannya kombinasi item D dan G yang di lihat dari nilai *support* tertinggi yaitu 0,502. Dari hasil ini dapat ditarik kesimpulan bahwasanya item yang paling banyak di beli konsumen atau yang mendominasi adalah kombinasi item celana panjang dan kemeja pendek dan sebaliknya. Pada percobaan ini juga didapatkan hasil barang yang memiliki hubungan keterkaitan atau kesamaan paling tinggi adalah item E dan D yaitu kombinasi item kemeja pendek dan kemeja panjang dengan dilihat dari nilai *confidence* yang tinggi yaitu 0,925. Dapat di tarik kesimpulan ketika konsumen membeli kemeja panjang ada indikasi konsumen juga akan membeli kemeja pendek

Gambar 9. Tabel hasil dan grafik perhitungan pada *rapitminner* 5.3.4 dengan confidence 40%

No.	Premises	Conclusion	Support	Confid...	LaPla...	Gain	p-s	Lift	Convi...
1	D	A	0.353	0.429	0.742	-1.293	-0.004	0.990	0.992
2	A	G	0.237	0.546	0.863	-0.631	-0.023	0.913	0.885
3	D	G	0.502	0.610	0.824	-1.145	0.009	1.019	1.029
4	A	D	0.353	0.815	0.944	-0.514	-0.004	0.990	0.954
5	G	D	0.502	0.839	0.940	-0.695	0.009	1.019	1.097
6	E	D	0.249	0.925	0.984	-0.289	0.027	1.124	2.368



Percobaan terakhir dilakukan dengan memasukkan nilai *support* 0,95 dan *confidence* 0,4. Dari percobaan terakhir hasil yang didapat serupa dengan percobaan pertama dan kedua yaitu tingkat dominasi item set yang paling tinggi adalah pada kombinasi item G dan D yaitu pada kombinasi item celana panjang dan kemeja pendek dan kebalikannya kombinasi item D dan G yang di lihat dari nilai *support* tertinggi yaitu 0,502. Dari hasil ini dapat ditarik kesimpulan bahwasanya item yang paling banyak di beli konsumen atau yang mendominasi adalah kombinasi item celana panjang dan kemeja pendek dan kombinasi sebaliknya kemeja pendek dan celana panjang. Pada percobaan ini juga di dapatkan hasil barang yang memiliki hubungan keterkaitan atau kesamaan paling tinggi adalah item E dan D yaitu kombinasi item kemeja pendek dan kemeja panjang dengan dilihat dari nilai *confidence* yang tinggi yaitu 0,925. Dapat di tarik

kesimpulan ketika konsumen membeli kemeja panjang ada indikasi konsumen juga akan membeli kemeja pendek.

5. Kesimpulan

Dari hasil analisa penelitian yang peneliti lakukan dengan memasukkan minimum support 95% atau 0,95 dan confidence 40%, 60%, dan 80% atau 0,4, 0,6, dan 0,8 dapat ditarik kesimpulan hasil yang didapat relatif sama yaitu tingkat dominasi item set yang paling tinggi adalah kombinasi item G dan D yaitu kombinasi item celana panjang dan kemeja pendek dan sebaliknya kombinasi D dan G yaitu kemeja pendek dan celana panjang yang dilihat dari nilai support tertinggi yaitu 0,502. Dari hasil ini dapat ditarik kesimpulan bahwasanya item yang paling banyak di beli konsumen atau yang mendominasi adalah kombinasi item celana panjang dan kemeja pendek. Dari hasil percobaan yang dilakukan item yang memiliki tingkat hubungan keterkaitan paling tinggi adalah item E dan D yaitu kemeja panjang dan kemeja pendek yang dilihat dari nilai confidence tertinggi 0,925. Dapat di Tarik kesimpulan letika konsumen membeli kemeja panjang ada indikasi konsumen juga akan membeli kemeja pendek.

Daftar Pustaka

- [1] Distro Coffeepark clothes, "*Data Penjualan Ramadhan*". 2016
- [2] Gunadi, Goldie,dkk. 2012. "*Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth(FP-Growth)*". Jurnal TELEMATIKA MKOM Vol.4 No.1, Maret 2012 ISSN: 2085-7225X
- [3] Jiawei Han dan Micheline Kamber, "*Data Mining; concepts and Techniques : chapter 6. Mining Association Rules in Large Databases*", Simon Fraser University, 2000.
- [4] Prasetyo, Eko "*Data Mining : Mengelola Data Menjadi Informasi Menggunakan Matlab*", Penerbit Andi, Yogyakarta, 2014.