



PANDA



Extract, Transform, Load (ETL) Modelling in BurgerStar Company

TEAM PANDA :

1. Afifah Nur Iswari
2. Dwi Smaradahana Indraloka
3. Nurrimah

OUTLINE :

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Evaluation

1. Business Understanding

BurgerStar merupakan salah satu perusahaan *fast food* tertua di Indonesia, toko pertamanya didirikan di Bandung pada tanggal 6 Maret 1981. BurgerStar selalu menjaga kualitas dari produk yang dihasilkan, dimana hanya bahan-bahan dengan kualitas terbaik yang digunakan dalam pembuatan produk dari BurgerStar. Sampai saat ini, BurgerStar telah memiliki 120 *outlet* yang tersebar di seluruh penjuru Indonesia dari Sabang hingga Marauke.

Dalam usaha meningkatkan pendapatan perusahaan, BurgerStar menjalankan program pengembangan dan perluasan menu produk. Program tersebut dijalankan dengan secara aktif meminta feedback dari pelanggan, sehingga BurgerStar dapat mengembangkan menu produk yang dimiliki. Saat ini menu produk dari BurgerStar telah mencakup beragam produk hamburger, ayam dan vegetarian yang juga dilengkapi dengan minuman dan makanan penutup. Selama bertahun-tahun program ini telah meningkatkan reputasi BurgerStar, sehingga



PANDA

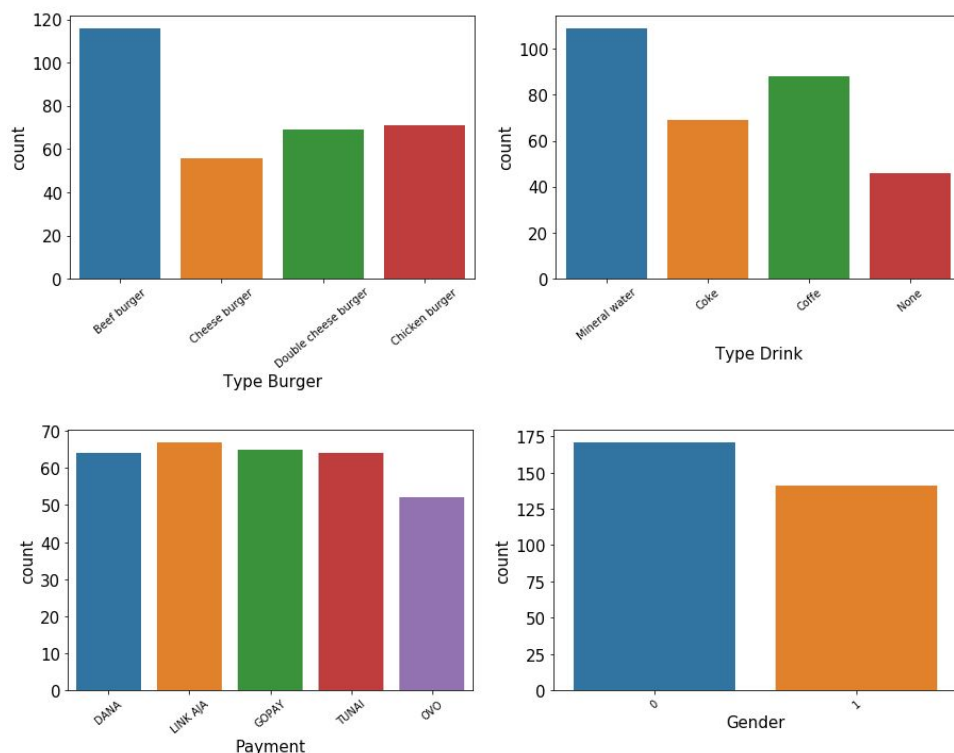
semua pelanggan telah memiliki kesetiaan pada berbagai menu produk dari BurgerStar.

Seiring dengan berjalannya waktu, gerai makanan cepat saji jumlahnya semakin berkembang secara pesat. Persaingan dari segi kualitas produk, harga produk dan kualitas pelayanan tidak dapat dihindarkan. Program yang telah diterapkan oleh BurgerStar tidak dapat lagi membantu untuk meningkatkan pendapatan perusahaan, sehingga manajemen BurgerStar melalui team Panda ingin melakukan optimasi penjualan menu produk khususnya produk hamburger dan minuman.

Solusi dari team Panda adalah dengan mencari asosiasi terhadap Burger Type dan Drink Type dari seluruh transaksi setiap jam-nya dan dengan melakukan *Clustering* jumlah pembelian burger dan minuman untuk mempersiapkan strategi marketing yang dapat digunakan sehingga dapat meningkatkan pendapatan perusahaan.

2. Data Understanding

Dataset diperoleh dari satu file penjualan seluruh item di BurgerStar dalam satu hari (10.00 s.d 21.00) yang terdiri dari sembilan variabel (TimeStamp, Type Burger, Code Burger, Amount Burger, Type Drink, Code Drink, Amount Drink, Gender dan Payment). Pembelian yang dilakukan selama satu hari di BurgerStar Company yang tercatat yaitu sebanyak 312 data.





PANDA

Berdasarkan *Bar Plot* dapat diketahui mayoritas kostumer dalam satu hari di BurgerStar secara keseluruhan membeli Beef Burger dan Mineral Water. Dan jika dilihat berdasarkan metode pembayaran, secara visualisasi tidak terdapat perbedaan yang signifikan jika dilihat lima metode pembayaran. Perbedaan tidak signifikan juga terjadi jika kostumer dilihat berdasarkan gender, jumlah kostumer laki-laki maupun perempuan dapat dikatakan tidak jauh berbeda dalam satu hari.

Berdasarkan visualisasi *Bar Plot* menunjukkan variabel *Type Burger* dan *Type Drink* pola yang cukup menarik. Berdasarkan data, dapat diketahui terdapat 4 Burger Type yang terdiri dari :

- Beef Burger
- Cheese Burger
- Double Cheese Burger
- Chicken Burger

Sedangkan untuk Drink Type, terdapat 3 Type, diantaranya adalah :

- Mineral Water
- Coke
- Coffee

3. Data Preparation

Data yang diberikan kepada tim Panda sudah siap untuk digunakan karena sudah melalui proses *cleansing* dan siap untuk dianalisis. Kemudian, seluruh data di *import* ke dalam *colab.research.google.com*

4. Modelling

Terdapat dua pemodelan yang digunakan dalam melakukan analisis pengoptimalan penjualan di BurgerStar Company, diantaranya adalah dengan menggunakan *Association* dan *Clustering*. Berikut penjelasan untuk kedua pemodelan tersebut.

Association: Apriori Algorithm

Association rule mining atau analisis asosiasi merupakan teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Atau dapat dikatakan association rule adalah suatu prosedur untuk mencari hubungan antar item dalam suatu dataset yang ditentukan. Dalam hal ini, Algoritma Apriori digunakan sebagai proses dalam menemukan *frequent-itemset* dengan melakukan iterasi pada data. Dimana *itemset* adalah himpunan item-item yang berada di dalam himpunan data, sedangkan *frequent-itemset* menunjukkan *itemset* yang memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan. Ukuran yang menentukan suatu item sering muncul dalam suatu data set disebut *support*. Semakin tinggi *support*, maka semakin sering item tersebut muncul (dalam data ini terjual).



PANDA

```
frequent_itemsets = apriori(df1, min_support=0.0005, use_colnames=True)  
print (frequent_itemsets)
```

```
      support      itemsets  
0  0.371795      (Beef burger)  
1  0.179487      (Cheese burger)  
2  0.227564      (Chicken burger)  
3  0.282051      (Coffe)  
4  0.221154      (Coke)  
5  0.221154      (Double cheese burger)  
6  0.349359      (Mineral water)  
7  0.147436      (None)  
8  0.089744      (Coffe, Beef burger)  
9  0.041667      (Beef burger, Coke)  
10 0.205128      (Mineral water, Beef burger)  
11 0.035256      (Beef burger, None)  
12 0.179487      (Coke, Cheese burger)  
13 0.115385      (Coffe, Chicken burger)  
14 0.112179      (None, Chicken burger)  
15 0.076923      (Coffe, Double cheese burger)  
16 0.144231      (Mineral water, Double cheese burger)
```

Berdasarkan Output, dapat diketahui *support* tertinggi secara keseluruhan :

No	Itemsets	Support
1	Beef Burger	0.37
2	Mineral Water	0.34
3	Coffee	0.28

Sedangkan jika dilihat berdasarkan *support* tertinggi pada saat *Burger Type* dikaitkan dengan *Drink Type* secara keseluruhan, maka diketahui sebagai berikut :

No	Itemsets	Support
1	Beef Burger & Mineral Water	0.20
2	Cheese Burger & Coke	0.17
3	Double Cheese Burger & Mineral Water	0.14

Dari hasil *running* di atas maka dapat diketahui barang item apa saja yang sering dibeli (nilai support) dari penjualan BurgerStar dalam satu hari. Jika dilihat berdasarkan tiga nilai support tertinggi pada satuan item, maka dapat diketahui terdapat tiga item yang memiliki probabilitas tinggi untuk dibeli. Diantaranya adalah Beef Burger, Mineral Water dan Coffee. Jika dilihat berdasarkan pasangan item antara Burger Type dan Drink Type, maka dapat diketahui probabilitas tertinggi pasangan item type yang dibeli diantaranya adalah Beef Burger-Mineral Water, Cheese Burger-Coke, dan Double Cheesee Burger-Mineral Water.



	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(Coffe)	(Beef burger)	0.282051	0.371795	0.089744	0.318182	0.855799	-0.015122	0.921368
1	(Beef burger)	(Coffe)	0.371795	0.282051	0.089744	0.241379	0.855799	-0.015122	0.946387
2	(Coke)	(Beef burger)	0.221154	0.371795	0.041667	0.188406	0.506747	-0.040557	0.774038
3	(Beef burger)	(Coke)	0.371795	0.221154	0.041667	0.112069	0.506747	-0.040557	0.877147
4	(Mineral water)	(Beef burger)	0.349359	0.371795	0.205128	0.587156	1.579247	0.075238	1.521652
5	(Beef burger)	(Mineral water)	0.371795	0.349359	0.205128	0.551724	1.579247	0.075238	1.451430

Sedangkan jika dilihat berdasarkan *output* di atas, maka dapat diketahui aturan asosiasi sebagai berikut :

- **[Confidence]** Jika terdapat satu customer yang membeli *Coffee* maka peluang customer itu juga akan membeli *Beef Burger* adalah 31%
- **[Lift]** Menunjukkan angka 0.85, bahwa terdapat asosiasi yang cukup, dikarenakan jika semakin besar angka lift maka semakin besar pula asosiasi yang terbentuk.
- **[Leverage]** Nilai ini menunjukkan perbedaan item set pada saat item dibeli bersamaan, serta menerangkan frekuensi yang diharapkan bila kedua item independen. Dikarenakan *Coffe* dan *Beef Burger* menunjukkan angka -0.01 pada nilai leverage maka dapat diketahui terdapat hubungan yang negatif antar kedua item tersebut.
- **[Conviction]** Nilai ini menunjukkan kuat tidaknya hubungan antara item set. Makin tinggi nilainya, maka semakin kuat asosiasi kedua item set. Nilai conviction tertulis 0.92, maka menunjukkan peluang kustomer membeli *Coffee* dan disertai membeli *Beef Burger* tidak cukup tinggi. Hal ini dikarenakan rentang nilai conviction ialah 0 s.d tak hingga.

Clustering: K-Means

Proses klastering dilakukan dengan mengambil dua variabel dari data yang ada yaitu "Amount Burger" dan "Amount Drink". Algoritma yang digunakan dalam proses klastering adalah K-Means dan untuk mencari jumlah klaster terbaik digunakan Silhouette Coefficient. Berikut adalah hasil yang di dapat:

1. Penentuan jumlah klaster terbaik:

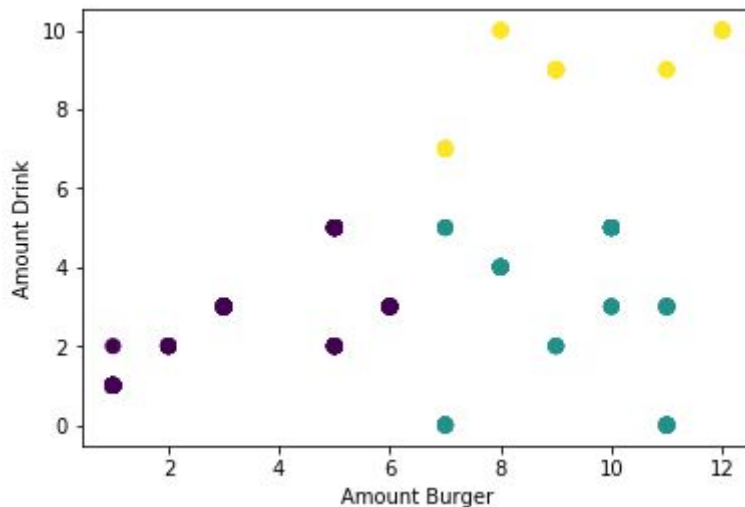
Jumlah Klaster	Nilai Silhouette Coefficient
2	0.501806
3	0.536418
4	0.522178



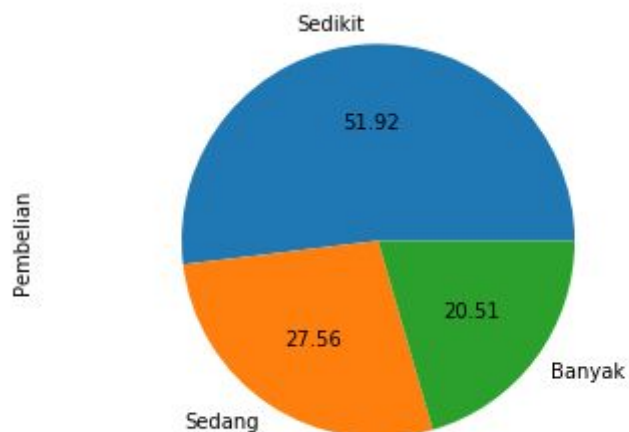
PANDA

Sehingga jumlah klaster yang digunakan adalah sebanyak tiga.

2. Klastering dengan K-Means:



Berdasarkan hasil klastering di atas, didapat tiga klaster dengan rincian klaster pertama adalah klaster dengan jumlah pembelian yang sedikit, klaster kedua adalah klaster dengan jumlah pembelian sedang (Burger banyak tetapi Minuman sedikit), klaster ketiga adalah klaster dengan jumlah pembelian banyak.



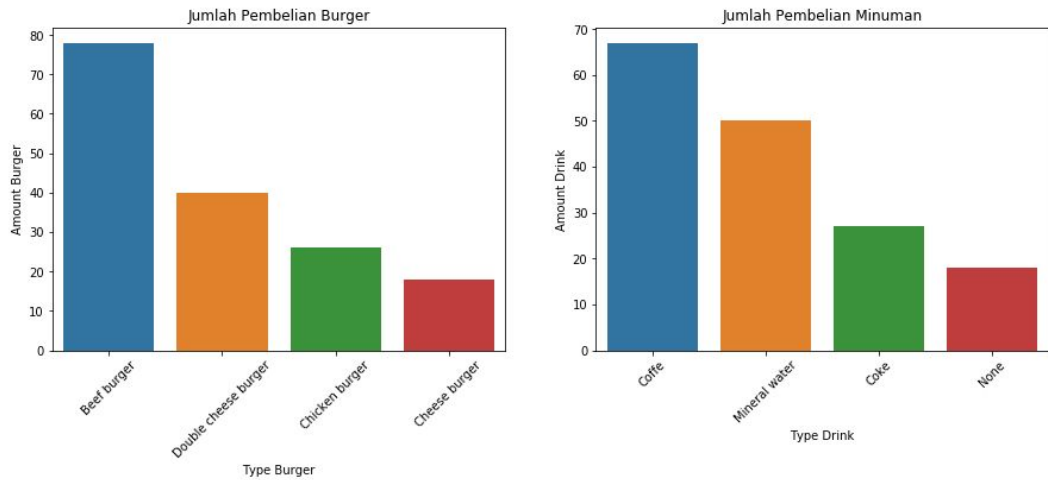
Dari hasil Klastering K-Means diatas, dapat dilihat persentase Jumlah Pembelian ketiga klaster dalam dataset terdiri dari Jumlah Pembelian Klaster “Sedikit” sebanyak 51.92%, Jumlah Pembelian Klaster “Sedang” sebanyak 27.56%, dan Jumlah Pembelian Klaster “Banyak” sebanyak 20.51%. Dapat disimpulkan bahwa Klaster dengan Jumlah Pembelian terbanyak adalah Klaster “Sedikit” dan Klaster dengan Jumlah Pembelian paling sedikit adalah Klaster “Banyak”.



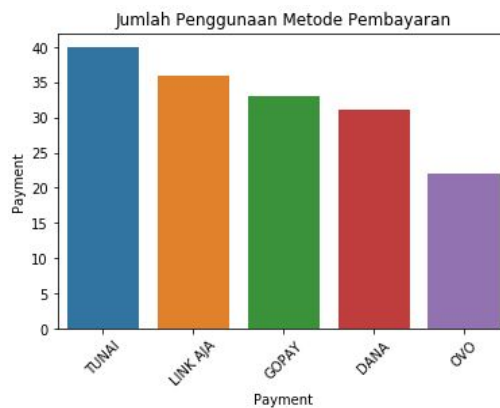
PANDA

3. Penjelasan pada tiap klaster:

- Klaster Pertama (Klaster dengan jumlah pembelian sedikit)

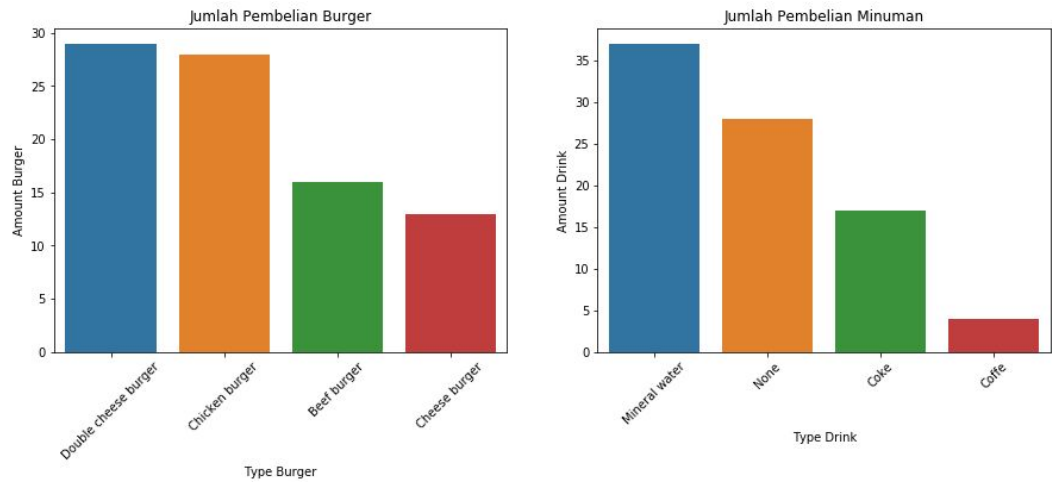


Pada klaster dengan jumlah pembelian sedikit, untuk tipe burger yang paling banyak dibeli adalah Beef Burger dan yang paling sedikit dibeli adalah Cheese Burger. Sedangkan untuk tipe minuman yang paling banyak dibeli adalah Coffee dan yang paling sedikit dibeli adalah Coke.

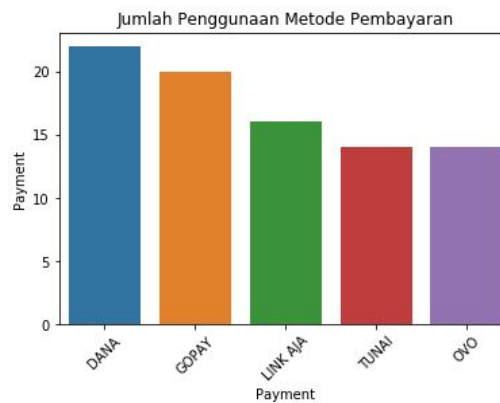


Untuk proses pembayaran paling banyak dilakukan dengan cara membayar secara tunai (cash).

- Klaster Kedua (Klaster dengan jumlah pembelian sedang - Burger banyak tetapi Minuman sedikit)

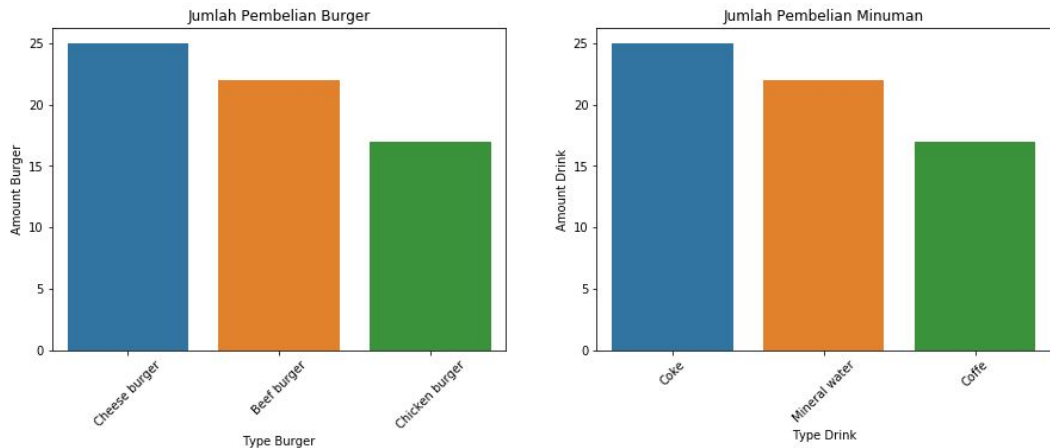


Pada klaster dengan jumlah pembelian sedang, untuk tipe burger yang paling banyak dibeli adalah Double Cheese Burger dan yang paling sedikit dibeli adalah Cheese Burger. Sedangkan untuk tipe minuman yang paling banyak dibeli adalah Mineral Water dan yang paling sedikit dibeli adalah Coffee.

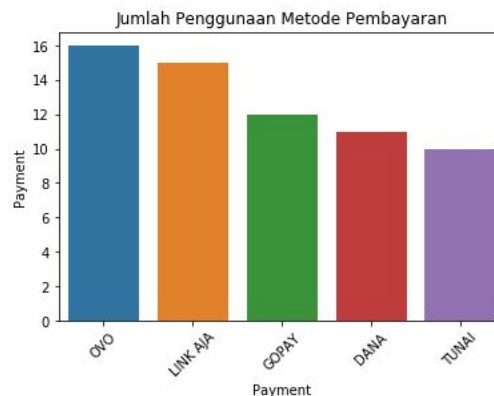


Untuk proses pembayaran sebagian besar dilakukan dengan cara membayar menggunakan e-payment.

- Klaster Ketiga (Klaster dengan jumlah pembelian banyak)



Pada klaster dengan jumlah pembelian banyak, untuk tipe burger yang paling banyak dibeli adalah Cheese Burger dan yang paling sedikit dibeli adalah Double Cheese Burger. Sedangkan untuk tipe minuman yang paling banyak dibeli adalah Coke dan yang paling sedikit dibeli adalah Coffee.



Untuk proses pembayaran sebagian besar dilakukan dengan cara membayar menggunakan e-payment dan pembayaran dengan cara tunai paling sedikit dilakukan.

5. Development (Evaluation)

Berdasarkan hasil analisis yang dilakukan, evaluasi dan saran dari Tim Panda kepada BurgerStar Company adalah sebagai berikut :

- Melakukan penjualan paket burger dan minuman.

1. Paket Single

Berdasarkan output dari pemodelan Asosiasi, didapatkan bahwa Beef Burger dan Mineral Water memberikan kontribusi terbanyak untuk masing-masing jenis burger dan minuman dalam jumlah penjualan per item. Menurut kami, keuntungan perusahaan akan lebih meningkat jika kedua item tersebut dijual secara bersamaan dan dengan harga jual yang lebih murah, atau dengan kata lain **pemberian diskon jika membeli Paket Single Beef Burger dan Mineral Water**. Hal ini dikarenakan pada saat proses modelling menunjukkan



PANDA

terdapat probabiliti yang cukup besar bahwa kombinasi dua item tersebut saling berketergantungan satu sama lain dan jika membeli mineral water besar kemungkinan setelahnya akan membeli Beef Burger, begitupun sebaliknya. Walaupun pada saat penerapan package ini terlihat tidak memberikan profit yang tinggi, namun dikarenakan kuantitas penjualan package yang diprediksi akan bertambah maka harapannya dengan menerapkan package ini juga dapat meningkatkan *value* atau pun pendapatan BurgerStar.

Berdasarkan output dari pemodelan Klastering diatas, diperoleh untuk klaster pertama dengan jumlah pembelian sedikit, jenis minuman dengan jumlah penjualan terendah adalah Coke, sedangkan jenis burger dengan jumlah penjualan tertinggi adalah Beef Burger. Dalam hal ini, untuk meningkatkan jumlah penjualan Coke, BurgerStar dapat menjual Coke dengan memasangkannya secara bersamaan dengan menjual Beef Burger yang memiliki banyak peminat, dengan kata lain dibuat **Paket Single Beef Burger dan Coke dengan harga yang lebih murah** juga tentunya.

2. Paket Medium

Secara metode, komposisi dari menu pada paket medium didapatkan dengan cara yang sama dengan paket single, yaitu **Paket Medium Beef Burger dan Coke, juga Paket Medium Beef Burger dan Mineral Water**. Paket medium dibuat untuk pelanggan yang menginginkan pembelian Burger dan Minuman dengan jumlah yang sedang. Plotting pada hasil klastering juga menunjukkan bahwa terdapat pelanggan yang melakukan pembelian Burger dan Minuman dengan jumlah yang sedang. Sehingga BurgerStar dapat meningkatkan jumlah penjualan secara lebih cepat.

3. Paket Big

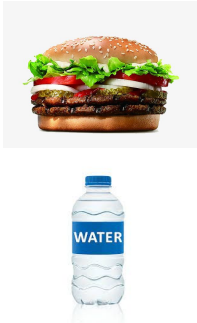
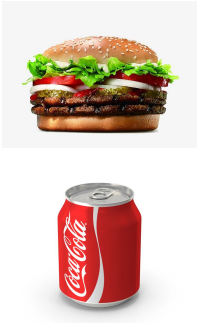
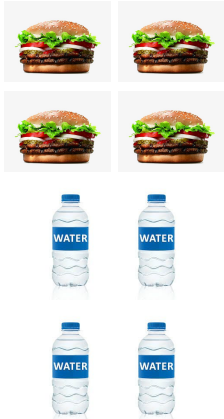
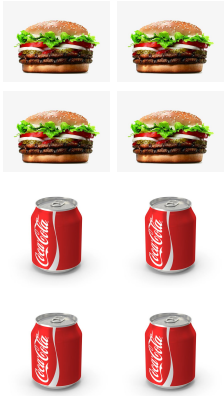
Beda halnya dengan paket Single dan Medium, pada paket Big ketiga tipe burger ditawarkan. Namun dipasangkan dengan tipe minuman yang ditentukan. Paket Big dilakukan agar terlihat menguntungkan para kostumer. Namun jika dilakukan pengamatan lebih lanjut, hal ini dapat dikatakan sangat menguntungkan BigStar. Penentuan pemasangan Burger dan Minuman diperoleh dari klaster kedua dan ketiga dengan juga memperhatikan nilai asosiasinya. Pada klaster kedua, jenis minuman dengan jumlah penjualan terendah adalah Coffee dan Coke, sedangkan jenis burger dengan jumlah penjualan tertinggi adalah Double Cheese Burger dan Chicken Burger. Pada klaster ketiga, jenis minuman dengan jumlah penjualan terendah adalah Coffee, sedangkan jenis burger dengan jumlah penjualan tertinggi adalah Cheese Burger. Dalam hal ini, untuk meningkatkan jumlah penjualan Coffee, BurgerStar dapat menjual Coffee dengan memasangkannya secara bersamaan dengan menjual Double Cheese Burger atau Cheese Burger yang



PANDA





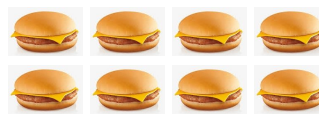

memiliki banyak peminat dan nilai asosiasinya rendah, sedangkan untuk meningkatkan jumlah penjualan Coke, BurgerStar dapat menjual Coke dengan memasangkannya secara bersamaan dengan menjual Chicken Burger yang memiliki banyak peminat dan nilai asosiasinya rendah. dengan kata lain dibuat **Paket Big Double Cheese Burger dan Coffee, Paket Big Chicken Burger dan Coke dan Paket Big Cheese Burger dan Coffee dengan harga yang lebih murah** tentunya.

Tidak hanya melakukan penerapan Package yang saling menguntungkan. Harapannya, kasir pada setiap *outlet* dapat lebih mengerti kustomer. Yaitu dengan cara lebih interaktif pada saat kustomer ingin melakukan pemesanan dan melakukan penawaran terhadap burger dan minuman yang jumlah pembeliannya masih relatif rendah, seperti Chicken Burger, Cheese Burger, dan Coffee. Selain itu, Coffee juga dapat diberikan diskon untuk meningkatkan jumlah penjualannya dilihat dari nilai support pada modeling *Asosiasi* nya yang paling rendah. Selanjutnya, proses pembayaran dengan menggunakan e-payment juga dapat dijadikan strategi untuk meningkatkan penjualan. Berdasarkan hasil klastering, pelanggan yang melakukan pembelian dengan jumlah besar banyak melakukan pembayaran dengan menggunakan e-payment, kita dapat memberikan diskon tambahan untuk pembelian Big Package dengan menggunakan e-payment (Dengan tetap memperhatikan jumlah keuntungan). Sehingga dikemudian hari, harapannya BurgerStar dapat meningkatkan penjualan sampai lebih dari 20%.

Single Package		Medium Package	
			
Beef Burger - Mineral Water	Beef Burger - Coke	Beef Burger - Mineral Water	Beef Burger - Coke



PANDA

Big Package		
 	 	 
Double Cheese Burger - Coffee	Chicken Burger - Coke	Cheese Burger - Coffee

Coding Link:

<https://github.com/nurrimahrima/Burger-Set-Modeling-by-Machine-Learning>

