CLUSTERING ITEM FAST MOVING DAN SLOW MOVING PADA PRODUK UNILEVER MENGGUNAKAN ALGORITMA K-PROTOTYPE (Studi Kasus: YOGYA PURWAKARTA)

Akhmad Subhan, Ahmad Faqih, Bambang Irawan
Program Studi Teknik Informatika S1, Fakultas Teknik Informatika
STMIK IKMI Cirebon, Jl. Perjuangan No.10B, Karyamulya Cirebon, Indonesia
akhmad.subhan.45@gmail.com

ABSTRAK

Pada supermarket terdapat dua kategori produk yaitu produk *fast moving* merupakan produk yang cepat terjualan sedangkan produk *slow moving* merupakan produk yang penjualannya lambat. Penentuan kategori produk sebenarnya sangat penting, namun penentuannya masih menggunakan pendekatan empiris sehingga penentuannya masih kurang tepat atau bahkan tidak dilakukan pengkategorian produk sama sekali. Padahal kesalahan dalam penentuan kategori suatu produk jelas mempengaruhi stok produk. Produk yang salah dikategorikan akan mengakibatkan stok produk habis ataupun sebaliknya stok produk menjadi menumpuk. Penentuan kategori produk harus memiliki akurasi tinggi sehingga meminimalkan kerugian. Tujuan penelitian ini adalah untuk mengelompokan produk termasuk ke dalam kategori *fast moving* atau *slow moving* menggunakan algoritma *K-prototype*. Metode pengelompokan menggunakan algoritma *K-prototype*. Algoritma ini adalah hasil pengembangan dari algoritma *K-Means* untuk menangani *clustering* pada data dengan atribut bertipe campuran numerik dan kategorikal. Hasil dari penelitian ini ialah *clustering* yang di peroleh dengan jumlah *cluster*, *cluster* k = 0 sampai k = 2 diperoleh item produk yang termasuk ke dalam *fast moving* dan *slow moving* di Yogya Purwakarta. *Cluster* yang mempunyai nilai *sellout* terbesar ada pada *cluster* 0 dan *cluster* 1 yaitu 1.219 – 23.761 pcs penjualannya, sedangkan *cluster* yang mempunyai *sellout* terkecil ada pada *cluster* 2 yaitu di bawah 1.000 pcs penjualannya.

Kata kunci : *K-prototype, K-Means, Kategori, Fast Moving, Slow Moving.*

1. PENDAHULUAN

Pada setiap supermarket diperlukan penanganan persediaan yang efisien. Pengelolaan persediaan barang yang efisien akan meningkatkan penjualan dan keuntungan, penanganan yang dimaksud yaitu pengelolaan barang yang perputarannya dapat dikendalikan sehingga tidak ada penumpukan barang kadaluarsa yang sia-sia. Pengelompokkan barang yang tidak tepat merupakan salah satu kendala yang akan mengakibatkan penumpukan barang kadaluarsa dan tidak ada penjualan.

Perputaran produk ada 2 jenis yaitu produk laku (*fast moving*) dan produk kurang laku (*slow moving*). Produk laku (*fast moving*) adalah barang yang dapat terjual secara cepat, dan biasanya merupakan kebutuhan sehari-hari. Produk kurang laku (*slow moving*) adalah barang yang tidak diminati atau dijual selama waktu hampir 90 hari sehingga menyebabkan perputaran lambat.

Keuntungan yang didapat dari pengelompokkan ini untuk produk laku (*fast moving*) pihak toko diharapkan bisa menambah stok penjualannya sesuai dengan rata-rata penjualan per pekan (*rpp*) guna menjaga supaya tidak terjadi kekosongan barang terhadap produk laku, sedangkan untuk produk kurang laku (*slow moving*) pihak toko bisa mempertimbangkan produk-produk tersebut untuk dihapuskan atau tidak jual lagi, caranya dengan diadakannya promo *flush out*.

Yogya Purwakarta merupakan toko retail yang menjual berbagai macam produk dengan format supermarket atau swalayan, salah satu produknya adalah dari Unilever yang memiliki banyak jenis produk dengan berbagai variant, tentunya dengan banyaknya produk sering kali lalai dalam pengecekan produk dikarenakan belum adanya metode yang tepat dalam melakukan pengelompokkan produk jual yang ada di Yogya Purwakarta, sehingga masih menggunakan cara manual yang memakan waktu yang cukup lama.

Penelitian ini bertujuan untuk mengelompokkan produk termasuk ke dalam kategori produk laku (*fast moving*) atau kurang laku (*slow moving*) menggunakan algoritma *K-prototype*, sehingga dapat mempermudah pihak toko dalam memonitoring produk yang menghasilkan penjualan bagus dengan produk yang bisa berdampak kepada nilai pendapatan toko. Diharapkan dengan adanya pengelompokkan tersebut dapat meningkatkan penjualan serta mengurangi kekosongan barang dan meminimalkan produk tidak terjual yang dapat merugikan toko.

Selain melakukan wawancara penulis juga mecari penelitian sebelumnya yang berhubungan dengan penelitian ini, diantaranya peneliatian yang dilakukan oleh [1], dengan judul segmentasi pelanggan PT. Telekomunikasi Seluler Indonesia menggunakan *clustering algoritma K-prototype* dan metode *elbow* sebagai perumusan strategi marketing. Hasil percobaan ini menghasilkan rekomendasi strategi pelanggan yang dapat diterapkan oleh pihak pemasaran Telkomsel dengan mendapatkan 4 jumlah

cluster terbaik. Selanjutnya penelitian yang dilakukan oleh [2], yang berjudul metode *cluster* menggunakan kombinasi *algoritma cluster K-prototype* dan *algoritma* genetika untuk data bertipe campuran. Hasil dari penelitian menunjukkan optimasi pusat cluster dengan algoritma genetika berhasil meningkatkan akurasi hasil cluster dengan *K-prototype*, nilainya lebih tinggi dari yang lainnya yaitu 0,0054 dibandingkan 0,0031.

Berdasarkan penelitian terdahulu diharapkan mampu mengetahui pengelompokan data penjualan produk Unilever yang ada di Yogya Purwakarta (PT Akur Pratama) dengan penerapan metode *clustering* menggunakan *algoritma K-prototype* berdasarkan data penjualan produk Unilever dari bulan Januari-Oktober 2021 di Yogya Purwakarta. Langkah awal yang harus dilakukan adalah mengumpulkan dataset. Data yang dipakai pada penelitian ini ialah data privat yang didapat dari pihak toko sebagai data yang dipakai dalam proses pengelompokkan memakai metode *clustering algoritma K-prototype* dilihat dari hasil cluster algoritma tersebut didapat untuk mengetahui informasi kategori produk jual sesuai dengan data penjualan.

2. TINJAUAN PUSTAKA

2.1. Sistem

Sistem adalah sebuah tatanan (keterpaduan) yang terdiri dari sejumlah komponen fungsional (dengan satuaan fungsi/tugas khusus) yang saling berhubungan dan secara bersama-sama bertujuan untuk memenuhi suatu proses/pekerjaan tertentu [3].

Sistem adalah kumpulan dari komponen-komponen yang memiliki unsur keterkaitan antara satu dan lainnya. Sehingga dapat dikatakan bahwa sistem adalah merupakan suatu hal yang saling terkait satu sama lain untuk mencapai sebuah tujuan yang sama [4].

2.2. Informasi

Data adalah fakta mentah mengenai orang, tempat, kejadian, dan hal-hal yang penting dalam organisasi. Informasi adalah data yang telah diproses atau diorganisasi ulang menjadi bentuk yang berarti [3].

Pengolahan data adalah proses perubahan bentuk data menjadi informasi yang memiliki kegunaan. Semakin banyak data dan kompleksnya aktivitas pengolahan data, maka metode pengelolaan data yang tepat sangat dibutuhkan [4].

2.3. Data Mining

Data mining merupakan satu kata yang digunakan buat mendapatkan pemahaman yang terdapat pada basis data. Data mining adalah metode semi mekanis yang memakai teknik statistik, matematika, kecerdasan buatan, dan machine learning serta ringkasan dan mengetahui pemahaman berguna terdapat pada basis data [5].

2.4. Clustering

Clustering mengacu pada pengelompokkan kemiripan dokumen, observasi, atau mencermati serta menciptakan berbagai materi yang mempunyai persamaan. *Cluster* merupakan kelompok asal data yang mempunyai persamaan satu sama lain, serta tidak sama menggunakan report pada *cluster* lain. *Clustering* mencoba untuk membagi seluruh kumpulan facts membentuk *cluster* yang cukup mempunyai kesamaan, di mana persamaan file pada satu *cluster* akan bernilai maksimal, sedangkan persamaan menggunakan arsip pada *cluster* lain akan bernilai minimal. *Clustering* pada facts mining bermanfaat untuk mendapatkan model pembagian di dalam sebuah facts set yang bermanfaat untuk metode analisis *records*. Persamaan objek umumnya di dapat dari relasi nilai-nilai atribut yang menyebutkan materi records, sedangkan materi statistics umumnya di sampaikan menjadi sama titik dalam ruang bersegi banyak [6].

2.5. Algoritma K-Prototype

Algoritma K-prototype adalah salah satu metode *Clustering* yang berbasis partitioning. *Algoritma* ini adalah hasil pengembangan dari *algoritma K-Means* untuk menangani *clustering* pada data dengan atribut bertipe campuran numerik dan kategorikal. Pengembangan yang dilakukan oleh Huang mempertahankan efisiensi algoritma *K-Means* dalam menghadapi data berukuran besar dan dapat diterapkan pada data numerik dan kategorikal. Pengembangan yang mendasar pada algoritma *K-prototype* terdapat pada pengukuran kesamaan (*similarity measure*) antara *object* dengan *centroid* (*prototype*)-nya [2].

2.6. Manajemen Persediaan

Manajemen persediaan (*inventory management*) atau pengendalian tingkat persediaan adalah kegiatan yang berhubungan dengan perencanaan, pelaksanaan dan pengawasan penentuan kebutuhan produk, sehingga kebutuhan produk dapat dipenuhi pada waktunya dan di lain pihak investasi persediaan produk dapat ditekan secara optimal. Pengendalian tingkat persediaan bertujuan mencapai efisiensi dan efektivitas optimal dalam penyediaan produk [7]. Beberapa istilah dalam persediaan yaitu:

1. Persediaan bahan baku di tangan (*stock on hand*).
2. Daftar persediaan secara fisik.
3. Produk kosong (*out of stock*).
4. Permintaan produk (*pre-order*).

Metode persediaan yang digunakan di Yogya Purwakarta adalah metode *Min-Max* (*minimum-maksimum*). Metode *Min-Max* merupakan metode persediaan dimana kuantitas pemesanan hanya dapat dilakukan diantara kuantitas minimum dan maksimum dari rata-rata penjualan. Rata-rata penjualan diperoleh data penjualan tahun lalu.

2.7. FSN (Fast Slow Non) Analysis

FSN merupakan kepanjangan dari fast, slow, non-moving. FSN analysis bertujuan dalam pengelompokkan produk yang berdasarkan atas pergerakan produk dari pendataan inventory penjualan. Setelah itu produk akan diklasifikasikan berdasarkan fast, slow dan non-moving [8].

Slow moving item adalah item yang memiliki permintaan yang sangat kecil, sedangkan *fast moving* item adalah item yang memiliki permintaan yang sangat besar baik dari besaran pesanan atau jumlah pesanan, permintaan merupakan salah satu hal yang menjadi tolak ukur dalam menentukan apakah sebuah item termasuk dalam *slow movin*g item atau tidak [9].

3. METODE PENELITIAN

Metode analisa data di pergunakan pada penelitian ini merupakan analisis deskriptif pada pemilihan data kuantitatif. Pada manfaatnya, analisa deskriptif dipergunakan buat mengambarkan dan menguraikan data yang dikumpulkan dari informasi-informasi yang ada, data yang di maksud merupakan data sekunder berupa data kuantitatif berupa angka-angka yang sanggup digunakan untuk operasi matematika [10].

Penelitian ini dilakukan di Yogya Purwakarta (PT. Akur Pratama). Sumber data dalam penelitian ini diperoleh dari data lapangan yaitu data tentang penjualan produk Unilever dari bulan Januari - Oktober 2021.

3.1. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini adalah penelitian lapangan dengan cara melakukan pengamatan langsung pada toko yang bersangkutan. Adapun teknik pengumpulan data dilakukan dengan cara:

1. *Interview*
   Yaitu metode untuk mendapatkan data dengan cara melakukan tanya jawab secara langsung dengan pihak-pihak yang bersangkutan guna mendapatkan data dan keterangan yang menunjang.
2. Observasi
   Observasi sebagai teknik pengumpulan data mempunyai ciri yang spesifik bila dibandingkan teknik yang lain. Observasi merupakan suatu proses yang kompleks, suatu proses yang tersusun dari berbagai proses biologis dan psikologis. Dua proses pengamatan dan ingatan.

3.2. Analisa Data

Terdapat proses dalam Data Mining yang dapat dijelaskan sebagai berikut:

1. Seleksi Data
   Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.
2. *Pre-processing / Cleaning* (Pemilihan Data)
   Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi focus KDD. Proses cleaning mencangkup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses "memperkaya" data yang sudah ada dengan data informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.
3. Transformasi.
   Coding adalah proses transformasi pada data yang telah dipilih, sehinga data tersebut sesuai untuk proses data mining. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.
4. Data Mining
   Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Terknik, metode, atau algoritma dalam data mining sangat bervariasi, pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan.
5. *Knowledge*
   Knowledge adalah informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

4. HASIL DAN PEMBAHASAN

Pada penelitian ini dataset yang digunakan tentang penjualan produk Unilever di Yogya Toserba pada area Jawa Barat. Data didapat dari internal HO Unilever dengan HO Yogya Toserba sebanyak 1069 item produk dari bulan Januari - Oktober 2021 di Yogya Purwakarta menggunakan 16 atribut antara lain no, art_code, category, sku_desc, harga, januari, februari, maret, april, mei, juni, juli, agustus, september, oktober, total.

1. Seleksi Data

Pada tahap ini dilakukan seleksi data pada proses penelitian menggunakan data dari bulan Januari - Oktober 2021 di Yogya Purwakarta yang digunakan dalam proses data mining. Terdapat 4 atribut setelah dilakukan seleksi data diantaranya art_code, category, sku_desc, sellout_pcs. Dapat dilihat pada gambar 1.

```
In [3]: yogya_purwakarta_df = pd.read_csv('C:/Users/Asus/Documents/KAMPUS/SKRIPSI/Latihan/yogya_purwakarta.csv'
In [4]: yogya_purwakarta_df
Out[4]:
```

	art_code	category	sku_desc	sellout_pcs
0	153438	SOY & FRUIT BEVERAGES	BUAVITA APPLE 1 LT_1C12P_20081902	165
1	146236	SOY & FRUIT BEVERAGES	BUAVITA APPLE 125ML_1C40P_20082350	766
2	146223	SOY & FRUIT BEVERAGES	BUAVITA APPLE TBA 250ML_1C24P_20082312	1452
3	219963	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 125ML_1C40P_20082448	1499
4	148639	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 1LT_1C12P_20082239	625
...	...	...	...	...
1064	224686	BABY CARE	ZWITSAL MINYAK TELON 60ML _1C24P	65
1065	20754	BABY CARE	ZWITSAL MINYAK TELON PLUS 100ML _1C24P_67660399	13
1066	892347	BABY CARE	ZWITSAL POWDER EXTRA CARE300GR _1C12P	166
1067	892346	BABY CARE	ZWITSAL POWDER NATURAL MILKHONEY300GR _1C12P_YM	0
1068	6449	BABY CARE	ZWITSALKIDS HAIRLOTSFTNMST100ML_1C24P_68145824...	14

1069 rows × 4 columns

Gambar 1. Seleksi Data

Data yang sudah dipilih terlebih dahulu dilakukan pemeriksaan tipe data dari bingkai data dengan metode dtypes, kita dapat melihat semua variabel pada data tersebut termasuk integer atau object.

```
In [2]: df.dtypes
Out[2]: art_code      int64
        category     object
        sku_desc     object
        sellout_pcs   int64
        dtype: object
```

Gambar 2. Tipe Data

2. Data Preprocessing

Setelah dilakukan proses seleksi data selanjutnya dilakukan proses *preprocessing* data, data yang diambil merupakan data penjualan produk Unilever dari bulan Januari – Oktober 2021 di Yogya Purwakarta. Tujuan dari data preprocessing ini adalah mengabaikan nilai-nilai/atribut data yang kurang berpengaruh pada proses pengelompokkan data, memperbaiki kekacauan data serta mempelajari data yang tidak selaras.

Pada tahap ini peneliti tidak akan mengelompokkan semua fitur, melainkan fokus pada fitur berikut ini :

1. fitur numerik: sellout_pcs
2. fitur categoris: category, sku_desc

Langkah selanjutnya menempatkan fitur numerik pada skala yang sama. Hal ini dilakukan untuk memastikan semua fitur dipertimbangkan secara setara dalam pengelompokan akhir (yaitu jarak antar titik disetiap fitur yang sama).

```
In [3]: # select columns to cluster
        cluster_columns = ['category', 'sku_desc', 'sellout_pcs']
        df = df[cluster_columns]

In [4]: from sklearn.preprocessing import StandardScaler
        # define numerical and categorical columns
        numerical_columns = ['sellout_pcs']
        categorical_columns = ['category', 'sku_desc']
        scaler = StandardScaler()
        # create a copy of our data to be scaled
        df_scale = df.copy()
        # standard scale numerical features
        for c in numerical_columns:
            df_scale[c] = scaler.fit_transform(df[[c]])

In [5]: df_scale.head()
Out[5]:
```

	category	sku_desc	sellout_pcs
0	SOY & FRUIT BEVERAGES	BUAVITA APPLE 1 LT_1C12P_20081902	-0.232806
1	SOY & FRUIT BEVERAGES	BUAVITA APPLE 125ML_1C40P_20082350	0.203621
2	SOY & FRUIT BEVERAGES	BUAVITA APPLE TBA 250ML_1C24P_20082312	0.701773
3	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 125ML_1C40P_20082448	0.735903
4	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 1LT_1C12P_20082239	0.101232

Gambar 3. Skala Data

Untuk menentukan algoritma *K-prototype* yang tidak terlalu terpengaruh oleh *outlier* ini harus dihapus. Di sini menghapus setiap titik yang terletak di luar 3 standar deviasi dari meannya. Dari total 1069 item setelah diproses menjadi 1052 item. Proses tersebut bisa dilihat pada gambar 4.

```
In [6]: df_out = df_scale[(df_scale['sellout_pcs'].abs()<=3)]
        df_out.shape
Out[6]: (1052, 3)
```

Gambar 4. Penghapusan Data

3. Menemukan Jumlah Cluster Optimal

Kita harus melihat berbagai metrik evaluasi cluster untuk mendapatkan saran tentang berapa banyak klaster yang bisa kita klasterkan datanya.

Untuk bagian ini kita akan melihat metrik evaluasi *cluster*:

- Metode Siku

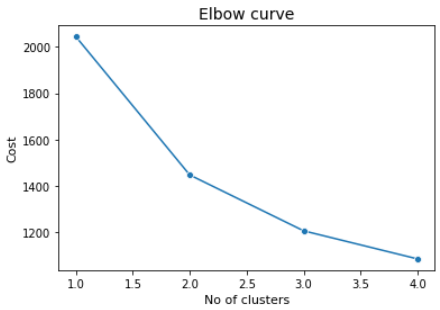
Untuk menerapkan metode evaluasi tersebut, kita harus menerapkan algoritma *K-prototype* dengan jumlah *cluster* yang berbeda. Algoritma ini membutuhkan posisi indeks dari variabel kategori:

```
In [7]: from kmodes.kprototypes import KPrototypes
        categorical_indexes = []
        for c in categorical_columns:
            categorical_indexes.append(df.columns.get_loc(c))
        categorical_indexes
Out[7]: [0, 1]
```

Gambar 5. Posisi Indeks Variabel

Metode Siku (*Elbow*) ini menghitung total *variant cluster* (biaya) untuk berbagai jumlah cluster. Saat kami meningkatkan jumlah cluster, kami berharap variant cluster total menurun. Metode Siku (*Elbow*) menunjukkan jumlah cluster dimana menambahkan lebih banyak cluster tidak akan secara signifikan mengurangi variant cluster total. Nomor ini ditemukan di "siku" plot. Kita dapat melihat bahwa siku terletak di sekitar 3 cluster dapat dilihat pada gambar 6.

```
Out[8]: Text(0, 0.5, 'Cost')
```



Gambar 6. Plot Cluster

4. Data Mining

Proses *data mining* yang diterapkan dalam penelitian ini yaitu *clustering algortima K-prototype*. Setelah melihat metode evaluasi *cluster*, sekarang kita dapat melanjutkan dan menerapkan *algoritma K-prototype* untuk mengelompokkan data. Berikut penerapannya:

- n_cluster = 3: atur jumlah cluster sama dengan 3
- init = "huang": inisialisasi centroid diatur menggunakan metode "huang"
- n_init = 25: jumlah inisialisasi centroid
- random_state = 42: metode inisialisasi centroid

```
In [10]: # check if clusters have been added to the dataframe
         df.head(1069)
```

Out[10]:

	category	sku_desc	sellout_pcs	cluster
0	SOY & FRUIT BEVERAGES	BUAVITA APPLE 1 LT_1C12P_20081902	165	2
1	SOY & FRUIT BEVERAGES	BUAVITA APPLE 125ML_1C40P_20082350	766	2
2	SOY & FRUIT BEVERAGES	BUAVITA APPLE TBA 250ML_1C24P_20082312	1452	0
3	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 125ML_1C40P_20082448	1499	0
4	SOY & FRUIT BEVERAGES	BUAVITA GUAVA 1LT_1C12P_20082239	625	2
...	...	...	...	...
1064	BABY CARE	ZWITSAL MINYAK TELON 60ML _1C24P	65	2
1065	BABY CARE	ZWITSAL MINYAK TELON PLUS 100ML _1C24P_67660399	13	2
1066	BABY CARE	ZWITSAL POWDER EXTRA CARE300GR _1C12P	166	2
1067	BABY CARE	ZWITSAL POWDER NATURAL MILKHONEY300GR _1C12P_YM	0	2
1068	BABY CARE	ZWITSALKIDS HAIRLOTSFTNMST100ML_1C24P_68145824...	14	2

1069 rows × 4 columns

Gambar 7. Plot Cluster Dataframe

5. Knowledge

Langkah berikutnya dari penelitian ini adalah mengeksplorasi *cluster* yang terbentuk. Dari sebelumnya, kita telah menambahkan kolom *cluster* pada dataframe yang terdiri dari 3 *cluster* untuk membantu pengelompokkan data yang kemudian kita harus tahu berapa banyak item di setiap masing-masing *cluster*.

```
In [11]: # size of each cluster
         df["cluster"].value_counts()

Out[11]: 2    978
         0     82
         1      9
         Name: cluster, dtype: int64
```

Gambar 8. Jumlah Anggota Cluster

Berdasarkan percobaan yang sudah dilakukan diperoleh 3 *cluster* terbaik. Yakni pada algoritma *K-prototypes* dengan *cluster* 0 yaitu 82 items serta *cluster* 1 yaitu 9 items, sedangkan *cluster* 2 yang paling banyak yaitu 978 items. Anggota masing-masing cluster sesuai pengelompokan data penjualan produk Unilever di Yogya Toserba.

6. Hasil Clustering

Setelah dilakukan percobaan *cluster* k = 0 sampai dengan k = 2 kita dapat menyimpulkan dari plot dan statistik *cluster* berikut ini:

Sebanyak 82 *record* item produk pada *cluster* 0 yang memiliki rata-rata penjualan sebesar 2.607 pcs dengan memiliki jumlah penjualan dari 1.219 - 6.065 pcs. Dapat dilihat pada gambar 9.

```
In [12]: df[df["cluster"] == 0].describe()
```

Out[12]:

	sellout_pcs	cluster
count	82.000000	82.0
mean	2607.670732	0.0
std	1148.414603	0.0
min	1219.000000	0.0
25%	1708.000000	0.0
50%	2209.500000	0.0
75%	3181.000000	0.0
max	6065.000000	0.0

Gambar 9. Plot Dan Statistik *Cluster* 0

Sebanyak 9 record item produk pada *cluster* 1 yang memiliki rata-rata penjualan sebesar 11.602 pcs dengan memiliki jumlah penjualan dari 7.082 – 23.761 pcs. Dapat dilihat pada gambar 10.

```
In [13]: df[df["cluster"] == 1].describe()
```

Out[13]:

	sellout_pcs	cluster
count	9.000000	9.0
mean	11602.777778	1.0
std	6001.060152	0.0
min	7082.000000	1.0
25%	7934.000000	1.0
50%	9155.000000	1.0
75%	10736.000000	1.0
max	23761.000000	1.0

Gambar 10. Plot Dan Statistik *Cluster* 1

Sebanyak 978 record item produk pada *cluster* 2 yang memiliki rata-rata penjualan sebesar 205 pcs dan memiliki jumlah penjualan dari 0-1.372 pcs. Dapat dilihat pada gambar 11.

```
In [14]: df[df["cluster"] == 2].describe()
```

Out[14]:

	sellout_pcs	cluster
count	978.000000	978.0
mean	205.365031	2.0
std	275.047089	0.0
min	0.000000	2.0
25%	17.000000	2.0
50%	95.000000	2.0
75%	285.000000	2.0
max	1372.000000	2.0

Gambar 11. Plot Dan Statistik *Cluster* 2

5. KESIMPULAN DAN SARAN

Dari hasil setelah di lakukan proses clustering dengan jumlah *cluster* k = 0 sampai k = 2 diperoleh item produk yang termasuk ke dalam *fast moving* dan *slow moving* di Yogya Toserba Purwakarta (PT. Akur Pratama) bisa dilihat dari hasil evaluasi *cluster* menggunakan *algortima K-prototype*. *Cluster* yang mempunyai nilai *sellout* terbesar ada pada *cluster* 0 dan *cluster* 1 sehingga kedua *cluster* tersebut termasuk ke dalam kategori item *fast moving* yaitu 1.219 – 23.761 pcs penjualannya, sedangkan *cluster* yang mempunyai *sellout* terkecil ada pada cluster 2 sehingga termasuk ke dalam kategori item *slow moving* yang mempunyai *sellout* di bawah 1.000 pcs penjualannya oleh karena itu item tersebut harus segera dilakukan proses *delisted* dan digantikan oleh produk baru.

Penelitian ini hanya sebagai bahan masukan kepada toko dalam menyempurnakan aplikasi yang sudah ada karena belum adanya metode yang tepat dalam mengelolah data penjualan berdasarkan karakteristik produk jual yang terbagi menjadi produk laku (*fast moving*) dan produk kurang laku (*slow moving*), diharapkan penelitiaan ini dapat mempersingkat waktu dalam pengolahan data untuk tercapainya penjualan yang maksimal.

DAFTAR PUSTAKA

[1] A. S. Sulthoni, "Segmentasi Pelanggan PT. Telekomunikasi Seluler Indonesia Menggunakan Clustering Algoritma K-prototypes Dan Metode Elbow," 2020.

[2] R. Nooraeni, S. Tinggi, and I. Statistik, "Metode Cluster Menggunakan Kombinasi Algoritma Cluster K-Prototype Dan Algoritma Genetika Untuk Data Bertipe Campuran Cluster Method Using a Combination of Cluster K-Prototype Algorithm and Genetic Algorithm for Mixed Data," *J. Apl. Stat. Komputasi Stat.*, vol. 7, no. 2, pp. 17–17, 2015, [Online]. Available: https://jurnal.stis.ac.id/index.php/jurnalasks/article/view/23.

[3] R. Rusinta, "Pengembangan Sistem Informasi Berbasis Web Kasongan," p. 194, 2007.

[4] M. J. Sablik *et al.*, "No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title," *Acta Mater.*, vol. 33, no. 10, pp. 348–352, 2012, [Online]. Available: http://dx.doi.org/10.1016/j.actamat.2015.12.003%0Ahttps://inis.iaea.org/collection/NCLCollectionStore/_Public/30/027/30027298.pdf?r=1&r=1%0Ahttp://dx.doi.org/10.1016/j.jmrt.2015.04.004.

[5] R. R. Arista, R. A. Asmara, and D. Puspitasari, "Pengelompokan Kejadian Gempa Bumi Menggunakan Fuzzy C-Means Clustering," *J. Teknol. Inf. dan Terap.*, vol. 4, no. 2, pp. 103–110, 2019, doi: 10.25047/jtit.v4i2.67.

[6] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: http://jim.teknokrat.ac.id/index.php/JTSI.

[7] D. Mulyanti, "Analisis Pengendalian Persediaan Buah Segar Hipermarket Giant Poins Lebak Bulus," p. 150, 2011.

[8] S. Kasna, Y. D. L. Widyasari, and Y. Fitrisia, "Implementasi Metode Time Series Dan Metode Fsn Pada Sistem Informasi Apotek Willy Farma," *JSI J. Sist. Inf.*, vol. 12, no. 1, pp. 1906–1916, 2020, doi: 10.36706/jsi.v12i1.9470.

[9] W. O. Zusnita, U. Kaltum, U. W. Pramudya, O. Zusnita, D. Manajemen, and F. Ekonomi, "Pengendalian Persediaan Slow Moving Item PT PLN (Persero) Area Bandung," vol. 5, pp. 412–424, 2018.

[10] F. A. M. Ali, Irfan, Arif Rinaldi Dikananda, "PENGELOMPOKAN JUMLAH PENDUDUK BERDASARKAN KATEGORI USIA 0-18 TAHUN DENGAN MENGGUNAKAN ALGORITMA K-MEANS UNTUK MENENTUKAN PENGEMBANGAN POTENSI DESA WISATA DI KABUPATEN CIREBON Irfan," *J. Manaj. Inform.*, vol. 4, no. 2, 2017, doi: 0.37600/tekinkom.v2i2.115.