

PERBANDINGAN PENGOLAHAN DATA PREDIKSI PERSEDIAAN GAS LPG 3KG PADA PT. BLORA MUSTIKA KOTABUMI MENGGUNAKAN REGRESI LINIER BERGANDA DAN K-MEANS

Annisa Rismanitanti¹, Rima Mawarni², Sidik Rahmatullah³, Dwi Marisa Efendi⁴, Sulis Nurbaiti⁵

^{1,2,3,4,5}STMIK Dian Cipta Cendikia Kotabumi

annisarismanitanti@gmail.com¹, rima@dcc.ac.id², sidik@dcc.ac.id³, dwimarisa89@gmail.com⁴, sulisnurbaiti004@gmail.com⁵

ABSTRACT

The oil and gas sector is a sector that is used with great importance for Indonesia's national development. An interesting commodity to watch out for in the oil and gas industry is liquefied petroleum gas (LPG). At this time PT. Blora Mustika North Lampung does not focus on when household demand will increase and when it will not, current LPG gas sales data are only recorded and not used properly. This of course makes PT Blora Mustika unable to predict demand from sub-distributors and results in frequent LPG gas supply shortages. Therefore, LPG gas sales data will be processed to determine predictions of gas availability and also to choose a more efficient method between Multiple Linear Regression and K-Means. The result is for the Multiple Linear Regression method the predictions are 15504.108 with an error value of 33.85. While the K-Means method produces predictions of 12889.429 with an error value of 86.76. With these results, it can be determined that using Multiple Linear Regression is more efficient.

Keywords—*Data Mining, Multiple Linear Regression, K-Means*

ABSTRAK

Sektor migas merupakan sektor yang digunakan dengan sangat penting bagi pembangunan nasional Indonesia. Komoditas menarik untuk diwaspadai dalam industri migas adalah bahan bakar gas cair (LPG). Pada saat ini PT. Blora Mustika Lampung Utara tidak fokus pada saat kapan kebutuhan rumah tangga akan meningkat dan kapan data-data penjualan gas LPG saat ini hanya sekedar dicatat saja dan tidak digunakan dengan baik. Hal ini tentunya membuat PT Blora Mustika tidak dapat memprediksi permintaan dari sub penyalur dan mengakibatkan sering terjadinya kekosongan persediaan gas LPG. Oleh karena itu data-data penjualan gas LPG akan diolah untuk menentukan prediksi ketersediaan gas dan juga pemilihan metode yang lebih efisien antara Regresi Linier Berganda dan K-Means. Hasilnya adalah untuk metode Regresi Linier Berganda prediksi yang didapat 15504,108 dengan nilai eror 33,85. Sedangkan metode K-Means menghasilkan prediksi 12889,429 dengan nilai eror 86,76. Dengan hasil ini dapat ditentukan bahwa menggunakan Regresi Linier Berganda lebih efisien.

Kata Kunci—*Data Mining, Regresi Linier Berganda, K-Means*

I. PENDAHULUAN

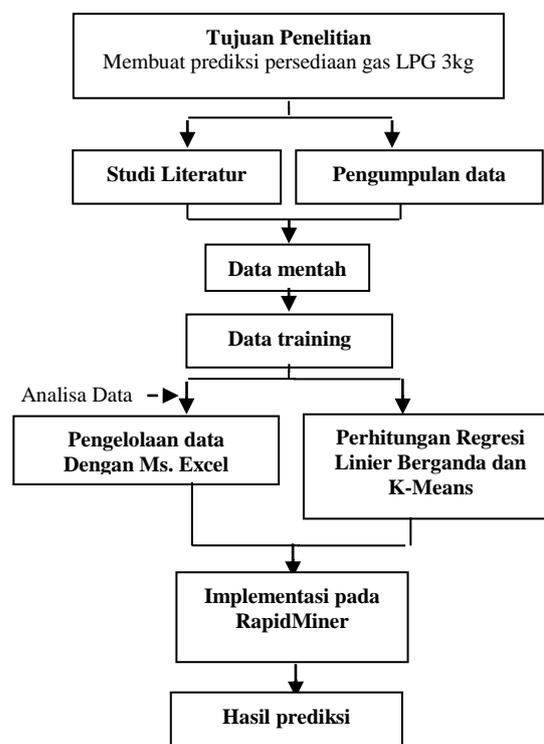
Seiring dengan program konversi energi yang terbagi menjadi LPG ukuran 3 kg (bersubsidi), ukuran 12 kg dan ukuran 50 kg (non subsidi). Keberadaan tiga varian LPG di pasar membawa dampak signifikan terhadap kenaikan permintaan LPG [1].

PT. BLORA MUSTIKA sebagai salah satu agen penyalur gas LPG 3kg di Lampung Utara memiliki banyak permintaan, tentunya agen LPG seperti PT. BLORA MUSTIKA perlu melakukan prediksi terhadap frekuensi/kekerapan pada pembelian guna mengetahui suatu penjualan yang sudah terjual dengan semaksimal mungkin dan juga mengetahui stok yang dapat tersedia sesuai dengan permintaan yang ada.

Persoalan yang ada saat ini yaitu pemilik tidak pernah memperhatikan berapa banyak dalam permintaan konsumen kapan waktunya meningkat dan tidak, artinya disini adalah data-data pembelian gas LPG tidak dipergunakan dengan baik dan hanya sekedar dicatat saja, hal itu tentunya membuat PT BLORA MUSTIKA tidak dapat memprediksi permintaan dari sub penyalur dan mengakibatkan sering terjadinya kekosongan persediaan gas LPG yang menyebabkan masyarakat kesulitan mendapatkan gas LPG 3 Kg.

Data mining adalah kemampuan untuk menganalisis data besar menjadi informasi dalam bentuk pola yang berarti bagi pengambil keputusan. Harus diingat bahwa kata mining itu sendiri berkonotasi dengan upaya mengekstraksi jumlah kecil komoditas berharga dari jumlah yang besar dalam bahan dasar. Maka dari itu, data mining ini sebenarnya berasal dari suatu bidang ilmiah yang melingkupi kecerdasan buatan, suatu pembelajaran tentang mesin, statistika, dan juga pada database. Data mining ini juga dapat diartikan dengan analisis secara otomatis dari suatu data yang besar untuk menemukan pola/tren yang penting dan seringkali luput dari perhatian [1].

II. METODE PENELITIAN



Gambar 1. Kerangka Berikir

Gambar 1 diatas menjelaskan tentang bagaimana tahapan-tahapan yang dilakukan dalam penelitian ini.

Adapun tahapan-tahapan yang dilakukan dengan metode Regresi Linier Berganda dan K-Means adalah sebagai berikut :

2.1 Pembersih Data

Pada tahap ini dilakukan dengan menggunakan data-data penjualan gas LPG 3 Kg selama tahun 2019, 2020, dan 2021 yang berasal dari PT. Blora Mustika Lampung Utara. Yang setelah dilakukan pembersihan data menghasilkan data yang lebih rinci agar dapat memudahkan untuk dihitung. Pada Tabel 1 yang ada dibawah ini merupakan tabel yang berisi data-data setelah dilakukannya pembersihan data:

Tabel 1. Data *Summar* Permintaan Gas LPG 3 Kg Tahun 2019-2021

Bulan	2019	2020	2021
Januari	14.365	13.734	12.388
Februari	10.853	13.148	12.557
Maret	14.824	11.540	12.660
April	14.993	11.500	13.210
Mei	14.758	12.044	12.220
Juni	14.640	12.385	12.660
Juli	14.941	12.140	12.425
Agustus	14.617	12.415	12.710
September	14.998	12.520	12.805
Oktober	14.982	12.620	12.485
November	14.621	12.065	12.840
Desember	14.595	12.540	13.085
JUMLAH	173.187	148.651	152.045

2.2 Analisis Data

Jenis analisa data yang digunakan adalah analisis kuantitatif. Dan data-data yang digunakan adalah data yang ada pada tabel 1 diatas, yang merupakan data-data permintaan gas pada bulan Januari sampai dengan Desember selama tahun 2019 sampai 2021.

2.3 Pengolahan Data

Pada tahap pengolahan data dilakukan dengan menggunakan 2 metode yaitu Regresi Linier Berganda dan juga K-Means.

2.3.1 Regresi Linier Berganda (*Linier Regeression*)

Regresi Linier Berganda adalah salah satu cara meramalkan pada pengaruh terhadap dua variabel prediktor atau lebih yang ada dalam satu variabel guna membuktikan ada atau tidaknya suatu hubungan yang fungsional antara dua variabel bebas (X) atau dengan lebih variabel terikat (Y) [2]. Adapun tahapan-tahapan yang harus dilakukan adalah sebagai berikut [3]:

1. Menganalisis regresi untuk mengukur kekuatan hubungan linier antara dua variabel atau lebih.
2. Menentukan vaiabel bebas dan menghitung persediaan pada tahun 2019, 2020, dan 2021.
3. Menghitung persamaan linier regresi.

4. Implementasikan pada Rapid Miner.

2.3.2 Algoritma K-Means

Algoritma K-Means adalah salah satu cara clustering/pengelompokkan berbasis jarak yang dapat membagi data menjadi beberapa kelompok, algoritma ini juga hanya bisa bekerja dengan atribut angka/numerik [4].

Langkah-langkah yang harus dilakukan pada algoritma K-Means adalah [5] [6] [7]:

1. Menentukan nilai K/Kluster yang ada pada dataset.
2. Menghitung nilai titik pusat/centroid. Penentuan nilai ini pertama kali akan dilakukan acak, dan pada tahap iterasi harus menggunakan rumus :

$$D(a, b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2}$$

Keterangan :

n = jumlah dimensi (atribut)

a_k dan b_k = atribut ke-k dari objek data p dan q

3. Perhitungkan jarak centroid dengan titik tiap objek dengan menggunakan rumus Euclidean Distance dengan rumus seperti yang dibawah ini :

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan :

De = Eulidean Distance

i = Banyaknya objek

(x, y) = Koordinat objek

(s,t) = Koordinat centroid

4. Mengelompokkan objek sesuai dengan jarak ke Centroid terdekat.
5. Ulangi langkah-langkah ke-2 sampai dengan ke-4, dan juga lakukan iterasi hingga nilai dari centroid memberi nilai yang optimal.

2.3.3 RapidMiner

RapidMiner adalah salah satu perangkat lunak yang berhubungan dengan penambangan data. Pekerjaan yang dilakukan oleh penambangan teks RapidMiner berkisar pada analisis teks, mengekstraksi pola dari sekumpulan database dan digabungkannya menggunakan metode *statistic*, *artificial intelligence*, dan *database* [8].

Langkah-langkah yang dilakukan pada RapidMiner antara lain [9][10]:

1. Siapkan *file microsoft excel* data yang akan digunakan mencari hasil agar dapat di hitung kembali untuk diprediksikan.
2. Buka aplikasi RapidMiner, lalu klik *New*.

3. klik *Repositories*, lalu *Import Read Excel*. Kemudian input file yang sudah disiapkan dengan cara *Import Configuration Wizard*.
4. *Select data* yang akan digunakan. Jika sudah lalu *Finish*.
5. Tambahkan atribut sesuai dengan metode yang dibutuhkan *drag* ke *main process*.
6. Untuk metode Regresi Linier Berganda, tambahkan atribut *Apply model*. Sedangkan untuk K-Means menggunakan atribut *Cluster Distance Performance*.
7. Sambungkan/relasikan masing-masing atribut. Kemudian *Running*.

III. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data

3.1.1 Regresi Linier Berganda

Pada tahap pengolahan data dilakukan menggunakan analisis regresi yang digunakan guna mengukur kekuatan suatu hubungan linier antara dua variabel atau lebih.

3.1.1.1 Perhitungan persediaan pada Tahun 2019

Untuk bulan desember tahun 2019 adalah sebagai berikut :

$$Y = b_1 + (b_2 * X_1) + (b_3 * X_2) + (b_4 * X_3)$$

$$Y = 13105,90692 + (0,908278773 * 14365) + (-0,102789014 * 13973) + (-0,035848012 * 18000)$$

$$Y = 13105,90692 + (13444,54343) + (-1438,976609) + (-634,297188)$$

$$Y = 24440,893868$$

$$Y = 24.440 \text{ Tabung gas}$$

3.1.1.2 Perhitungan persediaan pada tahun 2020

Untuk bulan desember tahun 2020 adalah sebagai berikut :

$$Y = b_1 + (b_2 * X_1) + (b_3 * X_2) + (b_4 * X_3)$$

$$Y = -8,742732185 + (0,08594761 * 12540) + (0,777177541 * 12350) + (0,253161938 * 19000)$$

$$Y = -8,742732185 + (1077,783035) + (9598,142626) + (4810,076816)$$

$$Y = 15477,25974$$

$$Y = 15.477 \text{ Tabung gas}$$

3.1.1.3 Perhitungan persediaan pada tahun 2021

Untuk bulan desember tahun 2021 adalah sebagai berikut :

$$Y = b_1 + (b_2 * X_1) + (b_3 * X_2) + (b_4 * X_3)$$

$$Y = -6,356977171 + (-9264,293807 * 13085) + (0,76954874 * 13010) + (0,245730422 * 20000)$$

$$Y = -6,356977171 + (-121223284,5) + (10011,8291) + (4914,608434)$$

$$Y = 12120,8364$$

$$Y = 12.120 \text{ Tabung gas}$$

3.1.1.4 Hasil Implementasi Data Mining

Dari data diatas, penulis akan melakukan perhitungan prediksi menggunakan aplikasi RapidMiner. Data uji yang digunakan untuk implementasi pada RapidMiner kali ini yaitu data bulan Desember dari tahun 2021 yang mana

hasilnya nanti dapat digunakan untuk memprediksi persediaan pada bulan berikutnya yaitu Januari 2022.

Pada Tabel 2 dibawah ini merupakan tabel data uji yang akan digunakan untuk pengujian di RapidMiner.

Tabel 2. Data uji untuk prediksi persediaan tahun 2022

Bulan	Per- min- taan (X1)	Ter- jual (X2)	Harga Jual (X3)	Perse- diaan (X4)
Des	13085	13010	20000	0

Berikut ini merupakan implementasi prediksi dalam RapidMiner menggunakan data training dan data uji diatas.

1. Rangkaikan proses RapidMiner untuk metode Regresi Linier Berganda sebagai berikut :



Gambar 2. Susunan Model Implementasi *Linier regression*

Pada Gambar 2 diatas merupakan gambar rangkaian untuk metode Regresi Linier Berganda dengan cara sambungkan data training ke operator *linier regression* lalu sambungkan ke *apply model*, pada data uji sambungkan ke operator *apply model* lalu hubungkan ke *result*.

2. Setelah seluruh tahapan dilakukan maka *running* proses dan akan menghasilkan hasil prediksi RapidMiner.

Row No.	Tahun	Persediaan	prediction	Permintaan	Terjual	Harga Jual
1	2022	0	15564.108	13085	13010	20000

Gambar 3. Hasil Prediksi *Regresi Linier Berganda*

Pada Gambar 3 diatas merupakan gambar hasil setelah gambar 1 di *running*. Yang menghasilkan prediksi 15504,108.

3.1.2 K-Means

Berdasarkan pada tahap mengelompokkan data-data yang ada pada Tabel 1. Data *Summar* Permintaan Gas LPG 3 Kg Tahun 2019-2021. Proses perhitungan data algoritma K-Means ini menggunakan *Microsoft Excell*. Berikut langkah-langkah penyelesaiannya:

1. Menentukan Jumlah Cluster

Jumlah cluster yaitu jumlah pada kelompok yang dihasilkan dengan menyesuaikan kebutuhan analisis. Dan dalam penelitian ini menggunakan 3 cluster.

2. Menentukan Centroid Awal

Pada tabel 3 dibawah ini merupakan tabel yang berisi Centroid Awal untuk memulai perhitungan metode K-Means.

Tabel 3. Centroid Awal

C1	14.365	13.734	12.388
C2	14.758	12.044	12.220
C3	14.998	12.520	12.805

3. Menghitung Jarak Data ke Titik Pusat Cluster

Pada tahap ini, jarak terdekat antara data dan cluster akan menentukan data mana saja yang termasuk dalam kelompok yang ada. Pada tabel 4 dibawah ini

merupakan hasil dari perhitungan jarak pada literasi ke-1.

Tabel 4. Hasil Literasi ke-1

Data ke-i	C1	C2	C3	Cluster
1	0,00	1743,21	1431,21	0
2	3564,56	4072,03	4199,63	0
3	2257,94	672,29	1005,83	2
4	2461,87	1153,80	1097,47	1
5	1743,21	0,00	791,45	2
6	1403,36	569,04	409,16	1
7	1695,28	291,08	540,42	2
8	1380,92	630,57	406,46	1
9	1431,21	791,45	0,00	1
10	1277,14	672,44	335,64	1
11	1747,97	635,30	591,93	1
12	1401,55	1010,35	491,13	1

4. Menghitung Titik Pusat Centroid Baru

Setelah perhitungan pertama telah menghasilkan keanggotaan cluster, selanjutnya yaitu menentukan titik pusat centroid baru dengan cara menghitung rata-rata dari keanggotaan cluster 1, cluster 2 dan cluster 3, maka akan diperoleh hasil seperti tabel dibawah ini, yang meruakan tabel Centroid literasi ke-2 :

Tabel 5. Centroid literasi ke-2

C1	12.609	13.441	12.473
C2	14.841	11.908	12.435
C3	14.778	12.292	12.828

5. Menghitung Literasi ke-2

pada tabel 6 dibawah ini merupakan hasil dari perhitungan literasi ke-2:

Tabel 6. Hasil Literasi ke-2

Data ke-i	C1	C2	C3	Cluster
1	1782,28	1887,61	1563,01	1
2	1782,28	4178,11	4026,35	0
3	2924,92	431,67	772,02	2
4	3161,46	888,93	905,40	2
5	2575,57	267,60	656,86	2
6	2296,79	564,41	236,31	1
7	2670,78	252,83	460,45	2
8	2267,41	618,75	234,32	1
9	2581,88	732,18	317,56	1
10	2511,04	727,55	516,39	1
11	2465,07	486,90	276,39	1
12	2265,20	939,38	401,30	1

6. Menghitung Titik Pusat Centroid Baru

Pada tabel 7 dibawah ini merupakan hasil perhitungan untuk menentukan Centroid baru guna menghitung literasi ke-3 :

Tabel 7. Centroid Literasi ke-3

C1	10.853	13.148	12.557
C2	14.879	11.806	12.629
C3	14.688	12.611	12.710

7. Menghitung Literasi ke-3

Tabel 8 dibawah ini merupakan hasil dari perhitungan literasi ke-3 :

Tabel 8. Hasil Literasi ke-3

Data ke-i	C1	C2	C3	Cluster
1	3564,56	2009,81	1212,01	1
2	0,00	4244,38	3875,70	0
3	4285,45	273,42	1081,02	2
4	4503,54	666,70	1255,94	2
5	4072,03	488,22	753,12	2
6	3864,47	627,17	236,81	1
7	4212,51	396,12	606,17	2

8	3837,76	667,93	208,83	1
9	4199,63	745,00	336,45	1
10	4163,25	832,99	370,35	1
11	3930,75	422,22	565,46	2
12	3827,66	909,71	392,54	1

8. Menghitung Titik Pusat Centroid Baru

Pada tabel 9 dibawah ini merupakan tabel yang berisi Centroid baru untuk menghitung literasi ke-4 :

Tabel 9. Centroid Literasi ke-4

C1	10.853	13.148	12.557
C2	14.827	11.858	12.671
C3	14.646	12.820	12.655

9. Menghitung Literasi ke-4

Pada tabel 10 dibawah ini merupakan hasil perhitungan pada literasi ke-4:

Tabel 10. Hasil Literasi ke-4

Data ke-i	C1	C2	C3	Cluster
1	3564,56	1952,95	992,75	1
2	0,00	4180,13	3807,93	0
3	4285,45	318,01	1292,15	2
4	4503,54	667,81	1473,46	2
5	4072,03	492,84	896,23	2
6	3864,47	559,62	434,82	1
7	4212,51	391,23	775,92	2
8	3837,76	596,88	409,53	1
9	4199,63	696,83	486,58	1
10	4163,25	799,65	426,45	1
11	3930,75	337,78	777,60	2
12	3827,66	831,15	515,89	1

3.1.2.1 Implementasi Algoritma K-Means pada RapidMiner

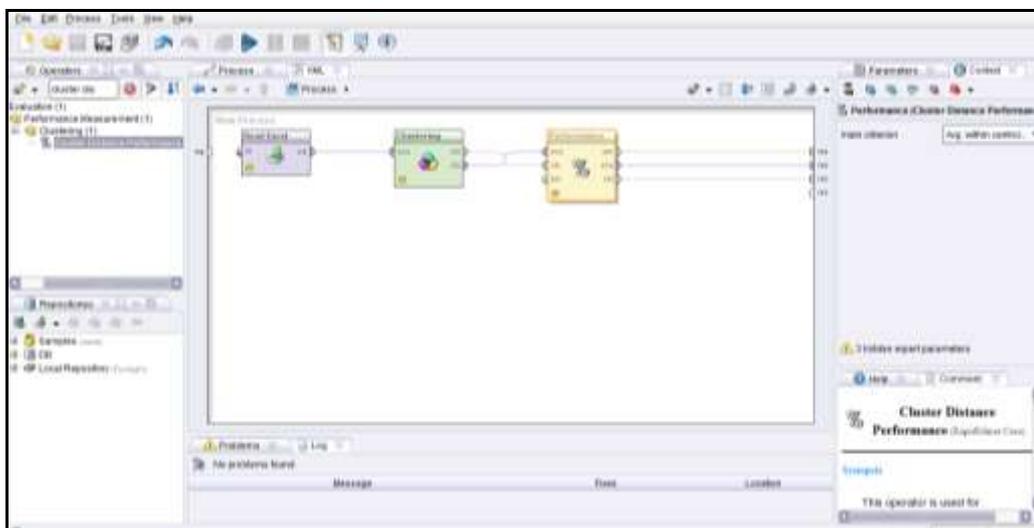
Setelah hasil dari literasi ke-4 tidak banyak berubah dibandingkan dengan

hasil literasi ke-3, maka perhitungannya dihentikan.

Selanjutnya yaitu pengimplementasian menggunakan aplikasi RapidMiner. Yang mana hasilnya nanti dapat diolah lagi sehingga menghasilkan hasil perhitungan pada tahun 2019 cluster 0 menghasilkan 10853, cluster 1 menghasilkan 14699, dan cluster 2 menghasilkan 14827. Pada tahun 2020 cluster 0 menghasilkan 13148,

cluster 1 menghasilkan 12702, dan cluster 2 menghasilkan 11857. Dan pada tahun 2021 cluster 0 menghasilkan 12557, cluster 1 menghasilkan 12688, dan cluster 2 menghasilkan 12671. Hasil-hasil tersebut didapat dari perhitungan pada RapidMiner yang ada pada tabel 11.

1. Rangkaikan proses RapidMiner untuk metode K-Means sebagai berikut:

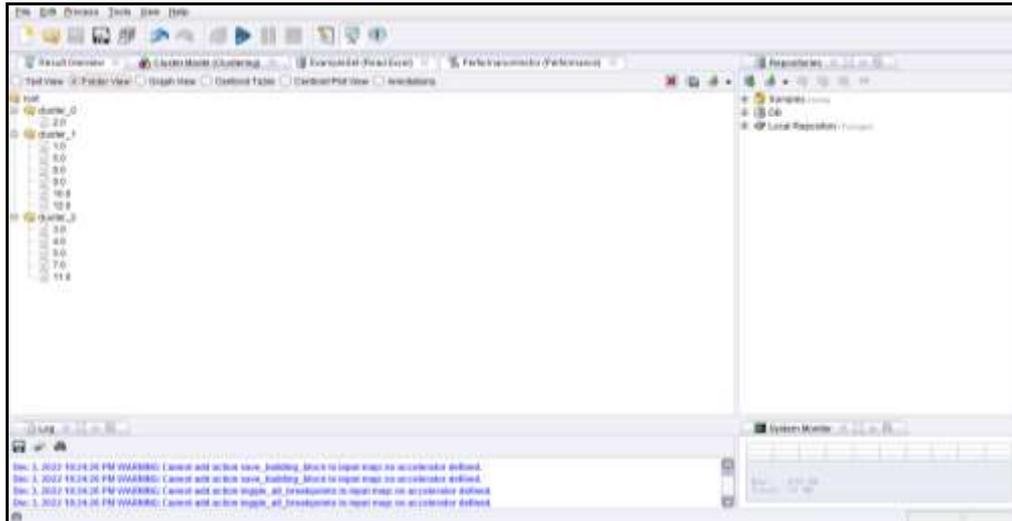


Gambar 4. Susunan jika Terkoneksi

Pada Gambar 4 di atas merupakan gambar rangkaian yang harus disambungkan agar bisa di *running*, dengan cara sambungkan *Read Excell*, *K-Means* dan juga operator

Cluster Distance Performance lalu hubungkan kedalam *result*

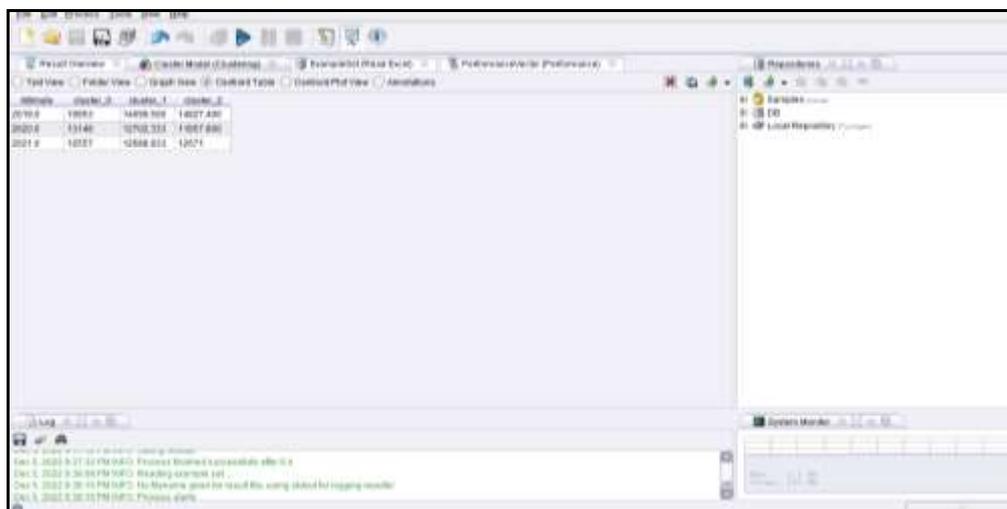
2. Setelah seluruh tahapan dilakukan maka *running* proses



Gambar 5. Hasil *Running* pada *Folder View*

Pada gambar 5 diatas merupakan gambar hasil *running* yang kalau di bandingkan dengan perhitungan *excel* sebelumnya maka hasilnya akan sama. Yang mana

tahapan-tahapan perhitungan excel maupun RapidMiner sudah berhasil dan juga benar.



Gambar 6. Hasil *Centroid Table*

Pada gambar 6 diatas merupakan gambar hasil pada RapidMiner. Karena tidak mendapatkan hasil prediksi langsung, maka dilakukan perhitungan

menggunakan excel maka akan mendapatkan hasil seperti Tabel 11 yang ada pada dibawah ini :

Tabel 11. Hasil Perhitungan Prediksi

	Cluster 0	Cluster 1	Cluster 2	Rata-rata pertahun
2019	10853	14699,5	14827,4	4486,655556
2020	13148	12702,33	11857,8	4189,792556
2021	12557	12688,83	12671	4212,981444
Rata-rata dari 3 Cluster			12889,42956	

3.1.3 Hasil

Pada perhitungan menggunakan regresi linier berganda didapat hasil prediksi 15504,108 yang ada pada hasil implementasi RapidMiner gambar 3. Hasil tersebut akan dikurangkan dengan perhitungan pada tahun 2021 untuk mengetahui nilai eror yang didapat dengan cara :

$$Ne = (\text{Hasil prediksi} - \text{Perhitungan Tahun 2021}) / 100$$

$$Ne = 15504,108 - 12120 = 3385 / 100 \\ = \mathbf{33,85}$$

Sedangkan pada metode K-Means diperoleh nilai prediksi 12889,429 dari hasil perhitungan prediksi pada tabel 11. Dan untuk mencari nilai eror akan dikurangkan dengan hasil perhitungan pada tahun 2021 yang ada pada tabel 11 dengan cara :

$$Ne = (\text{Hasil rata-rata 3 cluster} - \text{Rata-rata tahun 2021}) / 100$$

$$Ne = (12889,429 - 4212,981) / 100 \\ = \mathbf{86,76}$$

Tabel 12. Perbandingan Hasil

Metode	Hasil
<i>Regresi Linier Berganda</i>	33,85
K-Means	86,76

IV. KESIMPULAN

Perbandingan antara memprediksi persediaan gas lpg 3 kg pada PT. Blora Mustika menggunakan Regresi Linier Berganda dan K-Means menunjukkan bahwa memprediksi menggunakan metode Regresi Linier Berganda efisien karena nilai kesalahan (error) yang diperoleh 33,85 yaitu cenderung lebih kecil dari metode K-Means yang menghasilkan 37,81.

Adapun saran dari penulis untuk penelitian selanjutnya agar dapat menggunakan lebih banyak sumber referensi, dan dapat menghasilkan hasil penelitian yang lebih baik lagi.

DAFTAR PUSTAKA

- [1] P. Purwadi, P. S. Ramadhan, and N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk

- Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 1, p. 55, 2019, doi: 10.53513/jis.v18i1.104.
- [2] Y. Asohi and A. Andri, “Impelementasi Algoritma Regresi Linier Berganda Untuk Prediksi Penjualan,” *J. Nas. Ilmu Komput.*, vol. 1, no. 3, pp. 149–158, 2020, doi: 10.47747/jurnalnik.v1i3.161.
- [3] I. L. L. Gaol, S. Sinurat, and E. R. Siagian, “Implementasi Data Mining Dengan Metode Regresi Linear Berganda Untuk Memprediksi Data Persediaan Buku Pada Pt. Yudhistira Ghalia Indonesia Area Sumatera Utara,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 130–133, 2019, doi: 10.30865/komik.v3i1.1579.
- [4] R. A. Indraputra and R. Fitriana, “K-Means Clustering Data COVID-19,” *J. Tek. Ind.*, vol. 10, no. 3, pp. 275–282, 2020, [Online]. Available: <https://www.trijurnal.lemlit.trisakti.ac.id/index.php/tekin/article/view/8428/6033>
- [5] A. ett all Wanto, *Data Mining : Algoritma dan Implementasi*. Jakarta: Kita Menulis, 2020.
- [6] A. R. Isnain, J. Supriyanto, and M. P. Kharisma, “Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 2, p. 121, 2021, doi: 10.22146/ijccs.65176.
- [7] D. Alita, I. Sari, and A. Rahman Isnain, “Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa,” *Jdmsi*, vol. 2, no. 1, p. 702022, 2021.
- [8] M. Faid, M. Jasri, and T. Rahmawati, “Perbandingan Kinerja Tool Data Mining Weka dan RapidMiner Dalam Algoritma Klasifikasi,” *Teknika*, vol. 8, no. 1, pp. 11–16, 2019, doi: 10.34148/teknika.v8i1.95.
- [9] S. Wahyuni, K. Saputra, and M. I. Perangin-Angin, “Implementasi RapidMiner Dalam Menganalisa Data Mahasiswa Drop Out,” vol. 10, no. 2, pp. 2013–2016, 2017.
- [10] W. Sudrajat, I. Cholid, M. Rachmadi, and V. Amalia, “Penerapan Algoritma K-Means

Dalam Segmentasi Pelanggan Pada
Toko Sembako Menggunakan
RapidMiner Application of the K-
Means Algorithm in Customer
Segmentation at Grocery Stores
Using RapidMiner,” *J. Ilm. Bin.
STMIK Bina Nusant. Jaya*, vol. 0,
pp. 2657–2117, 2021.