1. **Business Understanding & Data Understanding**

Sebelum melihat business understanding mari kita lihat terlebih dahulu data yang diberikan, yaitu alerts, irregularties, dan jams di Cirebon Jawa Barat. Dalam data pertama yaitu **data alerts** termuat column dengan pengertian sbb:

|  |  |  |
| --- | --- | --- |
| Id | : | Serial |
| Time | : | time |
| kemendagri\_kabupaten\_kode | : | Kemendagri code of cities |
| kemendagri\_kabupaten\_nama | : | Kemendagri name of cities |
| **street** | **:** | **Nama jalan** |
| **Type** | **:** | **Jenis alert** |
| Avg\_location | : | Geometry data (spatial data) |
| **date** | **:** | **Tanggal** |
| **Total\_records** | **:** | **Total data yang terekam** |

Dalam data alerts untuk column yang penting adalah street, date, total\_records dan type. Kita dapat melihat insight seperti:

1. mana saja jalan yang memuat alerts tersebut?
2. Pada tanggal berapa alerts tersebut muncul?
3. Dan alerts yang muncul apakah JAM atau WEATHERAZARD atau ACCIDENT?

Dalam data kedua yaitu data irregularities, termuat column dengan pengertian sbb:

|  |  |  |
| --- | --- | --- |
| Time | : | time |
| kemendagri\_kabupaten\_kode | : | Kemendagri code of cities |
| kemendagri\_kabupaten\_nama | : | Kemendagri name of cities |
| street | : | Nama jalan |
| Jam\_level | : | Tingkat kemacetan |
| Median\_length | : | Median irregularities length |
| Median\_delay\_sec | : | Median delay sec from regular speed |
| Median\_regular\_speed | : | Median historical reg speed in kmh |
| Total\_records | : | Total daya yang terekam |
| Median\_sec | : | Minus the free flow time (second) |
| Median\_speed | : | Median traffic speed in irregularity |
| Date | : | Date |
| Median\_jam\_level | : | Median level kemacetan |
| id | : | Row id |
| geometry | : | Geometry data (spatial data) |

Dalam data irregulartieis, kita dapat melihat kejanggalan-kejanggalan pada hari normal, data yang dapat dilihat adalah jam\_level, total\_records, dan date.

Data yang ketiga adalah data jams, termuat column dengan pengertian sbb:

|  |  |  |
| --- | --- | --- |
| Time | : | time |
| kemendagri\_kabupaten\_kode | : | Kemendagri code of cities |
| kemendagri\_kabupaten\_nama | : | Kemendagri name of cities |
| **street** | **:** | **Nama jalan** |
| level | : | Level kemacetan |
| **Median\_length** | **:** | **Median irregularities length** |
| **Median\_delay\_sec** | **:** | **Median delay sec from regular speed** |
| **Median\_speed\_kmh** | **:** | **Median historical reg speed in kmh** |
| Total\_records | : | Total daya yang terekam |
| Id | : | Row id |
| Date | : | Tanggal |
| Median\_level | : | Median kemacetan |
| geometry | : | Spatial data |

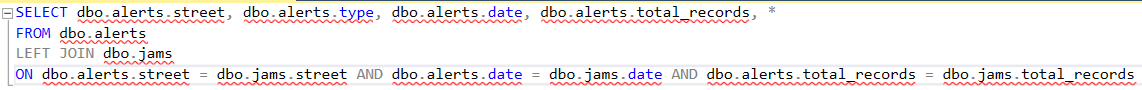
Setelah mendapat data ketiga tersebut diatas, kita langsung join kan saja untuk melihat data mana saja yang sama. Seperti data street, jam\_level, dan date dengan SMSS 19.

Dalam case kali ini, **bisnis understandingnya** adalah melihat alerts (kemacetan, kecelakaan, dan cuaca buruk) dan setelah melihat alerts tersebut manakah alerts yang banyak terjadi di cirebon, dan bagaimana pemangku kepentingan dapat mengambil kebijakan atas alerts terkait?

1. **Data Cleansing and Preprocessing (termasuk uji statistik, jika menggunakan)**

Data cleaning pertama menggunakan SMSS 19 untuk menjoinkan data alerts dan jams. Untuk data irregularities masih belum digunakan karena hipotesis penulis adalah irregularities jarang terjadi karena sifatnya tidak sering (irregular) dari hari-hari biasanya, maka dari itu yang dijoinkan hanya data alerts dan jams.

Join data alerts dan jams



Data yang sudah terjoin di save di csv (terlampir dengan nama dokument ‘alerts\_jams.csv’). CSV selanjutnya dibersihkan dengan menggunakan excel, delete colum yang duplicate, penghapusan null. Total data yang terjoin ada 708 rows dan 20 column Untuk colum yang duplicate dan tidak dibutuhkan ada 11 column (ada yang terduplicate dan ada yang tidak digunakan), yaitu:

|  |  |  |  |
| --- | --- | --- | --- |
| Kemendagri\_kab\_code  (**duplicate**) | Total records  (**duplicate**) | Street  (**duplicate**) | Date  (**duplicate**) |
| Kemendagri\_kab\_name  (**duplicate**) | Jam\_level  (**duplicate**) | ID1  (**tidak digunakan**) | ID2  (**tidak** **digunakan**) |
| Geometry  (**tidak digunakan**) | Avg\_location  (**tidak digunakan**) | Type  (**duplicate**) |  |

Sehingga kolom yang digunakan adalah 9 kolum yaitu:

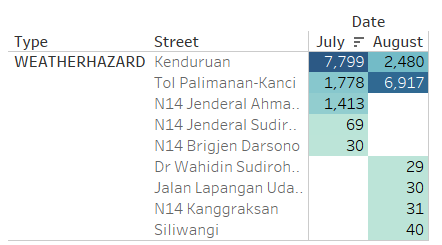
|  |  |  |  |
| --- | --- | --- | --- |
| Street | Type | Date | Total\_records |
| Time | Jam\_level | Median\_length | Median\_delay |
| Median\_speed\_kmh |  |  |  |

Akan tetapi jika menggunakan 9 column ada 599 data null, maka dari itu penulis akan memisahkan analisa untuk column **street, type, date**, dan **total records** untuk **analisa pertama.** Csv terlampir (nama dokumen analisa1\_alerts\_jam.csv).

Dalam analisa pertama untuk preprocessing terdapat 708 rows dengan data null sebanyak 25 data untuk kolom street, sehingga data row street dihapuskan, sehingga data yang digunakan sebanyak 683 rows dan 4 columns.

1. **Modelling**

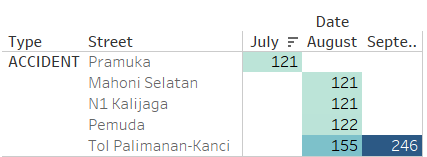
Dalam **analisa pertama** diketahui ada beberapa laporan mengenai alert terkait weatherharzard, jam, dan accident. Berikut laporan mengenai weatherhazard



Sumber: Data diolah (Tableu), 2023

Pada july-agustus 2022 ada 10.279 laporan mengenai cuaca buruk di Jl. kenduruan dan 8.695 laporan di tol palimanan-kanci, cuaca ini disebabkan karena sudah memasuki musim penghujan. Dalam laporan cuaca buruk sebaiknya para pengendara melihat atau mendengar berita-berita terkait cuaca di daerah cirebon dan prediksi cuaca sebelumnya, sebelum melakukan perjalanan. Cuaca adalah salah satu faktor yang tidak dapat dihindari. Setidaknya pemerintah daerah membuat peringatan melalui berita atau media sosial resmi untuk berhati-hati dalam mengendarai saat cuaca buruk.

Berikut laporan mengenai accident

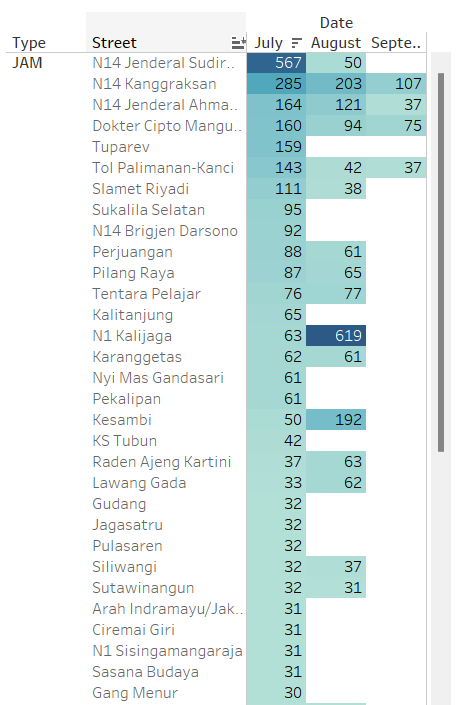


Sumber: Data diolah (Tableau), 2023

Pada july-september 2022 ada 121 laporan kecelakaan di jl pramuka, cirebon pada bulan juli 2022, 519 laporan kecelakaan di berbagai tempat seperti mahoni selatan, N1 kalijaga, pemuda, dan tol palikanci pada agustus 2022, dan 246 laporan kecelakaan di tol palikanci pada september 2022. Tol palikanci adalah lokasi yang paling sering mendapatkan laporan kecelakaan terlebih pada bulan september 2022, penulis melihat berita terkait kecelakaan dengan bulan dan tahun yang sama dengan sumber (Prihatnomo, 2022), jalan tol memang lebih berisiko mengalami kecelakaan dikarenakan berbagai macam faktor seperti kecepatan yang tinggi, cuaca, faktor eksternal seperti kabut atau kepulan asap dari kendaraan lain dan lain-lain. Sebaiknya pengendara melakukan hal-hal ini agar menimimalisir kecelakaan yang terjadi di jalan tol (Agustiar, 2022):

1. mematuhi aturan batas kecepatan maksimal
2. Jaga jarak kendaraan
3. Berhati-hati jika kontur jalan tidak rata atau bergelombang
4. Memperhatian Penyempitan jalan
5. Memperhatiakan tanjakan terjal dan/atau turunan curam yang panjang
6. Mewaspadai truk dengan muatan berlebih
7. Beristirahat di rest area jika kelelahan atau mengantuk
8. Memperhatikan kondisi kendaraan sebelum melakukan perjalanan

berikut laporan mengenai kemacetan

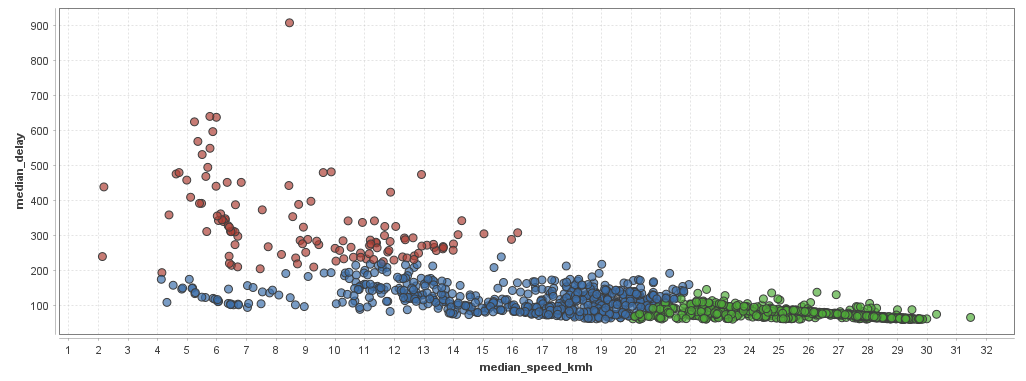


Sumber: Data diolah (Tableau), 2023

Pada july-september 2022 ada banyak laporan mengenai kemacetan di cirebon. Data menyebutkan bahwa setidaknya kemacetan yang sering terjadi ada di jalan N14 Jenderal Sudirman, N14 Kanggraksan, N14 Jenderal Ahmad Yani, dan N1 kalijaga. Pada bulan july 2022 ada 567 laporan mengenai kemacetan tertinggi ada di N14 Jenderal sudirman, dan pada bulan agustus 2022 kemacetan tertinggi di N1 Kalijaga dengan 619 laporan, dan bulan agustus 2022 ada 107 laporan dengan kemacetan tertinggi di jalan N14 Kanggraksan. Penulis melihat berita kemacetan di cirebon tidak ada berita yang relevan pada bulan juli 2022, sedangkan di bulan agustus 2022 ada berita yang relevan dengan sumber (Baehaqi, 2022), berita tersebut jl N1 kalijaga macet disebabkan karena adanya rekonstruksi jembatan kanci, sehingga jl kaligaja harus digunakan, jl sudirman pun terkena imbasnya karena jembatan kalijaga terusannya adalah jl jenderal sudirman. Kemacetan yang terjadi di jalan jenderal sudirman harus diteliti lebih lanjut karena lokasi tersebut memiliki kemacetan yang sering terjadi. Selanjutnya untuk analisa kedua adalah melihat lebih spesifik untuk alerts kemacetan di Jl. Jenderal Sudirman, data terlampir (dengan nama dokumen **‘analisa2\_jams\_sudirman.csv**’).

Berdasarkan data median\_jams penulis mengambil hanya jl jenderal sudriman yang selanjutnya dianalisis dengan cara clustering antara median\_delay dan median\_speed\_kmh untuk mengetahui seberapa lama kemacetan itu terjadi dan dengan kecepatan berapa kmh. disini penulis membersihkan data dengan menggunakan excel dengan cara membersihkan data yang kosong, row yang tidak tersusun rapih, dan menghilangkan outlier (data yang digunakan terlampir). Cara clustering yang digunakan adalah dengan model Kmeans. K-means adalah salah satu algoritma dengan sifat *unsupervised learning* dengan fungsi mengelompokakan data ke dalam data cluster (Dhuhita, 2015). Penulis disini menggunakan software RapidMiner Studio, **karena penulis tidak fasih dalam menggunakan programming tools** seperti python.

Berikut adalah hasil dari clustering menggunakan k-means dengan K=3.



Sumber: Data diolah (RapidMiner), 2023

Insight yang didapat adalah sebagai berikut:

1. Cluster 0 (berwarna biru) memiliki 564 data dengan **median\_speed\_kmh** rata-rata adalah 22.99% lebih kecil, dan **median\_delay** rata-rata 4.11% lebih besar.
2. Cluster 1 (berwarna hijau) memiliki 646 data dengan **median\_delay** rata-rata adalah 72.55% lebih kecil dan median **median\_speed\_kmh** rata-rata 31.15% lebih besar
3. Cluster 2 (berwarna merah) memiliki 118 data dengan median\_delay rata-rata 377.52% lebih besar median\_speed\_kmh rata-rata 60.66% lebih kecil

Berikut adalah tabel centroid dari hasil modelling

|  |  |  |
| --- | --- | --- |
| Cluster | Median\_delay | Median\_speed\_kmh |
| Cluster 0 | 118.824 | 16.008 |
| Cluster 1 | 75.509 | 25.764 |
| Cluster 2 | 329.797 | 9.219 |

Sumber: Data diolah, 2023

Hasil clustering k-means terlampir dengan nama dokumen kmeans-hasil.xlsx.

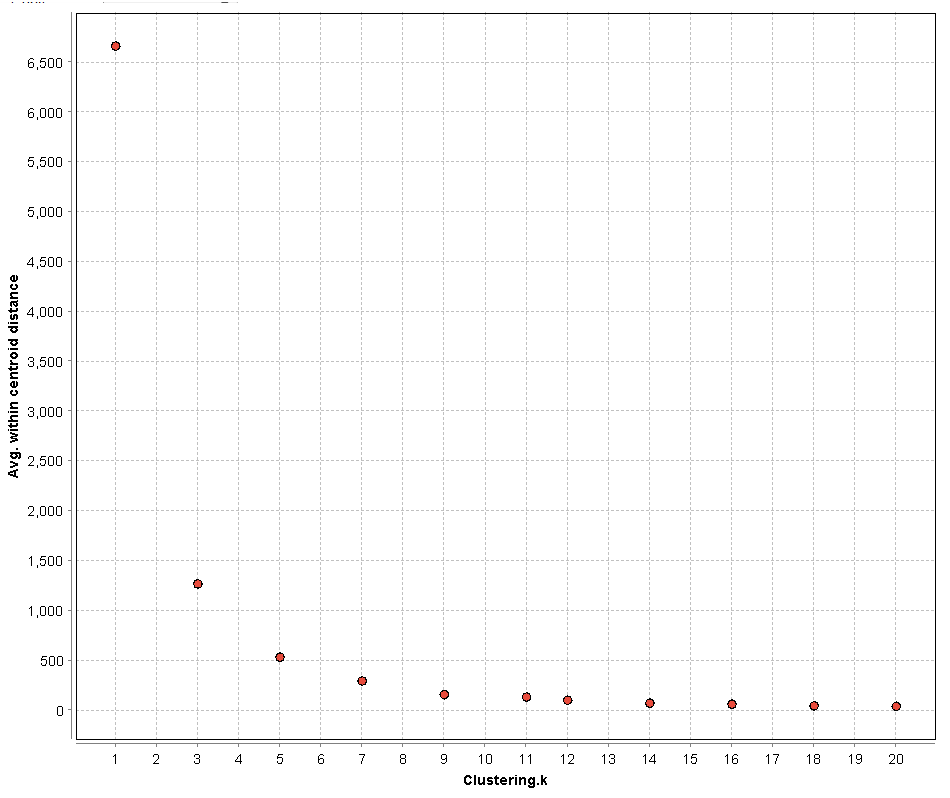
1. **Evaluation**

Berdasarkan (Dabbura, 2018), ‘*clustering analysis doesn’t have a solid evaluation metrics that we can used to evaluate the outcome of different clustering’,* dikarenakan kmeans membutuhkan *k* sebagai input dan tidak dapat mempelajari itu dari data, maka tidak ada jawaban yang benar dalam istilah *number of cluster.* Tetapi biasanya dalam evaluasi clustering, beberapa instituti menggunakan 2 metode evaluasi yaitu:

1. Elbow methods
2. Silhouette analysis

Dalam kasus yang ini penulis mengambil *elbow method.* Elbow method adalah metode yang digunakan untuk menentukan jumlah cluster (k) yang tempat melalui persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku (elbow) pada suatu titik (Dewi & Pramita, 2019).

Berikut adalah hasil **elbow method** :

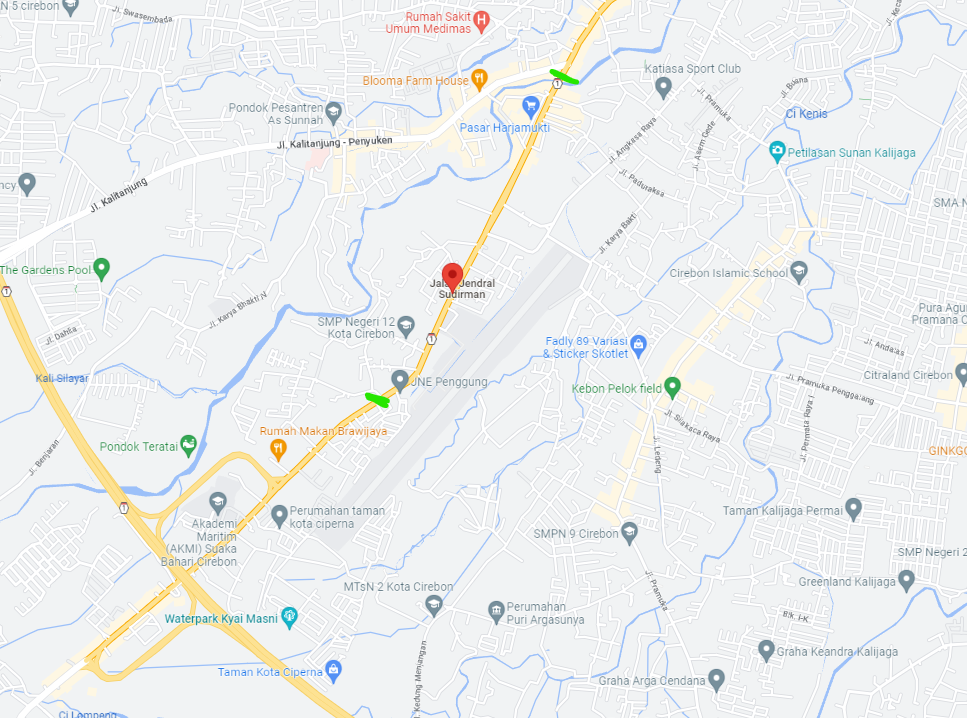


Sumber: Data diolah (RapidMiner), 2023

Dapat dilihat pada hasil diatas jumlah kluster (k) yang terbaik adalah 3 cluster. Dalam modelling juga penulis menggambil k=3 sebagai banyaknya cluster, sehingga jumlah cluster yang digunakan adalah tepat.

1. **Rekomendasi untuk pemerintah daerah adalah**

Berikut adalah jalan jenderal soedirman, kab cirebon.



Catatan: Jl Jenderal Sudirman adalah yang ditandai dengan warna hijau

Sumber: (Google, 2023)

1. Berdasarkan hasil analisa dan pengamatan jl jenderal sudirman rawan akan kemacetan karena beberapa hal seperti:
2. Dekat dari masuk/keluar jalan tol palimanan-kanci
3. Dekat dengan bandara penggung
4. Dekat dengan pasar harjamukti
5. Jalan utama masuk ke kota cirebon
6. Berdasarkan hasil analisa data jika suatu kemacetan sudah masuk dalam kategori macet jika kendaraan mempunyai kecepatan <31.5 kmh dengan jarak delay adalah 100 second (atau 1 menit 40 detik)
7. Berdasarkan (Bull, 2003), ada beberapa hal yang bisa dilakukan oleh pemerintah daerah dengan beberapa strategi yaitu dari sisi:

* *supply-side*
* Memperhatian *intersection* dan *intersection control systems*
* Melebarkan jalan
* Melakukan *contraflow*
* Jalur khusus kendaraan tertentu (seperti *bus-only lanes)*
* Demand-side
* *Controlling parking*
* *Flex time*
* Kendaraan yang dilarang untuk melalui jalan tersebut

1. Pemerintah harus mengkaji ulang apakah strategi diatas bisa untuk dieksekusi atau tidak melihat kebutuhan masyarakat.
2. **Referensi**

Agustiar, D. (2022). *Tips Menghindari Tabrakan Beruntun di Tol*. Idntimes.Com. https://www.idntimes.com/automotive/car/dwi-agustiar/tips-dan-cara-menghindari-kecelakaan-beruntun-di-jalan-tol

Baehaqi, A. I. (2022). *Antisipasi Macet Saat Rekonstruksi Jembatan Kanci, Satlantas Polresta Cirebon Siapkan Langkah Ini*. Tribuncirebon.Com. https://cirebon.tribunnews.com/2022/08/10/antisipasi-macet-saat-rekonstruksi-jembatan-kanci-satlantas-polresta-cirebon-siapkan-langkah-ini

Bull, A. (2003). *Traffic Congestion: The Problem and How to Deal With It*. United Nations Publication.

Dabbura, I. (2018). *K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks*. Towardsdatascience.Com. https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a#:~:text=Compute the sum of the,that belong to each cluster

Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix : Jurnal Manajemen Teknologi Dan Informatika*, *9*(3), 102–109. https://doi.org/10.31940/matrix.v9i3.1662

Dhuhita, W. (2015). Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita. *Jurnal Informatika Darmajaya*, *15*(2), 160–174.

Google. (2023). *Maps Jl. Jenderal Sudirman, Cirebon*. Google Maps. https://www.google.com/maps

Prihatnomo, J. (2022). *Kecelakaan Beruntun Terjadi di Tol Brebes, Akibat Kepulan Asap Menutupi Jalur Tol*. Suaramerdeka.Com. https://www.suaramerdeka.com/jawa-tengah/pr-044794543/kecelakaan-beruntun-terjadi-di-tol-brebes-akibat-kepulan-asap-menutupi-jalur-tol?page=2