

Analisis Data Penyakit DBD Dengan *K-Means Clustering* di Kabupaten Bantul Menggunakan Data mining

Rizki Khusna Utami¹, Syamsu Windarti^{2*}, Muhammad Muslim³

¹Prodi DIV Manajemen Informasi Kesehatan, STIKES AKBIDYO Yogyakarta

*email : windartisyamsu@gmail.com

ABSTRACT

The Bantul Health Service (Dinkes) noted that there will be an increase in dengue cases in 2021 and 2022, namely 410 cases and 885 cases. Clustering data mining techniques with Orange tools can be used to find clusters of DHF cases in Bantul Regency. Therefore, clustering has been carried out with the K-means clustering technique using Orange tools to determine the level of distribution of DHF cases in Bantul Regency. With -Means Clustering, 3 clusters are produced, namely high, medium and low. The Bantul Health Office has been trying to tackle DHF with PSN, larvicidation, fogging, counseling and Wolbachia 2022. The use of SISKLB is a reminder when DHF cases are continuously high. The results of the One Way Anova test showed a significant difference in the average number of DHF cases with a p value of 0.000.

Keywords : *Dengue Hemorrhagic Fever, Data mining, K-means clustering*

ABSTRAK

Dinas Kesehatan (Dinkes) Bantul mencatat kasus DBD tahun 2021 dan 2022 mengalami kenaikan yaitu 410 kasus dan 885 kasus. Teknik *clustering* data mining dengan *tools orange* dapat digunakan untuk menemukan *cluster* kasus DBD di Kabupaten Bantul. Maka dari itu telah dilakukan *clustering* dengan teknik *K-means clustering* menggunakan *tools orange* untuk mengetahui tingkat persebaran kasus DBD di Kabupaten Bantul. Dengan *K-means clustering* dihasilkan 3 *cluster* yaitu tinggi, sedang dan rendah. Dinkes Bantul sudah berupaya menanggulangi DBD dengan PSN, larvasidasi, *fogging*, penyuluhan dan *Wolbachia* 2022. Pemanfaatan SISKLB terdapat pengingat saat kasus DBD tinggi terus-menerus. Hasil uji *Anova one way* terdapat perbedaan rata-rata jumlah kasus DBD secara signifikan dengan *p value* 0.000.

Kata kunci: *Demam Berdarah Dengue, Data mining, K-means clustering*

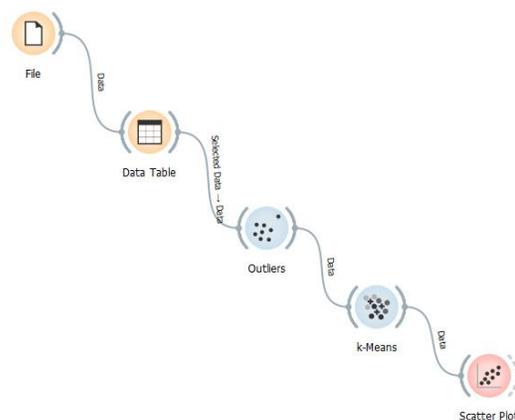
PENDAHULUAN

Penyakit Demam Berdarah *Dengue* (DBD) banyak ditemui pada saat musim hujan. Nyamuk *Aedes aegypti* merupakan vektor utama penyakit DBD yang berkembang di daerah tropis dan sub tropis (Sutriyawan dkk, 2020). Dinas Kesehatan (Dinkes) Bantul bagian Pencegahan dan Pengendalian Penyakit (P2P) mencatat kasus DBD pada tahun 2021 dan 2022 mengalami kenaikan yaitu 410 kasus dan 885 kasus. Data yang telah didapatkan dapat diolah kedalam data mining. Pemanfaatan data mining dapat digunakan bersamaan *clustering* yang dapat memanfaatkan *tools orange* yang merupakan salah satu aplikasi digunakan untuk data mining dan machine learning. *Tools orange* adalah teknologi pembelajaran berbasis open source atau perangkat lunak data mining (Nurhafida dan Sembiring, 2021). Penerapan metode *K-means clustering* memiliki keunggulan untuk menghasilkan informasi dengan *cluster* yang lebih rinci. *K-means clustering* akan mengelompokkan item data dalam suatu dataset ke suatu *cluster* berdasarkan jarak terdekat (Prianto dan Bunyamin, 2020). Fungsi *clustering* dalam rangka pengendalian penyakit DBD di Kabupaten Bantul digunakan untuk mengetahui daerah mana yang *cluster* persebaran penyakit DBD *cluster* tinggi, sedang atau rendah.

METODE PENELITIAN

Penelitian dilakukan dengan metode kuantitatif menggunakan teknik *K-means clustering*. Populasi penelitian adalah jumlah penderita penyakit DBD di Kabupaten Bantul. Sampel penelitian adalah data penderita penyakit DBD per kecamatan tahun 2020 sampai dengan tahun 2022 di Kabupaten Bantul. Data yang digunakan pada penelitian adalah data

sekunder dengan teknik pengumpulan data dokumentasi. Metode pengolahan data yaitu menggunakan *tools orange* seperti pada gambar 1.



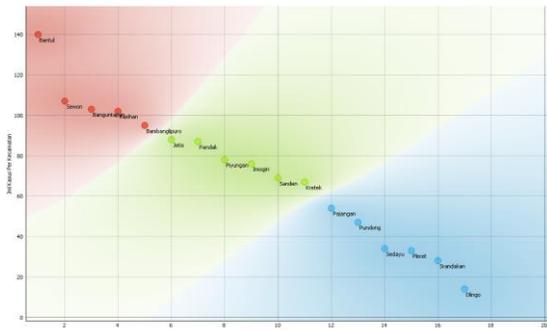
Gambar 1. *Widget tools orange*

Berdasarkan Gambar 1 dapat dilihat desain *widget K-means clustering*, pada *widget file* digunakan untuk memilih *dataset* dan membaca nilai atribut dari input file. *Widget Data Tabel* dapat menampilkan data atribut nilai dalam sebuah *spreadsheet*. *Widget Outlier* untuk mendeteksi outlier dengan membandingkan jarak. *Widget K-means* digunakan untuk mengelompokkan item/data menggunakan algoritma *K-means clustering*. Terakhir, *widget scatter plot* adalah bentuk visualisasi dengan kemampuan analisis eksploratif dan peningkatan visualisasi data cerdas. Analisis data yang digunakan dalam penelitian adalah *Anova one way* dengan menggunakan *Statistical Package for the Social Sciences (SPSS)* untuk mengetahui apakah ada perbedaan yang signifikan antar kelompok.

HASIL

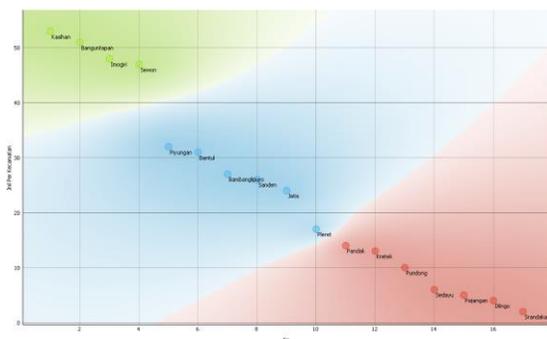
Hasil dari penelitian yang telah dilakukan, maka didapatkan hasil visualisasi *scatter plot* penyakit DBD

seperti gambar 2, 3 dan 4. Data visual *scatter plot* menampilkan persebaran kasus DBD di Kabupaten Bantul.



Gambar 2. Visualisasi *scatter plot* penyakit DBD tahun 2020 menggunakan *tools orange*

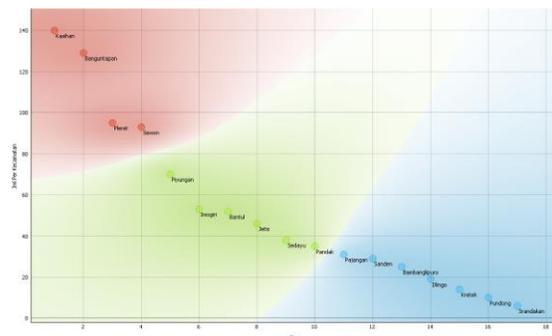
Berdasarkan gambar 2 terdapat 3 *cluster* yaitu *cluster 1* untuk skala tinggi, *cluster 2* untuk skala sedang dan *cluster 3* untuk skala rendah. *cluster 1* berwarna merah, *cluster 2* berwarna hijau dan *cluster 3* berwarna biru. *Cluster 1* terdapat Kecamatan Bantul, Sewon, Banguntapan, Kasihan dan Banganglipuro. Selanjutnya pada *cluster 2* terdapat Kecamatan Jetis, Pandak, Piyungan, Imogiri, Sanden dan Kretek. Kemudian *cluster 3* terdapat Kecamatan Pajangan, Pundong, Sedayu, Pleret, Srandakan dan Dlingo.



Gambar 3. Visualisasi *scatter plot* penyakit DBD tahun 2021 menggunakan *tools orange*

Berdasarkan gambar 3 terdapat 3 *cluster* yaitu *cluster 1* untuk skala tinggi,

cluster 2 untuk skala sedang dan *cluster 3* untuk skala rendah. *cluster 1* berwarna hijau, *cluster 2* berwarna biru dan *cluster 3* berwarna merah. *Cluster 1* terdapat Kecamatan Kasihan, Banguntapan, Imogiri dan Sewon. Selanjutnya pada *cluster 2* terdapat Kecamatan Piyungan, Bantul, Banganglipuro, Sanden, Jetis dan Pleret. Kemudian *cluster 3* terdapat Kecamatan Pandak, Kretek, Pundong, Sedayu, Pajangan, Dlingo dan Srandakan.



Gambar 4. Visualisasi *scatter plot* penyakit DBD Tahun 2022 menggunakan *tools orange*

Berdasarkan gambar 4 terdapat 3 *cluster* yaitu *cluster 1* untuk skala tinggi, *cluster 2* untuk skala sedang dan *cluster 3* untuk skala rendah. *cluster 1* berwarna hijau, *cluster 2* berwarna biru dan *cluster 3* berwarna merah. *Cluster 1* terdapat Kecamatan Kasihan, Banguntapan, Pleret dan Sewon. Selanjutnya pada *cluster 2* terdapat Kecamatan Piyungan, Imogiri, Bantul, Jetis, Sedayu dan Pandak. Kemudian *cluster 3* terdapat Kecamatan Pajangan, Sanden, Banganglipuro, Dlingo, Kretek, Pundong dan Srandakan.

Tabel 1. Skala *cluster* penyakit DBD di Kabupaten Bantul tahun 2020-2022

No	Kecamatan	Tahun		
		2020	2021	2022
1	Bantul	Tinggi	Sedang	Sedang
2	Sewon	Tinggi	Tinggi	Tinggi
3	Banguntapan	Tinggi	Tinggi	Tinggi
4	Kasihan	Tinggi	Tinggi	Tinggi
5	Banganglipuro	Tinggi	Sedang	Rendah

6	Jetis	Sedang	Sedang	Sedang
7	Pandak	Sedang	Rendah	Sedang
8	Piyungan	Sedang	Sedang	Sedang
9	Imogiri	Sedang	Tinggi	Sedang
10	Sanden	Sedang	Sedang	Rendah
11	Kretek	Sedang	Rendah	Rendah
12	Pajangan	Rendah	Rendah	Rendah
13	Pundong	Rendah	Rendah	Rendah
14	Sedayu	Rendah	Rendah	Sedang
15	Pleret	Rendah	Sedang	Tinggi
16	Srandakan	Rendah	Rendah	Rendah
17	Dlingo	Rendah	Rendah	Rendah

Berdasarkan tabel 1 dapat dilihat bahwa terdapat 3 kecamatan konsisten tinggi yaitu Kecamatan Sewon, Banguntapan dan Kasihan. Selanjutnya terdapat 4 kecamatan menurun yaitu Kecamatan Bantul, Bambanglipuro, Sanden dan Kretek. Kemudian terdapat 2 kecamatan konsisten sedang dalam 3 tahun yaitu Kecamatan Jetis dan Piyungan. Lalu terdapat 2 kecamatan mengalami kenaikan dan penurunan yaitu Kecamatan Pandak dan Imogiri. Selanjutnya terdapat 4 kecamatan yang konsisten rendah yaitu Kecamatan Pajangan, Pundong, Srandakan dan Dlingo. Terakhir, terdapat 2 kecamatan mengalami kenaikan dalam 3 tahun yaitu Kecamatan Sedayu dan Pleret. Uji *Anova one way* pada penelitian digunakan untuk melihat apakah ada perbedaan rata-rata dari ketiga tingkat/*cluster*.

Tabel 2. Uji Anova one way penyakit DBD di Kabupaten Bantul tahun 2020 menggunakan SPSS

Jumlah	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15389.065	2	7694.532	40.949	.000
Within Groups	2630.700	14	187.907		
Total	18019.765	16			

Berdasarkan tabel 2 dapat dilihat terdapat bahwa nilai signifikansi sebesar 0.000, jika dibandingkan nilainya dengan $\alpha = 0.05$ (alpha 5%) maka lebih kecil (0.000

< 0.05) maka H_0 ditolak. Sehingga dapat disimpulkan terdapat perbedaan rata-rata jumlah kasus DBD yang signifikan antara *cluster* 1, 2 dan 3.

Tabel 3. Uji Anova one way penyakit DBD di Kabupaten Bantul tahun 2021 menggunakan SPSS

Jumlah	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4536.753	2	2268.376	106.207	.000
Within Groups	299.012	14	21.358		
Total	4835.765	16			

Berdasarkan tabel 3 dapat dilihat terdapat bahwa nilai signifikansi sebesar 0.000, jika dibandingkan nilainya dengan $\alpha = 0.05$ (alpha 5%) maka lebih kecil (0.000 < 0.05) maka H_0 ditolak. Sehingga dapat disimpulkan terdapat perbedaan rata-rata jumlah kasus DBD yang signifikan antara *cluster* 1, 2 dan 3.

Tabel 4. Uji Anova one way penyakit DBD di Kabupaten Bantul tahun 2022 menggunakan SPSS

Jumlah	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	22997.644	2	11498.822	50.891	.000
Within Groups	3163.298	14	225.950		
Total	26160.941	16			

Berdasarkan tabel 4 dapat dilihat terdapat bahwa nilai signifikansi sebesar 0.000, jika dibandingkan nilainya dengan $\alpha = 0.05$ (alpha 5%) maka lebih kecil (0.000 < 0.05) maka H_0 ditolak. Sehingga dapat disimpulkan terdapat perbedaan rata-rata jumlah kasus DBD yang signifikan antara *cluster* 1, 2 dan 3.

PEMBAHASAN

1. K-Means Clustering Berbentuk Relative

K-means clustering Relative adalah teknik *Clustering* yang belum punya

standar atau acuan. Rata-rata *K-means clustering* dapat berubah-ubah. *Cluster* akan berubah jika datanya berbeda. Misalnya penentuan 3 *cluster* atau 2 *cluster* belum ada acuan dan standar nilai *cluster* tinggi, sedang atau rendah belum ada standar atau acuannya. Berdasarkan informasi yang didapat, dari pihak Dinkes Bantul tidak menerapkan *cluster* untuk kasus penyakit DBD. *Clustering* yang dilakukan menggunakan *cluster* tinggi, sedang atau rendah. guna memudahkan dalam mengambil keputusan untuk pengendalian serta pencegahan penyakit DBD. Kemudian data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lainnya (Talakua dkk, 2017).

2. Perbandingan 3 Tahun Kasus Penyakit DBD di Kabupaten Bantul Tahun 2020-2022

Berdasarkan tabel 1 menggunakan Microsoft Excel perbandingan 3 tahun kasus penyakit DBD di Kabupaten Bantul tahun 2020-2022 terdapat 3 kecamatan yang konsisten di *cluster* 1/tinggi. Tiga kecamatan tersebut adalah Kecamatan Sewon, Banguntapan dan Kasihan. Tiga kecamatan tersebut merupakan daerah endemis penyakit DBD karena kepadatan jumlah penduduk dan urbanisasi yang paling banyak terjadi pada kecamatan tersebut (Fauzi dan Winarni, 2018). Selanjutnya terdapat 4 kecamatan menurun yaitu Kecamatan Bantul, Bambanglipuro, Sanden dan Kretek. Empat kecamatan tersebut mengalami penurunan kasus penyakit DBD karena lahan pemukiman penduduk yang tidak terlalu padat (Yuniastuti dan Respati, 2013). Kemudian terdapat 2 kecamatan konsisten sedang dalam 3 tahun yaitu Kecamatan Jetis dan Piyungan. Dua

kecamatan tersebut merupakan daerah endemis penyakit DBD karena kepadatan penduduk di kecamatan tersebut (Hidayani, 2020). Lalu terdapat 2 kecamatan mengalami kenaikan dan penurunan yaitu Kecamatan Pandak dan Imogiri. Dua kecamatan tersebut mengalami kenaikan dan penurunan karena telah melakukan gerakan PSN dan akibat dari perubahan cuaca (Hasanudin, 2023). Selanjutnya terdapat 4 kecamatan yang konsisten rendah yaitu Kecamatan Pajangan, Pundong, Srandakan dan Dlingo. Empat kecamatan tersebut konsisten rendah karena merupakan daerah minim penduduk (Sari, 2016). Terakhir, terdapat 2 kecamatan mengalami kenaikan dalam 3 tahun yaitu Kecamatan Sedayu dan Pleret. Dua kecamatan tersebut merupakan daerah endemis penyakit DBD karena kepadatan penduduk di kecamatan tersebut (Hidayani, 2020). Berdasarkan informasi yang didapat, dari pihak Dinkes Bantul jika ada kecamatan yang kasus penyakit DBDnya tinggi akan muncul alert/peringatan di sistem SISKLB. Kemudian Dinkes Bantul melakukan penyuluhan dan PSN yang lebih sering lagi.

3. Analisis Data

Analisis data yang digunakan yaitu *Anova one way*. Dari hasil uji *Anova one way* diperoleh hasil bahwa nilai signifikansi sebesar 0.000, jika dibandingkan nilainya dengan $\alpha = 0.05$ (alpha 5%) maka lebih kecil ($0.000 < 0.05$) maka H_0 ditolak. Sehingga dapat disimpulkan terdapat perbedaan rata-rata jumlah kasus DBD yang signifikan.

4. Pengendalian yang Dilakukan Oleh Dinkes Bantul

Berdasarkan informasi yang didapat, dari pihak Dinkes Bantul melakukan pengendalian penyakit DBD yaitu dengan dilakukan PSN, Larvasidasi,

fogging, penyuluhan serta yang terbaru ada *Wolbachia* 2022. Upaya pencegahan penyakit DBD yang telah dilakukan antara lain dengan PSN, penaburan larvasida dan *fogging* (Werdiningsih, 2016). Dinkes Bantul telah menerapkan program *Wolbachia* Wes Masuk Bantul atau “Wow Mantul” yaitu program pelengkap pengendali DBD yang dilaksanakan Dinkes bekerja sama dengan World Mosquito Program (WMP) Yogyakarta dan didukung oleh Yayasan Tahija (Dinkes Bantul, 2020). Pada website Dinkes Bantul, dinyatakan bahwa “Wow Mantul” merupakan program pengendalian *dengue* dengan teknologi nyamuk *Aedes aegypti* ber-*Wolbachia* yang diimplementasikan di wilayah Bantul, mencakup 11 Kecamatan, 11 Kalurahan dan 519 Padukuhan. Walaupun sudah menerapkan program Wow Mantul kasus penyakit DBD dari tahun 2021 ke 2022 mengalami kenaikan, yakni dari 410 kasus menjadi 885 kasus. Pemanfaatan sistem SISKLB juga merupakan suatu upaya pengendalian penyakit DBD. Data kasus penyakit DBD dapat diambil dari sistem SISKLB. Berdasarkan informasi yang didapat, dari pihak Dinkes Bantul sistem SISKLB juga berguna untuk peringatan atau alert, misalnya dalam 3 tahun berturut-turut suatu kecamatan kasusnya tinggi maka akan muncul peringatan.

5. Kelebihan dan Kekurangan Menggunakan Teknik Clustering dengan K-means

Kelebihan menggunakan teknik *Clustering* dengan *K-means* adalah pengelompokan suatu objek dengan cepat (Novianto dan Goeirianto, 2019). Selain itu, *Clustering* dengan *K-means* mudah untuk diimplementasikan, waktu yang dibutuhkan *relative* cepat dan populer digunakan (Mulaab, 2017). Sedangkan kekurangan menggunakan

teknik *Clustering* dengan *K-means* adalah menentukan nilai awal *K* atau jumlah *cluster* (Purwati, Kurniawan dan Karnila, 2021).

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan di Dinas Kesehatan Kabupaten Bantul, menghasilkan visualisasi *scatter plot* dengan mendapatkan *clustering* penyakit DBD terdapat 3 *cluster* yaitu tinggi, sedang atau rendah dengan 3 kecamatan yang konsisten di *cluster* tinggi yaitu Kecamatan Sewon, Banguntapan dan Kasihan, lalu terdapat 2 kecamatan konsisten di *cluster* sedang yaitu Kecamatan Jetis dan Piyungan, kemudian 4 kecamatan yang konsisten di *cluster* rendah yaitu Kecamatan Pajangan, Pundong, Srandakan dan Dlingo.

SARAN

Berdasarkan penelitian yang telah dilakukan di Dinas Kesehatan Kabupaten Bantul, maka peneliti selanjutnya diharapkan dilakukan pengamatan lebih detail untuk tahun berikutnya dan Dinas Kesehatan Kabupaten Bantul diharapkan memperhatikan kasus penyakit DBD terutama yang konsisten tinggi.

DAFTAR PUSTAKA

- Buulolo, E. (2020). *Data mining Untuk Perguruan Tinggi*. Yogyakarta:Deepublish.
- Dinkes Bantul. (2021). *Profil Kesehatan Kabupaten Bantul Tahun 2021*. Yogyakarta:Dinas Kesehatan Bantul.
- Fauzi, M., & Winarni, F. (2018). *Efektivitas Program Pemberantasan Sarang Nyamuk Melalui Gertak Psn Di Desa Banguntapan Kecamatan Banguntapan, Bantul*. *Journal of Public Policy and Administration Research*, 7(4), 443-457.
- Hasanudin, U. (2023). *Kasus DBD di Bantul Menurun, Masyarakat*

- Diminta Tetap Waspada.*
<https://jogjapolitan.harianjogja.com>
. Diakses tanggal 18 Mei 2023
- Hidayani, W. R. (2021). *Demam berdarah dengue: Perilaku Rumah Tangga dalam Pemberantasan Sarang Nyamuk dan Program Penanggulangan Demam Berdarah Dengue*. Purwokerto:Pena Persada.
- Mulaab. (2017). *Data mining Konsep dan Aplikasi*.Malang:Media Nusa Creative.
- Novianto, R., & Goeirmanto, L. (2019). *Penerapan Data mining Menggunakan Algoritma K-means clustering untuk Menganalisa Bisnis Perusahaan Asuransi. Jurnal Teknik Informatika dan Sistem Informasi*, 85-95.
- Nurhafida, I. S., & Sembiring, F. (2021). *Analisis Text Clustering Masyarakat Di Twiter Mengenai Mcdonald'Sxbts Menggunakan Orange Data mining. SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, 28–35.
- Prianto, C., S. B. (2020). *Pembuatan Aplikasi Clustering Gangguan Jaringan Menggunakan Metode K-means clustering*. Bandung:Kreatif Industri Nusantara.
- Purwati, N., Kurniawan, H., & Karnila, S. (2021). *Data mining*. Banyumas:Zahira Media Publisher.
- Ramadhani, R. & Bina, N. S. (2021). *Statistika Penelitian Pendidikan: Analisis Perhitungan Matematis dan Aplikasi SPSS*. Jakarta:Kencana
- Sari, R. A. (2017). *Analisis Faktor Yang Berpengaruh Terhadap Penyakit Demam Berdarah Dengue (DBD) Di Kabupaten Bantul Menggunakan Regresi Poisson. Skripsi*. Fakultas Matematika dan Ilmu Pengetahuan Alam UII, Yogyakarta.
- Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). *Analisis Cluster dengan Menggunakan Metode K-means untuk Pengelompokkan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahyn 2014. Jurnal Ilmu Matematika dan Terapan*, 119-128.
- Yuniastuti, E., & Suryo, D. R. (2013). *Kajian Tingkat Kerawanan Penyakit Demam Berdarah Dengue (DBD) di Kecamatan Bantul. Geo Media: Majalah Ilmiah dan Informasi Kegeografian*, 11(1).