

## Data Kategorik dalam Penelitian: Review Bibliometrik

Ardiansyahroni<sup>1</sup>, Awaludin Tjalla<sup>2</sup>, Mahdiyah Mahdiyah<sup>3</sup>

Program S3 PEP Universitas Negeri Jakarta

---

### Article Info

#### Article history:

Accepted: 20 Oktober 2023

Publish: 28 Oktober 2023

---

#### Keywords:

data kategorik, *Analysis*

*Bibliometric*.

---

### Abstrak

Penelitian ini menggunakan pendekatan deskriptif dengan metode study literatur. Penelitian ini bertujuan untuk mengkaji dan menganalisis secara sistematis penggunaan data kategorik dalam penelitian. Teknik penelitian menggunakan *bibliometric analysis*. *Bibliometric Analysis* digunakan untuk melihat trend penelitian saat ini, mencari peluang penelitian yang belum pernah diteliti, melihat jumlah penelitian yang berkaitan dengan tema penelitian dan hubungannya dengan penelitian lain. Pencarian artikel menggunakan database *google scholar* serta dibatasi jurnal nasional terbitan tahun 2015 sampai 2022. Pengolahan *bibliometric analysis* menggunakan perangkat *Vos Viewer*. Berdasarkan pencarian artikel melalui *Publish and Perish* dihasilkan 500 artikel yang berkaitan dengan tema penelitian. Hasil penelitian menemukan bahwa Penggunaan data kategorik dalam penelitian banyak di gunakan dalam bidang pendidikan, kesehatan dan telekomunikasi.

---

*This is an open access article under the [Lisensi Creative Commons Atribusi-BerbagiSerupa 4.0 Internasional](#)*



---

### Corresponding Author:

**Ardiansyahroni**

Universitas Negeri Jakarta

Email: pelangiemas1974@mail.com

---

## 1. PENDAHULUAN

Seorang peneliti dalam melakukan sebuah penelitian akan dihadapkan dengan informasi apa yang akan dikumpulkan, bagaimana cara mengumpulkan informasi tersebut, bagaimana bentuk informasi tersebut dan bagaimana memperlakukan informasi tersebut. Informasi berhubungan dengan data penelitian, bagaimana cara mengolah informasi tersebut menjadi sebuah kesimpulan. Moh Nazir mengatakan bahwa pengumpulan informasi merupakan langkah yang amat penting karena informasi yang dikumpulkan akan menjadi sebuah jawaban atas sebuah permasalahan yang ingin kita teliti ( Moh, Nazir :1999).

Oleh sebab itu peneliti harus memahami tentang data penelitian. Seperti diketahui dalam bidang penelitian terdapat beberapa kategori data yaitu, data nominal data ordinal, data interval dan data rasio. Sidney Siegel (1992:27) menyebutnya dengan 4 tingkatan pengukuran. Masing masing data mempunyai karakteristik masing-masing sehingga berakibat pada teknik analisis statistik yang dipakai. Furqon (2013:6) mengatakan bahwa teknik analisis statistik adalah prosedur yang digunakan untuk memproses data dan menghasilkan ukuran tertentu mengenai kecenderungan dan karakteristik data yang bersangkutan. Artikel ini akan membahas penggunaan data kategorik dalam penelitian.

## 2. LITERATUR REVIEW

Sebelum membahas tentang data kategorik, baiknya kembali kita review tentang pengertian data. Menurut Sudjana (1996) data adalah keterangan atau ilustrasi mengenai suatu hal bisa berbentuk kategori atau bisa berbentuk bilangan. Data yang berbentuk

kategorik disebut dengan data kualitatif sementara data yang berbentuk bilangan disebut data kuantitatif. Menurut Arikunto data adalah semua fakta dan angka dari mana informasi dapat dikumpulkan

Dinamakan data kategori jika data tersebut memiliki skala pengukuran yang terdiri dari sekumpulan kategori tertentu. Seperti umur terdiri dari beberapa kategori, anak-anak, remaja, dan dewasa (Tiro 1999). Data kategorik dapat berasal dari pengamatan kualitatif yang diringkas sebagai penghitungan, atau dari pengamatan data kuantitatif yang dikelompokkan dalam kelompok tertentu. Biasanya data ini terdiri dari skala nominal dan ordinal.

Data nominal ini hanya bisa menghasilkan satu jenis kategori saja. Operasi matematika tidak bisa dilakukan pada data Nominal. Misalnya: laki laki : 0 perempuan : 1. Laki laki tidak bisa dikali dengan perempuan, atau laki laki tidak bisa dikuadratkan dengan perempuan. Sementara itu data ordinal adalah data yang memiliki sebuah tingkatan tertentu. Tingkatan yang ada pada data ordinal ini biasanya berurutan baik itu dari yang terendah hingga urutan tertinggi ataupun sebaliknya. Akan tetapi, meskipun memiliki jenjang urutan pada data ordinal, data ini pun masih tidak bisa dilakukan pengolahan dengan operasi matematika. Misalnya, Penghitungan suara pada sebuah pemilu. Total suara partai A 60%, partai B 30%, dan partai C 20%, maka suara tertinggi di pegang oleh partai A sebagai peringkat pertama, sehingga menjadi pemenang dalam pemilu tersebut. Namun secara operasional matematika partai A tidak bisa di kali dengan partai B atau C.

### 3. METODOLOGI PENELITIAN

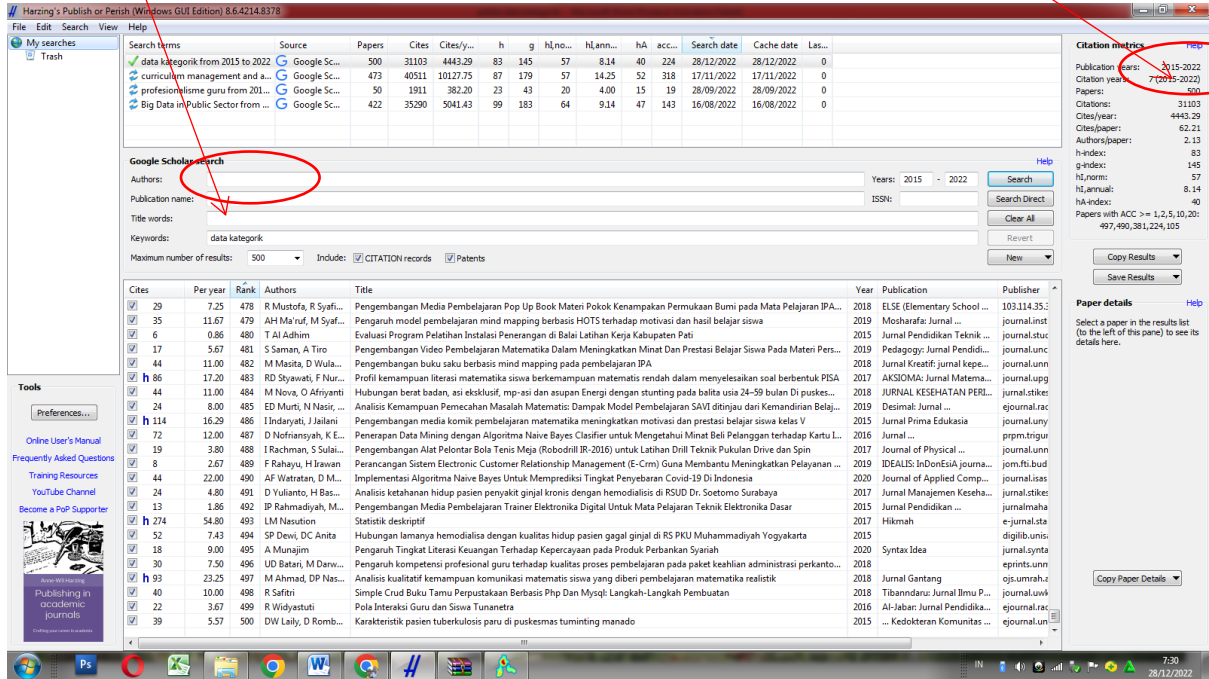
Penelitian ini menggunakan metode studi literatur. Penelitian kepustakaan atau study literatur (literature review, literature research) merupakan penelitian yang mengkaji atau meninjau secara kritis pengetahuan, gagasan, atau temuan yang terdapat di dalam tubuh literatur, berorientasi akademik (academic-oriented literature), serta merumuskan kontribusi teoritis dan metodologisnya untuk topik tertentu (Cooper, 1984).

Teknis analisa data melalui analisa bibliometrik yang berguna untuk mengetahui berbagai topik penelitian yang sedang tren, perkembangan jumlah serta ragam penelitian (T.Tupan N. W., 2020). Vos viewer juga dapat digunakan untuk mencari topik yang masih ada peluang diteliti, mencari referensi yang paling banyak digunakan pada bidang tertentu dan lainnya. Sumber data berupa artikel yang dianalisis dalam penelitian ini diperoleh dari database Google Scholar. Penelitian dilakukan dengan melakukan pencarian data artikel jurnal nasional secara daring dengan menggunakan Publish and Perish, melalui kata kunci, "data kategorik" (topik area penelitian). Tahun terbit artikel dibatasi dari tahun 2015- 2022. Jumlah artikel dibatasi sebanyak 500 artikel. Artikel yang sesuai diunduh menggunakan format RIS kemudian diimpor ke dalam perangkat lunak Vos Viewer yang menyimpan data referensi yang dimiliki oleh artikel, membuat pengelompokan berdasarkan kategori khusus, selanjutnya mengambil metadata yang ada di dalam dokumen artikel.

Dibawah ini adalah hasil pencarian artikel dengan Publish And Perish. sebagaimana gambar yang di lingkari dibawah terdapat keyword data kategorik dengan jumlah total pencairan 500 bahan bacaan.

Keyword : data kategorik

Jumlah pencarian : 500 sumber bahan bacaan



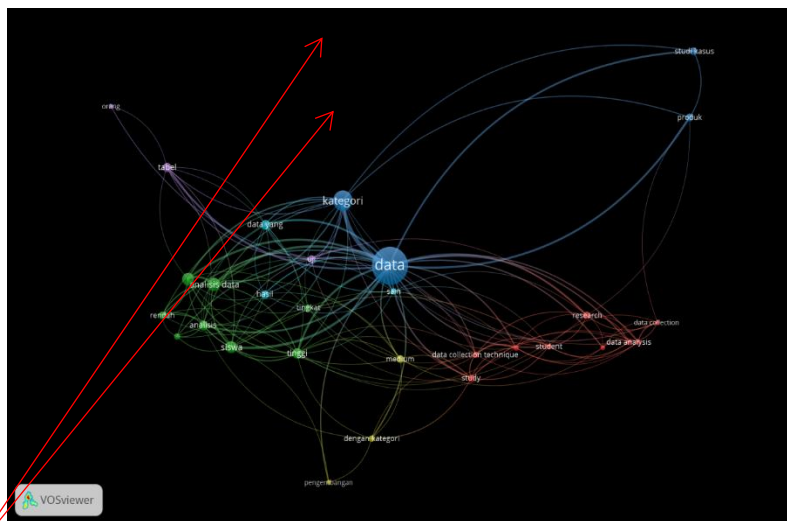
4. HASIL DAN PEMBAHASAN

a. Visualisasi Vos Viewer

Visualisasi vos viewer akan menghasilkan beberapa output yaitu *networking visualization, overlay visualization, density visualization dan data author* (Ranjbar-Sahraei, 2017). Di bawah ini akan dijelaskan sebagai berikut:

1. Networking Visualization

Sebagaimana pada gambar 1. menunjukkan hasil olahan vos viewer berupa *networking visualization*. Hal ini menjelaskan adanya hubungan antara variabel satu dengan variabel lainnya. Hubungan ditandai dengan warna jaringan yang sama, seperti dalam gambar terlihat kata kategorik dan data terhubung dalam warna yang sama (warna biru).

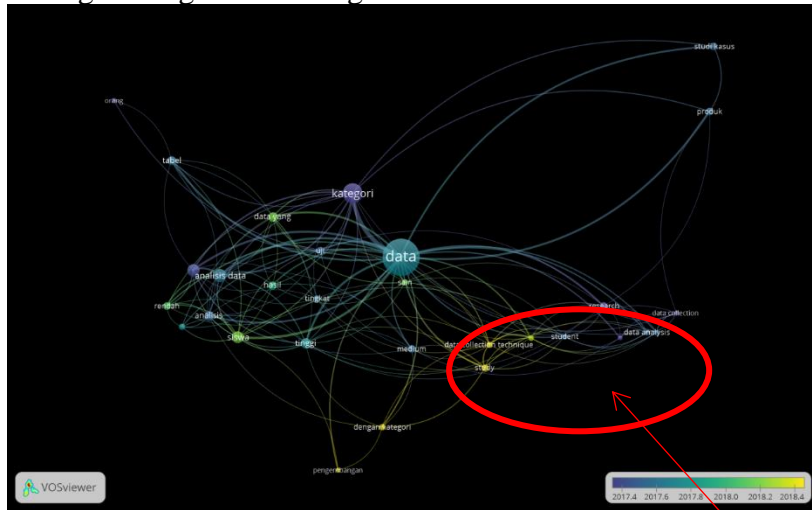


networking visualization

Gambar.1. Hasil *networking visualization*.

2. *Overlay Visualization*

Sebagaimana pada gambar 2. *Overlay visualization* menjelaskan trend topik penelitian yang muncul pada tahun 2015 sampai dengan 2022 yang berkaitan dengan data kategorik. Warna yang ditunjukkan adalah semakin kuning warna garis penghubung semakin muda tahun penelitian. Terlihat titik *data collectin*, *teknik* dan *student* adalah titik titik yang penelitian yang trending saat ini yang berhubungan dengan data kategorik.



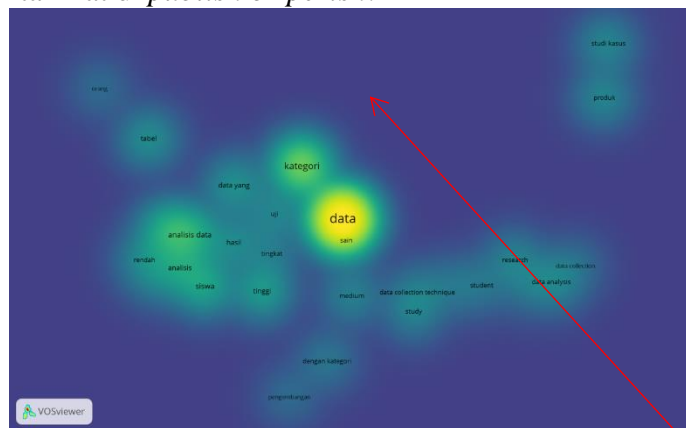
overlay visualization

Gambar.2. Hasil *overlay visualization*.

3. *Density Visualization*

Gambar 3 menunjukkan kedalaman dari density visualization, item diwakili oleh labelnya dengan cara yang sama seperti dalam visualisasi network dan visualisasi overlay. Setiap titik item dalam visualisasi density memiliki warna yang menunjukkan kepadatan item pada titik tersebut. Secara default, warna berkisar dari biru ke hijau hingga kuning. Semakin besar jumlah item di lingkungan titik dan semakin tinggi bobot item di sekitarnya, semakin dekat warna titik ke warna kuning.

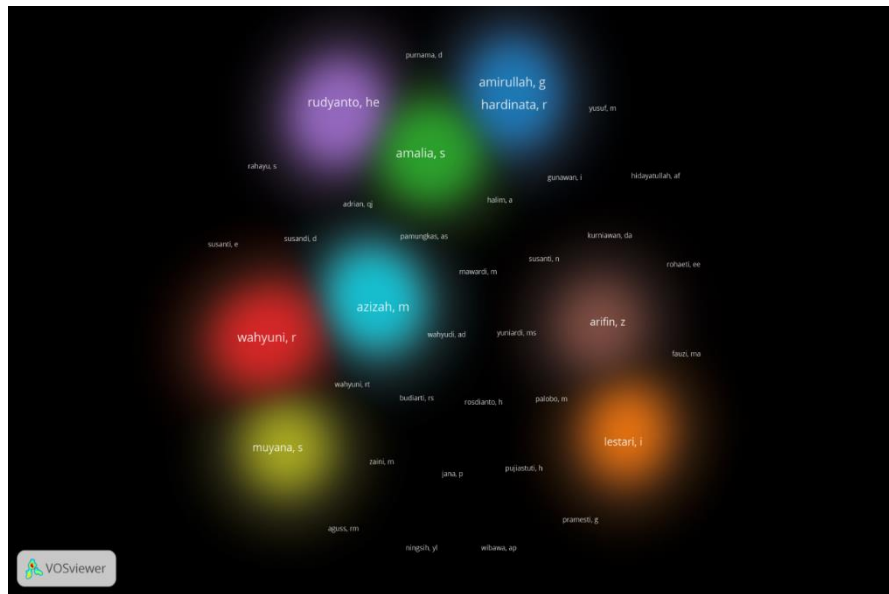
Item data, item kategorik dan item analisa data adalah item yang memiliki kepekatan warna yang tinggi (gambar 3). Hal ini berarti adalah sudah banyak penelitian tentang item tersebut. untuk melihat berapa jumlah penelitian tersebut dapat kita lihat di *publish or perish*.



density visualization

Gambar.3. Hasil *density visualizatio*

4. *Data author*



Gambar.4. author visualization

Gambar 4. diatas menunjukkan data para author/penulis yang memiliki link/hubungan dengan penulis lainnya lebih dari satu. Semakin tebal warna yang muncul maka semakin banyak hubungan penulis tersebut dengan penulis lainnya, artinya terdapat kedekatan kajian penelitian antar penulis. Sebut saja penulis yang memiliki banyak link amirullah, amalia, rudyanto, wahyuni dan mayana.

**1. Penggunaan data kategorik dalam penelitian.**

Beberapa penggunaan analisa data kategorik dalam penelitian yang sesuai dengan tema yang diinginkan oleh penulis di gambarkan dalam tabel dibawah ini:

No	Penulis	Judul	Hasil
1	Endah Wahyu RN, Apit Fathurohman, Sardianto MS	ANALISIS BUKU SISWA MATA PELAJARAN IPA KELAS VIII SMP/MTs BERDASARKAN KATEGORI LITERASI SAINS	Hasil penelitian menunjukkan bahwa kategori yang paling banyak muncul adalah kategori pengetahuan sains, yakni sebesar 46,3%. Kategori penyelidikan hakikat sains sebesar 30,2%, kategori sains sebagai cara berpikir sebesar 19,5% dan kategori interaksi sains, teknologi dan masyarakat sebesar 4,0%.
2	Nurul Rohmawati W, Sofi Defiyanti, Mohamad Jajuli	IMPLEMENTASI ALGORITMA K-MEANS DALAM PENGKLASTERAN MAHASISWA PELAMAR BEASISWA	Nilai purity pada dataset data kodifikasi sebagian untuk hasil cluster algoritma k-means sebesar 61.11%. Pada dataset kodifikasi keseluruhan nilai purity hasil cluster algoritma k-means sebesar 80.56%. Dan untuk dataset data asli nilai purity hasil cluster algoritma k-means sebesar 75%. Maka dapat di simpulkan bahwa algoritma k-means lebih cocok digunakan pada dataset dengan format atribut yang dikodifikasi keseluruhan.

3	Galih Kurniadi, Jayanti Putri Purwaningrum	KESALAHAN SISWA PADA KATEGORI KEMAMPUAN AWAL MATEMATIS RENDAH DALAM PENYELESAIAN TES KEMAMPUAN PEMECAHAN MASALAH MATEMATIS	Hasil penelitian menunjukkan bahwa penyebab terjadinya kesalahan yang dilakukan siswa dalam menyelesaikan tes kemampuan pemecahan masalah matematis pada kelompok siswa rendah adalah: (1) Siswa tidak membaca soal dengan seksama sehingga ada informasi yang terlewat; dan (2) Siswa tidak bisa menyebutkan apa yang diketahui dengan lengkap, tidak mengidentifikasi apa yang diketahui dengan tepat sehingga menyebabkan salah penafsiran.
4	Hidayah Karuniawati, Ika Gilar Hapsari, Marwiani Arum, Adiva Tanyas Aurora, Nungky Asmaraning, Wahyono	EVALUASI PELAKSANAAN STANDAR PELAYANAN MINIMAL (SPM) FARMASI KATEGORI LAMA WAKTU TUNGGU PELAYANAN RESEP PASIEN RAWAT JALAN DI RSUD KOTA SALATIGA	Waktu tunggu rata-rata obat racikan adalah 9,18 menit dan rata-rata waktu tunggu obat jadi atau obat non racikan adalah 5,70 menit. Hal tersebut sudah sesuai dengan standar pelayanan minimal yang dipersyaratkan oleh Kepmenkes No 129/Menkes/SK/II/2008 tentang pelayanan resep baik obat jadi maupun obat racikan yaitu lama waktu tunggu obat racikan $\leq 60$ menit dan obat non racikan atau obat jadi $\leq 30$ menit, dan dari semua sampel yang diteliti tidak ada yang melebihi lama waktu tunggu seperti yang dipersyaratkan.
5	Sila Prayoginingsih, Renny Pradina Kusumawardani	Klasifikasi Data Twitter Pelanggan Berdasarkan Kategori myTelkomsel Menggunakan Metode <i>Support Vector Machine</i> (SVM)	Data difilter untuk mendapatkan <i>tweet</i> terkait dengan informasi, keluhan, kritik, saran, dan pertanyaan seputar layanan atau produk Telkomsel. Menggunakan kernel RBF yang dioptimasi dengan metode <i>grid search</i> , didapatkan pengklasifikasi dengan performa yang baik, yaitu dengan akurasi dan <i>f-measure</i> sebesar 84,84% dan 84,88%.
6	Agustina Ayu Wulandari, Nurjazuli, M. Sakundarno Adi	Faktor Risiko dan Potensi Penularan Tuberkulosis Paru di Kabupaten Kendal, Jawa Tengah	Faktor risiko lingkungan fisik rumah pada kasus sebagian besar mempunyai proporsi yang tidak memenuhi syarat yaitu luas ventilasi sebesar 95,4%, kepadatan hunian sebesar 63,1%, suhu ruangan sebesar 70,8%, intensitas pencahayaan alami sebesar 84,6% dan kelembaban ruangan sebesar 52,3% kecuali jenis lantai rumah mempunyai proporsi

			memenuhi syarat ( jenis lantai kedap air ) lebih banyak sebesar 56,9%,
7	Nurul Munawaroh, Euis Eti Rohaeti, Usman Aripin	Analisis Kesalahan Siswa Berdasarkan Kategori Kesalahan Menurut Watson Dalam Menyelesaikan Soal Komunikasi Matematis Siswa SMP	Persentase jenis kesalahan siswa dalam menyelesaikan soal komunikasi matematis berdasarkan kategori kesalahan menurut Watson adalah kesalahan data tidak tepat ( <i>Inappropriate Data/ID</i> ) termasuk kedalam kesalahan yang cukup tinggi yaitu sebesar 38%, Kesalahan prosedur tidak tepat ( <i>Inappropriate Procedure/IP</i> ) termasuk kedalam kesalahan yang cukup tinggi juga yaitu sebesar 34%. Kesalahan kesimpulan hilang ( <i>Omitted Conclution/OC</i> ) termasuk kedalam kesalahan yang cukup rendah yaitu sebesar 26%.
8	Try Azisah Nurman	ANALISIS DATA KATEGORI DENGAN LOG LINIER MENGGUNAKAN PRINSIP HIRARKI (STUDI KASUS JUMLAH KECELAKAAN LALU LINTAS DI KOTA MAKASSAR TAHUN 2011).	Berdasarkan hasil analisis model log linier dalam kasus kecelakaan lalu lintas dengan tiga variabel, maka dapat diketahui bahwa variabel yang saling berinteraksi adalah jenis kendaraan dengan pendidikan dan umur dengan pendidikan. Hal ini berarti yang memiliki pengaruh dalam suatu kasus kecelakaan lalu lintas adalah jenis kendaraan dengan pendidikan terakhir dari pengendara, dan umur pengendara dengan pendidikan terakhir pengendara.

**5. KESIMPULAN**

Penggunaan data kategorik dalam penelitian digunakan diberbagai bidang, terutama dalam bidang pendidikan, kesehatan dan telekomunikasi. Benang merah yang didapat adalah dalam menggunakan data kategorik peneliti harus merumuskan kriteria atau kategorisasi terlebih dahulu. Kriteria ini akan dibandingkan dengan hasil empiris dilapangan. Sebagian besar data yang digunakan dalam data kateogri adalah data nominal dan data ordinal. Penggunaan pendekatan statistik non-parametrik digunakan untuk menganalisis kedua data tersebut. Proses perhitungan statistika non parametrik biasanya lebih sederhana dibandingkan pada statistika parametrik khususnya untuk ukuran sampel kecil.

**6. DAFTAR PUSTAKA**

Agustina Ayu Wulandari, Nurjazuli, M. Sakundarno Adi, FAKTOR RISIKO DAN POTENSI PENULARAN TUBERKULOSIS PARU di Kabupaten Kendal , Jawa Tengah, Jurnal Kesehatan Lingkungan Indonesia Vol. 14 No.1 / April 2015  
 Cooper, H. (1984). the integratif research review : A systematic approach.Social research method series. *journal educational researcher*, SAGE Publishe.

- Endah Wahyu RN, Apit Fathurohman, Sardianto MS, ANALISIS BUKU SISWA MATA PELAJARAN IPA KELAS VIII SMP/MTs BERDASARKAN KATEGORI LITERASI SAINS. JURNAL INOVASI DAN PEMBELAJARAN FISIKA, ISSN: 2355 – 7109 <http://fkip.unsri.ac.id/index.php/menu/1>
- Furqon, Statistika Terapan untuk Penelitian, Alfabeta, Bandung 2013.
- Galih Kurniadi, Jayanti, Putri Purwaningrum, KESALAHAN SISWA PADA KATEGORI KEMAMPUAN AWAL MATEMATIS RENDAH DALAM PENYELESAIAN TES KEMAMPUAN PEMECAHAN MASALAH MATEMATIS, JPPM Vol. 11 No. 2 (2018)
- Hidayah Karuniawati, Ika Gilar Hapsari, Marwiani Arum, Adiva Tanyas Aurora, Nungky Asmaraning, Wahyono, EVALUASI PELAKSANAAN STANDAR PELAYANAN MINIMAL (SPM) FARMASI KATEGORI LAMA WAKTU TUNGGU PELAYANAN RESEP PASIEN RAWAT JALAN DI RSUD KOTA SALATIGA, KARTIKA-JURNAL ILMIAH FARMASI, Jun 2016, 4(1), 20-25 p-ISSN 2354-6565 /e-ISSN 2502-3438
- Muh, Nazir, Metode Penelitian, Ghalia Indonesia, Jakarta 1999.
- Nurul Munawaroh, Euis Eti Rohaeti, Usman Aripin, ANALISIS KESALAHAN SISWA BERDASARKAN KATEGORI KESALAHAN MENURUT WATSON DALAM MENYELESAIKAN SOAL KOMUNIKASI MATEMATIS SISWA SMP, Jurnal Pembelajaran Matematika Inovatif Volume 1, No. 5, September 2018.
- Nurul Rohmawati W, Sofi Defiyanti, Mohamad Jajuli, IMPLEMENTASI ALGORITMA *K-MEANS* DALAM PENGKLASTERAN MAHASISWA PELAMAR BEASISWA, Jurnal Ilmiah Teknologi Informasi Terapan Volume I, No 2, 30 April 2015 ISSN : 2407 – 3911
- Sidney Siegel, Statistik Non Parametrik untuk Ilmu – Ilmu Sosial diterjemahkan oleh Zanzawi Suyuti dan Landung Simatupang, Gramedia Jakarta, 1992.
- Sila Prayoginingsih, Renny Pradina Kusumawardani, KLASIFIKASI DATA TWITTER PELANGGAN BERDASARKAN KATEGORI MYTELKOMSEL MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE* (SVM), Jurnal Sisfo, Vol. 07 No. 02 (2018) 83–98
- Sudjana, Metoda Statistika, Cetakan ke-enam, Tarsito , Bandung 1996.
- Suharsimi Arikunto, Pengembangan Instrumen Penelitian dan Penilaian Program, Pustaka Pelajar , Jakarta, 2017.
- T.Tupan, N. W. (2020). ANALISIS BIBLIOMETRIK PUBLIKASI ILMIAH TENTANG PREDIKSI GEMPA BUMI BERBASIS DATA SCOPUS periode 2015 - 2020 . *Libr.J.Perpust Vol 8 No.1*, 31-48.
- Try Azisah Nurman, ANALISIS DATA KATEGORI DENGAN LOG LINIER MENGGUNAKAN PRINSIP HIRARKI (STUDI KASUS JUMLAH KECELAKAAN LALU LINTAS DI KOTA MAKASSAR TAHUN 2011). Jurnal Teknosains, Volume 7 Nomor 1, Januari 2013, hlm: 99-110 DOI: <https://doi.org/10.24252/teknosains.v7i1.77>