KAJIAN PEMANFAATAN DATA GOOGLE MAPS DALAM OFFICIAL STATISTICS

Studi Kasus: Usaha Sektor Penyedia Makan Minum di Pulau Jawa dan Bali

Utilization of Google Maps Data for Official Statistics

Case Study: Food and Beverage Supply Sector Business in Java and Bali Islands

Cholifa Fitri Annisa¹, Setia Pramana²

Politeknik Statistika STIS¹

Badan Pusat Statistik RI²

E-mail: 16.9053@stis.ac.id

ABSTRAK

Publikasi statistik usaha penyediaan makan minum yang diterbitkan oleh BPS tidak bisa memfasilitasi pebisnis dalam mengidentifikasi daerah yang berpotensi memiliki kemampuan untuk dikembangkan usaha pada sektor penyediaan makan dan minum. Selain itu, adanya keterbatasan waktu, biaya, dan tenaga dalam pengumpulan data oleh Subdirektorat Pariwisata BPS pada survei VREST sehingga, menyebabkan statistik penyediaan makan minum tidak bisa diterbitkan sesuai metodologi yaitu setiap tahun. Penelitian ini memanfaatkan metode web scraping untuk mendapatkan data usaha penyedia makan minum dari situs web google maps. Jumlah data yang terkumpul sebanyak 34.526 usaha penyedia makan minum di Pulau Jawa dan Bali. Hasil nilai pencocokan data hasil web scraping dengan data frame BPS menunjukkan persentase kemiripan (match) sebesar 68,22%. Provinsi Bali adalah daerah yang memiliki potensi untuk mengembangkan usaha penyediaan makanan minuman terkhusus pada Kota/Kabupaten Jembrana, Buleleng, Tabanan, Karangasem, dan Klungkung. Sedangkan, provinsi Jawa Tengah adalah daerah yang memiliki potensi untuk mengembangkan usaha akomodasi terkhusus pada Kota/Kabupaten Cilacap, Blora, Grobogan, Batang dan Kendal.

Kata kunci: official statistics, web scraping, google maps, data matching, potensi daerah.

ABSTRACT

Publication of food and beverage supply business statistics published by BPS cannot facilitate business people in identifying areas that have the potential to develop businesses in the food and drink supply sector. In addition, there are limitations in time, cost, and manpower in data collection by the Tourism Sub-directorate of BPS on the VREST survey, so that the food and beverage supply statistics cannot be published according to the methodology, namely every year. This study utilizes the web scraping method to obtain business data on food and beverage providers from the google maps website. The amount of data collected is as many as 34,526 food and beverage providers in Java and Bali. The results of the matching value of web scraping data with BPS data frames showed a match percentage of 68.22%. Bali Province is an area that has the potential to develop food and beverage supply businesses, especially in the City / Regency of Jembrana, Buleleng, Tabanan, Karangasem, and Klungkung. Meanwhile, Central Java province is an area that has the potential to develop accommodation businesses, especially in the cities / regencies of Cilacap, Blora, Grobogan, Batang, and Kendal.

Keywords: official statistics, web scraping, google maps, data matching, regional potential

PENDAHULUAN

Badan Pusat Statistik sebagai penyelenggara statistik sektoral bekerjasama dengan Kementerian Pariwisata dalam menghasilkan statistik penyedia makan minum yang setiap tahun diterbitkan melalui website BPS. Berdasarkan Peraturan Presiden No. 69 Tahun 2019, tugas utama dari Kementerian Pariwisata dan Ekonomi Kreatif RI (Kemenparekraf) adalah memfasilitasi ketersediaan penunjang kegiatan kepariwisataan yang meliputi jasa penyedia akomodasi, penyedia makan minum, dan agen perjalanan wisata. Beberapa program seperti pariwisata berkelanjutan dilakukan oleh Kemenparekraf untuk membantu wirausaha/pebisnis dalam mengelola usahanya

agar ketersediaan fasilitas kegiatan pariwisata tetap tersedia dan menjadi daya tarik tourism/wisatawan.

Pada Mei 2020, BPS telah menerbitkan publikasi statistik penyedia makan minum Tahun 2018 pada halaman website bps.go.id. Publikasi ini memuat profil usaha yang dirinci menurut provinsi, banyaknya pekerja menurut jenis kelamin dan status kewarganegaraan, serta balas jasa pekerja, pengeluaran perusahaan, dan pendapatan yang diterima. Melalui publikasi ini, Kemenparekraf menyatakan bahwa adanya kendala oleh wirausaha/pebisnis dalam mengidentifikasi daerah yang memiliki potensi untuk membangun usahanya jika menggunakan publikasi statistik penyedia makan minum dari BPS.

Publikasi statistik penyedia makan minum hanya tersedia pada tahun tertentu yaitu publikasi statistik penyedia makan minum Tahun 2007-2009, publikasi tahun 2011-2015, dan publikasi tahun 2017 dan 2018 yang diterbitkan BPS pada awal tahun 2020. Sedangkan metodologi pada *website bps.go.id*, publikasi seharusnya diterbitkan setiap tahun oleh Subdirektorat Pariwisata BPS sehingga secara tidak langsung terdapat beberapa kendala yang dialami BPS dalam penyelenggaraan survei.

Dalam mengumpulkan data statistik penyedia makan minum, BPS menggunakan survei VREST yaitu daftar usaha menengah besar (UMB) yang meliputi restoran, rumah makan, catering, dan penyediaan makan minum lainnya sebagai sumber data statistik. Pengumpulan data dilakukan dengan cara pengisian kuesioner oleh setiap responden usaha sehingga memerlukan biaya, waktu, dan tenaga yang tidak sedikit. Selain itu, pengumpulan data dengan cara pengisian kuesioner memungkinkan adanya *non sampling error*, yaitu kesalahan yang diakibatkan oleh responden. *Non sampling error* memungkinkan data yang terkumpul tidak sesuai dengan kondisi yang semestinya yang menyebabkan ketidakmampuan parameter dalam menggambarkan populasi secara akurat.

Berdasarkan identifikasi masalah tersebut, maka tujuan dari penelitian ini yaitu, 1) mengumpulkan data usaha penyedia makan minum dengan metode webscraping, 2) mencocokkan data hasil web scraping dengan data BPS, 3) dan menganalisis daerah yang memiliki potensi usaha penyedia makan minum.

METODE

Gambaran kerangka pikir pada penelitian ini dijelaskan pada **Gambar 1**.

Gambar 1. Kerangka Pikir Penelitian.

Penelitian ini diawali dari pengidentifikasian masalah yang ditemukan bahwa publikasi statistik yang diterbitkan oleh BPS mengenai statistik usaha penyediaan makanan dan minuman tidak diterbitkan sesuai jadwal pada metodologi yaitu setiap tahun. Hal ini disebabkan adanya keterbatasan BPS pada pengumpulan data yang membutuhkan biaya, waktu, tenaga yang besar dalam melaksanakan kegiatannya. Selain itu, adanya kesulitan dari pengguna publikasi untuk menentukan daerah potensi usaha. Oleh karena itu, penelitian ini melakukan pemanfaatan data *webscraping* yang selanjutnya dilakukan pencocokan data untuk mendapatkan presentase *matching* guna menginterpretasikan apakah pemanfaatan data dapat digunakan sebagai pelengkap data di survei VREST BPS. Selanjutnya dilakukan analisis potensi usaha untuk mendapatkan daerah yang berpotensi untuk mengembangkan usaha penyediaan makan dan minum.

Penelitian ini memanfaatkan metode web scraping untuk mendapatkan data usaha penyedia makan minum dari situs web google maps. Cakupan region yang digunakan dalam pengumpulan data adalah pulau Jawa dan Bali yang terdiri dari provinsi Jawa Timur, Jawa Tengah, Jawa Barat, DKI Jakarta, Banten, DI Yogyakarta dan Bali. Alasan pemilihan cakupan region pulau Jawa dan Bali karena sebagai daerah penyumbang devisa pendapatan paling besar melalui sektor pariwisatanya (Publikasi BPS Proporsi Kontribusi Pariwisata terhadap PDB, 2015). Pengumpulan data dilakukan dengan cara *webscraping* dimulai pada tanggal 16 April hingga 28 Mei 2020 menggunakan bantuan software Botsol. Botsol yang digunakan adalah botsol berlisensi versi 8.0 dengan chrome version 83.0.4.4103.116. Software ini digunakan untuk membantu dalam mengekstrak seluruh informasi dari web yang dibutuhkan tanpa adanya pemblokiran serta dapat dilakukan pengambilan data secara otomatisasi menggunakan robot. Atribut data yang diambil yaitu daftar data nama usaha, alamat lokasi, kode plus, dan alamat web sebagai variabel pembanding dengan data frame BPS. Serta pengambilan data jumlah rating dan reviews (penilaian user) untuk mengindikasikan kualitas dari setiap usaha sebagai pertimbangan analisis.

Penelitian ini juga menggunakan data direktori usaha penyedia makan minum yang dimiliki oleh BPS dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2017. Pengumpulan data sekunder digunakan sebagai data pembanding dengan hasil *webscraping*. Serta menggunakan data dari direktori penelitian yang dilakukan oleh Adhinugroho & Pramana (2019) mengenai direktori data akomodasi yang diambil dari web travel Pegi-peggi dan Agoda untuk membantu dalam analisis pengidentifikasian potensi daerah usaha.

Dalam melakukan pengolahan data, penelitian ini menggunakan *python* 3 dengan bantuan dari beberapa library. Proses data *preprocessing* menggunakan bantuan library *pandas* untuk pembersihan data duplikat, data NA, dan pembersihan data yang belum terkonfirmasi oleh google. Proses *geocoding* menggunakan bantuan library *openlocationcode* versi 1.0.1 dan library *geopy* versi 1.22.0. Setelah itu proses *matching* data antara data hasil webscraping dengan frame BPS dengan bantuan algoritma jarak jaro-winkler dan haversine menggunakan bantuan library *jellyfish* dan *math*. Proses selanjutnya, visualisasi data menggunakan bantuan aplikasi *Tableau Public* untuk membantu dalam menganalisis data. Pengidentifikasian daerah yang memiliki potensi usaha didasarkan pada jumlah dari konsep sarana kepariwisataan di suatu daerah. Pada penelitian ini, dalam mengidentifikasikan potensi daerah tersebut menggunakan perbandingan jumlah sarana kepariwisataan pada jasa akomodasi dan usaha penyedia makan minum.

HASIL DAN PEMBAHASAN

Dari proses pengumpulan data menggunakan metode *webscraping* yang telah melalui proses *data preprocessing* dan *geocoding*, didapatkan data yang berjumlah 34.526 data usaha penyedia makan minum di Pulau Jawa dan Bali yang dijelaskan pada **Tabel 1** dan **Gambar 2**.

Tabel 1. Jumlah Usaha Penyedia Makan Minum Metode Webscraping

Provinsi	Jumlah Usaha Penyedia Makan Minum
Bali	2.052
DKI Jakarta	9.103
Jawa Barat	5.135
Jawa Timur	7.357
Jawa Tengah	7.625
Banten	1.032
DI Yogyakarta	2.222
Total	34.526

Gambar 2. Visualisasi Data Usaha Penyedia Makan Minum dengan Metode Webscraping

Berdasarkan **Tabel 1** dan **Gambar 2** persebaran menunjukkan bahwa provinsi DKI Jakarta adalah provinsi yang memiliki jumlah usaha penyedia makan minum tertinggi yaitu sebesar 9.103 usaha. Sedangkan provinsi Banten adalah provinsi yang memiliki jumlah usaha penyedia makan minum terendah yaitu sebesar 1.032 usaha.

Dari data yang telah dikumpulkan menggunakan metode *web scraping* dapat dianalisis kualitas usaha penyediaan makan minum di Pulau Jawa dan Bali berdasarkan kualitas usahanya. Analisis ini menggunakan informasi jumlah *rating* dan *reviews* dari usaha penyedia makan minum. Suatu usaha yang memiliki jumlah *rating* dan *reviews* yang tinggi mengindikasikan bahwa daerah memiliki usaha penyedia makan minum yang berkualitas dari sisi penilaian dan kepuasan kunjungan.

Gambar 3. Grafik tiga daerah yang memiliki rata-rata jumlah rating tertinggi (**kiri**) dan grafik tiga daerah yang memiliki jumlah akun/user yang memberikan rating tertinggi (**kanan**)

Berdasarkan grafik pada **Gambar 3 (kiri)**, dapat diketahui bahwa sub-region yang memiliki rata-rata penilaian tertinggi adalah daerah Tangerang dengan jumlah rata-rata penilaian sebesar 4,709 bintang. Setelah itu, daerah Buleleng dengan jumlah rata-rata penilaian sebesar 4,393 bintang. Dan daerah Banyuwangi dengan jumlah rata-rata penilaian sebesar 4,358 bintang. Daerah yang memiliki jumlah rata-rata penilaian tertinggi dapat diindikasikan bahwa usaha penyediaan makan minum pada suatu daerah tersebut berkualitas.

Berdasarkan grafik pada **Gambar 3 (kanan)**, juga dapat diketahui bahwa sub-region yang memiliki jumlah akun/user yang melakukan penilaian terhadap usaha penyediaan makan minum paling banyak adalah daerah Badung dengan jumlah akun/user sebanyak 3.035. Setelah itu, daerah Surabaya dengan jumlah akun/user yang melakukan penilaian sebanyak 2.331 akun/user. Daerah Bandung memiliki jumlah akun/user yang melakukan penilaian sebanyak 2.257 akun/user. Daerah yang memiliki jumlah akun/user yang melakukan penilaian terhadap usaha penyediaan makan minum paling banyak dapat diindikasikan bahwa adanya kepekaan wisatawan/*tourism* untuk melakukan penilaian kunjungan.

Gambar 4. Grafik tiga daerah yang memiliki rata-rata jumlah reviews tertinggi (**kiri**) dan grafik tiga daerah yang memiliki jumlah akun/user yang memberikan reviews tertinggi (**kanan**)

Berdasarkan grafik pada **Gambar 4 (kiri)**, dapat diketahui bahwa sub-region yang memiliki rata-rata komentar tertinggi adalah daerah Yogyakarta dengan jumlah rata-rata komentar sebesar 2.317 komentar. Setelah itu, daerah Sleman dengan jumlah rata-rata komentar sebesar 1.673 komentar. Dan daerah Surabaya dengan jumlah rata-rata komentar sebesar 1.642 komentar. Daerah yang memiliki jumlah rata-rata komentar tertinggi dapat diindikasikan bahwa usaha penyediaan makan minum pada suatu daerah tersebut memiliki ulasan yang tinggi baik itu positif atau negatif.

Berdasarkan grafik pada **Gambar 4 (kanan)**, juga dapat diketahui bahwa sub-region yang memiliki jumlah akun/user yang melakukan komentar terhadap usaha penyediaan makan minum paling banyak adalah daerah Badung dengan jumlah akun/user sebanyak 1.701. Setelah itu, daerah Surabaya dengan jumlah akun/user yang melakukan komentar sebanyak 1.543 akun/user. Daerah Bandung memiliki jumlah akun/user yang melakukan komentar sebanyak 1.528 akun/user. Daerah yang memiliki jumlah akun/user yang melakukan komentar terhadap usaha penyediaan makan minum paling banyak dapat diindikasikan bahwa adanya kepekaan wisatawan/*tourism* untuk melakukan ulasan *review* atau komentar dari kunjungan.

Dari proses pengumpulan data sekunder yaitu data frame BPS yang bersumber dari data direktori usaha penyedia makan minum oleh BPS dari hasil Survei Sosial Ekonomi Nasional (Susenas) tahun 2017 yang telah melalui proses *data preprosesing* dan *geocoding*, didapatkan data yang berjumlah 47.387 data usaha penyedia makan minum di Pulau Jawa dan Bali yang dijelaskan pada tabel sebagai berikut:

Tabel 2. Jumlah Usaha Penyedia Makan Minum frame BPS

Provinsi	Jumlah Usaha Penyedia Makan Minum
Bali	3.953
DKI Jakarta	10.684

Jawa Barat	6.876
Jawa Timur	10.025
Jawa Tengah	10.267
Banten	2.005
DI Yogyakarta	3.577
Total	47.387

Gambar 5. Visualisasi Data Usaha Penyedia Makan Minum frame BPS

Berdasarkan **Tabel 2** dan **Gambar 5** persebaran menunjukkan bahwa provinsi DKI Jakarta adalah provinsi yang memiliki jumlah usaha penyedia makan minum tertinggi yaitu sebesar 3.953 usaha. Sedangkan provinsi Banten adalah provinsi yang memiliki jumlah usaha penyedia makan minum terendah yaitu sebesar 2.005 usaha.

Data yang telah terkumpul dilakukan perbandingan dari jumlah data antara data yang dikumpulkan dengan metode *webscraping* dan data frame BPS untuk melihat pola dari jumlahnya. Hasil perbandingan data terdapat pada **Gambar 6**.

Gambar 6. Perbandingan Jumlah Data Frame BPS dengan Data Hasil *Web Scraping*

Dari hasil perbandingan pada **Gambar 6** menunjukkan bahwa jumlah data hasil *webscraping* lebih sedikit dari data frame BPS. Namun, dari perbandingan data menunjukkan pola yang sama dari jumlah data usaha penyedia makan minum pada setiap provinsi. Analisis selanjutnya yaitu proses pencocokan data hasil dari pencocokan dengan penggunaan kombinasi algoritma *jaro-winkler* dan *haversine*. Hasil dari proses pencocokan data (*matching data*) dapat dilihat dari hasil diagram pada **Gambar 7**.

Gambar 7. Diagram Venn Hasil Pencocokan Data

Hasil nilai pencocokan data hasil *web scraping* dengan data frame BPS menunjukkan persentase kemiripan (*match*) sebesar 68,22%. Melalui hasil kemiripan menunjukkan bahwa data yang dihasilkan melalui pengumpulan *webscraping* pada situs *google maps* bisa melengkapi data BPS dengan penambahan variabel penilaian (*rating* dan *reviews*) sebagai indikator kualitas usaha sektor serta penambahan variabel koordinat (*longitude* dan *latitude*) sebagai alamat koordinat lokasi usaha sektor. Persentase sisa data hasil *webscraping* yang tidak mirip (*unmatch*) sebesar 2.68% dapat digunakan sebagai penambahan data baru. Sehingga, dapat disimpulkan bahwa metode *webscraping* ini dapat dijadikan sebagai variabel dan data pelengkap (*compliment*) dari pengumpulan data usaha sektor penyedia makan minum di BPS sebagai *official statistics*.

Pada penelitian ini, dalam mengidentifikasi potensi daerah tersebut menggunakan perbandingan jumlah sarana kepariwisataan pada jasa akomodasi dan usaha penyedia makan minum. Data akomodasi yang digunakan pada penelitian ini adalah daftar hotel dan penginapan dari beberapa situs web yaitu Agoda dan Pegi-peggi yang diambil dari directory data pada penelitian sebelumnya (Adhinugroho & Pramana, 2019). Perbandingan data jumlah akomodasi dengan data jumlah usaha penyedia makan minum dapat dilihat pada **Tabel 3**.

Tabel 3. Perbandingan jumlah akomodasi dengan usaha penyedia makan minum di Pulau Jawa dan Bali

Provinsi	Jumlah Akomodasi	Jumlah Usaha Penyedia Makan Minum
Bali	6.674	2.052
DKI Jakarta	6.793	9.103
Jawa Barat	2.347	5.135
Jawa Timur	1.800	7.357
Jawa Tengah	909	7.625
Banten	342	1.032
DI Yogyakarta	1.394	2.222
Total	20.259	34.526

Berdasarkan **Tabel 3** menunjukkan bahwa provinsi Bali memiliki ketidakseimbangan sarana kepariwisataan yang mana memiliki jumlah akomodasi yang lebih tinggi dibandingkan dengan jumlah usaha penyediaan makanan dan minuman. Jumlah akomodasi yang tinggi mengindikasikan bahwa provinsi Bali memiliki kemudahan dalam pemenuhan kebutuhan menginap wisatawan dari suatu region. Jumlah usaha penyedia makan minum yang rendah mengindikasikan bahwa adanya ketidakmampuan dalam memenuhi kebutuhan makan terhadap kunjungan wisatawan/tourism. Data usaha penyediaan makan minum pada provinsi Bali berjumlah sedikit dikarenakan adanya usaha akomodasi yang merangkap menjadi usaha penyediaan makan dan minum, yaitu adanya layanan restaurant pada setiap hotel di Bali. Berdasarkan perbandingan tersebut dapat disimpulkan bahwa Provinsi Bali adalah daerah yang memiliki potensi untuk mengembangkan usaha penyediaan makanan minuman

Untuk mendapatkan daerah yang memiliki potensi untuk dikembangkan usaha sektor penyediaan makan minum di Provinsi Bali dapat dilihat melalui visualisasi perbandingan data. Perbandingan persebaran data akomodasi dan persebaran data usaha penyediaan makan minum di Provinsi Bali terdapat pada **Gambar 8**.

Gambar 8. Visualisasi Persebaran Data Akomodasi di Bali (Kiri) dan Visualisasi Persebaran Data Usaha Penyediaan Makan Minum di Bali (Kanan)

Visualisasi data persebaran (lihat **Gambar 8**) didapatkan bahwa usaha penyediaan makan minum pada provinsi Bali yang masih kosong ada pada Kota/Kabupaten Jembrana, Buleleng, Tabanan, Karangasem, dan Klungkung. Visualisasi persebaran data ini dapat dilakukan identifikasi pada level terendah yaitu desa/kelurahan sehingga, dapat membantu pebisnis/wirausahawan dalam mengidentifikasi daerah yang berpotensi untuk dikembangkan usaha akomodasi dan usaha penyediaan makan minum pada level daerah yang dibutuhkan

Berdasarkan **Tabel 3** juga didapatkan bahwa provinsi Jawa Tengah memiliki ketidakseimbangan sarana kepariwisataan yang mana memiliki jumlah usaha penyediaan makan minum yang lebih tinggi dibandingkan dengan jumlah usaha akomodasinya. Jumlah usaha penyediaan makan minum yang tinggi mengindikasikan bahwa provinsi Jawa Tengah memiliki kemudahan dalam pemenuhan kebutuhan makan dan minum oleh wisatawan di suatu region. Jumlah akomodasi yang rendah mengindikasikan bahwa provinsi Jawa Tengah belum mampu memenuhi kebutuhan menginap dari kunjungan wisatawan/tourism. Ketidakseimbangan ini menyimpulkan bahwa provinsi Jawa Tengah adalah daerah yang memiliki kemampuan untuk mengembangkan usaha di bidang akomodasi.

Untuk mendapatkan daerah yang memiliki potensi untuk dikembangkan usaha sektor akomodasi di Provinsi Jawa Tengah dapat dilihat melalui visualisasi perbandingan data. Perbandingan persebaran data akomodasi dan persebaran data usaha penyediaan makan minum di Provinsi Jawa Tengah terdapat pada **Gambar 8**.

Gambar 9. Visualisasi Persebaran Data Akomodasi di Jawa Tengah (Kiri) dan Visualisasi Persebaran Data Usaha Penyediaan Maka Minum di Jawa Tengah (Kanan)

Berdasarkan visualisasi data persebaran dapat dilihat bahwa usaha akomodasi pada provinsi Jawa Tengah yang masih kosong ada pada Kota/Kabupaten Cilacap, Blora, Grobogan, Batang, dan Kendal. Visualisasi persebaran data ini dapat dilakukan identifikasi pada level terendah yaitu desa/kelurahan sehingga, dapat membantu pebisnis/wirausahawan dalam mengidentifikasi daerah yang berpotensi untuk dikembangkan usaha akomodasi dan usaha penyediaan makan minum pada level daerah yang dibutuhkan.

KESIMPULAN

Pengumpulan data dengan metode *webscraping* pada situs *web google maps* berhasil dilakukan. Jumlah keseluruhan data yang telah melalui tahapan preprosesing data dan geocoding berjumlah 34.526 usaha penyediaan makan minum di Pulau Jawa dan Bali. Hasil nilai persentase *matching* yang dihasilkan dari perbandingan antara data hasil webscraping dengan data BPS sebesar 68,22%. Provinsi Bali adalah daerah yang memiliki potensi untuk mengembangkan usaha penyediaan makanan minuman terkhusus pada Kota/Kabupaten Jembrana, Buleleng, Tabanan, Karangasem, dan Klungkung. Sedangkan, provinsi Jawa Tengah adalah daerah yang memiliki potensi untuk mengembangkan usaha akomodasi terkhusus pada Kota/Kabupaten Cilacap, Blora, Grobogan, Batang, dan Kendal.

DAFTAR PUSTAKA

- Amanda Pratama Putra & Heny Wulandari. (2018). *Using Big Data for Accomodation Statistics: A New Challenge in Producing Tourism Statistics*. Asia-Pacific Economic Statistics Week 2018 Closing the gaps in economic statistics for sustainable development, 3-19.
- Amanda Pratama Putra. (2019). *Measuring Commuting Statistics in Indonesia Using Mobile Positioning Data*. Bangkok, Thailand: UNESCAP
- Anna Kurniawati, S.P. (2018). *Implementasi Algoritma Jaro-Winkler Distance untuk Membandingkan Kesamaan Dokumen Berbahasa Indonesia*. Depok: Universitas Gunadharma.
- Arkka Dhiratara, (2016). *Social Media Data Analytics for Tourism : A Preliminary Study*. Mekelweg, Netherlands : Delft University of Technology.
- Arifin, (2016). *Evaluasi Pembelajaran: Prinsip, Teknik, Dan Prosedur*. Bandung: PT Remaja Rosdakarya.
- BPS. (2015). *Proporsi Kontribusi Pariwisata Terhadap PDB*. Jakarta: Badan Pusat Statistik
- BPS. (2018). *Jumlah Kunjungan Wisman Menurut Kebangsaan dan Bulan Kedatangan Tahun*. Jakarta: Badan Pusat Statistik.
- BPS. (2017). *Statistik Penyediaan Makanan Dan Minuman*. Jakarta: Badan Pusat Statistik.
- Christen, (2012). *Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection*. Springer Science & Business Media.

- Chopde & Nichat (2013). *Landmark based shortest path detection by using A* and Haversine formula*. International Journal of Innovative Research in Computer and Communication Engineering, 1(2), 298-302.
- Deepak Kumar Mahto, L. S. (2016). *A Dive into Web Scraper World*. International Conference on Computing for Sustainable Global Development (INDIACom), 689-693.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining Concepts and Techniques Third Edition*. Waltham: Elsevier Inc.
- Indriantoro, Nur & Bambang Supomo. (2014). *Metodologi Penelitian Bisnis Untuk Akuntansi & Manajemen. Edisi 1*. Cetakan ke-12. Yogyakarta: BPFE.
- Ismanto, (2010). *Pemanfaatan Teknologi Informasi Berpengaruh terhadap Kinerja Individu Mahasiswa Jurusan Akuntansi Perbanas Surabaya*. E-Jurnal Akuntansi Universitas Udayana, 354-369.
- Kementrian Pariwisata. (2015). *Laporan Akhir Kajian Pengembangan Wisata Syariah*. Kementrian Pariwisata: Jakarta.
- Kitchin, Rob & McArdle, Gavin. (2016). *What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets*. Big Data & Society. 3. 10.1177/2053951716631130.
- Mahdia, F., & Noviyanto, F. (2013). *Pemanfaatan Google Maps API untuk pembangunan sistem informasi manajemen bantuan logistik pasca bencana alam berbasis mobile web (studi kasus: badan penanggulangan bencana daerah Kota Yogyakarta)*. Doctoral dissertation, Universitas Ahmad Dahlan.
- Narendra, A. P. (2015). *Big Data, Data Analyst, and Improving the Competence of Librarian*. Salatiga
- Pramana, S., Yuniarto, B., Kurniawan, R., Yordani, R., Lee, J., Amin, I., Satyaning P.P.N. L, Riyadi. Y., Hasyati A.N., Indriani, R. "Big data for government policy: Potential implementations of bigdata for official statistics in Indonesia," 2017 International Workshop on Big Data and Information Security (IWBIS), IEEE. Jakarta, 2017, pp. 17-21.
- Pearson, M. (2008). *Manajemen, Sistem Informasi* (Vol. 10). Jakarta: Salemba Empat.
- Rumata, V. M. (2016). *Peluang dan Tantangan Big Data dalam Penelitian Ilmu Sosial: sebuah Kajian Literatur*. Jurnal Penelitian Komunikasi dan Opini Publik, 20(1), 155-67.
- Sanjaya, B. A., & Sulisty, S. (2015). Big Data: *Inkonsistensi Data Dan Solusinya*. Semnasteknomedia Online, 3(1), 1-2.
- Setia Pramana, B. Y. (2017). *Big Data for Government Policy: Potential Implementations of BigData for Official Statistics in Indonesia*. Jakarta:IEEE.
- Sugiyono, (2015). *Metode penelitian Pendidikan (Pendekatan Kualitatif, Kuantitatif, dan R an D)*.
- Sujana, A. P. (2013). *Memfaatkan Big Data Untuk Mendeteksi Emosi*. Jurnal Teknik Komputer Unikom, 2(2), 1-4.
- Suwantoro, G. Sh. 2004. *Dasar-Dasar Pariwisata*. Jakarta: andi publisher.
- Top, P., Dowla, F., & Gansemer, J. (2007). *A dynamic programming algorithm for name matching*. In 2007 IEEE Symposium on Computational Intelligence and Data Mining (pp. 547-551). IEEE.
- Tukey, J. (1977). *Exploratory Data Analysis*. Philippines: Addison-Wesley Company, Inc.
- Wa Ode Zuhayeni Madjida & Takdir, Takdir. (2020). *Big Data for Official Statistics: Administrative Area Identification from Plain Text Address*.
- Yiea-Funk Te, et all. (2018). *Predicting The Growth of Restaurant using web data*. Switzerland.
- Yuri Demchenko, P. G. (2013). *Addressing Big Data Issues in Scientific Data Infrastructure*. Amsterdam.
- Yustiar Adhinugroho & Setia Pramana. (2019). *Pemanfaatan data scraping webtravel sebagai proyeksi potensi wisata*. Jakarta: STIS
- Yustiar Adhinugroho dkk, (2020). *Development of Online Travel Web Scraping for Tourism Statistics in Indonesia*, ISSN 1368-1613
- Yoeti, O. A. (1991). *Pengantar Ilmu Pariwisata*. Bandung: Angkasa
- Yoeti, O. A. (1996). *Pengantar Ilmu Pariwisata*. Bandung: Angkasa.
- Zaharuddin, H. (2009). *Menangkap Peluang Usaha*. Jakarta: Dian Anugrah Perkasa.
- Zainal Arifin, M. R. (2016). *Nearest Tourism Site Searching Using Haversine Method*. Samarinda: ICITACEE.