

94% Unique

Total 20163 chars, 2682 words, 127 unique sentence(s).

Custom Writing Services - Paper writing service you can trust. Your assignment is our priority! Papers ready in 3 hours! Proficient writing: top academic writers at your service 24/7! Receive a premium level paper!

STORE YOUR DOCUMENTS IN THE CLOUD - 1GB of private storage for free on our new file hosting!

Results	Query	Domains (original links)
Unique	These statistics are displayed in map visualization for ease of use	-
Unique	Penelitian ini bertujuan untuk mengembangkan visualisasi peta persebaran toko daring yang ada di Indonesia	-
Unique	Statistik ini ditampilkan dalam bentuk visualisasi peta untuk kemudahan penggunaan	-
Unique	Diharapkan hasil penelitian ini dapat dijadikan dasar untuk pemerataan perkembangan perdagangan elektronik di Indonesia	-
Unique	Diharapkan hasil penelitian ini dapat dijadikan dasar untuk pemerataan perkembangan perdagangan elektronik di Indonesia	-
Unique	Hal ini dilakukan karena tidak ada data publik yang siap untuk digunakan	-
Unique	Permasalahan ini dapat diatasi dengan memilih pasar daring yang menyediakan lokasi toko pada lamannya	-
Unique	Resource yang dibutuhkan cukup besar, mengingat banyaknya data yang tersedia	-
Unique	Tidak semua data dapat diambil	-
Unique	Beberapa pasar daring membatasi banyak data yang dapat diakses	-
Unique	Dengan demikian, tidak bisa didapatkan statistik yang 100% akurat	-
Unique	Untuk mengatasi permasalahan ketiga dan keempat di atas maka digunakan metode random sampling	-
Unique	Hanya saja masih belum ada penelitian yang fokus dalam memvisualisasikan persebaran toko daring	-
Unique	Dengan demikian, terdapat celah penelitian yang dapat diambil disini	-
Unique	Deteksi Outlier Beberapa metode untuk mendeteksi outlier antara lain sebagai berikut	-
Unique	Tukey's Test [4] Outlier dideteksi dengan menggunakan interquartile range	-
Unique	Dixon's Q Test [6] Data diurutkan secara naik:	-

Unique	Metodologi Tahapan yang penelitian sebagaimana ditunjukkan pada Gambar 1 antara lain sebagai berikut	-
Unique	Selain itu, digunakan juga 20,699,536 data dasar produk yang dijual oleh toko-toko tersebut	-
Unique	Produk-produk tersebut dapat dikelompokkan lebih lanjut ke dalam 1,462 kategori	-
Unique	Data tersebut selanjutnya disimpan dalam database untuk kemudian diproses lebih lanjut	-
Unique	Skema dari database ini ditunjukkan pada Gambar	-
1 results	Pembersihan data dilakukan untuk menghilangkan data-data yang bersifat outlier	scribd.com
Unique	Pada Tukey's Test, outlier dideteksi dengan menggunakan interquartile range (IQR)	-
Unique	Prosedur penerapan Tukey's test pada proses pembersihan data ini ditunjukkan pada pseudocode berikut	-
Unique	Namun, dalam visualisasinya, jumlah toko pada masing-masing kategori dihitung secara distinct berdasarkan ID-nya	-
Unique	Setelah perancangan dilakukan, software dibangun mengacu pada rancangan yang telah dibuat sebelumnya	-
Unique	Uji coba dan Evaluasi Visualisasi yang telah dibuat selanjutnya akan dievaluasi dari segi fungsi serta performanya	-
Unique	Tabel Data Komponen Toko Daring di Indonesia NoDataJumlah1Kota4602Kategori1.4623Toko Daring1.019.4934Produk20.699.536Gambar	-
2,370 results	Proses awal ditunjukkan pada Gambar	repository.uksw.edu lugman-aquarius.blogspot.com zh.scribd.com belajarlistrik456.blogspot.com wahyu2265.blogspot.com materikuliahteknik.info risetprojects.blogspot.com edoc.site hendeso.blogspot.com id.scribd.com
Unique	Server terlebih dahulu memeriksa apakah data yang diminta client telah tersimpan di cache	-
Unique	Jika ada, maka data langsung dikembalikan ke client sehingga dapat mempersingkat waktu pemrosesan	-
Unique	Jika data masih belum tersedia di cache, maka server melakukan query seperti alur awal	-
Unique	id/index.php/content/detail/6385/P/0/infografis	-
Unique	[2] PT Tokopedia, "Tokopedia," 2017	-
Unique	[3] PT Bukalapak.com, "Bukalapak," 2017	-
Unique	Tukey, Exploratory Data Analysis, 1st ed., New York: Pearson, 1977	-
1 results	Chauvenet, A Manual of Spherical and Practical Astronomy Vol	it.wikipedia.org
Unique	II, 5th ed., New York: Dover, 1863	-

13 results	Dixon, "Simplified Statistics for Small Numbers of Observations." Anal	en.wikipedia.org en.wikipedia.org sciencedirect.com backreaction.blogspot.com ijera.com returnofkings.com archive.org ideadiez.com lelive.free.fr
Unique	Starnes, The Practice of Statistics, 3rd ed., New York:	-
Unique	Friedman, "Data Visualization and Infographics," in Graphics, Monday Inspiration, 2008	-
Unique	Liu, "A survey on information visualization: recent advances and challenges." The Visual Computer, vol	-
Unique	Peikert, Eds., Vienna, Springer Vienna, 2001, pp	-
Unique	Bradie, A Friendly Introduction to Numerical Analysis, New York: Pearson Education, 2006	-
Unique	Arlinghaus, Practical Handbook of Curve Fitting, 1st ed., Florida: CRC Press, 1994	-
Unique	distribution, both in terms of the quantity of online stores and the types of goods	-
Unique	This study aims to develop a visualization of the distribution of online stores in	-
Unique	The data used in this study is crowdsourced data taken from one of online	-
71 results	The data is analyzed and processed to produce statistics on the distribution of online	patents.google.com patents.google.com mafiadoc.com archive.org dl.acm.org digitalhumanities.org mafiadoc.com pt.scribd.com dl.acm.org issuu.com
Unique	This visualization can be utilized by the government or stakeholders to monitor the development	-
Unique	types of goods sold online, comparing them to ordinary trade and typical products from each	-
Unique	Underdeveloped regions can then be motivated to develop their electronic commerce to their full	-
24,100 results	This research is also expected to be used as a reference for even distribution	researchgate.net patents.google.com researchgate.net sk.sagepub.com mdpi.com doi.acm.org en.wikipedia.org ar.scribd.com dl.acm.org rd.springer.com
Unique	saat ini sebagian besar hanya menunjukkan persebaran jumlah dan nilai transaksi, namun tidak menunjukkan persebaran	-
Unique	Data yang digunakan dalam penelitian ini adalah hasil dari crowdsourced data yang bersumber dari	-
Unique	Data dianalisis dan diproses untuk menghasilkan statistik tentang distribusi toko daring berdasarkan wilayah dan	-
Unique	Dengan adanya visualisasi tersebut, pemerintah atau pemegang kepentingan akan dapat memantau perkembangan toko daring	-
Unique	Daerah yang memiliki potensi namun belum berkembang secara maksimal dapat didorong untuk mengembangkan perdagangan	-
	Visualisasi tersebut dapat digunakan lebih lanjut untuk mendapatkan statistik	

Unique	maupun persebaran jenis-jenis barang yang	-
Unique	Kata kunci: visualisasi, toko daring, crowdsourced data, perdagangan elektronik© 2019 Jurnal SISFOHistori Artikel: Disubmit 25	-
Unique	Tersedia PendahuluanPerkembangan teknologi informasi dan komunikasi (TIK) yang pesat dalam dekade terakhir turut berperan	-
Unique	pada tahun 2015 sebanyak 93,4 juta jiwa dimana 77%-nya menggunakan internet untuk mencari informasi produk	-
Unique	Pada tahun yang sama, jumlah belanja daring di Indonesia mencapai 7,4 juta orang dengan	-
Unique	Angka ini diperkirakan akan naik secara signifikan pada tahun 2016 sehingga mencapai 8,7 juta	-
Unique	pembangunan infrastruktur TIK dan banyaknya jumlah penduduk, sektor perdagangan elektronik di Indonesia memiliki potensi yang	-
Unique	Untuk mengetahui potensi tersebut diperlukan statistik persebaran toko daring di Indonesia agar dapat menentukan	-
Unique	menunjukkan persebaran jumlah toko daring, jenis barang dan nilai transaksinyaPenelitian ini bertujuan untuk mengembangkan visualisasi	-
Unique	Dengan adanya visualisasi tersebut, pemerintah atau pemegang kepentingan akan dapat memantau perkembangan toko daring	-
Unique	Daerah yang memiliki potensi namun belum berkembang secara maksimal dapat didorong untuk mengembangkan perdagangan	-
Unique	Visualisasi tersebut dapat digunakan lebih lanjut untuk mendapatkan statistik maupun persebaran jenis-jenis barang yang	-
Unique	Metode yang gunakan dalam penelitian ini adalah ekstraksi data web (web scraping) untuk mengambil data	-
Unique	Metode ini banyak digunakan sebagai komponen dari pengindeksan web, integrasi data antar website, data	-
Unique	Beberapa permasalahan yang dihadapi dalam proses ekstraksi data ini antara lainTidak semua pasar daring mencantumkan	-
Unique	Sehingga untuk mendapatkan peta persebaran toko daring di Indonesia, hanya beberapa pasar daring yang	-
Unique	Beberapa pasar daring yang memenuhi kriteria tersebut antara lain adalah Tokopedia [2] dan Bukalapak	-
Unique	Karena data pasar daring merupakan crowdsourced data yang bersumber dari pengguna (toko daring), maka beberapa	-
Unique	Toko dapat mencantumkan harga produk yang jauh lebih rendah daripada yang sebenarnya untuk menarik	-
Unique	Permasalahan ini diduga dapat diatasi dengan menggunakan berbagai metode deteksi outlier, seperti Tukey's Test [4],	-
Unique	Hal ini selain menjadi beban bagi peneliti, juga membebani server pasar daring, terutama jika	-
Unique	Pada random sampling, beberapa data yang diambil sepenuhnya secara acak digunakan untuk mendapatkan statistik	-

Unique	Visualisasi Data Visualisasi data merupakan salah satu bagian dari sains data yang bertujuan untuk mengkomunikasikan	-
Unique	Visualisasi data yang baik , tidak hanya dapat mengkomunikasikan informasi dengan lebih baik, namun juga	-
Unique	Untuk menyampaikan gagasan secara efektif , bentuk visual dan fungsional harus seimbang, memberikan pengetahuan yang	-
Unique	Beberapa contoh diagram yang dapat digunakan untuk visualisasi data antara lain : diagram batang, histogram,	-
Unique	karena perannya dalam membantu manusia dalam memahami dan menganalisis data yang semakin meningkat baik dari	-
Unique	[10], visualisasi graf misalnya di bidang sosial media [11], teks [12], peta (data spasial) seperti	-
Unique	Di bidang perdagangan elektronik sendiri, visualisasi telah banyak dimanfaatkan, di antaranya untuk keperluan data	-
Unique	Visualisasi yang digunakan dalam penelitian ini berupa visualisasi peta karena melibatkan data-data wilayah yang	-
Unique	Jika Q1 dan Q3 melambangkan kuartil 1 dan 3, maka outlier adalah data dengan	-
Unique	Chauvenet's Criterion [5] Jika dan melambangkan rata-rata dan standar deviasi sampel, maka sebuah nilai dianggap outlier	-
Unique	Nilai dihitung sebagai rasio perbedaan dari nilai yang dicurigai sebagai outlier dan nilai terdekatnya	-
Unique	Metode Penelitian Pengumpulan Data Data yang digunakan dalam penelitian ini didapatkan dari salah satu pasar elektronik	-
Unique	Data ini mencakup informasi dasar dari 1,019,493 toko daring yang tersebar dalam 34 provinsi	-
Unique	Skema Database Awal Analisis dan Pra-pemrosesan Data Pada tahap ini, dilakukan pembersihan untuk data yang telah	-
Unique	Karena data pasar daring merupakan crowdsourced data yang bersumber dari pengguna (toko daring), maka	-
Unique	Toko dapat mencantumkan harga produk yang jauh lebih rendah daripada yang sebenarnya untuk menarik	-
Unique	Beberapa metode yang dapat digunakan dalam mendeteksi data outlier ini antara lain adalah Tukey's	-
Unique	Pada penelitian ini, Tukey's Test digunakan untuk mengidentifikasi dan menghilangkan data produk dengan harga	-
Unique	Jika Q1 dan Q3 melambangkan kuartil 1 dan 3, maka outlier adalah data dengan	-
Unique	Transformasi Data Data Transformation Data yang telah dibersihkan pada langkah sebelumnya selanjutnya diubah ke dalam	-
Unique	Skema Database Setelah Transformasi Dari skema ini terlihat bahwa sebuah toko daring dapat memiliki beberapa	-
Unique	dan & merupakan subkategori dari , maka dalam visualisasinya, jumlah toko dengan kategori adalah	-

Unique	Ilustrasi Cara Perhitungan Jumlah Toko Dalam Kategori Desain dan Pembangunan Software Visualisasi Perancangan visualisasi dilakukan	-
Unique	Software visualisasi yang dikembangkan berbasis web dan dilakukan dengan menggunakan Apache sebagai web servernya.	-
Unique	Performa software visualisasi akan diukur menggunakan waktu akses yang terdiri atas tiga kategori, yaitu	-
Unique	Hasil dan Pembahasan 4.1 Data Tabel 1 adalah data yang sudah dikumpulkan melalui proses	-
Unique	Hasil Deteksi Outlier Gambar 5 adalah data yang dihasilkan dari proses pembersihan untuk menghilangkan data-data	-
Unique	Persebaran Toko Daring di Indonesia Penelitian ini menghasilkan sebuah Visualisasi persebaran toko daring di Indonesia	-
Unique	Persebaran toko daring dapat dilihat dengan mudah melalui bentuk peta persebaran yang dihasilkan dalam	-
Unique	Persebaran dan jumlah toko daring dapat dilihat di setiap propinsi dengan cara menyentuh masing-masing	-
Unique	Visualisasi Peta Persebaran Toko Daring di Indonesia Gambar 7 menyajikan statistik persebaran toko daring berdasarkan	-
Unique	elektronik, mengetahui produk khas dari masing-masing wilayah sehingga dapat dijadikan dasar untuk pengambilan keputusan dan	-
Unique	Visualisasi Peta Persebaran Toko Daring Berdasarkan Daerah dan Ketogori 4.3 Pengujian dan Perbaikan Pengujian dilakukan untuk	-
Unique	Rata-rata waktu akses untuk proses data di ShopVis adalah lebih 5 detik, sebagaimana terlihat	-
Unique	Jika dilihat dari komposisi waktu aksesnya, lamanya akses halaman ini didominasi pada proses pengambilan	-
Unique	Sedangkan akses komponen halaman lain yaitu HTML Document dan Assets (CSS, Javascript) hanya memakan	-
Unique	Gambar 8 Rata-rata durasi akses halaman sebelum perbaikan Untuk mempercepat akses halaman ini, maka dilakukan	-
Unique	Pada saat client melakukan request ke server ShopVis, maka server akan melakukan query ke	-
Unique	Proses query ke database ini dapat memakan waktu yang lama karena banyaknya data yang	-
Unique	sama, maka dilakukan perbaikan pada proses dengan menambahkan mekanisme caching menggunakan file seperti yang ditunjukkan	-
Unique	Gambar 10 Alur Akses Setelah Dilakukan Perbaikan Setelah dilakukan penambahan cache, server tidak langsung melakukan	-
Unique	Namun, sebelum data dikembalikan ke client, data terlebih dahulu disimpan dalam cache sehingga waktu	-
Unique	Proses perbaikan ini mampu mempercepat waktu akses rata-rata menjadi kurang dari 0.3 detik seperti	-
Unique	Gambar 11 Rata-rata durasi akses halaman setelah perbaikan Kesimpulan 5.1 Simpulan Penelitian ini menghasilkan visualisasi peta persebaran	-

Unique	Pemerintah atau pemegang kepentingan dapat menggunakan visualisasi tersebut untuk memantau kondisi toko daring di	
Unique	Hasil pengujian menunjukkan bahwa software visualisasi yang telah dibangun berfungsi baik dengan performa yang	
Unique	5.2 Saran Hasil penelitian ini sangat mungkin untuk dilanjutkan untuk riset pasar bisnis, pengelompokan (clustering) daerah-daerah	
Unique	Penelitian selanjutnya dapat dilakukan dengan menggabungkan data dari beberapa pasar daring yang ada di	
Unique	Daftar Rujukan[1] Kementerian Komunikasi dan Informatika. "Kementerian Komunikasi dan Informatika: Potensi Besar untuk Pemasukan	
100 results	Grandjean, "Introduction à la visualisation de données, l'analyse de réseau en histoire." Geschichte und	martingrandjean.ch academia.edu martingrandjean.ch researchgate.net halshs.archives-ouvertes.fr blog.ahc-ch.ch github.com stonybrook.edu en.wikipedia.org unil.academia.edu
Unique	Nick, "Asymmetric Relations in Longitudinal Social Networks." IEEE Transactions on Visualization and Computer Graphics,	
Unique	Ebert, "Spatial Text Visualization Using Automatic Typographic Maps." IEEE Transactions on Visualization and Computer	
Unique	Wood, "Exploring Uncertainty in Geodemographics with Interactive Graphics." IEEE Transactions on Visualization and Computer	
1 results	Qu, "DICON: Interactive Visual Analysis of Multidimensional Clusters." IEEE Transactions on Visualization and Computer	huamin.org
212 results	of the Joint Eurographics --- IEEE TCVG Symposium on Visualization in Ascona, Switzerland, May 28--30,	institute-of-health-and-wellbeing.org.uk sites.google.com worldcat.org jet.com ebookphp.com firemarkeducation.com britishsolarrenewables.com infibeam.com link.springer.com oxfordrefugeecampaign.org

Top plagiarizing domains: **en.wikipedia.org** (4 matches); **patents.google.com** (3 matches); **dl.acm.org** (3 matches); **researchgate.net** (3 matches); **mafiadoc.com** (2 matches); **archive.org** (2 matches); **martingrandjean.ch** (2 matches); **doi.acm.org** (1 matches); **blog.ahc-ch.ch** (1 matches); **halshs.archives-ouvertes.fr** (1 matches); **ar.scribd.com** (1 matches); **github.com** (1 matches); **rd.springer.com** (1 matches); **academia.edu** (1 matches); **unil.academia.edu** (1 matches); **britishsolarrenewables.com** (1 matches); **firemarkeducation.com** (1 matches); **infibeam.com** (1 matches); **link.springer.com** (1 matches); **oxfordrefugeecampaign.org** (1 matches); **ebookphp.com** (1 matches); **jet.com** (1 matches); **huamin.org** (1 matches); **mdpi.com** (1 matches); **institute-of-health-and-wellbeing.org.uk** (1 matches); **sites.google.com** (1 matches); **worldcat.org** (1 matches); **stonybrook.edu** (1 matches); **pt.scribd.com** (1 matches); **materikuliahteknik.info** (1 matches); **risetprojects.blogspot.com** (1 matches); **edoc.site** (1 matches); **hendeso.blogspot.com** (1 matches); **wahyu2265.blogspot.com** (1 matches); **belajarlistrik456.blogspot.com** (1 matches); **repository.uksw.edu** (1 matches); **luqman-aquarius.blogspot.com** (1 matches); **zh.scribd.com** (1 matches); **id.scribd.com** (1 matches); **it.wikipedia.org** (1 matches); **lalive.free.fr** (1 matches); **digitalhumanities.org** (1 matches); **scribd.com** (1 matches); **issuu.com** (1 matches); **ideadiez.com** (1 matches); **returnofkings.com** (1 matches); **sciencedirect.com** (1 matches); **backreaction.blogspot.com** (1 matches); **ijera.com** (1 matches); **sk.sagepub.com** (1 matches);

number and value of transactions, but not the distribution, both in terms of the quantity of online stores and the types of goods sold. This study aims to develop a visualization of the distribution of online stores in Indonesia. The data used in this study is crowdsourced data taken from one of online marketplaces in Indonesia. The data is analyzed and processed to produce statistics on the distribution of online stores by region and type of goods. These statistics are displayed in map visualization for ease of use. This visualization can be utilized by the government or stakeholders to monitor the development of online stores in each region and make strategic decisions. The visualization can be used further to obtain statistics and the distribution of the types of goods sold online, comparing them to ordinary trade and typical products from each region. Underdeveloped regions can then be motivated to develop their electronic commerce to their full potential. This research is also expected to be used as a

reference for even distribution of electronic commerce in Indonesia. Keywords: crowdsourced data, online marketplaces, map visualization, electronic commerce

Abstrak Statistik distribusi perdagangan elektronik di Indonesia saat ini sebagian besar hanya menunjukkan persebaran jumlah dan nilai transaksi, namun tidak menunjukkan persebaran jumlah toko daring, jenis barang dan nilai transaksinya. Penelitian ini bertujuan untuk mengembangkan visualisasi peta persebaran toko daring yang ada di Indonesia. Data yang digunakan dalam penelitian ini adalah hasil dari crowdsourced data yang bersumber dari toko daring di Indonesia. Data dianalisis dan diproses untuk menghasilkan statistik tentang distribusi toko daring berdasarkan wilayah dan jenis barang. Statistik ini ditampilkan dalam bentuk visualisasi peta untuk kemudahan penggunaan. Dengan adanya visualisasi tersebut, pemerintah atau pemegang kepentingan akan dapat memantau perkembangan toko daring di masing-masing wilayah. Daerah yang memiliki potensi namun belum berkembang secara maksimal dapat didorong untuk mengembangkan perdagangan elektroniknya. Visualisasi tersebut dapat digunakan lebih lanjut untuk mendapatkan statistik maupun persebaran jenis-jenis barang yang dijual secara daring, membandingkannya dengan perdagangan biasa serta produk khas dari masing-masing wilayah. Diharapkan hasil penelitian ini dapat dijadikan dasar untuk pemerataan perkembangan perdagangan elektronik di Indonesia. Kata kunci: visualisasi, toko daring, crowdsourced data, perdagangan elektronik © 2019 Jurnal SISFOH Histori Artikel: Disubmit 25 Januari 2019; Diterima ; Tersedia

Pendahuluan Perkembangan teknologi informasi dan komunikasi (TIK) yang pesat dalam dekade terakhir turut berperan dalam berkembangnya sektor perdagangan elektronik (e-commerce) di Indonesia. Berdasarkan data Kementerian Komunikasi dan Informatika (Kominfo) Indonesia [1], jumlah pengguna internet di Indonesia pada tahun 2015 sebanyak 93,4 juta jiwa dimana 77%-nya menggunakan internet untuk mencari informasi produk dan berbelanja secara daring. Pada tahun yang sama, jumlah belanja daring di Indonesia mencapai 7,4 juta orang dengan nilai USD 3,56 milyar. Angka ini diperkirakan akan naik secara signifikan pada tahun 2016 sehingga mencapai 8,7 juta orang dan total nilai belanja sebesar USD 4,89 milyar. Dengan masih rendahnya persentase pengguna internet yang melakukan belanja daring serta didukung semakin meratanya pembangunan infrastruktur TIK dan banyaknya jumlah penduduk, sektor perdagangan elektronik di Indonesia memiliki potensi yang sangat besar untuk dikembangkan. Untuk mengetahui potensi tersebut diperlukan statistik persebaran toko daring di Indonesia agar dapat menentukan potensi setiap daerah. Statistik yang ada saat ini hanya menunjukkan persebaran jumlah dan nilai transaksi, namun tidak menunjukkan persebaran jumlah toko daring, jenis barang dan nilai transaksinya. Penelitian ini bertujuan untuk mengembangkan visualisasi peta persebaran toko daring yang ada di Indonesia. Dengan adanya visualisasi tersebut, pemerintah atau pemegang kepentingan akan dapat memantau perkembangan toko daring di masing-masing wilayah. Daerah yang memiliki potensi namun belum berkembang secara maksimal dapat didorong untuk mengembangkan perdagangan elektroniknya. Visualisasi tersebut dapat digunakan lebih lanjut untuk mendapatkan statistik maupun persebaran jenis-jenis barang yang dijual secara daring, membandingkannya dengan perdagangan biasa serta produk khas dari masing-masing wilayah. Diharapkan hasil penelitian ini dapat dijadikan dasar untuk pemerataan perkembangan perdagangan elektronik di Indonesia. Metode yang digunakan dalam penelitian ini adalah ekstraksi data web (web scraping) untuk mengambil data dari pasar daring (marketplace) yang ada di Indonesia. Hal ini dilakukan karena tidak ada data publik yang siap untuk digunakan. Metode ini banyak digunakan sebagai komponen dari pengindeksan web, integrasi data antar website, data mining, dan aplikasi-aplikasi lain seperti deteksi perubahan dan perbandingan harga produk daring. Beberapa permasalahan yang dihadapi dalam proses ekstraksi data ini antara lain Tidak semua pasar daring mencantumkan lokasi toko. Sehingga untuk mendapatkan peta persebaran toko daring di Indonesia, hanya beberapa pasar daring yang dapat digunakan sebagai sumber data. Permasalahan ini dapat diatasi dengan memilih pasar daring yang menyediakan lokasi toko pada lamannya. Beberapa pasar daring yang memenuhi kriteria tersebut antara lain adalah Tokopedia [2] dan Bukalapak [3]. Karena data pasar daring merupakan crowdsourced data yang bersumber dari pengguna (toko daring), maka beberapa data mungkin tidak akurat. Toko dapat mencantumkan harga produk yang jauh lebih rendah daripada yang sebenarnya untuk menarik minat pelanggan. Permasalahan ini diduga dapat diatasi dengan menggunakan berbagai metode deteksi outlier, seperti Tukey's Test [4], Chauvenet's criterion [5], dan

Dixon's Q test [6]. Resource yang dibutuhkan cukup besar, mengingat banyaknya data yang tersedia. Hal ini selain menjadi beban bagi peneliti, juga membebani server pasar daring, terutama jika data dengan volume besar diakses dalam waktu yang singkat. Tidak semua data dapat diambil. Beberapa pasar daring membatasi banyak data yang dapat diakses. Dengan demikian, tidak bisa didapatkan statistik yang 100% akurat. Untuk mengatasi permasalahan ketiga dan keempat di atas maka digunakan metode random sampling. Pada random sampling, beberapa data yang diambil sepenuhnya secara acak digunakan untuk mendapatkan statistik yang mewakili keseluruhan data yang ada [7].

2.1. Visualisasi Data
Visualisasi data merupakan salah satu bagian dari sains data yang bertujuan untuk mengkomunikasikan informasi secara jelas dan efisien menggunakan grafik, plot, dan infografis. Visualisasi data yang baik, tidak hanya dapat mengkomunikasikan informasi dengan lebih baik, namun juga akan membantu untuk memahami data secara lebih jelas dan mendalam [8]. Untuk menyampaikan gagasan secara efektif, bentuk visual dan fungsional harus seimbang, memberikan pengetahuan yang mendalam mengenai data yang kompleks dengan mengkomunikasikan aspek-aspek utamanya dengan cara yang lebih intuitif [9]. Beberapa contoh diagram yang dapat digunakan untuk visualisasi data antara lain: diagram batang, histogram, scatter plot, diagram jaring, streamgraph, treemap, gantt chart, dan heat map. Dalam tahun-tahun terakhir, visualisasi telah menjadi salah satu bidang penelitian yang aktif dan vital karena perannya dalam membantu manusia dalam memahami dan menganalisis data yang semakin meningkat baik dari segi volume maupun availabilitasnya.

Visualisasi telah banyak diterapkan dalam berbagai macam bidang yang dapat dikelompokkan menjadi empat kategori [10], visualisasi graf misalnya di bidang sosial media [11], teks [12], peta (data spasial) seperti demografi [13], dan data multivariat [14]. Di bidang perdagangan elektronik sendiri, visualisasi telah banyak dimanfaatkan, di antaranya untuk keperluan data mining [15]. Hanya saja masih belum ada penelitian yang fokus dalam memvisualisasikan persebaran toko daring. Dengan demikian, terdapat celah penelitian yang dapat diambil disini. Visualisasi yang digunakan dalam penelitian ini berupa visualisasi peta karena melibatkan data-data wilayah yang ada di Indonesia sehingga untuk memudahkan pengguna dalam mengakses wilayah tersebut secara langsung melalui peta.

2.2. Deteksi Outlier
Beberapa metode untuk mendeteksi outlier antara lain sebagai berikut. Tukey's Test [4] Outlier dideteksi dengan menggunakan interquartile range. Jika Q1 dan Q3 melambungkan kuartil 1 dan 3, maka outlier adalah data dengan nilai diluar range. Chauvenet's Criterion [5] Jika dan melambungkan rata-rata dan standar deviasi sampel, maka sebuah nilai dianggap outlier jika lebih besar dari nilai yang ada pada tabel nilai kriteria Chauvenet. Dixon's Q Test [6] Data diurutkan secara naik: . Nilai dihitung sebagai rasio perbedaan dari nilai yang dicurigai sebagai outlier dan nilai terdekatnya dengan range kemudian hasilnya dibandingkan dengan tabel nilai kritis.

3. Metodologi Tahapan yang penelitian sebagaimana ditunjukkan pada Gambar 1 antara lain sebagai berikut. Gambar 1. Metode Penelitian
Pengumpulan Data
Data yang digunakan dalam penelitian ini didapatkan dari salah satu pasar elektronik yang ada di Indonesia. Data ini mencakup informasi dasar dari 1,019,493 toko daring yang tersebar dalam 34 provinsi dan 460 kota serta kabupaten yang ada di seluruh Indonesia pada bulan April 2018.

Selain itu, digunakan juga 20,699,536 data dasar produk yang dijual oleh toko-toko tersebut. Produk-produk tersebut dapat dikelompokkan lebih lanjut ke dalam 1,462 kategori. Data tersebut selanjutnya disimpan dalam database untuk kemudian diproses lebih lanjut. Skema dari database ini ditunjukkan pada Gambar 2. Gambar 2. Skema Database Awal
Analisis dan Pra-pemrosesan Data
Pada tahap ini, dilakukan pembersihan untuk data yang telah diperoleh sebelumnya. **Pembersihan data dilakukan untuk menghilangkan data-data yang bersifat outlier.**

Karena data pasar daring merupakan crowdsourced data yang bersumber dari pengguna (toko daring), maka beberapa data mungkin tidak akurat. Toko dapat mencantumkan harga produk yang jauh lebih rendah daripada yang sebenarnya untuk menarik minat pelanggan. Beberapa metode yang dapat digunakan dalam mendeteksi data outlier ini antara lain adalah Tukey's Test [4], Chauvenet's criterion [5], dan Dixon's Q test [6]. Pada penelitian ini, Tukey's Test digunakan untuk mengidentifikasi dan menghilangkan data produk dengan harga di luar batas wajar pada masing-masing kategori. Pada Tukey's Test, outlier dideteksi dengan menggunakan interquartile range (IQR).

Jika Q1 dan Q3 melambungkan kuartil 1 dan 3, maka outlier adalah data dengan nilai diluar range. Prosedur penerapan Tukey's test pada proses pembersihan data ini ditunjukkan pada pseudocode berikut. Transformasi Data /

Data Transformation
Data yang telah dibersihkan pada langkah sebelumnya selanjutnya diubah ke dalam format yang mudah untuk dapat digunakan dalam tahapan selanjutnya sebagaimana ditunjukkan pada gambar 3. Gambar 3. Skema Database Setelah Transformasi
Dari skema ini terlihat bahwa sebuah toko daring dapat memiliki beberapa kategori berbeda. Namun, dalam visualisasinya, jumlah toko pada masing-masing kategori dihitung secara distinct berdasarkan ID-nya. Misal, toko menjual barang dengan kategori dan , toko menjual barang dengan kategori , dan & merupakan subkategori dari , maka dalam visualisasinya, jumlah toko dengan kategori adalah 2 seperti diilustrasikan pada gambar 4 berikut. Gambar 4. Ilustrasi Cara Perhitungan Jumlah Toko Dalam Kategori

Desain dan Pembangunan Software
Visualisasi Perancangan visualisasi dilakukan untuk memilih antarmuka yang dapat mengkomunikasikan informasi dengan lebih efektif. Setelah perancangan dilakukan, software dibangun mengacu

pada rancangan yang telah dibuat sebelumnya. Software visualisasi yang dikembangkan berbasis web dan dilakukan dengan menggunakan Apache sebagai web servernya, PHP sebagai bahasa pemrograman, dan MySQL untuk databasenya. Uji coba dan Evaluasi Visualisasi yang telah dibuat selanjutnya akan dievaluasi dari segi fungsi serta performanya. Performa software visualisasi akan diukur menggunakan waktu akses yang terdiri atas tiga kategori, yaitu document (base HTML), assets (CSS, javascript), dan data yang divisualisasikan (JSON).

4. Hasil dan Pembahasan

4.1 Data Tabel 1 adalah data yang sudah dikumpulkan melalui proses crowdsourced data yang bersumber dari pengguna (toko daring) di Indonesia. Tabel 1. Tabel Data Komponen Toko Daring di Indonesia

No	Data	Jumlah	Kota	Kategori
1	4623	Toko Daring	1.019.493	4
2	699.536	Produk	20.699.536	6

Gambar 5. Hasil Deteksi Outlier

Gambar 5 adalah data yang dihasilkan dari proses pembersihan untuk menghilangkan data-data yang bersifat outlier (data yang tidak akurat). Setelah dikurangi dengan toko yang "non aktif" maka data yang dihasilkan lebih efektif dibanding data sebelumnya yaitu: (1) Produk sebanyak 20,363,623 dan (2) toko sebanyak 544,107.

4.2 Hasil Visualisasi Peta Persebaran Toko Daring di Indonesia

Penelitian ini menghasilkan sebuah Visualisasi persebaran toko daring di Indonesia berbasis website dan selanjutnya disebut ShopVis. Persebaran toko daring dapat dilihat dengan mudah melalui bentuk peta persebaran yang dihasilkan dalam penelitian ini. Persebaran dan jumlah toko daring dapat dilihat di setiap propinsi dengan cara menyentuh masing-masing propinsi di peta, sebagaimana tampak pada Gambar 6. Gambar 6. Visualisasi Peta Persebaran Toko Daring di Indonesia

Gambar 7 menyajikan statistik persebaran toko daring berdasarkan daerah dan kategori toko. Hal ini dapat digunakan lebih lanjut untuk membandingkan jenis perdagangan biasa dan perdagangan secara elektronik, mengetahui produk khas dari masing-masing wilayah sehingga dapat dijadikan dasar untuk pengambilan keputusan dan pemerataan perkembangan perdagangan elektronik di Indonesia.

4.3 Pengujian dan Perbaikan

Pengujian dilakukan untuk memaksimalkan kinerja ShopVis. Rata-rata waktu akses untuk proses data di ShopVis adalah lebih 5 detik, sebagaimana terlihat pada Gambar 7. Jika dilihat dari komposisi waktu aksesnya, lamanya akses halaman ini didominasi pada proses pengambilan data dari web server yang mencapai 96.4% (5639.5 ms). Sedangkan akses komponen halaman lain yaitu HTML Document dan Assets (CSS, Javascript) hanya memakan 3.6% dari total akses waktu. Gambar 8 Rata-rata durasi akses halaman sebelum perbaikan

Untuk mempercepat akses halaman ini, maka dilakukan perubahan pada metode pemrosesan request pengunjung laman visualisasi. Proses awal ditunjukkan pada Gambar 9. Pada saat client melakukan request ke server ShopVis, maka server akan melakukan query ke database, mengolahnya, kemudian menampilkan hasilnya ke pengguna. Proses query ke database ini dapat memakan waktu yang lama karena banyaknya data yang disimpan dalam database. Gambar 9 Alur Akses Awal

Dengan pertimbangan data yang diakses masing-masing pengunjung adalah data yang sama, maka dilakukan perbaikan pada proses dengan menambahkan mekanisme caching menggunakan file seperti yang ditunjukkan pada Gambar 10. Gambar 10 Alur Akses Setelah Dilakukan Perbaikan

Setelah dilakukan penambahan cache, server tidak langsung melakukan query ke database ketika menerima request dari client. Server terlebih dahulu memeriksa apakah data yang diminta client telah tersimpan di cache. Jika ada, maka data langsung dikembalikan ke client sehingga dapat mempersingkat waktu pemrosesan. Jika data masih belum tersedia di cache, maka server melakukan query seperti alur awal. Namun, sebelum data dikembalikan ke client, data terlebih dahulu disimpan dalam cache sehingga waktu proses berikutnya akan lebih singkat. Proses perbaikan ini mampu mempercepat waktu akses rata-rata menjadi kurang dari 0.3 detik seperti yang ditunjukkan pada Gambar 11. Gambar 11 Rata-rata durasi akses halaman setelah perbaikan

5. Kesimpulan

5.1 Simpulan Penelitian ini menghasilkan visualisasi peta persebaran toko daring yang ada di Indonesia berdasarkan daerah dan kategori produk. Pemerintah atau pemegang kepentingan dapat menggunakan visualisasi tersebut untuk memantau kondisi toko daring di masing-masing wilayah di Indonesia meliputi jumlah persebaran toko dan jenis-jenis barang yang dijual. Hasil pengujian menunjukkan bahwa software visualisasi yang telah dibangun berfungsi baik dengan performa yang cukup memuaskan dengan waktu akses rata-rata sebesar kurang dari 0.3 detik untuk masing-masing request.

5.2 Saran

Hasil penelitian ini sangat mungkin untuk dilanjutkan untuk riset pasar bisnis, pengelompokan (clustering) daerah-daerah dengan karakteristik yang sama. Penelitian selanjutnya dapat dilakukan dengan menggabungkan data dari beberapa pasar daring yang ada di Indonesia.

Daftar Rujukan

- [1] Kementerian Komunikasi dan Informatika, "Kementerian Komunikasi dan Informatika: Potensi Besar untuk Pemasukan Negara," 12 11 2015. [Online]. Available: <https://kominfo.go.id/index.php/content/detail/6385/P/0/infografis>. [Accessed 8 Juni 2017].
- [2] PT Tokopedia, "Tokopedia," 2017. [Online]. Available: <https://www.tokopedia.com/>. [Accessed 8 June 2017].
- [3] PT Bukalapak.com, "Bukalapak," 2017. [Online]. Available: <https://www.bukalapak.com/>. [Accessed 8 Juni 2017].
- [4] J. W. Tukey, *Exploratory Data Analysis*, 1st ed., New York: Pearson, 1977.
- [5] W. Chauvenet, *A Manual of Spherical and Practical Astronomy Vol. II*, 5th ed., New York: Dover, 1863.
- [6] R. B. Dean and W. J. Dixon, "Simplified Statistics for Small Numbers of Observations," *Anal. Chem.*, vol. 23, no. 4, pp. 636-638, 1951.
- [7] D.

Yates, D. S. Moore and . D. S. Starnes, *The Practice of Statistics*, 3rd ed., New York: W. H. Freeman, 2008. [8] M. Grandjean, "Introduction à la visualisation de données, l'analyse de réseau en histoire," *Geschichte und Informatik*, vol. 18, no. 19, pp. 109-128, 2015. [9] V. Friedman, "Data Visualization and Infographics," in *Graphics, Monday Inspiration*, 2008. [10] S. Liu, W. Cui, Y. Wu and M. Liu, "A survey on information visualization: recent advances and challenges," *The Visual Computer*, vol. 30, no. 12, p. 1373–1393, 2014. [11] U. Brandes and B. Nick, "Asymmetric Relations in Longitudinal Social Networks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2283-2290, 2011. [12] S. Afzal, R. Maciejewski, Y. Jang, N. Elmqvist and D. S. Ebert, "Spatial Text Visualization Using Automatic Typographic Maps," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2556-2564, 2012. [13] A. Slingsby, J. Dykes and J. Wood, "Exploring Uncertainty in Geodemographics with Interactive Graphics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2545-2554, 2011. [14] N. Cao, D. Gotz, J. Sun and H. Qu, "DICON: Interactive Visual Analysis of Multidimensional Clusters," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2581-2590, 2011. [15] M. C. Hao, U. Dayal, M. Hsu, T. Sprenger and M. H. Gross, "Visualization of directed associations in e-commerce transaction data," in *Data Visualization 2001: Proceedings of the Joint Eurographics --- IEEE TCVG Symposium on Visualization in Ascona, Switzerland, May 28--30, 2001*, D. S. Ebert, J. M. Favre and R. Peikert, Eds., Vienna, Springer Vienna, 2001, pp. 185-192. [16] B. Bradie, *A Friendly Introduction to Numerical Analysis*, New York: Pearson Education, 2006. [17] S. Arlinghaus, *Practical Handbook of Curve Fitting*, 1st ed., Florida: CRC Press, 1994.