

Pendekatan Eksploratif dalam Analisis Data Harga Pangan Nasional dengan Tableau

Dzul Fadli Rahman^{1,*}, Agus Qomaruddin Munir²

^{1,2}Departemen Pendidikan Teknik Elektronika dan Informatika, Universitas Negeri Yogyakarta, Indonesia

¹dzulfadlirahman@uny.ac.id; ²agusqomaruddin@uny.ac.id

* penulis korespondensi

INFO ARTIKEL

Sejarah Artikel

Diterima: 6 Agustus 2024
Direvisi: 17 Agustus 2024
Diterbitkan: 30 Agustus 2024

Kata Kunci

Exploratory Data Analysis
Harga Pangan
Python
Tableau
Visualisasi Data

ABSTRAK

Penelitian ini menganalisis data Harga Pangan Nasional di Indonesia dari tahun 2020 hingga 2023 menggunakan metode *Exploratory Data Analysis* (EDA). Python, *Visual Studio Code* dan Tableau dipilih untuk kemudahan dalam penggunaan *library* statistik dan visualisasi data. Analisis dilakukan dalam tiga bentuk: univariat, bivariat, dan multivariat. Hasil univariat menunjukkan distribusi pasar dan sumber pangan, dengan dominasi bahan nabati. Analisis bivariat menunjukkan harga pangan lebih tinggi di wilayah timur, dengan daging sebagai komoditas termahal dan beras yang termurah. Analisis multivariat membandingkan harga pangan di Yogyakarta dan Merauke, menunjukkan harga lebih tinggi di Merauke kecuali beras yang stabil di kedua kota. Studi ini juga menekankan pentingnya visualisasi data dalam mempermudah pemahaman informasi kompleks, menggunakan berbagai jenis grafik untuk menunjukkan pola dan tren yang membantu pengambilan keputusan strategis. Hasil penelitian ini dapat dikembangkan untuk memprediksi perubahan harga pangan dan memberikan wawasan yang lebih dalam.

PENDAHULUAN

Di Abad ke-21, data dianggap sebagai minyak bumi, bahan bakar kontemporer yang kaya dan bervolume besar, yang memiliki fitur bawaan yang meningkatkan kemampuan *intelligent decision making* dalam segala aspek, seperti bisnis atau layanan [1]. Berdasarkan Data Harga Pangan Nasional di Indonesia, yang bersumber dari *World Food Programme Price Database*, diketahui bahwa beberapa komoditas pangan mengalami perubahan harga yang fluktuatif setiap bulannya, dan terdapat perbedaan harga pangan yang cukup signifikan antara satu daerah dengan daerah lainnya. Fluktuatif dan perbedaan harga pangan tentu dapat menjadi masalah serius yang perlu dianalisis untuk mencapai ketahanan pangan nasional yang maksimal [2]. Ketahanan pangan sendiri selalu menjadi isu dalam pembangunan nasional Indonesia dari jaman orde lama, orde baru hingga sekarang [3].

Dengan sumber daya alam yang melimpah, Indonesia dikenal sebagai negara agraris, tetapi ini tidak selalu berarti bahwa orang-orang di Indonesia dapat mendapatkan makanan dengan mudah dan murah [4]. Menurut data impor BPS (Badan Pusat Statistik), Indonesia mengimpor beras kurang lebih sebanyak 2,2 juta ton pada tahun 2015 dengan nilai sekitar 1 milyar USD (BPS 2020). Sejak tahun 2000, Indonesia telah mengimpor beras dalam jumlah yang bervariasi dengan nilai jutaan USD [5]. Fakta ini menunjukkan bahwa ketergantungan pada impor pangan dapat mempengaruhi ketimpangan harga pangan di dalam negeri, sehingga memperbesar tantangan dalam mewujudkan ketahanan pangan yang kuat di seluruh wilayah Indonesia.

Pendekatan eksploratif dengan menggunakan Tableau diperlukan untuk *dataset* harga pangan nasional di Indonesia, karena memungkinkan analisis yang lebih mendalam, cepat, dan interaktif dalam memahami dan mengatasi masalah ketahanan pangan di Indonesia, serta menjadikan analisis data sebagai faktor penting dalam pengembangan *intelligent decision making* [6].

Exploratory Data Analysis (EDA) merupakan pendekatan yang biasa digunakan untuk menganalisis dan menginterpolasi informasi yang berguna melalui grafik statistik dan visualisasi, untuk meningkatkan pemahaman terhadap sebuah *dataset* yang diteliti. EDA membantu analis mengenali pola, hubungan, dan kejadian tak terduga [7]. EDA telah berkembang dan digunakan secara luas di berbagai bidang untuk menyelesaikan masalah yang ada. EDA membantu menganalisis kumpulan data untuk meringkas karakteristik statistiknya yang berfokus pada empat aspek utama: *measures of central tendency* (*mean*, *mode* dan *median*), *measures of spread* (*standard deviation* dan *variance*), *shape of distribution* dan *existence of outliers* [8]. Terdapat berbagai teknik EDA, kebanyakan bersifat grafis dengan beberapa teknik kuantitatif untuk tujuan yang berbeda [9]. Analisis bisa dilakukan secara fleksibel tanpa aturan yang ketat [10]. EDA bersifat subjektif dan memerlukan pengetahuan serta pengalaman untuk mendapatkan hasil yang optimal, termasuk memahami asal-usul data [7].

Pemahaman manusia terhadap data visual 60.000 kali lebih cepat dibandingkan teks. Saat ini, 90% informasi yang diterima oleh otak bersifat visual [11][12]. Visualisasi data membantu organisasi modern memahami lebih baik informasi yang dihasilkan dari dalam dan luar perusahaan. Memeriksa berbagai lembar kerja, tabel, atau dokumen bisa menjadi melelahkan, sementara melihat bagan dan grafik selalu lebih sederhana dan mudah untuk mata [13]. Salah satu platform visualisasi data adalah Tableau, sebuah alat yang memungkinkan pengguna membuat visualisasi data interaktif yang mudah dipahami dan dianalisis dengan cepat [14]. Dengan Tableau, data dalam bentuk tabel yang mungkin terkesan kaku, dapat diubah menjadi berbagai format visual seperti diagram, grafik, peta geografis, dan lain-lain yang membantu memvisualisasikan perubahan dan perbedaan data secara efektif [15].

Pada penelitian sejenis dengan judul “*Exploratory Data Analysis (EDA): A Study of Olympic Medallist*”, pendekatan *Exploratory Data Analysis* (EDA) digunakan oleh penulis pada *dataset* peraih medali olimpiade, yang mencakup visualisasi dan penjelasan tentang perubahan yang ditemukan selama bertahun-tahun, yang dapat digunakan untuk membantu penyelenggaraan Olimpiade di masa depan. Sebagai salah satu acara olahraga paling penting di dunia, setiap negara dan atlet akan berusaha memberikan penampilan terbaiknya. Diharapkan dengan analisis tersebut, negara dan atlet dapat meningkatkan persiapan dalam partisipasi pada setiap Olimpiade [3].

Penelitian lainnya yang berjudul “*Heart Failure Risk Prediction and Medicine Recommendation using Exploratory Data Analysis*”, penulis menggunakan *Exploratory Data Analysis* (EDA) pada *dataset* medis dari Austria. Tujuan utamanya adalah untuk membantu dokter mengambil keputusan mengenai kondisi jantung pasien. EDA membantu menganalisis dan membentuk gambaran umum tentang *dataset* serta kemungkinan yang terjadi berdasarkan kondisi tertentu. Mereka melakukan EDA dan pemrosesan data, setelah itu mereka mengidentifikasi atribut dan baris yang paling penting untuk memilih kandidat terbaik dalam membentuk sistem prediktif [16]. Selain itu, EDA membantu menyelesaikan masalah keamanan siber dengan memahami korelasi antara atribut dan membedakan jenis serta karakteristik serangan. Hasilnya dapat digunakan untuk memperkirakan prioritas serangan dan mengusulkan kebijakan keamanan atau tindakan penanggulangan serangan siber di masa depan [17]. Sedangkan pada penelitian [14], Tableau digunakan sebagai media

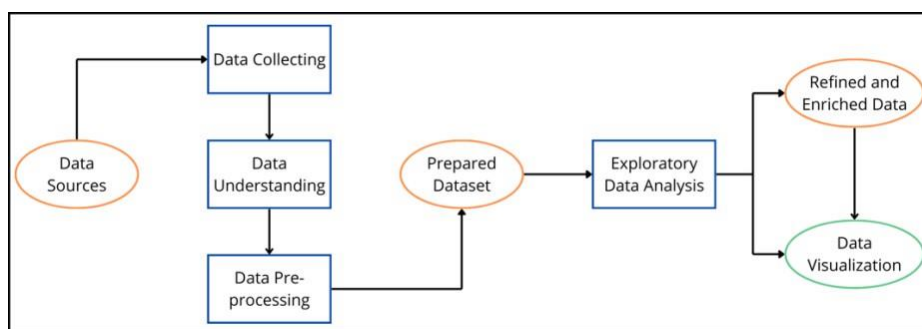
untuk visualisasi data penderita stroke di daerah Kalimantan Selatan dengan pendekatan EDA.

Adapun perbedaan penelitian ini dengan penelitian sejenis lainnya, terletak pada *dataset* yang digunakan. Fokus utama penelitian ini adalah untuk menganalisis dan mengungkap berbagai fakta serta wawasan menarik dari *dataset* yang berisi data Harga Pangan Nasional di Indonesia, yang bersumber dari *World Food Programme Price Database*.

Harga pangan merupakan indikator penting yang mempengaruhi stabilitas ekonomi dan kesejahteraan masyarakat di Indonesia. Perubahan harga pangan dapat berdampak signifikan terhadap daya beli masyarakat, tingkat inflasi, serta kebijakan pemerintah dalam mengatur distribusi dan ketersediaan pangan [18]. Oleh karena itu, pemantauan dan analisis harga pangan nasional menjadi sangat krusial. Dengan pendekatan *Exploratory Data Analysis* (EDA) dan penggunaan Tableau untuk visualisasi data harga pangan, informasi kompleks dapat disajikan dalam bentuk yang lebih mudah dipahami dan dianalisis. Pola dan tren harga pangan dapat terlihat lebih jelas, dengan menggunakan grafik, peta, dan diagram, sehingga memudahkan para pengambil kebijakan, peneliti, dan masyarakat umum dalam membuat keputusan yang tepat, yang pada akhirnya berkontribusi pada stabilitas ekonomi dan ketahanan pangan nasional.

METODE

Pendekatan yang jelas dan terstruktur sangat penting dalam penelitian agar prosesnya lebih fokus, lebih mudah dipahami, dan membawa hasil yang lebih baik dari segi tujuan. Selain itu, dengan menggunakan metode yang sistematis dan bertahap, kemajuan akan jelas dan ringkas, serta mengurangi kemungkinan terjadinya kesalahpahaman. Penelitian ini bertujuan untuk menganalisis data harga pangan nasional di Indonesia, menemukan fakta menarik, dan wawasan sebaran harga selama ini, serta memahami hubungan antar variabel tersebut. Analisis akan didasarkan pada diagram alir yang dapat dilihat pada Gambar 1 untuk mencapai tujuan penelitian ini; Setiap langkah yang tercantum akan dijelaskan secara rinci dalam penelitian ini.



Gambar 1. Langkah-langkah Analisis

Data Collection

Data yang digunakan pada penelitian ini merupakan data Harga Pangan Nasional di Indonesia, yang bersumber dari *World Food Programme Price Database* (Database Harga Program Pangan Dunia) untuk tahun 2007 hingga 2024.

Data Understanding

Data diperoleh dari *data.humdata.org* berupa file dengan nama *food_prices_idn.csv*. *Dataset* tersebut terdiri dari 206.116 baris dan 14 kolom. Pada Tabel 1 berikut ditampilkan daftar kolom yang menjadi variabel awal sebelum dilakukan analisis.

Tabel 1. Variabel Awal *Dataset*

Nama	Tipe	Deskripsi
date	numerik	Keterangan tanggal, bulan dan tahun setiap transaksi
admin1	kategorik	Keterangan provinsi setiap transaksi
admin2	kategorik	Keterangan kota setiap transaksi
market	kategorik	Keterangan nama pasar setiap transaksi
latitude	numerik	Titik berdasarkan garis lintang yang menunjukkan lokasi pasar
longitude	numerik	Titik berdasarkan garis bujur yang menunjukkan lokasi pasar
category	kategorik	Pengelompokan jenis pangan yang dibeli pada setiap transaksi
commodity	kategorik	Jenis pangan yang dibeli pada setiap transaksi
unit	kategorik	KG (Kilogram)
priceflag	kategorik	Aggregate (Nilai Tunggal)
pricetype	kategorik	Retail (Sumber Komoditas)
currency	kategorik	IDR (Rupiah)
price	numerik	Harga dalam Rupiah untuk setiap komoditas yang dibeli pada setiap transaksi
usdprice	numerik	Harga dalam Dollar untuk setiap komoditas yang dibeli pada setiap transaksi

Berdasarkan Tabel 1, nama variabel yang masuk dalam tipe numerik adalah data kuantitatif yang dapat diukur dan dihitung, seperti waktu atau koordinat geografis. Sedangkan variabel kategorik adalah data yang mengelompokkan transaksi berdasarkan kategori tertentu, seperti lokasi, jenis komoditas, atau satuan. Variabel kategorik membantu dalam pengelompokan dan klasifikasi, sedangkan variabel numerik memungkinkan analisis kuantitatif yang lebih dalam, seperti menghitung rata-rata atau mengukur jarak [19].

Data Pre-processing

Pra-pemrosesan data (*data pre-processing*) adalah proses mengubah data mentah ke dalam bentuk yang lebih mudah dipahami. Proses ini diperlukan untuk memperbaiki kesalahan pada data mentah yang seringkali tidak lengkap dan memiliki format yang tidak teratur. *Pre-processing* melibatkan proses validasi dan imputasi data. Validasi bertujuan untuk menilai tingkat kelengkapan dan akurasi data yang tersaring. Sedangkan imputasi bertujuan memperbaiki kesalahan dan memasukkan nilai yang hilang.

Dalam penelitian ini, dilakukan pembersihan dan transformasi data pada kumpulan data. Pembersihan data bertujuan untuk menghapus nilai yang hilang agar visualisasi data lebih mudah dipahami, sementara transformasi data dilakukan untuk mengubah kumpulan data menjadi bentuk yang dapat memberikan informasi spesifik dan berguna.

Perubahan nama variabel dilakukan pada *admin1* dan *admin2* menjadi *province* dan *city*, karena lebih mewakili nilainya. Penghapusan variabel dilakukan pada variabel *unit*, *priceflag*, dan *pricetype* karena hanya memiliki satu *unique value*, serta *usdprice* karena sudah diwakilkan oleh variabel *price*. Adapun penambahan variabel baru (*feature engineering*) yaitu *sources* dan *taste* yang dikembangkan dari variabel *commodity*, serta *year* dan *month* yang diambil dari variabel *date*, dilakukan untuk membantu menciptakan representasi yang lebih bermakna dari data untuk analisis lebih lanjut. Terakhir penghapusan baris dengan *missing value*, karena terdapat 1227 *missing value* pada variabel *admin1*, *admin2*, *latitude*, dan *longitude*, dan penghapusan baris pada variabel *date* untuk data pada tahun 2016, 2017 dan 2024, karena ketidaklengkapan data akhirnya hanya dipilih tahun 2020 – 2023 saja. Maka, berdasarkan tahap pra-pemrosesan yang dilakukan diperoleh 178.707 baris dan 13 kolom. Berikut pada Tabel 2 ditampilkan daftar variabel lanjutan setelah tahap pra-pemrosesan data untuk analisis pada penelitian ini.

Tabel 2. Variabel Lanjutan *Dataset*

Nama	Tipe	Deskripsi
date	numerik	Keterangan tanggal, bulan dan tahun setiap transaksi
province	kategorik	Keterangan provinsi setiap transaksi
city	kategorik	Keterangan kota setiap transaksi
market	kategorik	Keterangan nama pasar setiap transaksi
latitude	numerik	Titik berdasarkan garis lintang yang menunjukkan lokasi pasar
longitude	numerik	Titik berdasarkan garis bujur yang menunjukkan lokasi pasar
category	kategorik	Pengelompokan jenis pangan berdasarkan kandungannya
commodity	kategorik	Jenis pangan yang dibeli pada setiap transaksi
price	numerik	Harga dalam Rupiah untuk setiap komoditas yang dibeli pada setiap transaksi
sources	kategorik	Pengelompokan jenis pangan berdasarkan bahannya
taste	kategorik	Pengelompokan jenis pangan berdasarkan rasanya
year	numerik	Keterangan tahun setiap transaksi
month	numerik	Keterangan bulan setiap transaksi

Exploratory Data Analysis (EDA)

Pendekatan Eksploratif atau yang biasa disebut *Exploratory Data Analysis (EDA)* adalah pendekatan untuk menganalisis kumpulan data guna meringkas karakteristik utama pada data yang diteliti, umumnya menggunakan grafik statistik dan metode visualisasi data [20]. Tujuan EDA adalah untuk mengusulkan hipotesis tentang penyebab fenomena yang diamati, mengevaluasi asumsi yang menjadi dasar inferensi statistik, mendukung pemilihan alat dan teknik statistik yang tepat, memberikan dasar untuk pengumpulan data lebih lanjut melalui survei atau eksperimen, dan dalam penelitian ini akan fokus pada visualisasi data untuk membantu melihat data sebelum membuat asumsi apa pun serta menemukan wawasan dan hubungan menarik dari data. Analisis data eksploratif dibagi menjadi 3 kategori:

- a. Analisis univariat adalah bentuk analisis data yang paling sederhana. Kata "uni" yang berarti "satu" menunjukkan bahwa analisis ini hanya melibatkan satu variabel. Analisis ini tidak memeriksa penyebab atau hubungan, melainkan fokus untuk mendeskripsikan data, meringkasnya, dan menemukan pola.
- b. Analisis bivariat melibatkan dua variabel dengan tujuan menentukan hubungan empiris di antara keduanya.
- c. Analisis multivariat mencakup berbagai teknik untuk menganalisis kumpulan data yang mengandung lebih dari satu variabel. Teknik-teknik ini sangat berguna saat bekerja dengan variabel yang saling berkorelasi.

Pada penelitian ini Tableau akan digunakan sebagai alat visualisasi data untuk menemukan informasi spesifik dan memberikan visualisasi yang lebih baik, sederhana, dan informatif, dengan bentuk Diagram, Plot, Peta, dan bentuk lainnya.

Data Visualization

Data yang telah dihaluskan dan diperkaya selanjutnya diimplementasikan ke dalam platform Tableau yang merupakan *tool* untuk visualisasi data. Proses yang ada pada tahapan ini diantaranya ialah proses mengimpor data ke dalam tableau, kemudian setelah dilakukan impor data selanjutnya data dieksekusi untuk dimasukkan ke worksheet baru. Visualisasi data menggunakan tableau memudahkan dalam merepresentasikan simbol dan membuat hubungan antara bit informasi dengan lebih jelas.



Gambar 2. Manfaat Visualisasi Data

Dari Gambar 2 terlihat bahwa manfaat visualisasi data antara lain adalah menganalisis sejumlah besar data yang mungkin tidak berarti atau sulit dipahami. Visualisasi data juga dapat membantu organisasi dan bisnis yang memiliki data dalam jumlah besar untuk mengenali tren dan akhirnya membantu dalam pengambilan keputusan. Selain itu, visualisasi dapat mengidentifikasi kesamaan dan hubungan dalam data serta menyajikan informasi kepada orang lain dalam bahasa yang lebih sederhana [21].

HASIL DAN PEMBAHASAN

Pada penelitian ini, Python digunakan untuk *Exploratory Data Analysis* (EDA) dengan menggunakan *Visual Studio Code* sebagai platformnya. Pemilihan Python karena merupakan bahasa pemrograman serbaguna yang ringkas dan mudah dibaca, terutama dalam bidang analisis data, sesuai dengan tujuan penelitian untuk mengolah *dataset* menjadi data yang sesuai dengan kebutuhan visualisasi, yang dalam penelitian ini menggunakan Tableau, kemampuan untuk menampilkan berbagai jenis visualisasi menjadi alasan kenapa Tableau dipilih. Analisis dilakukan dalam bentuk *Univariate*, *Bivariate*, dan *Multivariate*, untuk menghasilkan wawasan menarik yang disajikan dalam berbagai bentuk visualisasi. Gambar 3 dan 4 merupakan tampilan deskripsi *dataset* awal dan lanjutan untuk proses EDA:

```
RangeIndex: 206116 entries, 0 to 206115
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  ---
0    date        206116 non-null object
1    admin1      204889 non-null object
2    admin2      204889 non-null object
3    market      206116 non-null object
4    latitude    204889 non-null object
5    longitude   204889 non-null object
6    category    206116 non-null object
7    commodity   206116 non-null object
8    unit        206116 non-null object
9    priceflag   206116 non-null object
10   pricetype   206116 non-null object
11   currency    206116 non-null object
12   price       206116 non-null object
13   usdprice    206116 non-null object
dtypes: object(14)
```

Gambar 3. Deskripsi *dataset* awal

```
RangeIndex: 178707 entries, 0 to 178706
Data columns (total 13 columns):
#   Column      Non-Null Count  Dtype
---  ---
0    date        178707 non-null datetime64[ns]
1    province    178707 non-null object
2    city        178707 non-null object
3    market      178707 non-null object
4    latitude    178707 non-null float64
5    longitude   178707 non-null float64
6    category    178707 non-null object
7    commodity   178707 non-null object
8    price       178707 non-null float64
9    sources     178707 non-null object
10   taste       178707 non-null object
11   year        178707 non-null int64
12   month       178707 non-null int64
dtypes: datetime64[ns](1), float64(3), int64(2), object(7)
```

Gambar 4. Deskripsi *dataset* lanjutan

Dalam proses merubah *dataset* awal menjadi *dataset* lanjutan dilakukan:

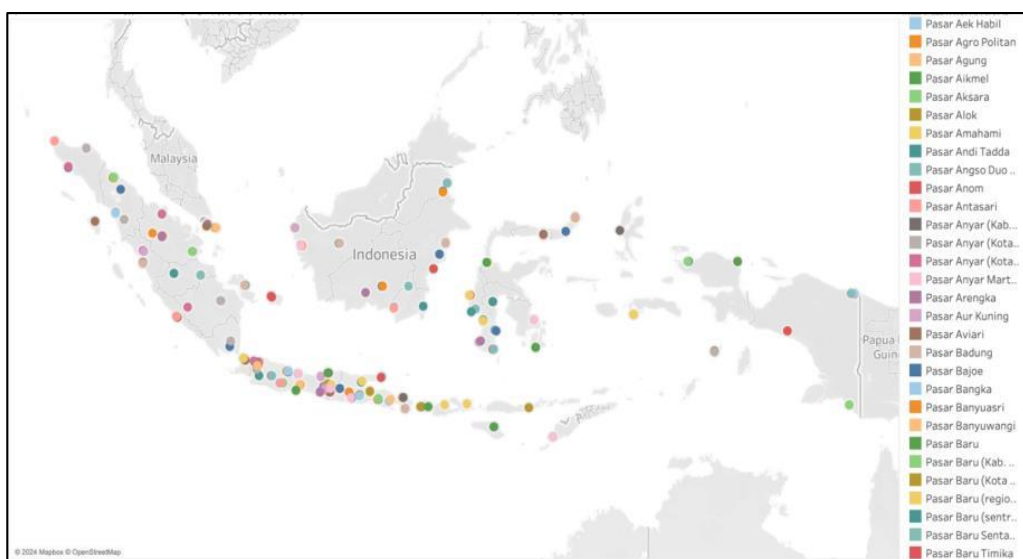
- Menghapus baris yang memiliki *null values* dan ketidaklengkapan data (*admin1*, *admin2*, *latitude*, *longitude*, dan *date*).
- Menghapus kolom yang hanya memiliki satu *unique value* (*unit*, *priceflag*, dan *pricetype*).
- Menghapus kolom yang memiliki *similar values* (*usdprice*).
- Merubah nama kolom yang kurang representatif (*admin1* dan *admin2*, menjadi *province* dan *city*).
- Membuat kolom baru dari kolom yang sudah ada atau *Feature Engineering* (*sources* dan *taste* dari *commodity*, serta *year* dan *month* dari *date*).
- Merubah tipe data beberapa kolom (*date*, *latitude*, *longitude*, dan *price*).

Proses Visualisasi Data

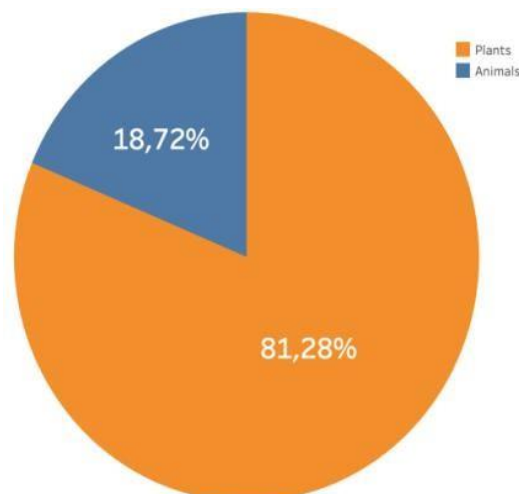
Visualisasi data dalam penelitian ini dilakukan menggunakan aplikasi Tableau. Pemilihan Tableau karena kemudahan penggunaannya dan kemampuannya untuk menampilkan berbagai jenis visualisasi. Berikut ini hasil visualisasi yang dibagi menjadi 3 kategori sesuai analisis data eksploratif:

1. Univariate Analysis

Analisis univariat (*Univariate Analysis*) adalah analisis yang dilakukan pada satu variabel dengan tujuan mengetahui dan mengidentifikasi karakteristik variabel tersebut. Variabel *Market* (Pasar) dan *Sources* (Bahan) masing-masing dipilih untuk mengidentifikasi komposisi dan sebaran data titik lokasi pasar serta sumber komoditas pangan di wilayah Indonesia dari tahun 2020 - 2023. Hasilnya ditampilkan pada Gambar 5 dan Gambar 6.



Gambar 5. Sebaran Data Titik Lokasi Pasar

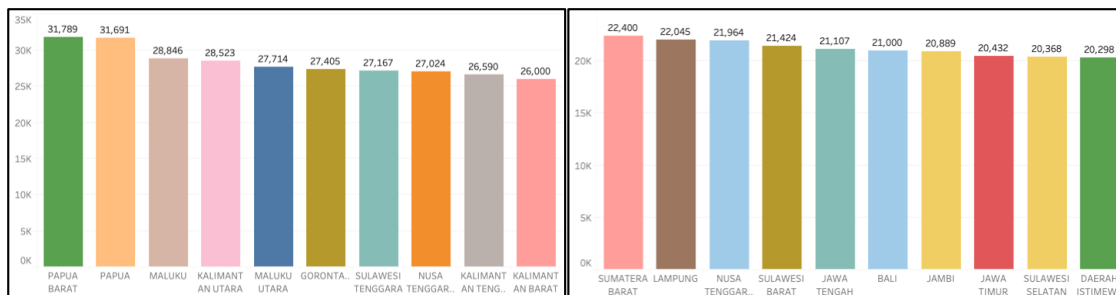


Gambar 6. Komposisi Data Bahan Pangan

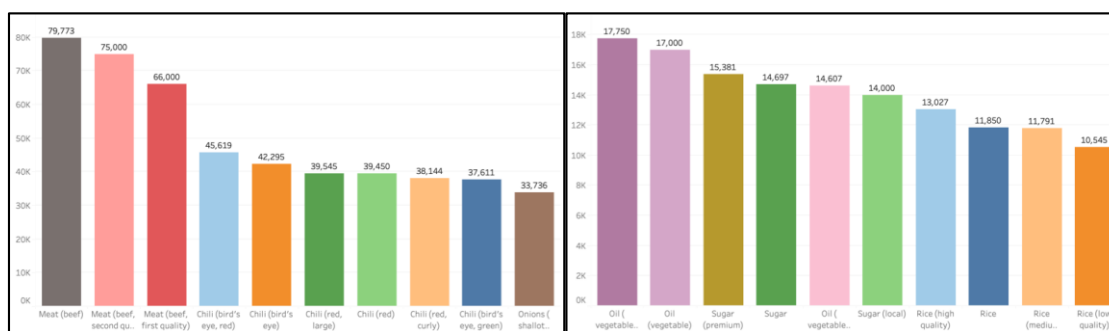
Gambar 5 menunjukkan bahwa data titik lokasi pasar cukup merata hampir mewakili seluruh wilayah di Indonesia, sedangkan dari Gambar 6 diketahui bahwa bahan pangan nabati memiliki sebaran data lebih banyak dibandingkan dengan bahan pangan hewani, dengan masing-masing jumlahnya adalah 131.969 dan 30.389.

2. *Bivariate Analysis*

Analisis bivariat (*Bivariate Analysis*) adalah analisis yang melibatkan dua variabel yang berbeda, dan analisis jenis data ini berfokus pada pemahaman hubungan atau asosiasi antara kedua variabel tersebut. Pada penelitian ini, hubungan antara variabel *Province* (Provinsi) dan *Commodity* (Komoditas) masing-masing terhadap *Price* (Harga) dari tahun 2020 - 2023 diidentifikasi. Secara lebih jelas dapat dilihat pada Gambar 7 dan Gambar 8.



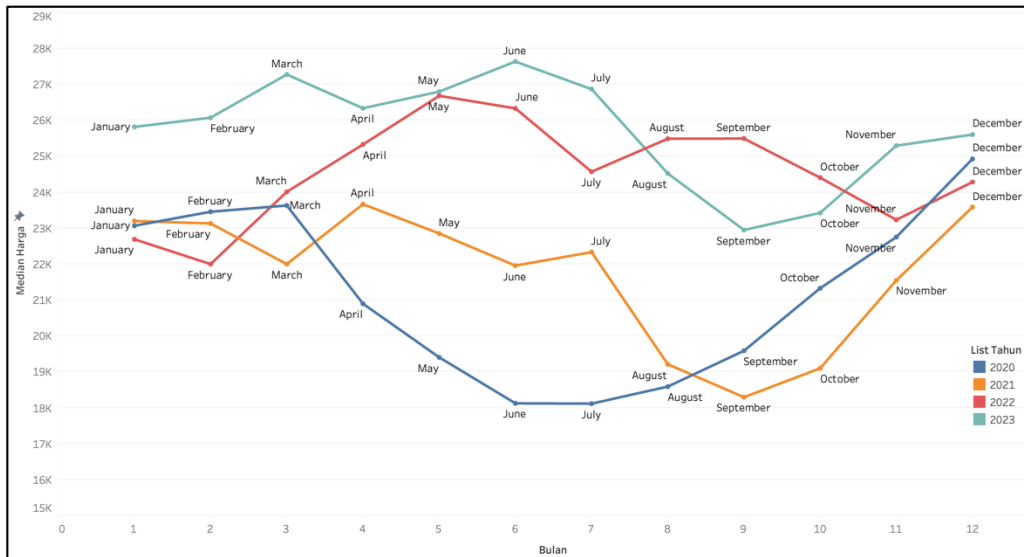
Gambar 7. Harga Median Pangan Nasional berdasarkan Provinsi



Gambar 8. Harga Median Pangan Nasional berdasarkan Komoditas

Berdasarkan Gambar 7, provinsi dengan harga median pangan nasional tertinggi adalah Papua Barat sebesar Rp31.789 sedangkan yang terendah adalah DI Yogyakarta dengan harga median sebesar Rp20.298. Jika dilihat, rata-rata provinsi dengan harga median tertinggi berada di wilayah timur Indonesia. Sedangkan provinsi dengan harga median terendah berada di wilayah barat Indonesia. Hal ini memperjelas bahwa sebaran harga pangan nasional antar wilayah di Indonesia kurang merata.

Kemudian berdasarkan komoditas pangan pada Gambar 8, *meat* (daging) menjadi komoditas dengan harga median tertinggi sebesar Rp79.773 dan *rice* (beras) menjadi yang terendah dengan harga median sebesar Rp10.545. Artinya bahan pangan hewani menjadi bahan pangan yang mahal di Indonesia, sedangkan beras yang merupakan bahan pangan nabati menjadi yang paling murah, hal ini cukup baik karena beras merupakan bahan makanan pokok masyarakat Indonesia. Selanjutnya pada Gambar 9 berikut ini akan ditampilkan tren harga pangan nasional per bulan antara tahun 2020 - 2023.

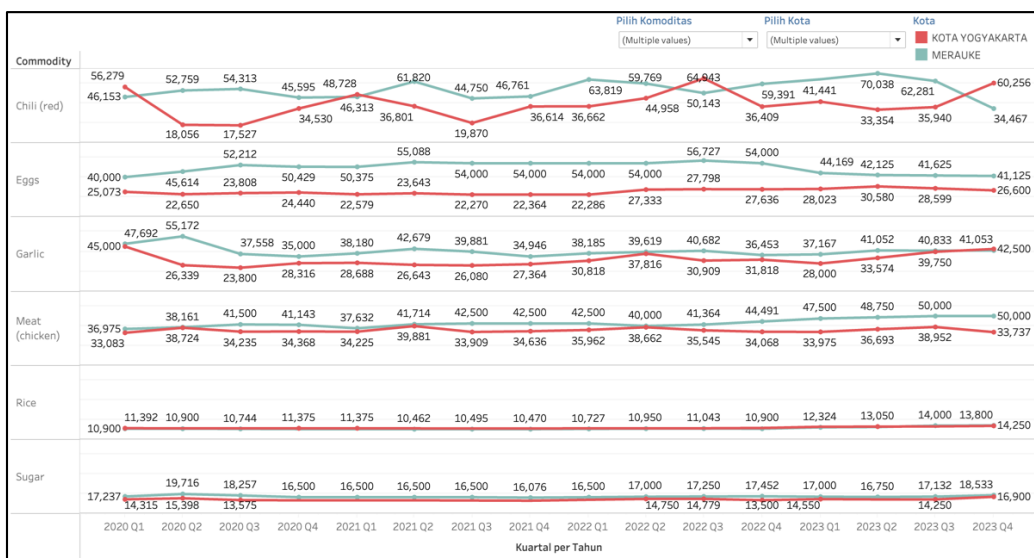


Gambar 9. Harga Median Pangan Nasional berdasarkan Tahun

Harga pangan cenderung fluktuatif setiap tahunnya seperti yang ditunjukkan pada Gambar 9, namun pada akhir tahun rata - rata selalu sama mengalami kenaikan harga. Pada akhir tahun kenaikan harga biasanya dipicu karena mendekati perayaan natal dan tahun baru yang umumnya mendorong tingkat kebutuhan masyarakat akan bahan pokok semakin tinggi.

3. Multivariate Analysis

Analisis multivariat (*Multivariate Analysis*) merujuk pada kumpulan data di mana setiap observasi atau titik sampel terdiri dari beberapa variabel atau fitur. Variabel-variabel ini dapat mewakili berbagai aspek, karakteristik, atau pengukuran yang berkaitan dengan fenomena yang diamati. Ketika data melibatkan tiga atau lebih variabel, maka data tersebut secara khusus dikategorikan sebagai multivariat. Untuk tujuan penelitian ini, hubungan antara variabel *City* (Kota), *Commodity* (Komoditas) dan *Price* (Harga) terhadap *Quarter* (Kuartal) dari tahun 2020 – 2023 dianalisis. Gambar 10 berikut ini menjelaskan hubungan antara setiap variabel yang dianalisis.



Gambar 10. Harga Median Pangan Nasional berdasarkan Komoditas dan Kota per Kuartal

Dari Gambar 10 dapat dilihat perbandingan tren harga pangan berdasarkan komoditasnya antara dua kota yaitu Kota Yogyakarta dan Merauke. Komoditas pangan yang dibandingkan dari kedua kota tersebut yaitu *red chili* (cabai merah), *eggs* (telur), *garlic* (bawang putih), *chicken meat* (daging ayam), *rice* (beras), dan *sugar* (gula). Pada komoditas cabai merah harga pangan di kedua kota sangat fluktuatif per kuartalnya, sedangkan pada komoditas lainnya harga pangan cenderung stabil. Hal lainnya yang dapat dilihat dari Gambar 10, kebanyakan harga komoditas pangan di Merauke lebih tinggi jika dibandingkan dengan harga komoditas pangan di Kota Yogyakarta, meskipun untuk harga beras terlihat sama antara kedua kota tersebut.

Setiap visualisasi data dengan grafik, peta, dan diagram yang dibuat pada Tableau dapat difilter sesuai dengan kebutuhan informasi yang diinginkan, dan lebih baik lagi jika menggunakan *dataset* yang sudah dihaluskan dan diperkaya.

KESIMPULAN

Dari penelitian ini diketahui bahwa analisis univariat mengidentifikasi distribusi data lokasi pasar dan sumber komoditas pangan di Indonesia, dengan sebaran yang cukup merata untuk pasar dan dominasi bahan pangan nabati dibandingkan dengan hewani. Analisis Bivariat mengevaluasi hubungan antara provinsi dan komoditas terhadap harga pangan, menunjukkan bahwa harga pangan di wilayah timur Indonesia lebih tinggi dibandingkan dengan wilayah barat, dengan daging sebagai komoditas termahal dan beras sebagai yang termurah. Terakhir, analisis Multivariat mempelajari hubungan antara kota, komoditas, dan harga terhadap kuartal, dengan perbandingan antara Yogyakarta dan Merauke yang menunjukkan harga komoditas pangan lebih tinggi di Merauke, kecuali untuk beras yang harganya stabil di kedua kota.

Hasil analisis dalam penelitian ini bisa dimanfaatkan untuk keperluan masa depan, seperti membangun model *machine learning* untuk memprediksi perubahan harga pangan nasional dengan lebih tepat. Hal ini akan membantu para pengambil kebijakan, peneliti, dan masyarakat umum dalam membuat keputusan yang lebih baik, yang pada gilirannya akan mendukung stabilitas ekonomi dan ketahanan pangan nasional.

REFERENSI

- [1] J. DSouza, "Using exploratory data analysis for generating inferences on the correlation of COVID-19 cases," *2020 11th Int. Conf. on Comp., Comm. and Net. Tech. (ICCCNT)*, no. July, pp. 1-6, 2020, doi: 10.1109/ICCCNT49239.2020.9225621.
- [2] S. Arif, W. Isdijoso, A.R. Fatah, and A.R. Tamyis, "Tinjauan Strategis Ketahanan Pangan dan Gizi di Indonesia," *Jakarta: SMERU Research Institute*, 2020. [E-book] Available: https://smeru.or.id/sites/default/files/publication/rr_fsn_indonesia_finalreport_ind.pdf
- [3] A.R. Salasa, "Paradigma dan Dimensi Strategi Ketahanan Pangan Indonesia," *Jejaring Administrasi Publik*, vol. 13, no. 1, pp. 35-48, 2021, doi: 10.20473/jap.v13i1.29357
- [4] G. Prayitno, A.W. Hasyim, A. Subagiyo, D. Dinanti, and F. Roziqin, "Ruang Berketahanan Pangan: Menjawab Tantangan Produksi Pangan Berkelanjutan dengan Optimasi Keruangan Menuju Indonesia Berdaulat," *Universitas Brawijaya Press*, 2022. [E-book] Available: <https://books.google.co.id/>.
- [5] F.A.A. Kurniawan, I. Ramdhani, and R.A. Fitrie, "Analisis Pengambilan Keputusan Kebijakan Impor Beras dalam Konteks Pemenuhan Pangan Nasional di Indonesia," *Indonesian Journal of Public Administration Review*, vol. 1, no. 2, pp.15-15, 2024, doi: 10.47134/par.v1i2.2394
- [6] A. S. Rao, B. V. Vardhan, and H. Shaik, "Role of Exploratory Data Analysis in Data Science," *Proc. 6th Int. Conf. Commun. Electron. Syst. ICCES 2021*, no. July, pp. 1457-1461, 2021, doi: 10.1109/ICCES51350.2021.9488986.
- [7] South Dakota State University, "Using exploratory data analysis (bivariate)," no. September, 2020, [Online]. Available: <http://bioinformatics.sdstate.edu/users/gex/index/indexfiles/ch2.pdf>.
- [8] K. Sahoo, A. K. Samal, J. Pramanik, and S. K. Pani, "Exploratory Data Analysis using Python," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 4727-4735, 2019, doi: 10.35940/ijitee.L3591.1081219.

- [9] M. Komorowski, D. C. Marshall, J. D. Saliccioli, and Y. Crutain, "Secondary Analysis of Electronic Health Records," *Second. Anal. Electron. Heal. Rec.*, September, pp. 1–427, 2016, doi: 10.1007/978-3-319-43742-2.
- [10] F. Paper, M. Rahmany, A. M. Zin, and A. Elankovan, "Comparing Tools Provided By Python," no. i, pp. 131–142, 2020.
- [11] H. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. Patel, R. Ramakrishnan and C. Shahabi, "Big data and its technical challenges," *Communications of the ACM*, vol. 57, no. 7, pp. 86-94, 2014, doi: 10.1145/2611567.
- [12] M. Mani and S. Fei, "Effective Big Data Visualization," *Proceedings of the 21st International Database Engineering & Applications Symposium on - IDEAS 2017*, 2017, doi: 10.1145/3105831.3105857.
- [13] W. Yafooz, S. Abidin, N. Omar and S. Hilles, "Interactive Big Data Visualization Model Based on Hot Issues (Online News Articles)," *Communications in Computer and Information Science*, pp. 89-99, 2016, doi: 10.1007/978-981-10-2777-2_8.
- [14] I.N. Rizki, D., Prayoga, M.L. Puspita, and M.Q. Huda, "Implementasi Exploratory Data Analysis untuk Analisis dan Visualisasi Data Penderita Stroke Kalimantan Selatan Menggunakan Platform Tableau," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, 2024, doi: 10.23960/jitet.v12i1.3856.
- [15] D. Saepuloh, "Visualisasi Data Covid 19 Provinsi DKI Menggunakan Tableau," *J. Ris. Jakarta*, vol. 13, no. 2, pp. 55–64, 2020, doi: 10.37439/jurnaldrd.v13i2.37.
- [16] N. Kaieski, L.P.L. de Oliveira, and M.B. Villamil, "Vis-health: Exploratory analysis and visualization of dengue cases in brazil," *2016 49th Hawaii International Conference on System Sciences (HICSS)*, no. January, pp. 3063-3072, 2016, doi: 10.1109/HICSS.2016.385.
- [17] J. D. Miranda-Calle, V. Reddy C, P. Dhawan, and P. Churi, "Exploratory Data Analysis for Cybersecurity," *World J. Eng.*, vol. 18, no. 5, pp. 734–749, 2021, doi: 10.1108/WJE-11-2020- 0560.
- [18] L. Hakim, "Analysis of the Impact of the General Election on Consumption Levels, Investment, Inflation and Stock Price Index to Maintain Economic Stability and Anticipate the Impact," *Prosiding Seminar Nasional Ilmu Manajemen, Ekonomi, Keuangan dan Bisnis*, vol. 3, no. 1, pp. 107-118, 2024.
- [19] C. Pan, A. Poddar, R. Mukherjee, and A.K. Ray, "Impact of categorical and numerical features in ensemble machine learning frameworks for heart disease prediction," *Biomedical Signal Processing and Control*, vol. 76, p.103666, 2022, doi: 10.1016/j.bspc.2022.103666
- [20] S.K. Mukhiya, and U. Ahmed, "Hands-On Exploratory Data Analysis with Python: Perform EDA techniques to understand, summarize, and investigate your data," *Packt Publishing Ltd*, 2020. [E-book] Available: <https://books.google.co.id/>.
- [21] S. K. A. Fahad and A. E. Yahya, "Big Data Visualization: Allotting by R and Python with GUI Tools," *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, Shah Alam, pp. 1-8, 2018, doi: 10.1109/ICSCEE.2018.8538413.