



Struktur Presentasi

01

Pendahuluan

02

Tinjauan Pustaka

03

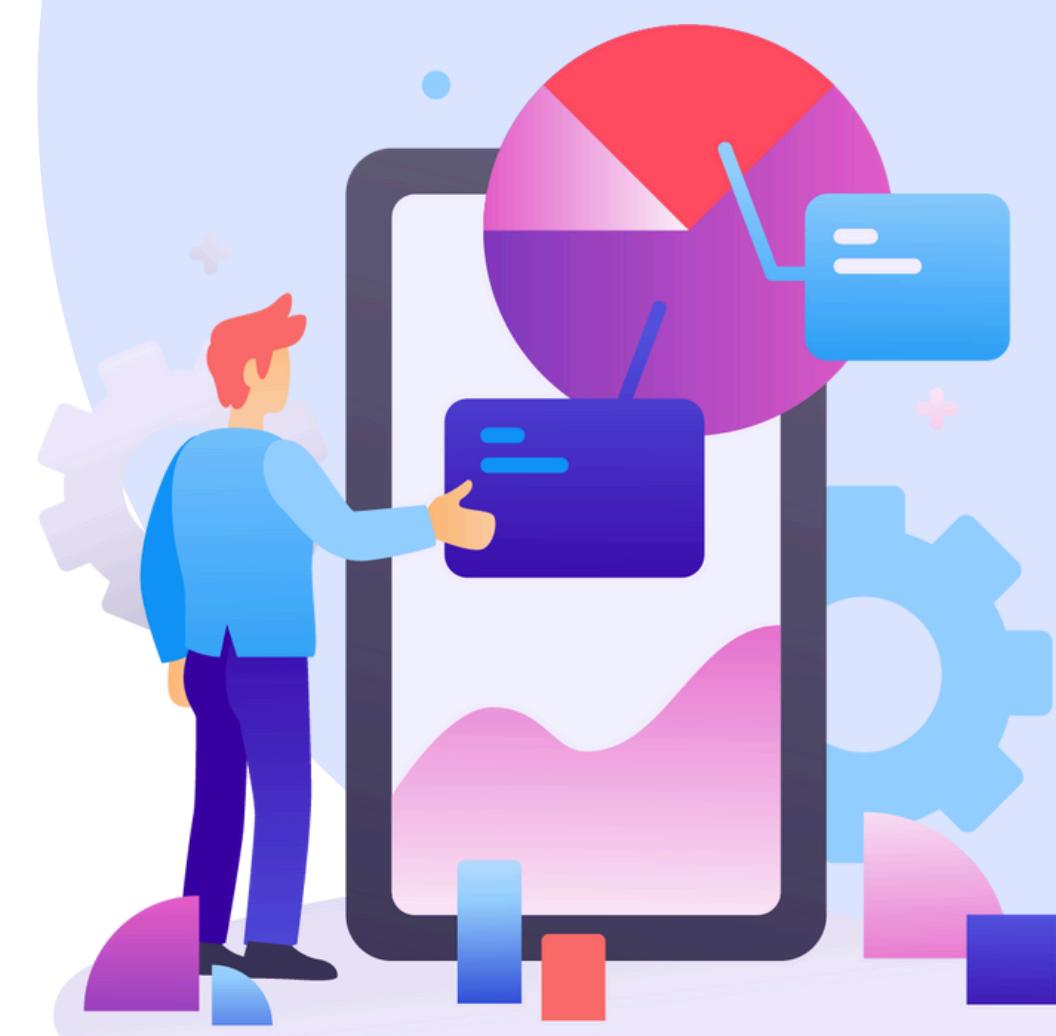
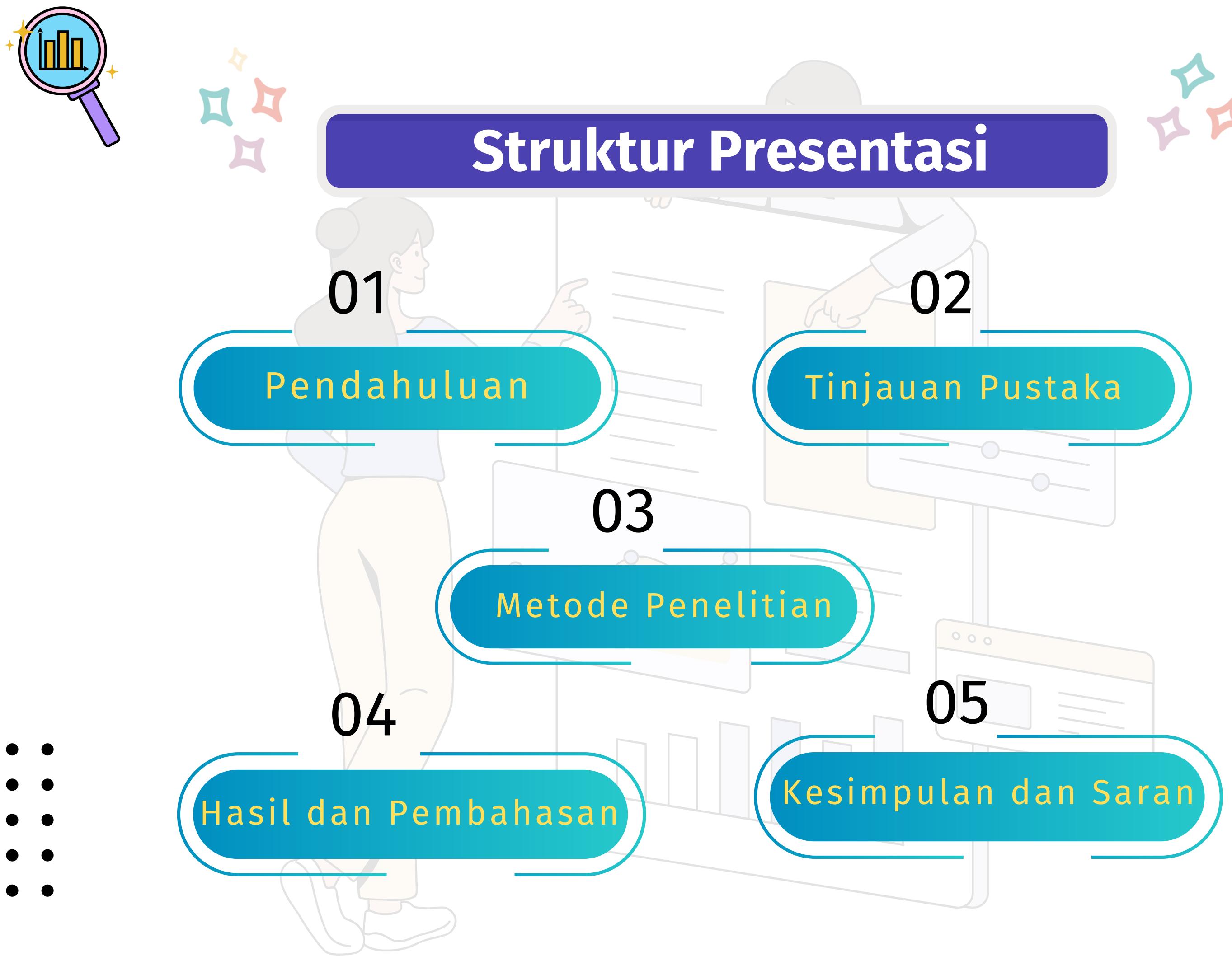
Metode Penelitian

04

Hasil dan Pembahasan

05

Kesimpulan dan Saran



PENDAHULUAN

Data kependudukan adalah komponen penting dalam perencanaan pembangunan karena menggambarkan kondisi sosial, ekonomi, dan demografi setiap wilayah. Provinsi Lampung memiliki karakteristik penduduk yang beragam, sehingga diperlukan analisis yang rapi dan akurat. Namun, data mentah sering tidak konsisten, memiliki format berbeda, dan mengandung nilai kosong. Karena itu, proses Data Munging/Wrangling diperlukan untuk membersihkan, menstandarkan, dan menyiapkan data agar layak dianalisis

Teknik wrangling digunakan untuk membersihkan, menata ulang, dan mentransformasi data agar menjadi lebih terstruktur. Selanjutnya, visualisasi statistik membantu menampilkan pola penting sehingga data lebih mudah dipahami dan diinterpretasikan. Melalui kombinasi wrangling dan visualisasi, penelitian ini memberikan gambaran yang lebih jelas tentang kondisi kependudukan Provinsi Lampung Tahun 2024.”



TINJAUAN PUSTAKA

Data wrangling merupakan proses penyiapan data melalui pembersihan dan penataan data mentah. Tujuannya adalah memastikan kualitas, konsistensi, dan keterbacaan data sebelum dianalisis.

Komponen Utama

Integrasi Data – menggabungkan data dari berbagai sumber

Pembersihan Data – menangani missing value, duplikasi, outlier

Transformasi Data – mengubah format, tipe, atau struktur data

Pengayaan Data – menambah fitur/variabel baru

Validasi Data – memastikan data final akurat & logis

Data Wrangling merupakan tahap awal yang sangat penting sebelum eksplorasi data, visualisasi, maupun pemodelan statistik agar hasil analisis tidak bias dan lebih akurat.

METODE PENELITIAN



Jenis Data

Menggunakan data sekunder dari BPS Provinsi Lampung tahun 2024 yang berisi jumlah penduduk, laju pertumbuhan, persentase penduduk, kepadatan, dan rasio jenis kelamin. Data dipilih karena sudah melalui proses validasi resmi sehingga akurat dan kredibel.

Variabel Penelitian

- Variabel Utama: Jumlah penduduk, Laju pertumbuhan (%), Persentase penduduk, Kepadatan (jiwa/km^2), Rasio jenis kelamin.
- Variabel Turunan: KepadatanGroup, GrowthGroup, Dominasi_JK, DominasiFlag

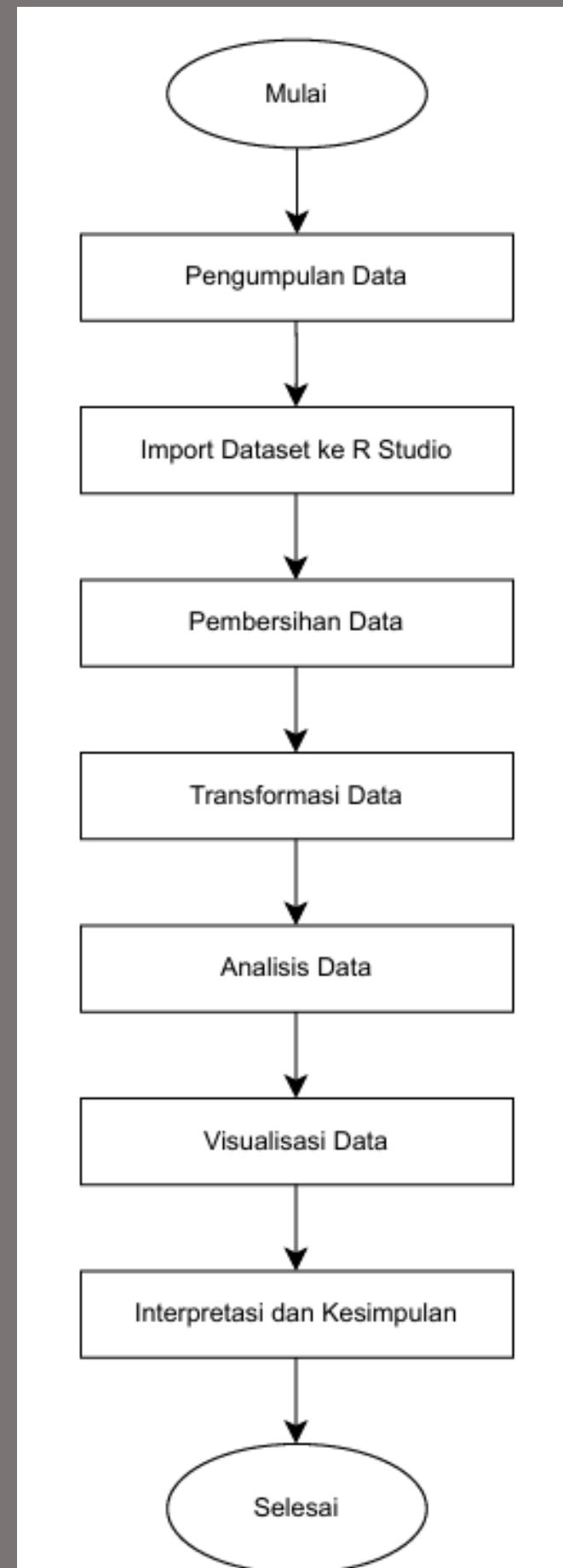
Teknik Pengumpulan Data

- Dataset diunduh dari situs resmi BPS dalam format CSV.
- Data diimpor ke R Studio menggunakan `read.csv()`.
- Dilakukan pembersihan awal: menghapus nilai hilang, duplikasi, dan penyesuaian nama kolom serta tipe data.

Teknik Pengolahan Data (Wrangling)

- Cleaning: hapus missing value & duplikasi, konversi tipe data, standarisasi format.
- Transformation: membuat variabel kategori dengan `mutate()` dan `case_when()`.
- Filtering & Sorting: memilih data relevan dan mengurutkan untuk analisis.

Diagram Alir





HASIL DAN PEMBAHASAN

Struktur Data Mentah Kependudukan Provinsi Lampung Tahun 2024

Kabupaten/Kota	Jumlah Penduduk (Ribu)	Laju Pertumbuhan Penduduk per Tahun	Persentase Penduduk	Kepadatan Penduduk per km persegi (Km2)	Rasio Jenis Kelamin Penduduk
Lampung Barat	315.9	1.12	3.35	150	106.9
Tanggamus	670.4	1.15	7.12	227	106.6
Lampung Selatan	1119.3	1.27	11.88	503	104.1
Lampung Timur	1153.8	0.96	12.25	299	103
Lampung Tengah	1525.1	1.1	16.19	334	103.9
Lampung Utara	659.9	1.04	7.01	247	103.2
Way Kanan	497.3	1.23	5.28	141	104.4
Tulangbawang	450.3	1.16	4.78	145	106.1
Pesawaran	500.3	1.17	5.31	388	105.4
Pringsewu	424.7	1.16	4.51	688	105
Mesuji	241.6	1.51	2.56	110	105.6
Tulang Bawang Barat	298.7	1.08	3.17	238	103.7
Pesisir Barat	172.3	1.45	1.83	59	107.5
Kota Bandar Lampung	1214.3	1.02	12.89	6609	103.5
Kota Metro	175.7	1.03	1.87	2400	100.8
Lampung	9419.6	1.12	100	281	104.3

1. Dataset berisi 16 kabupaten/kota di Provinsi Lampung.

Dengan 5 variabel utama: Jumlah Penduduk, Laju Pertumbuhan, Persentase Penduduk, Kepadatan, dan Rasio Jenis Kelamin.

2. Tidak ada missing value pada data mentah.
3. Format angka masih bervariasi

Beberapa variabel numerik masih menggunakan koma, titik ribuan, dan perlu dikonversi ke format numerik bersih saat wrangling

4. Nama kolom masih panjang sehingga perlu distandardkan.
5. Data masih dalam urutan asli (belum diurutkan atau dikelompokkan).

Sehingga pola awal belum terlihat jelas sebelum dilakukan wrangling.

HASIL DAN PEMBAHASAN

Perbandingan Sebelum dan Setelah Subsetting

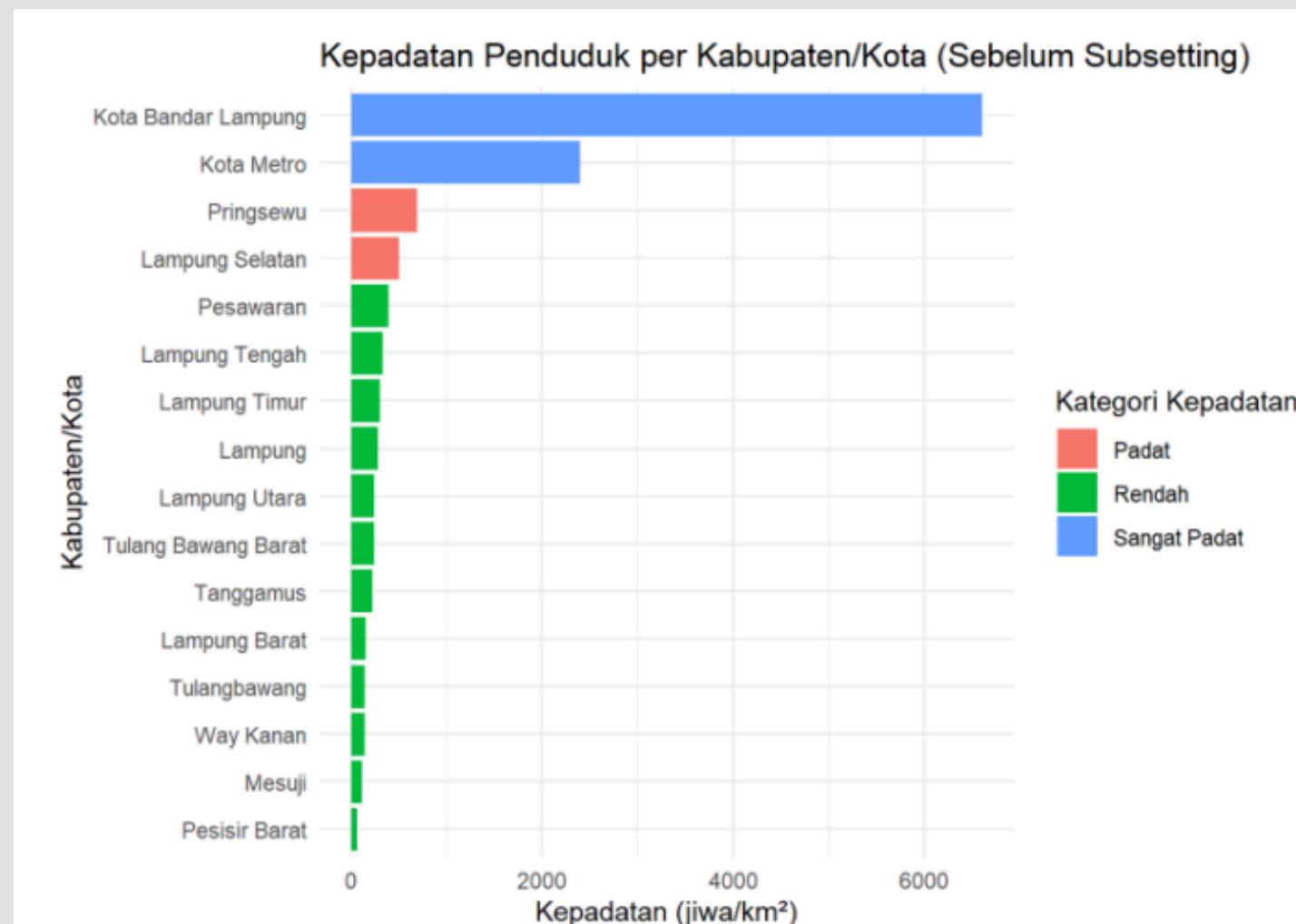


Diagram batang kepadatan penduduk per Kabupaten/Kota sebelum dilakukan subsetting.

Sebelum subsetting, grafik menampilkan seluruh kabupaten/kota sehingga pola kepadatan penduduk kurang terlihat jelas. Wilayah sangat padat bercampur dengan wilayah ber-kepadatan rendah, sehingga nilai ekstrem seperti Kota Bandar Lampung dan Kota Metro belum tampak dominan.

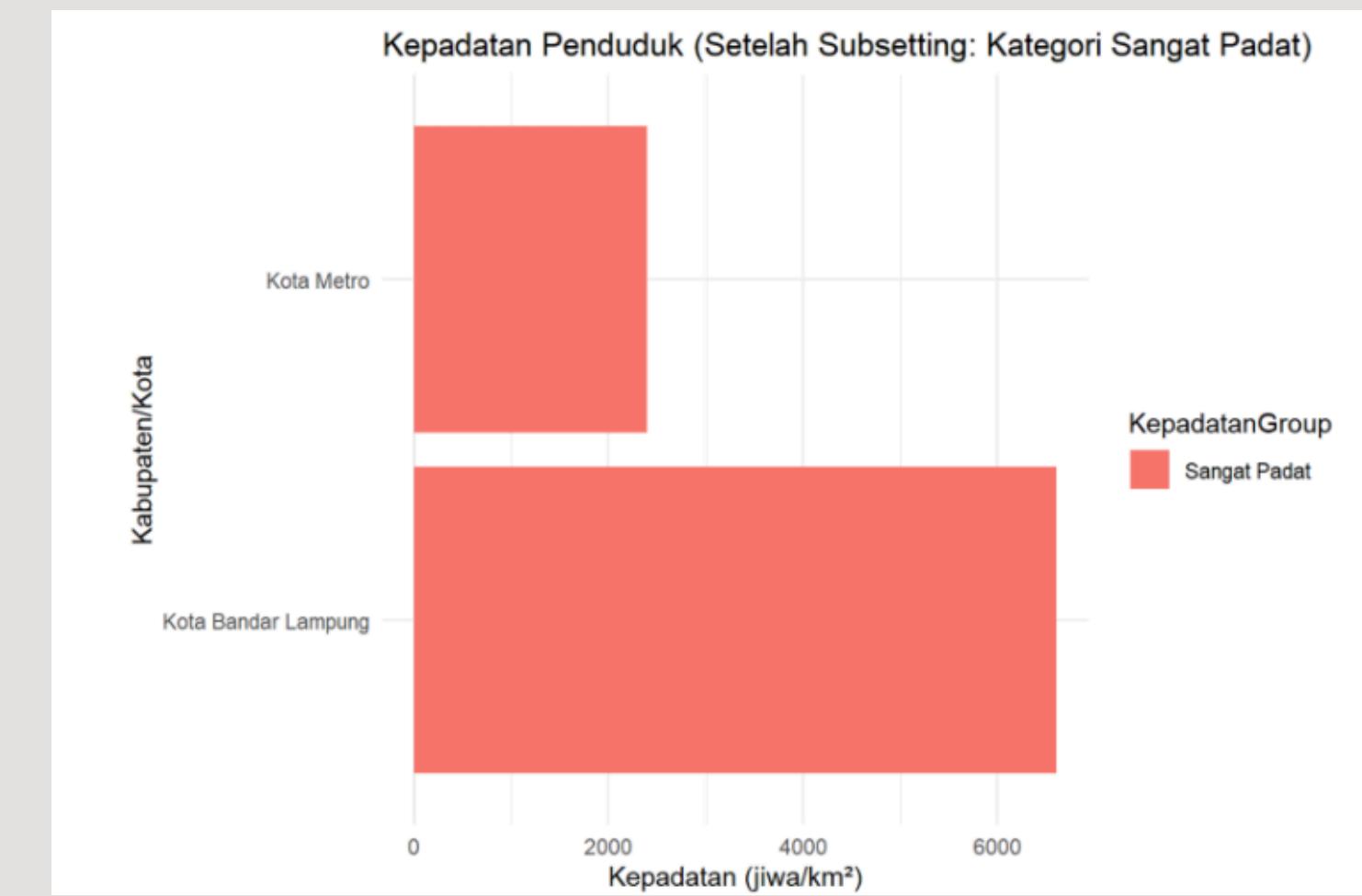


Diagram batang kepadatan penduduk per Kabupaten/Kota setelah dilakukan subsetting

Setelah subsetting hanya pada kategori "Sangat Padat", grafik menjadi lebih fokus dan bersih. Perbandingan antarwilayah terlihat jauh lebih jelas, memudahkan identifikasi pola dan analisis karakteristik wilayah ber-kepadatan tinggi. Subsetting membantu meningkatkan ketepatan interpretasi dan membuat hasil analisis lebih informatif.

HASIL DAN PEMBAHASAN

Perbandingan Sebelum dan Setelah Sorting

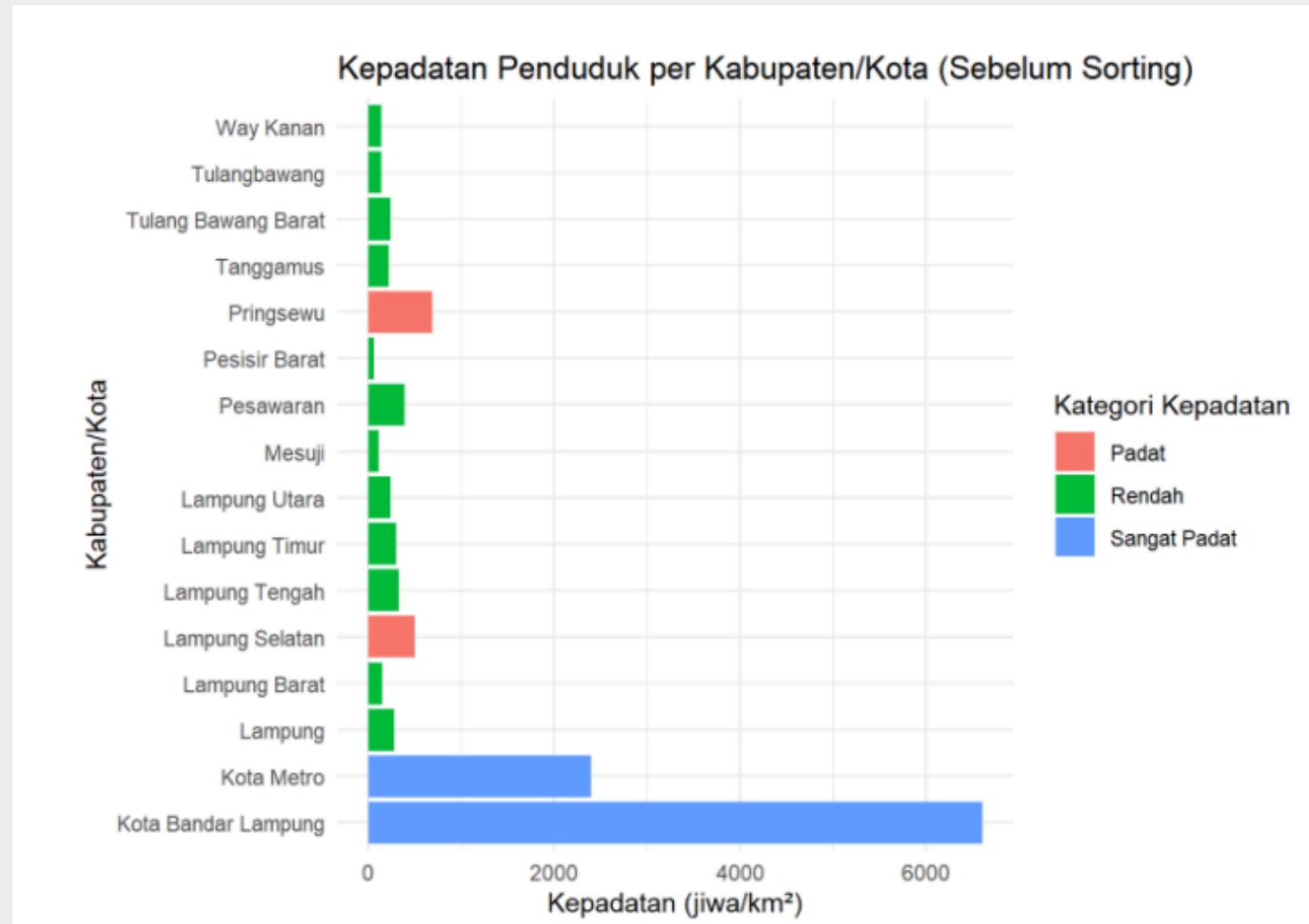


Diagram batang Kepadatan Penduduk per Kabupaten/Kota sebelum dilakukan proses Sorting.

Sebelum Sorting: urutan kabupaten/kota masih sesuai urutan asli dataset sehingga pola kepadatan tidak terlihat jelas. Wilayah kepadatan tinggi bercampur dengan wilayah kepadatan rendah.

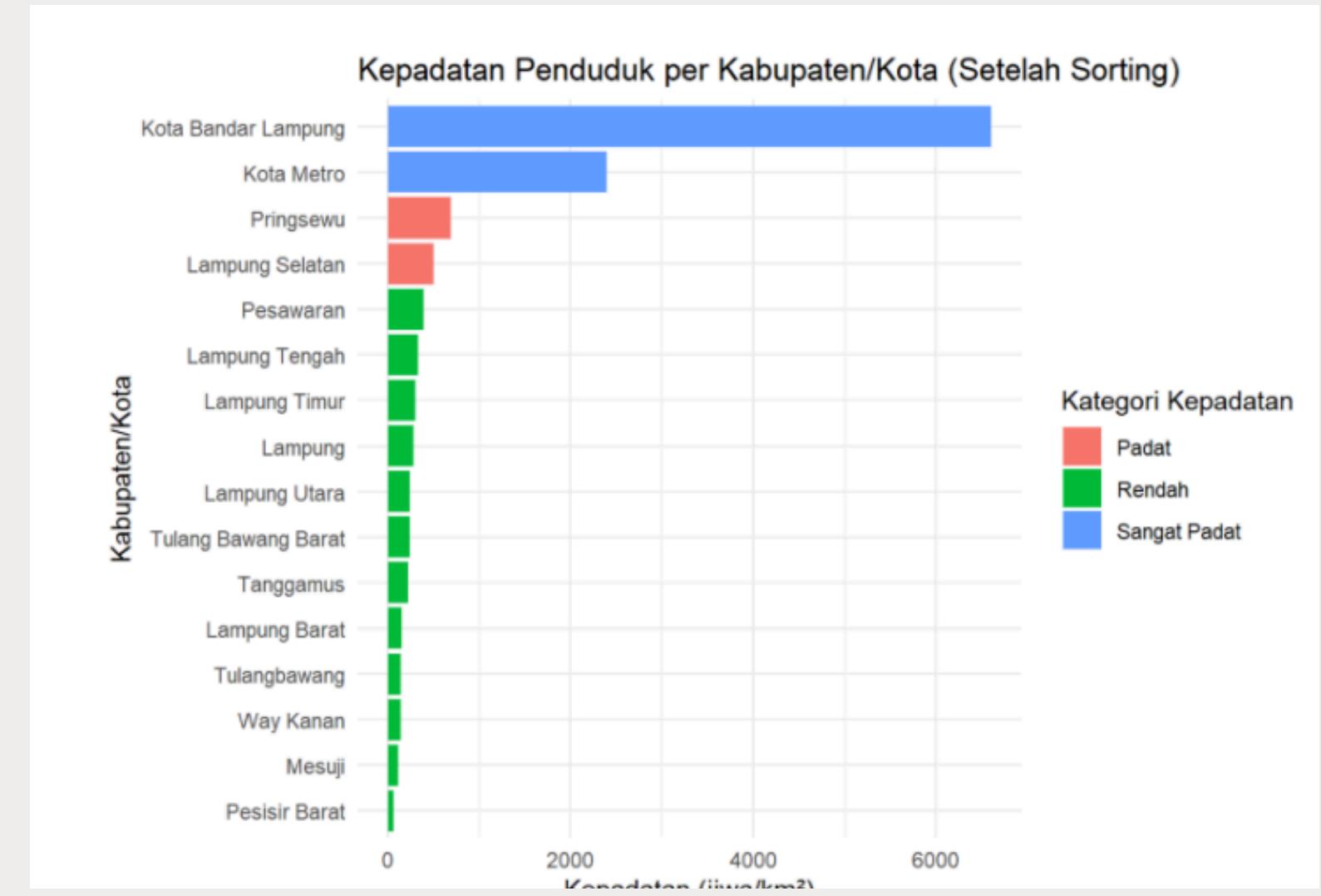
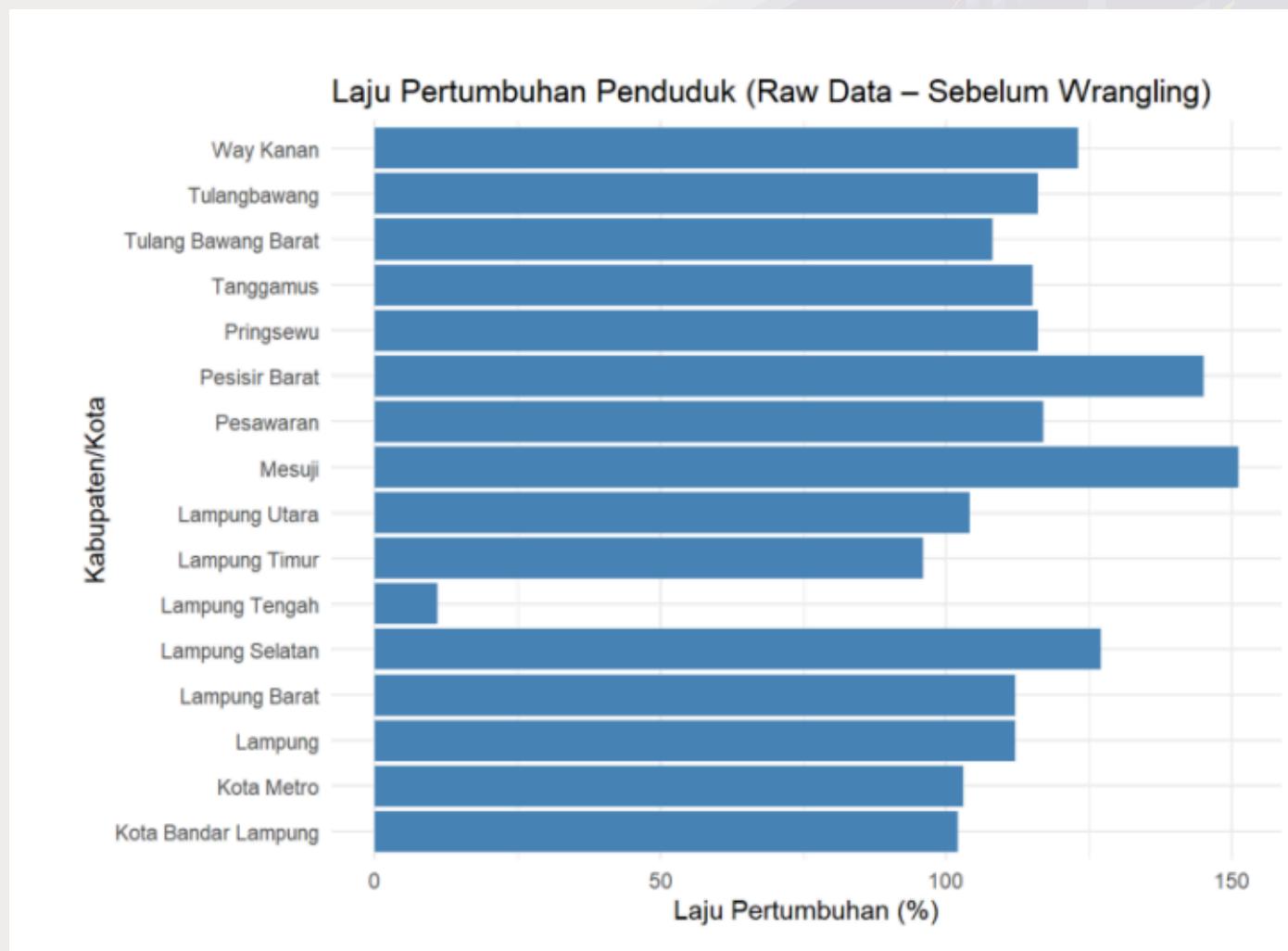


Diagram batang Kepadatan Penduduk per Kabupaten/Kota setelah dilakukan proses Sorting

Setelah Sorting: data diurutkan dari kepadatan tertinggi ke terendah (descending), sehingga pola distribusi lebih jelas dan nilai ekstrem (Bandar Lampung & Metro) langsung terlihat.

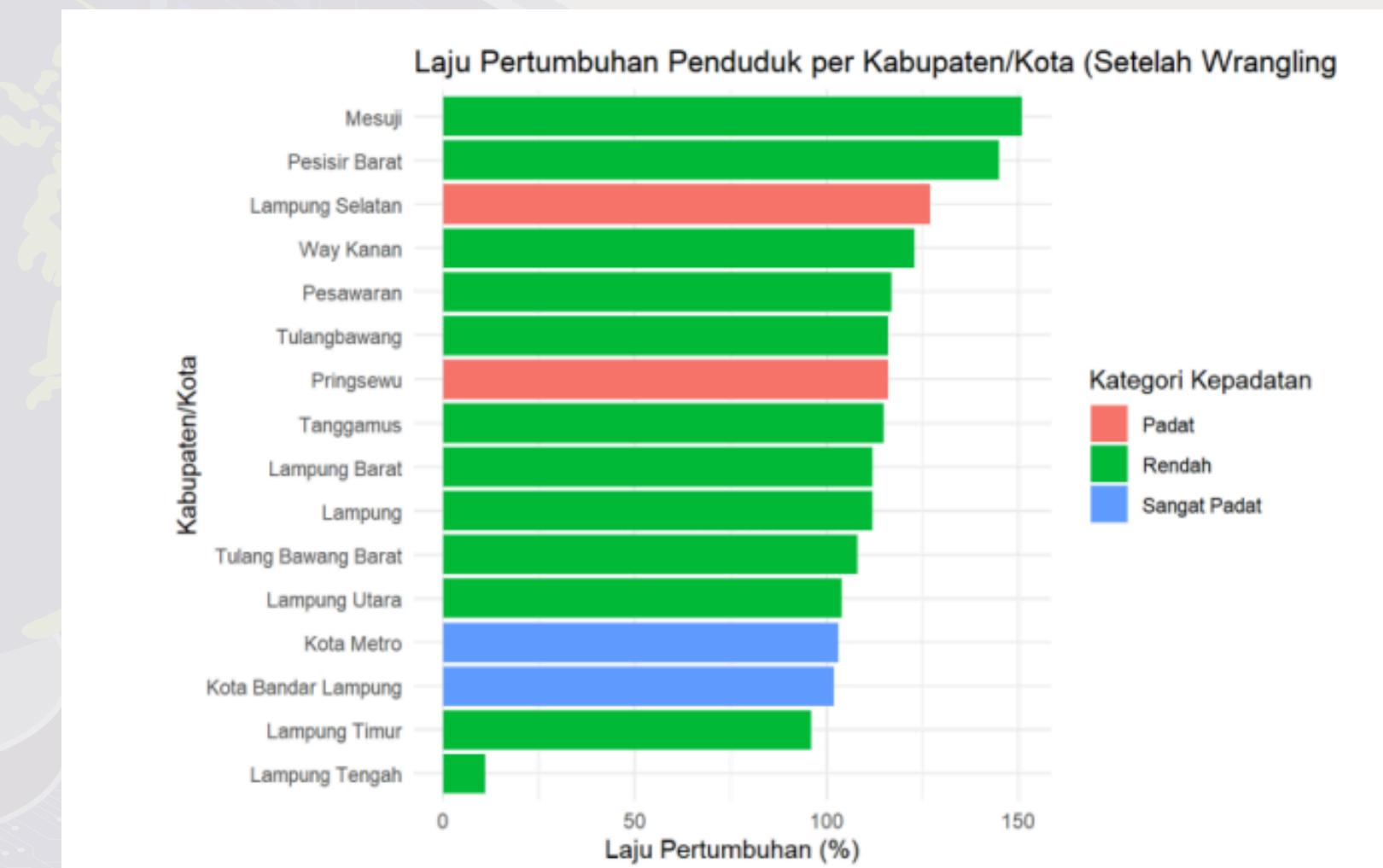
HASIL DAN PEMBAHASAN

Perbandingan Sebelum dan Setelah Proses Wrangling



Visualisasi laju pertumbuhan penduduk berdasarkan data mentah sebelum dilakukan proses Wrangling.

Grafik sebelum wrangling menunjukkan data yang masih berantakan, seperti format angka yang tidak seragam, urutan kabupaten/kota acak, dan belum adanya kategori tambahan. Kondisi ini membuat pola pertumbuhan penduduk sulit dibaca.



Visualisasi laju pertumbuhan penduduk berdasarkan data mentah sebelum dilakukan proses Wrangling

Setelah proses wrangling, data berhasil dibersihkan dan distandarisasi. Urutan kabupaten/kota telah tersortir, format numerik seragam, dan kategori baru seperti KepadatanGroup dan Dominasi_JK ditambahkan. Hal ini membuat visualisasi lebih informatif dan perbandingan antarwilayah menjadi jauh lebih jelas.

KESIMPULAN

Proses data munging membuat dataset lebih bersih dan konsisten, serta variabel baru yang dibentuk membantu analisis perbandingan antarwilayah. Hasil menunjukkan kepadatan tertinggi berada di kota besar, laju pertumbuhan penduduk relatif seragam, dan rasio jenis kelamin didominasi laki-laki tipis. Visualisasi juga memperjelas pola dan mendukung interpretasi hasil.

SARAN

Penelitian selanjutnya disarankan menambahkan data time series, luas wilayah resmi, serta uji korelasi untuk analisis lebih mendalam. Dokumentasi proses wrangling dan penggunaan otomatisasi seperti R Markdown juga perlu diterapkan agar analisis dapat direplikasi dengan mudah dan lebih efisien.