

Membedah Faktor-Faktor yang Memengaruhi Efisiensi Bahan Bakar (Dataset MPG)

Disusun untuk Memenuhi Tugas Mata Kuliah Data Science

Dosen Pengampu: Khaerul Anam, M.Kom.



Disusun oleh:

- | | |
|--------------------------|------------|
| 1. Ayu Aulia | (43240354) |
| 2. Putri Tri Setyawati | (43240462) |
| 3. Tion Hermawan | (43240445) |
| 4. Zacky Muhammad Dinata | (43240450) |

PROGRAM STUDI SISTEM INFORMASI

STMIK IKMI CIREBON

2025

KATA PENGANTAR

Puji syukur kami panjatkan kehadiran Tuhan Yang Maha Esa, karena atas rahmat dan karunia-Nya, kami dapat menyelesaikan penyusunan laporan tugas kelompok untuk mata kuliah Data Science dengan judul “Membedah Faktor-Faktor yang Mempengaruhi Efisiensi Bahan Bakar (Dataset MPG)” ini dengan baik dan tepat waktu.

Penyusunan laporan ini tidak akan terwujud tanpa bimbingan dan dukungan dari berbagai pihak. Oleh karena itu, kami ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

- Bapak Khaerul Anam, M.Kom., selaku dosen pengampu mata kuliah Data Science yang telah memberikan bimbingan, arahan, serta ilmu yang sangat bermanfaat selama proses pengerjaan tugas ini.
- Seluruh anggota kelompok: Ayu Aulia, Putri Tri Setyawati, Tion Hermawan, dan Zacky Muhammad Dinata, atas kerja sama, kontribusi, dan semangat kolaboratif yang solid sehingga laporan ini dapat terselesaikan.
- Semua pihak yang tidak dapat kami sebutkan satu per satu, yang telah memberikan dukungan baik secara moril maupun materil.

Kami menyadari bahwa laporan ini masih jauh dari kata sempurna karena keterbatasan pengetahuan dan pengalaman kami. Oleh karena itu, kami sangat mengharapkan kritik dan saran yang membangun dari pembaca demi perbaikan di masa mendatang.

Semoga laporan ini dapat memberikan manfaat serta menambah wawasan, khususnya dalam bidang analisis data, bagi kita semua.

Cirebon, 15 September 2025

Penyusun,
Kelompok Bismillah Bisa Ngerjainnya

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI.....	ii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	1
1.3 Tujuan Analisis	1
1.4 Metodologi dan Dataset.....	2
BAB 2 DESKRIPSI DAN VALIDASI DATA	3
2.1 Deskripsi Variabel dan Tipe Data	3
2.2 Analisis Statistik Deskriptif	4
2.2.1 Data Ratio/Interval	4
2.2.2 Data Nominal/Ordinal	5
2.3 Analisis Korelasi.....	5
BAB 3 EKSPLORASI DATA MELALUI VISUALISASI (EDA)	7
3.1 Histogram Visual.....	7
3.1.1 Histogram HWY (Highway Efficiency)	7
3.1.2 Histogram CTY (City Efficiency)	8
3.2 Scatter Plot Visual.....	8
3.2.1 Scatter Plot HWY (Highway Efficiency)	9
3.2.2 Scatter Plot CTY (City Efficiency)	10
3.3 Box Plot Visual	10
3.3.1 Box Plot HWY (Highway Efficiency)	11
3.3.2 Box Plot CTY (City Efficiency).....	12
BAB 4 KESIMPULAN DAN SARAN	13
LAMPIRAN.....	14

DAFTAR GAMBAR

Gambar 1 Histogram HWY (Highway Efficiency).....	7
Gambar 2 Histogram CTY (City Efficiency).....	8
Gambar 3 Scatter Plot HWY (Highway Efficiency)	9
Gambar 4 Scatter Plot CTY (City Efficiency).....	10
Gambar 5 Box Plot HWY (Highway Efficiency)	11
Gambar 6 Box Plot CTY (City Efficiency)	12

DAFTAR TABEL

Table 1 Deskripsi Variabel dan Tipe Data	3
Table 2 Data Ratio/Interval.....	4
Table 3 Data Nominal/Ordinal.....	5
Table 4 Data Nominal/Ordinal.....	5
Table 5 Analisis Korelasi	6

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Di tengah isu kenaikan harga bahan bakar dan meningkatnya kesadaran akan dampak lingkungan, efisiensi konsumsi bahan bakar telah menjadi salah satu pertimbangan utama bagi konsumen dalam memilih kendaraan. Produsen mobil pun berlomba-lomba menciptakan teknologi untuk menghasilkan kendaraan yang lebih irit tanpa mengorbankan performa. Namun, faktor apa saja yang sesungguhnya paling berpengaruh terhadap tingkat efisiensi sebuah mobil? Apakah kapasitas mesin yang besar selalu berarti boros? Dan apakah kelas mobil tertentu secara inheren lebih efisien daripada yang lain? Fenomena inilah yang melatarbelakangi analisis kami untuk membedah data dan menemukan jawaban yang didukung oleh bukti.

1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut, kami merumuskan beberapa pertanyaan analisis sebagai berikut:

1. Bagaimana distribusi umum tingkat efisiensi bahan bakar mobil, baik di dalam kota (city) maupun di jalan tol (hwy)?
2. Apakah terdapat hubungan antara ukuran mesin (displ) dengan efisiensi bahan bakar?
3. Bagaimana perbandingan tingkat efisiensi bahan bakar antar kelas mobil yang berbeda (misalnya suv vs compact)?
4. Kelas mobil mana yang dapat diidentifikasi sebagai yang paling efisien dan paling boros?

1.3 Tujuan Analisis

Analisis ini bertujuan untuk:

1. Memahami karakteristik dan pola sebaran data efisiensi bahan bakar.
2. Mengidentifikasi dan mengukur hubungan antara variabel-variabel kunci seperti ukuran mesin dan efisiensi.
3. Membandingkan performa efisiensi antar kelas mobil untuk memberikan wawasan yang jelas.

4. Menarik kesimpulan berbasis data mengenai faktor-faktor yang paling signifikan dalam menentukan efisiensi bahan bakar sebuah mobil.

1.4 Metodologi dan Dataset

Analisis ini menggunakan dataset mpg yang merupakan bagian dari paket ggplot2 di R. Dataset ini berisi informasi mengenai efisiensi bahan bakar dari 234 mobil dari tahun 1999 hingga 2008.

Kerangka kerja analisis yang digunakan mengacu pada proses **Data Understanding**, yang mencakup tahapan validasi kualitas data, deskripsi data melalui statistik, dan eksplorasi data menggunakan teknik visualisasi untuk menemukan pola, tren, dan anomali.

BAB 2

DESKRIPSI DAN VALIDASI DATA

2.1 Deskripsi Variabel dan Tipe Data

Variabel	Tipe Data	Justifikasi
Manufacturer	Nominal	Kategori nama pabrikan tanpa ada urutan tingkatan.
Model	Nominal	Kategori nama model mobil tanpa ada urutan tingkatan.
Displ	Ratio	Ukuran mesin (liter) dengan nilai nol mutlak dan interval yang jelas.
Year	Interval	Tahun pembuatan, intervalnya bermakna tapi tidak punya nol mutlak.
CYL	Ordinal	Jumlah silinder memiliki urutan ($4 < 6 < 8$), tapi jaraknya tidak selalu bermakna sama.
Trans	Nominal	Kategori jenis transmisi tanpa ada urutan tingkatan.
DRV	Nominal	Kategori jenis penggerak roda (f, r, 4) tanpa urutan.
CTY	Ratio	Ukuran efisiensi (MPG) dengan nilai nol mutlak dan interval yang jelas.
HWY	Ratio	Ukuran efisiensi (MPG) dengan nilai nol mutlak dan interval yang jelas.
FL	Nominal	Kategori jenis bahan bakar (r, p, d) tanpa urutan tingkatan.
Class	Ordinal	Kategori kelas mobil yang memiliki urutan berdasarkan ukuran (compact < midsize < suv).

Table 1 Deskripsi Variabel dan Tipe Data

Setiap variabel dalam dataset diklasifikasikan tipe datanya untuk menentukan metode analisis yang tepat

2.2 Analisis Statistik Deskriptif

Setelah memahami tipe datanya, kami melakukan analisis statistik deskriptif untuk mendapatkan gambaran awal dari data.

2.2.1 Data Ratio/Interval

hwy		cty		displ	
min.	:12.00	Min.	:9.00	Min.	:1.600
1st Qu.	:18.00	1st. Qu.	:14.00	1st Qu.	:2.400
Median	:24.00	Median	:17.00	Median	:3.300
Mean	:23.44	Mean	:16.86	Mean	:3.472
3rd Qu.	:27.00	3rd Qu.	:19.00	3rd Qu.	:4.600
Max.	:44.00	Max.	:35.00	Max.	:7.000

Table 2 Data Ratio/Interval

Untuk variabel numerik yang bersifat kontinu, kami menghitung ukuran pemusatan dan sebaran datanya:

- **Efisiensi di Tol (hwy):**

Rata-rata (Mean): 23.44 MPG

Nilai Tengah (Median): 24 MPG

Standar Deviasi: 5.95 MPG

- **Efisiensi di Kota (cty):**

Rata-rata (Mean): 16.86 MPG

Nilai Tengah (Median): 17 MPG

Standar Deviasi: 4.26 MPG

- **Ukuran Mesin (displ):**

Rata-rata (Mean): 3.47 Liter

Nilai Tengah (Median): 3.3 Liter

Standar Deviasi: 1.29 Liter

Dari data ini, terlihat bahwa rata-rata efisiensi di jalan tol (hwy) jauh lebih tinggi daripada di kota (cty). Standar deviasi hwy juga lebih besar, menandakan variasi efisiensi di jalan tol lebih beragam.

2.2.2 Data Nominal/Ordinal

Untuk variabel kategorikal, kami menghitung frekuensi atau jumlah kemunculan setiap kategori untuk melihat komposisi data:

- **Kelas Mobil (class):** Kelas mobil yang paling umum dalam dataset ini adalah suv (62 mobil), diikuti oleh compact (47 mobil) dan midsize (41 mobil).

2seater	compact	midsize	minivan	pickup	subcompact	suv
5	47	41	11	33	35	62

Table 3 Data Nominal/Ordinal

- **Pabrikan (manufacturer):** Pabrikan yang paling banyak muncul dalam data adalah dodge (37 mobil), toyota (34 mobil), dan volkswagen (27 mobil).

audi	chevrolet	dodge	ford	honda	hyundai	jeep	land rover
18	19	37	25	9	14	8	4
lincoln	mercury	nissan	pontiac	subaru	toyota	volswagen	
3	4	13	5	14	34	27	

Table 4 Data Nominal/Ordinal

2.3 Analisis Korelasi

	displ	year	cyl	cty	hwy
displ	1	0.147842816	0.9302271	-0.79852397	-0.766020021

year	0.1478428	1	0.1222453	-0.03723229	0.002157643
cyl	0.9302271	0.122245347	1	-0.80577141	-0.761912354
cty	-0.798524	-0.037232291	-0.8057714	1	0.955915914
hwy	-0.76602	0.002157643	-0.7619124	0.95591591	1

Table 5 Analisis Korelasi

Untuk melihat kekuatan hubungan antar variabel numerik, kami menghitung koefisien korelasi.

- Korelasi antara displ dan cty: **-0.80**
- Korelasi antara displ dan hwy: **-0.77**

Nilai korelasi yang mendekati -1 ini menunjukkan adanya **hubungan negatif yang sangat kuat**. Artinya, semakin besar nilai displ (ukuran mesin), semakin kecil nilai cty dan hwy (semakin boros).

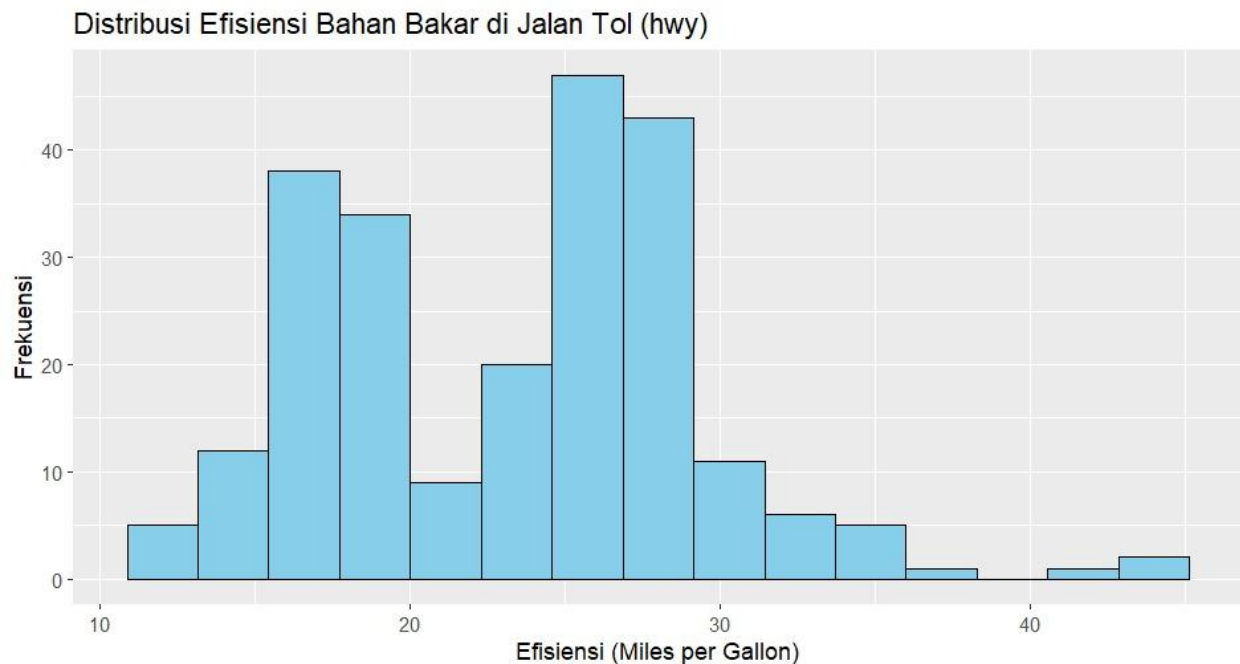
BAB 3

EKSPLORASI DATA MELALUI VISUALISASI (EDA)

3.1 Histogram Visual

Histogram digunakan untuk melihat distribusi atau sebaran data dari variabel efisiensi bahan bakar.

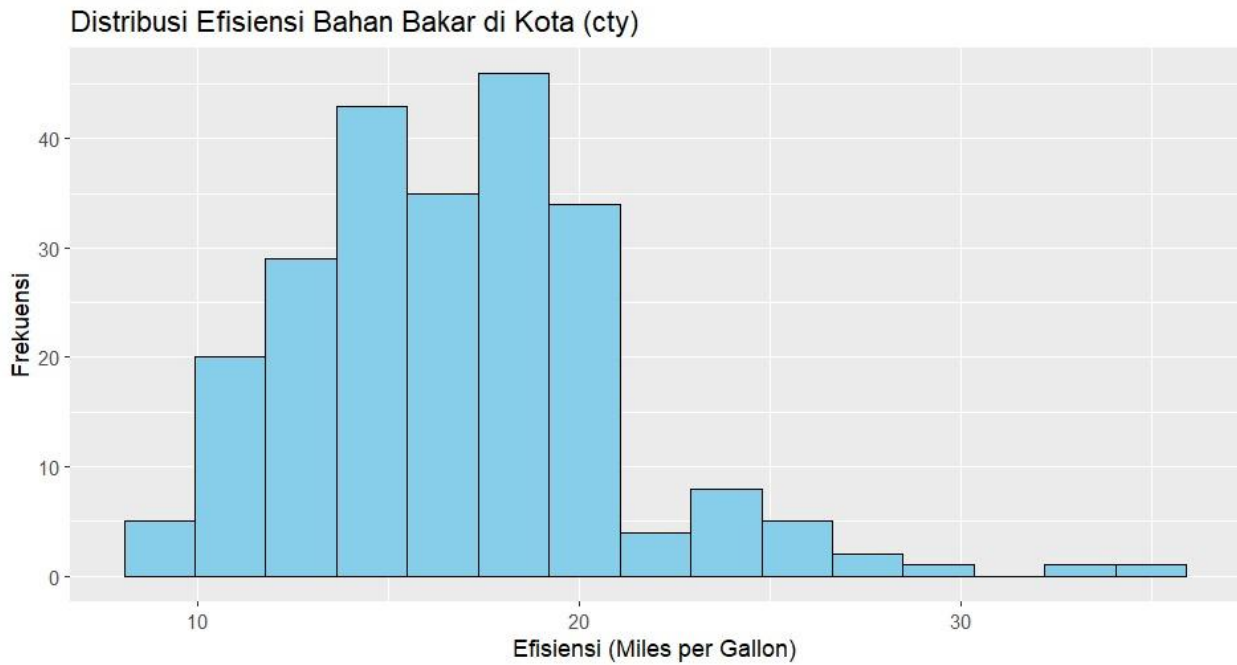
3.1.1 Histogram HWY (Highway Efficiency)



Gambar 1 Histogram HWY (Highway Efficiency)

Distribusi efisiensi di jalan tol (hwy) menunjukkan bahwa mayoritas mobil memiliki efisiensi di rentang 20 hingga 30 MPG, dengan puncak frekuensi (modus) berada di sekitar 25-30 MPG. Sebaran datanya cenderung miring ke kanan, menandakan ada beberapa mobil yang sangat efisien.

3.1.2 Histogram CTY (City Efficiency)



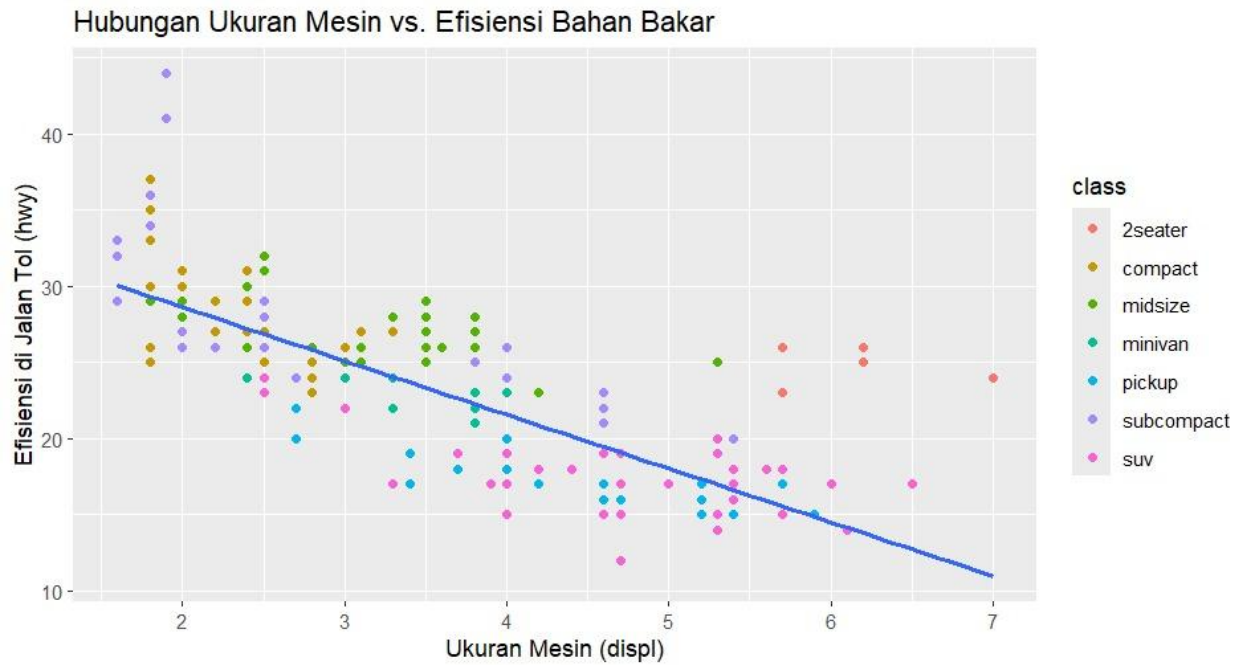
Gambar 2 Histogram CTY (City Efficiency)

Untuk efisiensi di dalam kota (cty), distribusinya bergeser ke kiri (nilai lebih rendah). Puncak frekuensi berada di rentang 15-20 MPG. Perbandingan kedua histogram ini secara visual mengkonfirmasi bahwa mobil secara umum lebih boros saat digunakan di perkotaan.

3.2 Scatter Plot Visual

Scatter plot digunakan untuk melihat hubungan antara dua variabel numerik, dalam hal ini antara ukuran mesin (displ) dan efisiensi bahan bakar.

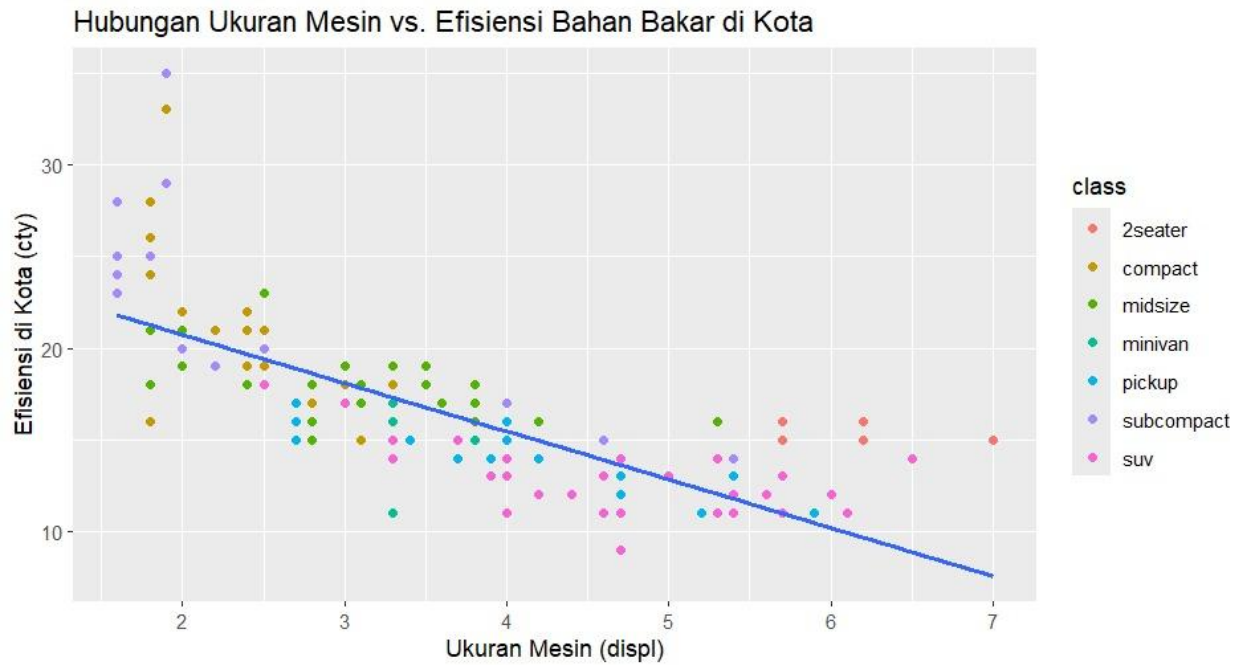
3.2.1 Scatter Plot HWY (Highway Efficiency)



Gambar 3 Scatter Plot HWY (Highway Efficiency)

Plot ini secara visual membuktikan temuan dari analisis korelasi. Pola titik-titik yang menurun dari kiri atas ke kanan bawah, ditambah dengan garis tren, menunjukkan hubungan negatif yang kuat antara ukuran mesin (displ) dan efisiensi di jalan tol (hwy).

3.2.2 Scatter Plot CTY (City Efficiency)



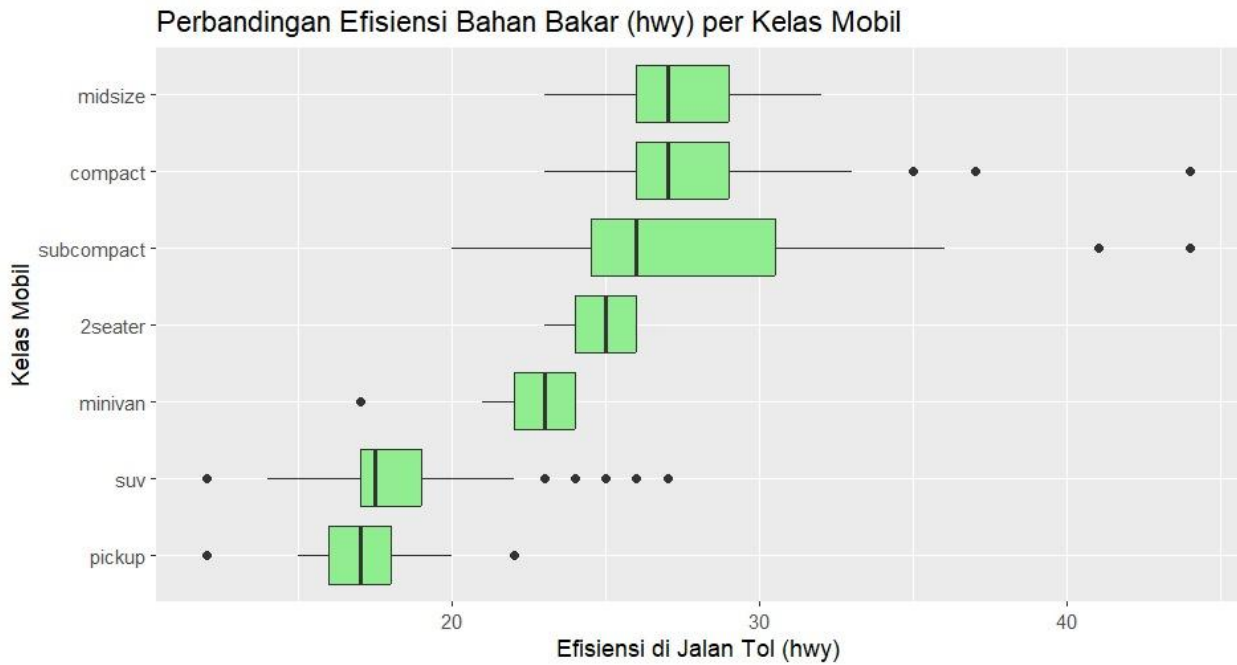
Gambar 4 Scatter Plot CTY (City Efficiency)

Pola yang sama juga terlihat pada efisiensi di kota. Sederhananya, baik untuk hwy maupun cty, semakin besar kapasitas mesin sebuah mobil, maka ia akan cenderung semakin boros.

3.3 Box Plot Visual

Box plot digunakan untuk membandingkan distribusi data antar kategori, dalam hal ini membandingkan efisiensi bahan bakar antar kelas mobil.

3.3.1 Box Plot HWY (Highway Efficiency)

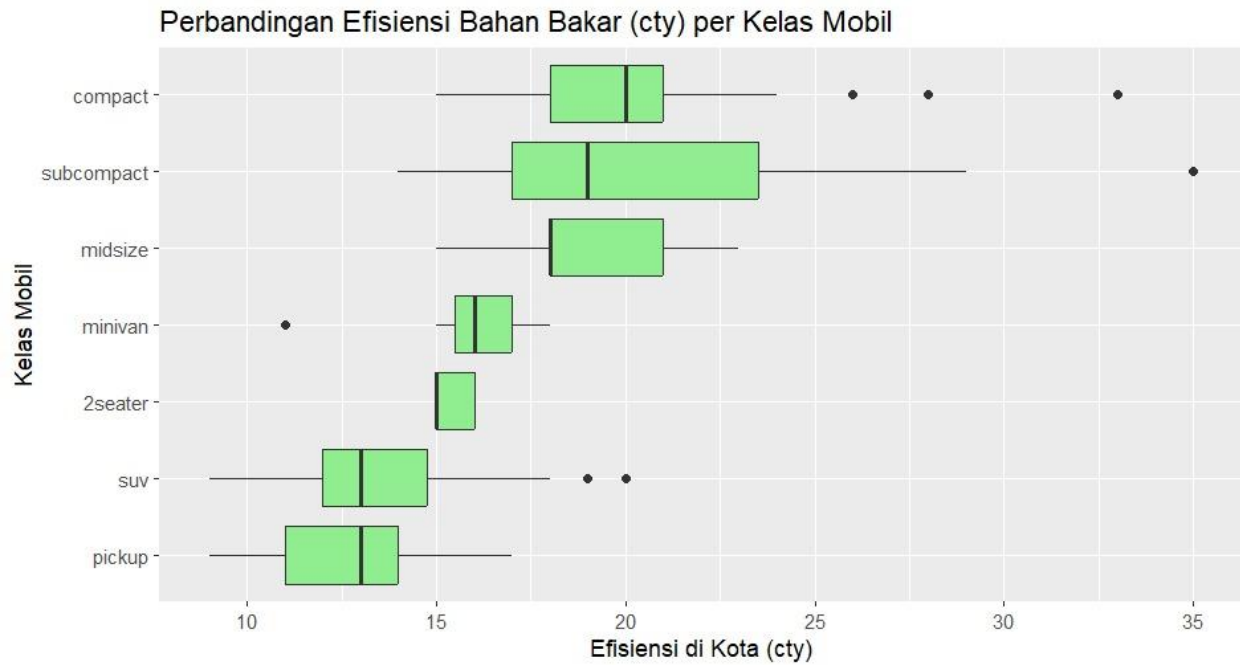


Gambar 5 Box Plot HWY (Highway Efficiency)

Box plot untuk hwy memberikan wawasan yang kaya:

- **Kelas Paling Efisien:** compact dan subcompact secara konsisten menempati posisi teratas dengan median efisiensi tertinggi.
- **Kelas Paling Boros:** pickup dan suv berada di posisi terbawah.
- **Identifikasi Outlier:** Terlihat ada beberapa mobil subcompact yang efisiensinya sangat tinggi (di atas 40 MPG), menjadikannya kasus spesial yang sangat irit.

3.3.2 Box Plot CTY (City Efficiency)



Gambar 6 Box Plot CTY (City Efficiency)

Pola perbandingan antar kelas mobil di dalam kota (cty) sangat mirip dengan di jalan tol. Kelas compact dan subcompact tetap menjadi yang paling irit, sementara pickup dan suv tetap yang paling boros, memperkuat kesimpulan mengenai pengaruh kelas mobil terhadap efisiensi.

BAB 4

KESIMPULAN DAN SARAN

Setelah melalui serangkaian tahap analisis, mulai dari validasi data hingga eksplorasi visual, kami berhasil menarik beberapa benang merah yang menjawab pertanyaan-pertanyaan utama penelitian ini. Terbukti bahwa efisiensi bahan bakar sebuah mobil bukanlah sebuah misteri, melainkan hasil dari interaksi beberapa faktor kunci yang dapat diukur.

Kesimpulan utama dari analisis kami adalah bahwa kondisi berkendara, ukuran mesin, dan kelas mobil secara signifikan menentukan tingkat efisiensi bahan bakar. Data secara konsisten menunjukkan bahwa mobil jauh lebih irit ketika dipacu di jalan tol (hwy) dibandingkan saat berjibaku dengan lalu lintas perkotaan (cty). Faktor paling dominan yang ditemukan adalah ukuran mesin (displ), di mana terdapat hubungan negatif yang sangat kuat artinya semakin besar kapasitas mesin, hampir dapat dipastikan konsumsi bahan bakarnya akan semakin boros. Temuan ini diperkuat dengan adanya hierarki yang jelas antar kelas mobil. Kelas compact dan subcompact yang umumnya bermesin kecil terbukti menjadi yang paling efisien, sementara kelas pickup dan suv yang cenderung memiliki mesin besar berada di spektrum paling tidak efisien.

Berdasarkan wawasan yang didapatkan, kami ingin memberikan beberapa saran. Bagi konsumen yang menjadikan keiritan bahan bakar sebagai prioritas utama, data dengan jelas merekomendasikan untuk memilih kendaraan dari kelas compact atau subcompact dengan kapasitas mesin yang tidak terlalu besar. Lebih jauh lagi, untuk pengembangan analisis di masa depan, penelitian ini dapat diperdalam dengan menginvestigasi pengaruh variabel lain seperti jenis transmisi (trans) atau sistem penggerak roda (drv) terhadap efisiensi bahan bakar untuk mendapatkan pemahaman yang lebih holistik dan komprehensif mengenai dinamika efisiensi bahan bakar.

LAMPIRAN

