

# UTS Analisis Data Media Sosial

Kelas B



Dennis Imanuel

NPM: 6161901066

UNIVERSITAS KATOLIK PARAHYANGAN

BANDUNG

2023

# 1. Dataset Review Motor

Dataset ‘Review\_Motor.csv’ adalah dataset yang telah saya kumpulkan terkait hasil ulasan-ulasan pembeli sepeda motor melalui *marketplace* online seperti Tokopedia, Shopee, dan beberapa situs belanja online lainnya. Dataset tersebut disimpan dalam bentuk file *comma separated value* (csv) dan dapat diakses melalui *respository* Github saya menggunakan link: [https://github.com/aiightvert/Data-UTS-Andat-Medsos/blob/0371e5926cf11d7f09e30e71053fe70da8efcf4c/Review\\_Motor.csv](https://github.com/aiightvert/Data-UTS-Andat-Medsos/blob/0371e5926cf11d7f09e30e71053fe70da8efcf4c/Review_Motor.csv)

Dataset tersebut mengandung 5 kolom parameter yaitu:

- No : penomoran baris dataset
- Nama : nama atau deskripsi produk
- Rating : nilai ulasan pembeli terhadap produk dalam skala 1 sampai 5
- Comments : komentar pembeli terhadap produk
- Links : tautan menuju laman penjualan produk

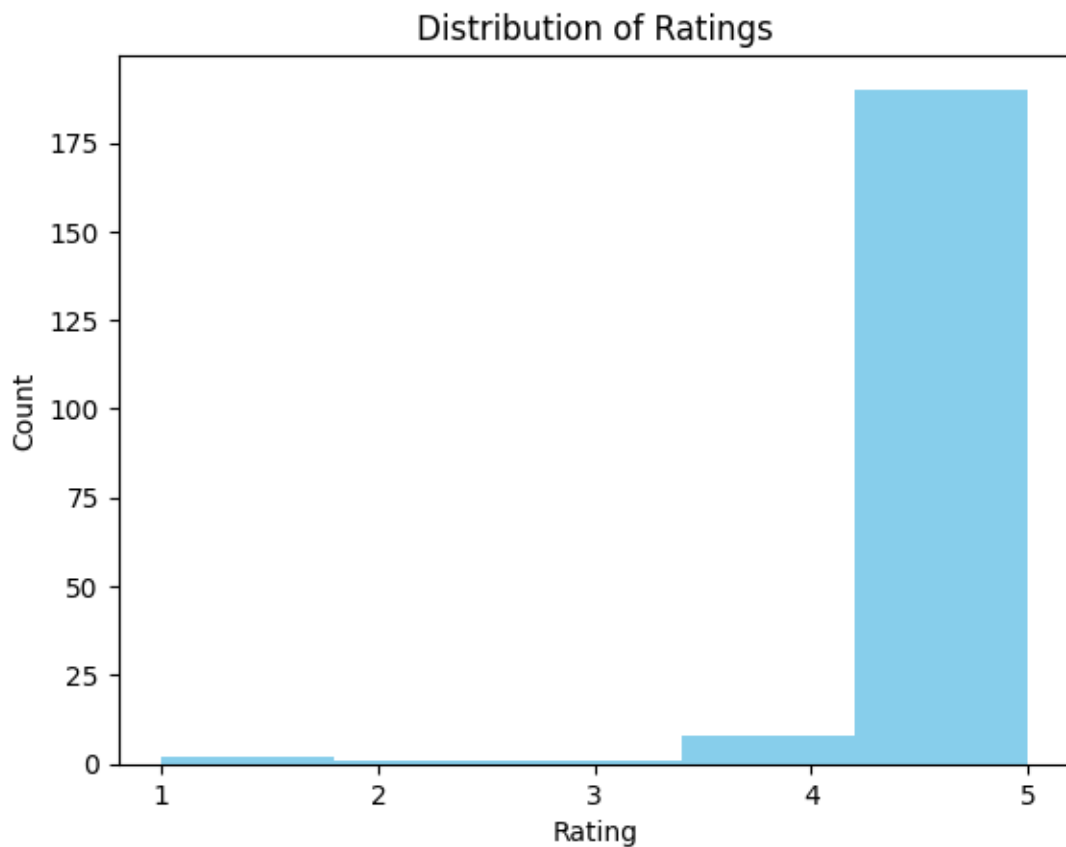
Dataset tersebut memiliki 202 baris data yang telah berhasil saya kumpulkan. Berikut adalah tampilan *screenshoot* tabel data yang ditampilkan dengan Python untuk 20 baris data teratas.

df.head(20)

	No	Nama	Rating	Comments	Links
0	1	Honda BeAT CBS	5	seller fast respon, berintegritas dan helpful ...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
1	2	Honda BeAT CBS	5	Semua layanan dan service toko nya bagus, komu...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
2	3	Honda BeAT CBS	5	Kendaraan sangat bagus, tepat waktu, toko nya ...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
3	4	Honda BeAT CBS	5	Pelayanan cepet dan respon cepet pokoknya mantap	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
4	5	Honda BeAT CBS	5	Proses dan pengiriman cepat sampai sesuai deng...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
5	6	Honda BeAT CBS	5	Puas Berkali-kali beli motor di sini nih: PUAS...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
6	7	Honda BeAT CBS	5	Pengiriman lebih cepat dari janji semula. Kond...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
7	8	Honda BeAT CBS	1	sudah hampir 1 bulan sejak motor dikirim, tp p...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
8	9	Honda BeAT CBS	5	Barang datang nya cepat. Sudah sesuai pesanan. ...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
9	10	Honda BeAT CBS	5	motor ok, pengiriman ok, seller terpercaya nih...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
10	11	Honda BeAT CBS	5	Respon penjual ramah , ga sampai 3 hari barang...	https://www.tokopedia.com/ajaya-179/sepeda-mot...
11	12	Honda BeAT CBS	5	fast respon, pengiriman cepat. maksih numa motor.	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
12	13	Honda BeAT CBS	5	Good seller pesan malam motor besok nya langsun...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
13	14	Honda BeAT CBS	5	Alhamdulillah motor nya dah datang. Proses nya...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
14	15	Honda BeAT CBS	5	Alhamdulillah motor sampe dg selamat sesuai yg...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
15	16	Honda BeAT CBS	5	alhamdulillah proses nya cepat hanya 2 hari mot...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
16	17	Honda BeAT CBS	5	Seller responsive, pengiriman cepat, barang se...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
17	18	Honda BeAT CBS	5	Cepet sampe, cucok lah pokok nya.. tq...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
18	19	Honda BeAT CBS	4	Mantap zooos motornya	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...
19	20	Honda BeAT CBS	5	Alhamdulillah.. Tetap Semangat.. Tetap Berbuat...	https://shopee.co.id/HONDA-BEAT-CBS-SEPEDA-MOT...

# 2. Distribusi Rating

Analisis pertama yang ingin dilakukan adalah mengetahui distribusi rating dari data yang telah dikumpulkan menggunakan bantuan plot histogram untuk memvisualisasikan distribusinya.



```
rating_counts = df['Rating'].value_counts().sort_index()  
rating_counts
```

```
1      2  
2      1  
3      1  
4      8  
5     190  
Name: Rating, dtype: int64
```

Berdasarkan hasil visualisasi distribusi rating, dapat dilihat bahwa kebanyakan pembeli memberi hasil ulasan sangat baik atau maksimal yaitu nilai 5. Hanya beberapa pembeli yang tidak puas dengan hasil pembelian motornya yaitu 2 orang yang memberikan nilai 1, dan 1 orang yang memberikan nilai 2.

### 3. Data Pre-processing

Pada bagian berikutnya, perlu dilakukan *pre-processing* dan pembersihan data terlebih dahulu sebelum dapat lanjut ke analisis berikutnya. Pada tahap ini, kita akan fokus kepada kolom ‘Comments’ untuk disiapkan agar bisa dapat dijadikan sebagai sumber informasi yang berguna untuk data *text analysis* dan *visualization*. Adapun langkah-langkah yang dilakukan pada data *pre-processing* dalam proyek ini sebagaimana telah dikerjakan dalam Google Collab Python Notebook adalah:

1. Mengubah seluruh karakter dalam masing-masing kolom ‘Comments’ menjadi tipe data *string*.
2. Menghilangkan baris data yang kosong (NULL) dalam kolom ‘Comments’.
3. Menghilangkan karakter-karakter spesial yang mungkin tidak berbentuk alfabetis dan angka baku atau umum yang tidak akan berguna dalam analisis teks.
4. Mengubah seluruh kata menjadi huruf kecil dalam kolom ‘Comments’.

5. Menghilangkan tanda baca dari kolom ‘Comments’.
6. Menghilangkan kata-kata langka yang hanya muncul sesekali saja pada kolom ‘Comments’.
7. Mengkoreksi pengejaan yang dapat dikoreksi pada masing-masing kata dalam kolom ‘Comments’.
8. Melakukan *stemming* dan *lemmatization* untuk mengubah dan mengklasifikasi masing-masing kata dari kolom ‘Comments’ ke dalam bentuk paling sederhana-nya tanpa imbuhan.

## 4. Data Analysis

Sebelum melakukan analisis teks dan visualisasi menggunakan *wordcloud*, akan ditambahkan beberapa kolom yaitu:

1. Panjang\_Comments : untuk menghitung jumlah huruf dalam masing-masing baris untuk kolom 'Comments'.
2. Jumlah\_Kata : untuk menghitung jumlah kata dalam masing-masing baris untuk kolom 'Comments'.
3. Polarity : untuk menentukan nilai sentimen positif atau negatif terhadap komentar yang disampaikan pembeli dalam kolom 'Comments' dengan skala -1 sampai dengan 1.

Walaupun parameter-parameter di atas tidak akan digunakan dalam analisis untuk proyek ini, namun penambahan parameter ini dapat mempermudah pengamat untuk melakukan analisis dan visualisasi lain terhadap analisis data teks review motor ini serta menambah informasi keterangan masing-masing baris dataset.

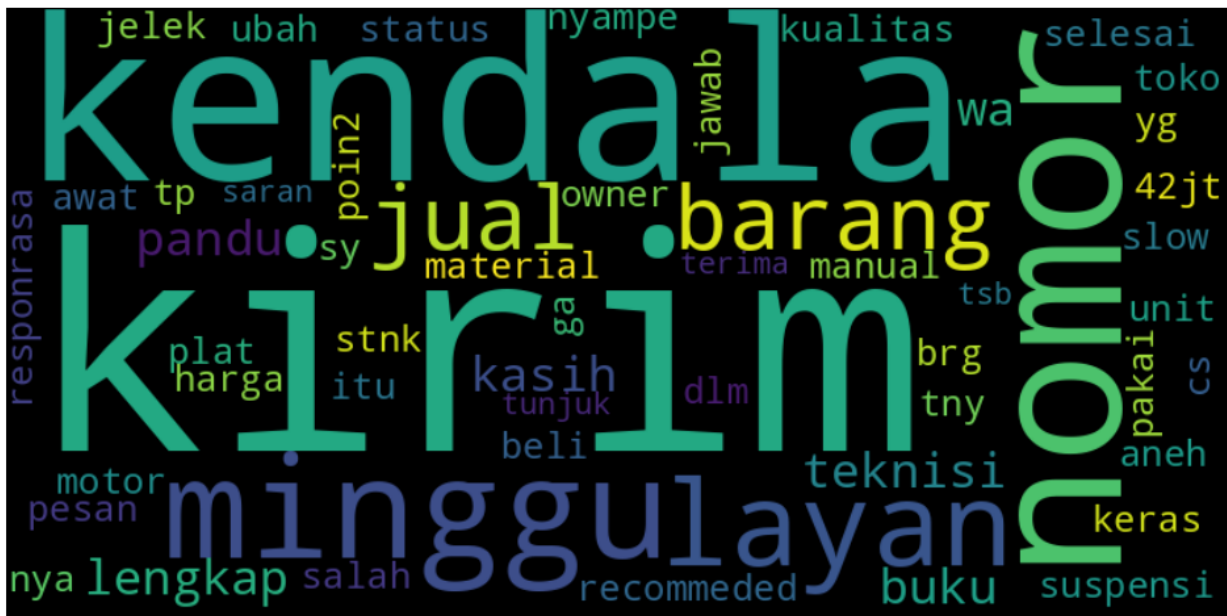
4.1. Buatkan *wordcloud* untuk semua komentar yang didapatkan setelah dilakukan *pre-processing* data dan jelaskan hasil yang didapatkan!



Komentar yang diberikan pembeli yang umumnya kebanyakan merujuk kepada pelayanan penjual ketimbang terhadap produk motor itu sendiri. Karena mayoritas ulasan merupakan ulasan positif yang disertai nilai ulasan 5, maka kata-kata yang muncul pada *wordcloud* ini merupakan kata-kata dengan sentimen baik. Kata-kata kunci dapat dilihat pada kata-kata berukuran besar

dalam *wordcloud* yang menandakan frekuensi tersebut sering muncul yaitu “seller”, “kirim cepat”, “ramah”, “mantap”, “bagus,” “proses cepat”, dan lain-lain. Kata-kata tersebut menandakan bahwa pembeli lebih sering mengungkapkan kepuasan terhadap layanan penjual ketimbang membahas kualitas produk motor itu sendiri. Hal tersebut juga menunjukkan bahwa situs laman jual motor online yang ditarik ini merupakan situs-situs penjual atau *dealer* sepeda motor yang dapat dipercaya untuk transaksi secara daring.

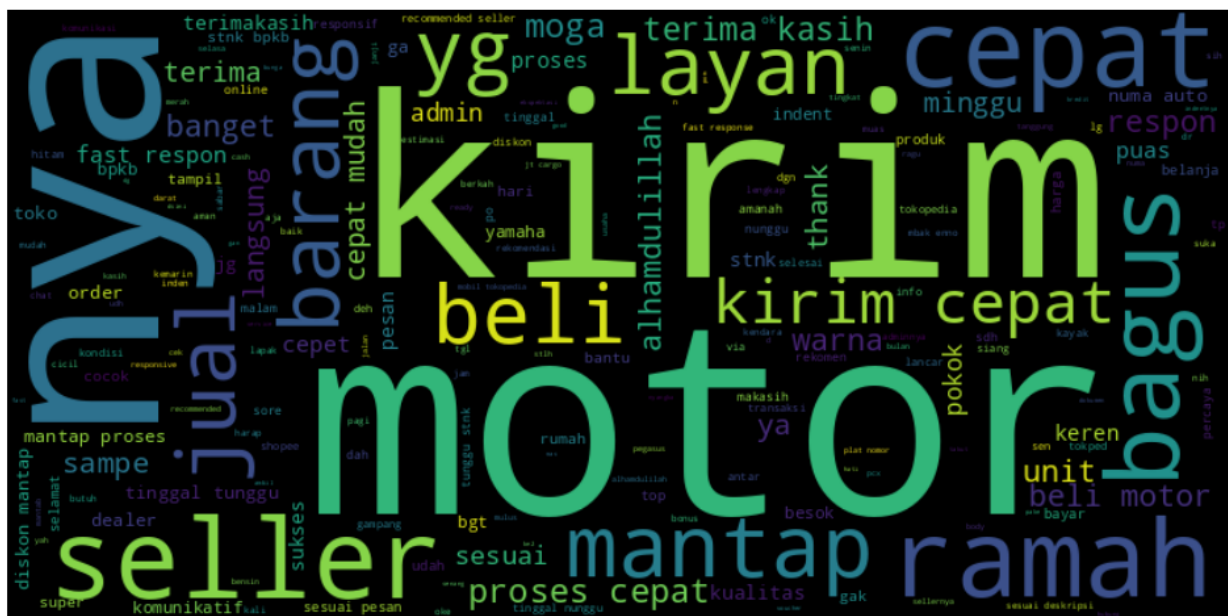
4.2. Buatlah *wordcloud* untuk komentar pada rating 3 ke bawah yang didapatkan setelah dilakukan *pre-processing* data dan jelaskan hasil yang didapatkan!



Pada komentar yang diberikan pembeli yang memberikan nilai ulasan 3 ke bawah, pembeli merasa tidak puas dengan hasil pembelian yang ditandakan dengan kata kunci “kendala” dan “kirim” yang cukup sering muncul. Artinya, pembeli yang merasa tidak puas dan memberi ulasan tersebut pasti mengalami suatu kendala baik dari produk itu sendiri atau disebabkan oleh pengirimannya. Beberapa kata yang mengandung sentimen buruk lain yang frekuensinya lebih sedikit ada seperti pada kata “jelek”, “salah”, “aneh”, dan “keras” yang bisa merujuk kepada kualitas dari produk motor itu sendiri.

4.3. Buatlah *wordcloud* untuk komentar pada rating 4 ke atas yang didapatkan setelah dilakukan *pre-processing* data dan jelaskan hasil yang didapatkan!





Karena dataset kebanyakan dan umumnya terdiri dari pembeli yang memberikan nilai ulasan 5 atau sempurna dan merasa puas, sehingga analisis *wordcloud* untuk komentar dengan ulasan 4 ke atas memiliki hasil analisis yang kurang lebih serupa dengan *wordcloud* pada nomor **4.1**. Kata kunci kepuasan pembeli ada pada kata “cepat”, “mantap”, “ramah”, dan “bagus” yang menandakan kepuasan pembeli terhadap pelayanan penjual. Sehingga kunci dari penjualan produk motor ini sendiri terdapat pada bagaimana penjual dapat membangun kepercayaan dengan pembeli dengan memberikan pelayanan berkualitas dan cepat.