

# Scraping data mamikos dari twitter

Keyword

1. Mamikos
2. Mamikost
3. Mami kos
4. Mami kost

# Hasil scraping "mamikost" 9 data

4]: df\_mamikost

5	mamikost, filter area dan opsi ca...	2021-09-17 03:03:50	1438700194283995141	zaa	kamualexya	1410840461044903940	Slaman, Indonesia	None	setengah kemenangan
6	@enzaaofc_ Jakselnya dmn kak? Mamikost coba ad...	2021-09-16 15:46:46	1438529804014215170	Vinjo	Vinjo09270500	1395542809164148738		None	hi
7	Mamikost kepahe setelah gw delete appnya, tiba...	2021-09-15 15:00:55	1438155879224803329	KAL	arrInteresting	2743559767		None	
8	@Detatika Eh ini yg dipake di mamikost bukan s...	2021-09-14 03:15:04	1437615854871408645		smir_iam	1314331849380950021		None	
9	@Askrlfess cari di mamikost	2021-09-13 01:06:03	1437221002728001536	kdj🌟	kindpers0n	325911324		None	?????

# Hasil scraping “mamikos” 142 data

[6]: df\_mamikos

[6]:

	text	created_at	tweet_id	user_profile_name	username	user_id	user_location	user_url	user_d
0	Rekomendasi kosan hari ini adala Kos Parman Mi...	2021-09-21 05:00:00	1440178979340759040	Mamikos	mamikosapp	3460456292	Yogyakarta, Indonesia	https://t.co/qXHQsQYpga	Aplil Anak No. 1 Indone Satu a
1	Haloo Ges, rekomendasi kosan hari ini adalah ...	2021-09-21 03:00:01	1440148783149420555	Mamikos	mamikosapp	3460456292	Yogyakarta, Indonesia	https://t.co/qXHQsQYpga	Aplil Anak No. 1 Indone Satu a
2	@ngalamfess Cek mamikos aja nder banyak bgt wk...	2021-09-21 02:39:43	1440143674940157954	nam	namminnmi	1624506618	20 t kpop and weeb behavior ⚠	None	cli
3	Coba dulu maba uda tau apk mamikos	2021-09-21 01:26:11	1440125169440215050	Nidoo	nidanido19	1255662640195502080		None	

# Hasil scraping “mami kos” 171 data

8]: df\_mami\_kos

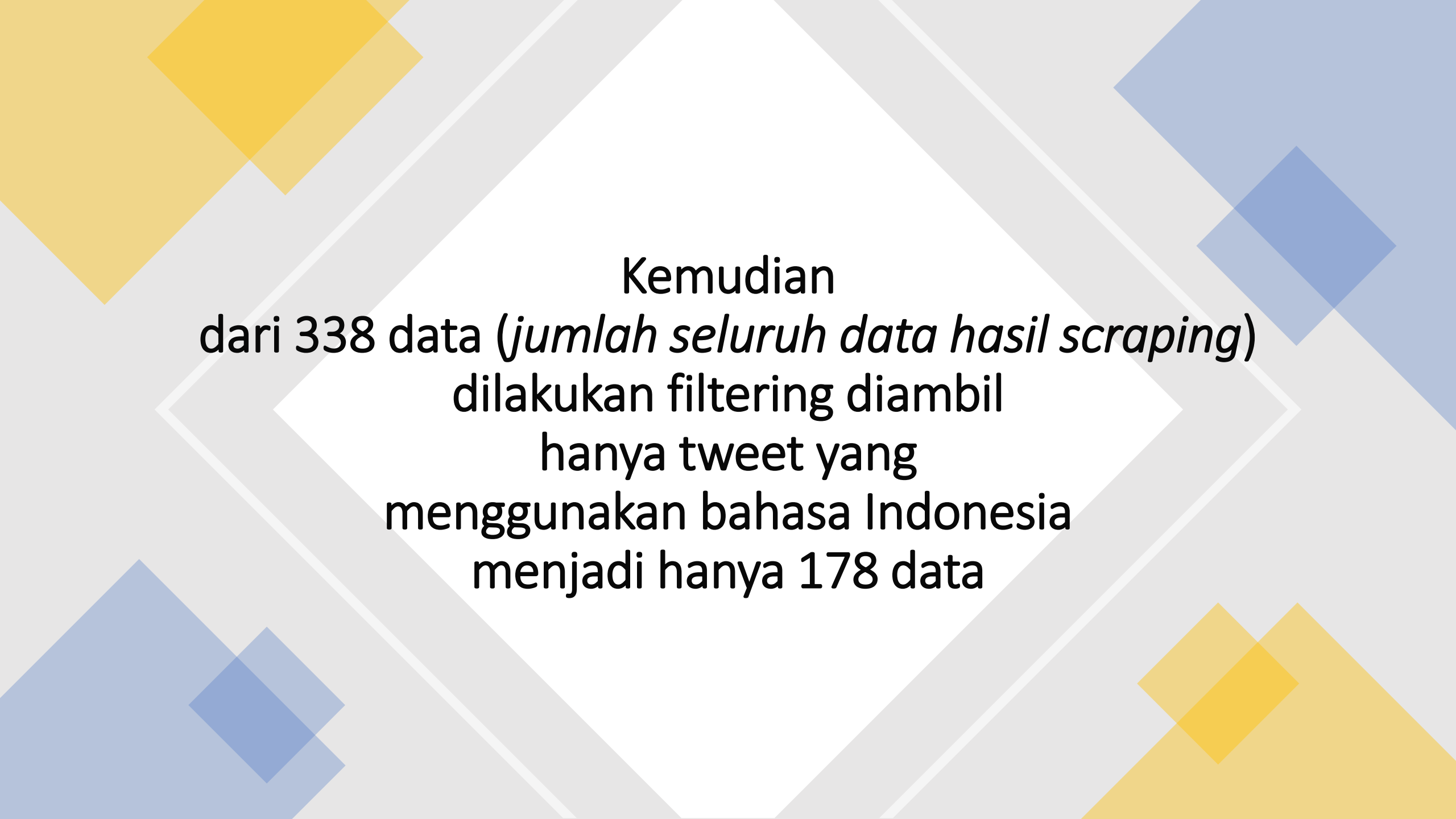
166	@keniaos mami manda la ubicación ya tienes mi ...	2021-09-12 20:25:20	1437150358011367424	DIOSA_KOS	DIOSA_KOS	1307881703982104581	KENIA KE	<a href="https://t.co/ICF8U2TzrN">https://t.co/ICF8U2TzrN</a>	actualizada de inform s
167	@Diego_KOs Obvio que la opinión de su mami cue...	2021-09-12 19:55:31	1437142852211990540	Alejandra ❤️.	unakeninifeliz	1354199705656373253	Durango, México	None	
168	tabangan jud kos mga boang oh founder jud nis ...	2021-09-12 17:21:07	1437103994380251139	prollie	igopnehs	1266721922101809152	PAGSULAY RANI	None	lahi ang Y
169	@keniaos AAAAAAA Mami la vas a romper &lt;33	2021-09-12 16:57:35	1437098073856090117	Emily ♥'s biyin	biyin_KOs	1277451964041900032	s!her ♦ EC	<a href="https://t.co/UidrWX7QZO">https://t.co/UidrWX7QZO</a>	🏠- Biyin, K
170	Keinget pas itu. Mamiku mau ke kos, trus aku d...	2021-09-12 16:37:33	1437093033808252928	d_αβ 🐼	winsknoyjy	895684632921784321	she/her. 99.	None	juykyurenwii ♦ genshin - v for g

171 rows x 17 columns

# Hasil scraping "mami kost" 15 data

df\_mami\_kost

12	ayok lagi ❤️ ❤️ awalnya sebel maca...	2021-09-13 16:13:08	1437449276553003010	아밍(茗)🐱	crispypork27	955035158121398273	Nurhalimah	<a href="https://t.co/AADXDPAlCs">https://t.co/AADXDPAlCs</a>	capit   Gm   cats lovers ❤️   Sing...
13	@hrdbacot Mami kost segera daftar akun linked ...	2021-09-13 15:19:37	1437435808089186308	yamet kudasi	ara_arakimochi	1280902018383532032	None		cek inbox gan. cita- cita big win di tw: makji,...
14	Ada yg pernah sewa kost lewat mami kost tidak?	2021-09-13 10:46:56	1437367185656057856	🌸	floweryliy	1301601226052624384	minki's	None	i love baso ikan
15	jog ada cewe yang kost di bantul ga ya? kalo a...	2021-09-12 22:50:37	1437186919226429440	BASE JOGJA /delvote on	jogmfs	420821783	RULES CEK LINK	<a href="https://t.co/5qEhswH8lv">https://t.co/5qEhswH8lv</a>	— AUTO BASE untuk membahas tentang Jogja dan b...



Kemudian  
dari 338 data (*jumlah seluruh data hasil scraping*)  
dilakukan filtering diambil  
hanya tweet yang  
menggunakan bahasa Indonesia  
menjadi hanya 178 data

# Preprocessing -> Tokenisasi -> dipisah per kata

```
df['tokenized_sents'] = df.apply(lambda row: nltk.word_tokenize(row['text']), axis=1)
```

```
df.loc[:, "tokenized_sents"]
```

```
0      [rekomendasi, kosan, hari, ini, adala, kos, parman, minara, bandara, soekarno, hatta, taman, mahkota, benda, tangerang,
912pa, kosa..., https, :, //t.co/7xfgwzwsjr]
1      [haloo, ges, ,, rekomendasi, kosan, hari, ini, adalah, kos, pegangsaan, tipe, a, kelapa, dua, jakarta, utara, kosannya,
sudah, isian..., https, :, //t.co/h6agqi10go]
2      [@, ngalamfess, cek, mamikos, aja, nder, banyak, bgt, wkwk, :, (]
3      [coba, dulu, maba, uda, tau, apk, mamikos]
4      [haloo, ges, ,, rekomendasi, kosan, hari, ini, adalah, kos, hans, sungai, bambu, tanjung, priuk, jakarta, utara, kosanny
a, sudah, isia..., https, :, //t.co/8lhjqtbovn]
...
173     [@, um_fess, cek, di, apk, mami, kost, coba]
174     [@, ntet__, haha, ayok, lagi, ❤️❤️, awalnya, sebel, macam, di, buat, anak, mami, sama, security, nya, kost, aku, ternya
ta, wkwkwkwk, ada, udang, di, balik, udang, 🤔]
175     [@, hrdbacot, mami, kost, segera, daftar, akun, linked, wkwk]
176     [ada, yg, pernah, sewa, kost, lewat, mami, kost, tidak, ?]
177     [jog, ada, cewe, yang, kost, di, bantu, ga, ya, ?, kalo, ada, rep, dong, kalo, cocok, mau, kost, disitu, jg, nyari, di,
mami, kost, gada, yang, cocok, :, (]
Name: tokenized_sents, Length: 178, dtype: object
```

# Menghapus StopWords

*(Kata/emot/tanda baca yang tidak perlu)*

```
df["clean_text"]
```

```
0      [rekomendasi, kosan, adala, kos, parman, minara, bandara, soekarno, hatta, taman, mahkota, benda, tangerang, https]
1      [haloo, ges, rekomendasi, kosan, kos, pegangsaan, tipe, kelapa, jakarta, utara, kosannya, https]
2      [ngalamfess, cek, mamikos, aja, nder, bgt, wkwk]
3      [coba, maba, uda, tau, apk, mamikos]
4      [haloo, ges, rekomendasi, kosan, kos, hans, sungai, bambu, tanjung, priuk, jakarta, utara, kosannya, https]
...
173     [cek, apk, mami, kost, coba]
174     [haha, ayok, sebel, anak, mami, security, kost, wkwwkwwk, udang, udang]
175     [hrdbacot, mami, kost, daftar, akun, linked, wkwk]
176     [yg, sewa, kost, mami, kost]
177     [jog, cewe, kost, bantul, ga, kalo, rep, kalo, cocok, kost, disitu, jg, nyari, mami, kost, gada, cocok]
Name: clean_text, Length: 178, dtype: object
```



# Wordcloud

## 1. Gabungkan Semua kata menjadi satu

'rekomendasi kosan adala kos parman minara bandara soekarno hatta taman mahkota benda tangerang haloo ges rekomendasi kosan ko s pegangsaan tipe kelapa jakaa utara kosannya ngalamfess cek mamikos aja nder bgt coba maba uda tau apk mamikos haloo ges rek omentasi kosan kos hans sungai bambu tanjung priuk jakaa utara kosannya syauur cari mamikos cin rekomendasi kosan kos blessing house tipe suko manunggal surabaya kosannya haloo ges rekomendasi kosan kos bintaro plaza pondok aren tangerang selatan kosann ya haloo ges rekomendasi kosan kos campur jakaa pusat namanya kos brantas gambir jakaa undipmenfess kamar kosong kost go gree n banyumanik semarang cek review amp survei kost online via viual info kos deket its upn kamar mandi dalem max drop cari mamik os ga nemu be mamikos membikin orangtua suudzon info cowok campur jogja budget maks udah internetnya gak indihome karawangfe ss cek aja mamikos ythsangyeon ngeledak kakak jeje yaa kakak aya gigit udah yg pny kontaknya kos cowo wisma jalan timoho tlg r ep dongg ntr ak dm nomer yg mamikos gadibales makasii ubsansfess deh kayanya kamar mandi coba cari mamikos aja nder cek mamiko s aja nder senjasungjin akuu udh nemu mamikos bru liat nnti minggu dipposauruss cek mamikos coba ca sdf info kos putri daerah waru budget plis yg parkir mobil kamar mandi dalem cari mamikos udah full justhiewra malfaiteurr bismillah bu ginaa butuh inf o kos mamikos info kost putri nyari mamikos susah deket ub kalo gaiss cek mamikos aja bang percaya mamikos mah nyediain kosan ngurus utbk lu nape liat info sbm mamikos fathiathau mamikos cenayang utbk pleaseeee mamikos tu cari kos gaada utbk harga ko s deket ugm tunggu info resmi beanyarl pengalaman cari kos mamikos nampaknya bagus engga terimacakihh jogmfs mamikos tu app web deh jogmfs coba cek mamikos biar daerahnya cek manual biar murah cek percaya mamikos mah nyediain kosan ngurus utbk jog okto ber jogja rencana sih offline kuliahnya tpi ttp aja suruh offline ga offline hrus adik adik tahun utbk menggurui sayangku perc aya info percaya mamikos mah nyediain kosan ngurus utbk hehhh tu website mamikos jgn prcaya dl klo ltmtpt yg baumulutt gws y g percaya berita utbk dr ltmtpt tp mamikos undipmenfess mamikos nomor pemiliknya mamikos emang sesat stanmenfess si tergantung jakaa fasilitasnya seadanya cek mamikos aja yg kos kontak pemilik kosnya kalo reply dm mamikos percaya mamikos mah nyediain k osan ngurus utbk jog oktober jogja kuliah offline tpi tekan saiki sek bingung luru kos kating angel annogerson kalo biaya hid up sih relatif mas yg aja kalo harga yg scrolling mamikos asik scrolling ig kos harian temen niatnya buka mamikos buka tokped anjir coba cek mamikos nder tau coba cek mamikos atur filter kost harian adambahduqun cek mamikos gw dapet kost dr mamikos co ba cari aplikasi mamikos deh bun jeongnindy mamikos tuh aplikasi bandeel iyaa mamikos masang iklan mamikos repotttnyaaa ampunnn aslii kosan deket rs bethesda ga bingung nyari mamikos ga deket undipmenfess mamikos sih nunggu temennya diboletin gw sampe udh



# Sentimen Analisis

1. Hapus Kata yang tidak perlu seperti *undipmenfess*, *https*, *jogmfs*, dll

```
df = (pd.read_feather("dataset/data_kalimat_clean.feather").loc[:, ["text", "kalimat_clean"]])
df["kalimat_clean"] = df["kalimat_clean"].str.replace("rt", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("sbmptnfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("tegalfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("https", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("sbyfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("keniaos", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("wkwk", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("collegemenfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("que", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("melankos", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("jogmfs", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("undipmenfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("ngalamfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("ubsansfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("karawangfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("stanmenfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("schfess", "")
df["kalimat_clean"] = df["kalimat_clean"].str.replace("utbkfess", "")
df = df.drop(19, axis="rows").reset_index(drop=True)
```

# Karena menggunakan TextBlob, data harus ditranslate ke bahasa inggris

```
df["kalimat_en"] = ""
for i in range(len(df["kalimat_clean"])):
    df["kalimat_en"][i] = GoogleTranslator(source="auto", target="en").translate(df["kalimat_clean"][i])

df['sentiment'] = df['kalimat_en'].apply(lambda tweet: TextBlob(tweet).sentiment.polarity)
```



# Hasil sentiment analysis menggunakan TextBlob

	text	kalimat_clean	kalimat_en	sentiment	sentiment_label
0	rekomendasi kosan hari ini adala kos parman minara bandara soekarno hatta taman mahkota benda tangerang 912pa\n\nkos... <a href="https://t.co/7xfgwzwsjr">https://t.co/7xfgwzwsjr</a>	rekomendasi kosan adala kos parman minara bandara soekarno hatta taman mahkota benda tangerang	The recommendation for boarding houses is Parman Minara boarding house Soekarno Hatta Airport, Taman Mahkota Benda, Tangerang	0.00	netral
1	haloo ges, rekomendasi kosan hari ini adalah kos pegangsaan tipe a kelapa dua jakarta utara\n\nkosannya sudah isian... <a href="https://t.co/h6agqi10go">https://t.co/h6agqi10go</a>	haloo ges rekomendasi kosan kos pegangsaan tipe kelapa jakaa utara kosannya	hello guys, recommendations for the boarding house for the Kelapa Jaka type North Jakarta boarding house	0.00	netral
2	@ngalamfess cek mamikos aja nder banyak bgt wkwk :(	cek mamikos aja nder bgt	check mamikos, it's great	0.80	positif
3	coba dulu maba uda tau apk mamikos	coba maba uda tau apk mamikos	try mamikos already know apk mamikos	0.00	netral
4	haloo ges, rekomendasi kosan hari ini adalah kos hans sungai bambu tanjung priuk jakarta utara\n\nkosannya sudah isia... <a href="https://t.co/8lhjqtbovn">https://t.co/8lhjqtbovn</a>	haloo ges rekomendasi kosan kos hans sungai bambu tanjung priuk jakaa utara kosannya	hello ges recommendation for Hans boarding house, bamboo river, cape priuk, jakaa north of his boarding house	0.00	netral
...	...	...	...	...	...
172	@um_fess cek di apk mami kost coba	cek apk mami kost coba	check the apk mami boarding try	0.00	netral
173	@ntet__ haha ayok lagi ❤️❤️ awalnya sebel macam di buat anak mami sama security nya kost aku ternyata wkwwkwkwk ada udang di balik udang 🤔	haha ayok sebel anak mami security kost udang udang	haha, let's be annoying, mami's child security, prawn boarding house	-0.30	negatif
174	@hrdbacot mami kost segera daftar akun linked wkwk	hrdbacot mami kost daftar akun linked	hrdbacot mami boarding account list linked	0.00	netral
175	ada yg pernah sewa kost lewat mami kost tidak?	yg sewa kost mami kost	the one who rents the boarding house, mommy boarding house	0.00	netral
176	jog ada cewe yang kost di bantu ga ya? kalo ada rep dong kalo cocok mau kost disitu jg nyari di mami kost gada yang cocok:(	jog cewe kost bantu ga kalo rep kalo cocok kost disitu jg nyari mami kost gada cocok	Jog, a boarding house girl, can you help if the rep is suitable, the boarding house is there, looking for a boarding house mami, doesn't match	0.55	positif

# Hasil Perbandingan Sentimen

Netral: 124 data  
Negatif: 11 data  
Positif: 42 data

