INDONESIAN ABUSIVE AND HATE SPECH TWITTER By Firdaus Romandhanu

PENDAHULUAN

<u>Ujaran kebencian</u> (hate speech) bisa berarti tindakan komunikasi yang dilakukan oleh suatu individu atau kelompok dalam bentuk <u>provokasi</u>, hasutan, ataupun hinaan kepada individu atau kelompok yang lain. Ujaran kebencian biasanya menyangkut aspek ras, warna kulit, gender, cacat, orientasi seksual, kewarganegaraan, agama dan lain-lain.

Pada umumnya ujaran kebencian sangat banyak di temukan di media sosial. Menurut pakar medsos mengatakan ujaran kebencian, caci maki, hinaan dan hoaks terjadi paling tinggi membanjiri ruang-ruang media sosial pada gelaran pemilihan presiden 2019. Saat ini ujaran kebencian masih banyak ditemukan di sejumlah platform media sosial. Selain itu, menurut beliau Twitter menjadi medsos yang paling banyak ujaran kebenciannya.

RUMUSAN MASALAH

- 1. Seberapa banyak tweet yang mengandung hate speech beserta abusive?
- 2. Berapa banyak kata kata kasar yang ada di tweet?
- 3. Kata apa saja yang banyak muncul?

TUJUAN

- 1. Mengetahui banyak tweet yang mengandung hate speech beserta abusive
- 2. Mengetahui banyak kata kata kasar yang ada di tweet
- 3. Mengetahui Kata apa saja yang banyak muncul
- 4. Cleansing Data



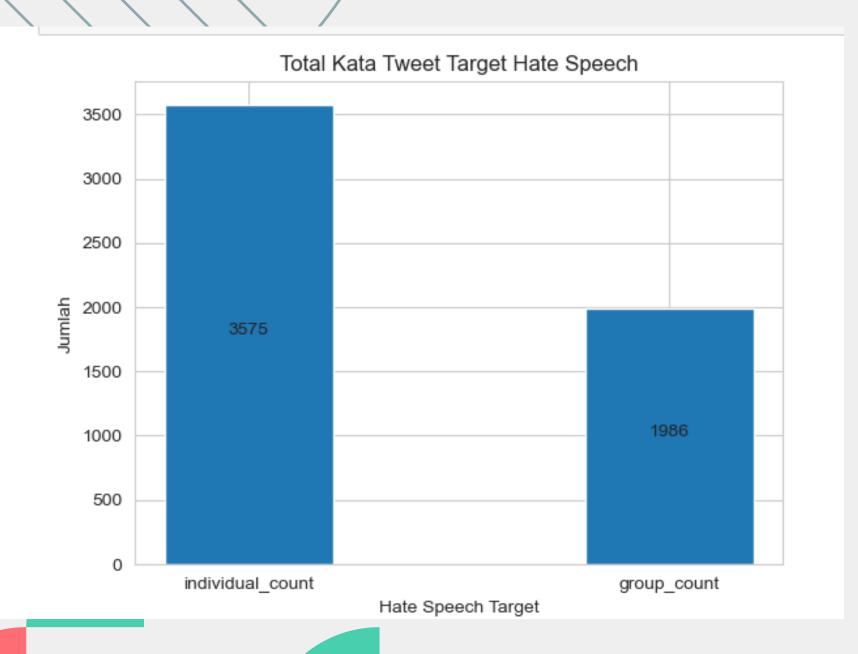
METODE PENELITIAN

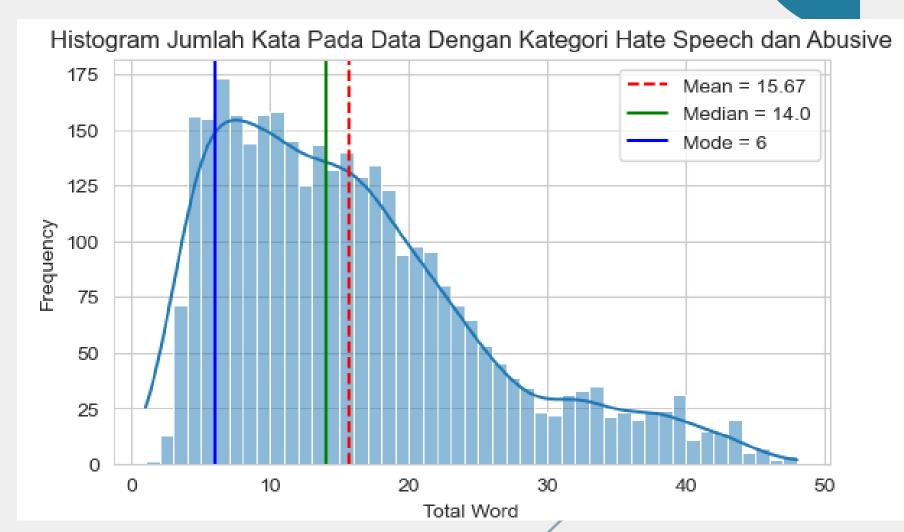
Metode Penelitian ini menggunakan data sekunder yang bersumber dari akun Kaggle Ilham Firdausi

Proses Cleansing Data

```
II no moderace
                                15109 HOH-HULL THOU
           12 HS Strong
                                13169 non-null int64
          dtypes: int64(12), object(1)
          memory usage: 1.3+ MB
In [6]:
         def cleansing data(tweets):
              tweets= tweets.lower()
              tweets= re.sub(r'&amp', ' ', str(tweets)) #menghapus &amp
tweets= re.sub(r'user', ' ', str(tweets)) #menghilangkan kata yang double
              tweets= re.sub(r'#\w+',' ', str(tweets)) #menghapus hastag
              tweets= re.sub(r'\\[0-9A-z]\{2,\}',' ', str(tweets)) # \\khusus
              tweets= re.sub(r'\\[0-9A-z]',' ', str(tweets)) #menghapus \n
              tweets= re.sub(r'[^a-zA-Z\s]',' ', str(tweets)) #menghilangkan huruf dengan karakter khusus
              tweets= re.sub(r'[^\w\s]',' ', str(tweets)) #menghilangkan tanda baca
              tweets= re.sub(r'\s+', '', str(tweets)) #menghapus spasi
              tweets= re.sub(r'\t+', '\t ', str(tweets)) #menghapus tab
              tweets= re.sub(r'RT[\s]+', ' ', str(tweets)) #menghapus RT
tweets= re.sub(r'rt[\s]+', ' ', str(tweets)) #menghapus rt
              tweets= tweets.strip()
              return tweets
```

PROSES ANALIS DATA





HASIL

- Terdapat 25 kata kasar yang paling umum di gunakan
- Kebanyakan Ujaran Kebencian Ditunjukan Ke Perseorangan

KESIMPULAN

 Tingginya Tingkat Abusive and Hate Speech: Analisis data menunjukkan adanya tingkat yang signifikan dari konten abusive dan hate speech di platform Twitter di Indonesia. Hal ini menunjukkan adanya masalah serius terkait bahasa kasar, penghinaan, pelecehan, dan kebencian yang diungkapkan dalam lingkungan online.

