

# Data Caleg Pemilu 2019

oleh:

Nur Aini Rakhmawati, Ahmad Choirun Najib, Eunike  
Andriani Kardinata, Nurrida Aini Zuhroh, Akhmad  
Irsyad, Herlambang Haryo Putro, Ashr Hafiizh Tantri

Pilihlah caleg bermutu dan merakyat

# Overview

Dalam rangka menyambut Pemilihan Umum Anggota Legislatif yang diadakan pada tahun 2019, masyarakat semakin gencar mencari informasi tentang calon perwakilan suara rakyat di pemerintahan. Berbagai macam analisis banyak dilakukan menjelang hari Pemilihan Umum. Tantangan yang dihadapi masyarakat sekarang bukan hanya mengenai penyebaran informasi, namun juga penyaringan informasi. Sangat mudah untuk membuat viral sebuah informasi, terutama yang kontroversial; tetapi sulit untuk menyaring informasi yang tidak benar atau hoax. Pemerintah dituntut untuk bersaing dengan penyebar hoax untuk memberikan klarifikasi. Pun setelah diberikan, ada kalanya klarifikasi tersebut tidak diterima oleh semua orang terlanjur mempercayai hoax.

Salah satu langkah yang dapat dilakukan untuk meminimalisir dampak negatif hoax adalah dengan memberikan pembekalan informasi yang akurat dan netral kepada masyarakat. Masyarakat juga dapat dilatih untuk mengamati indikator tertentu yang dapat dipercayai untuk membuat penilaian.

Oleh karena itu, kami mencoba mengambil dari KPU dengan teknik web scraping dan menganalisa hasil scraping untuk mendapatkan informasi lebih mendalam terkait kondisi calon legislatif baik DPR maupun DPRD. Proses analisa data yang kami sajikan dapat dilihat secara terbuka di setiap slide data yang kami tampilkan.

Semua kode dapat dilihat di halaman terakhir presentasi ini

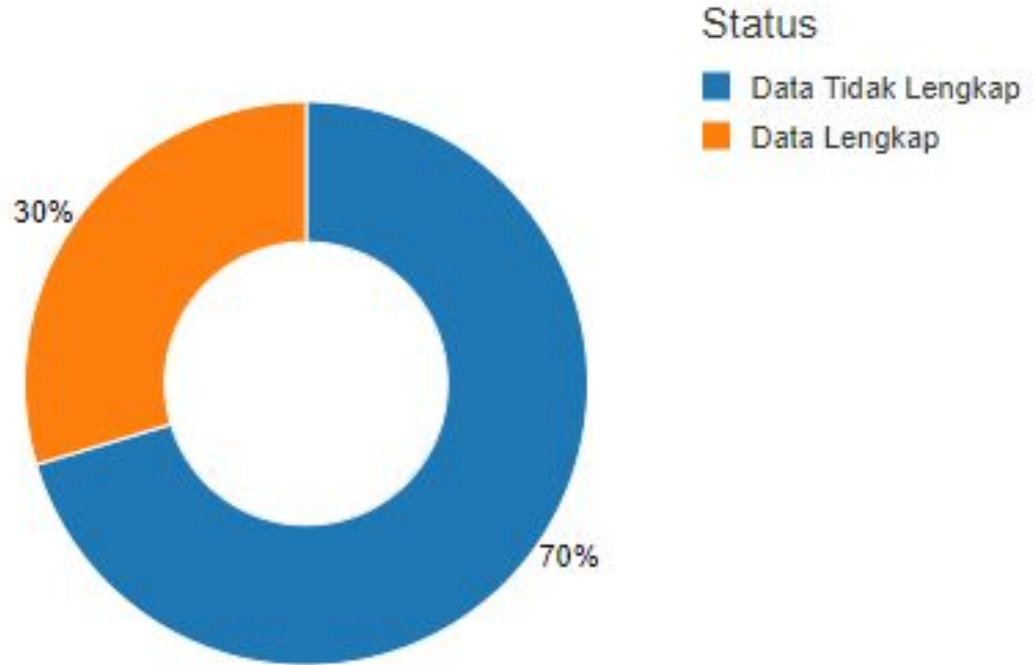
# Pengambilan data

1. Data diambil dari web KPU
2. Kode untuk akusisi data dapat dilihat  
<https://github.com/nurainir/caleg2019>
3. Ada 36000 data caleg yang terambil

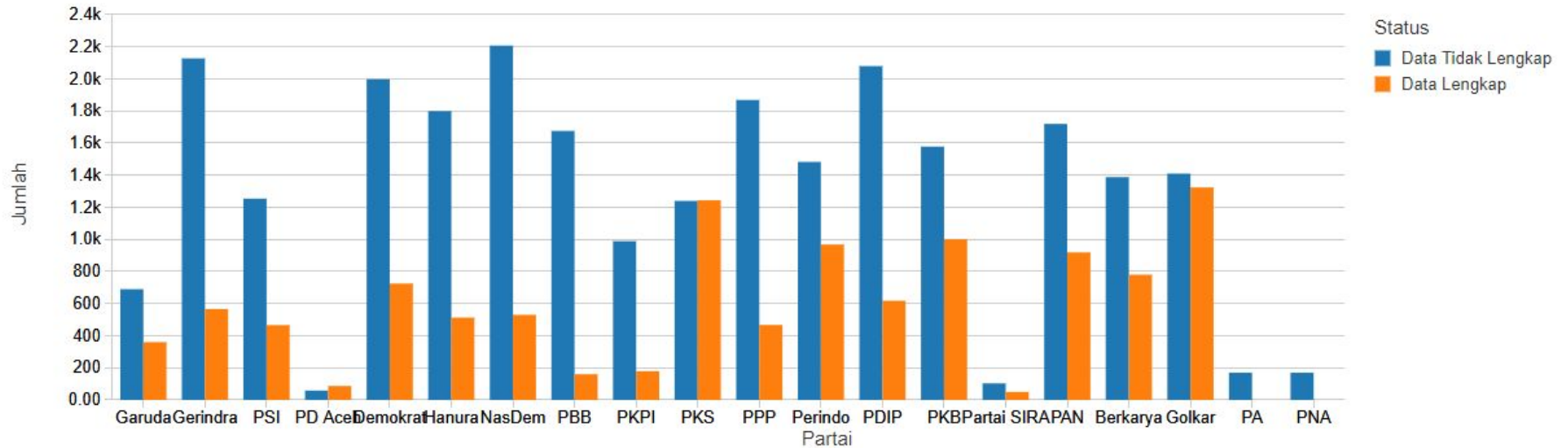
# Kelengkapan Data Caleg

Dari sekitar 36000 data caleg yang diproses, ditemukan sekitar

**26000** caleg  
*tidak melengkapi data*  
pada website resmi KPU

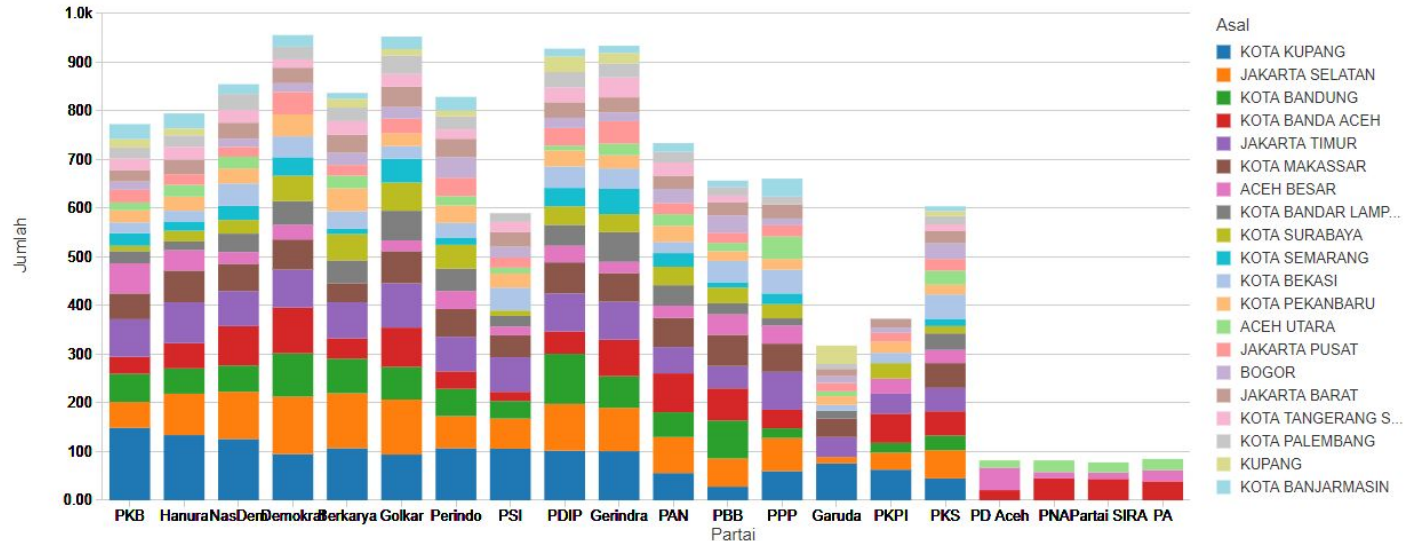


# Kelengkapan Data Caleg



**TIDAK ADA** partai yang seluruh data calegnya lengkap  
partai yang persentase data lengkap dan tidak lengkap kira-kira sebanding **HANYA 2**

# Partai Pengusung Caleg



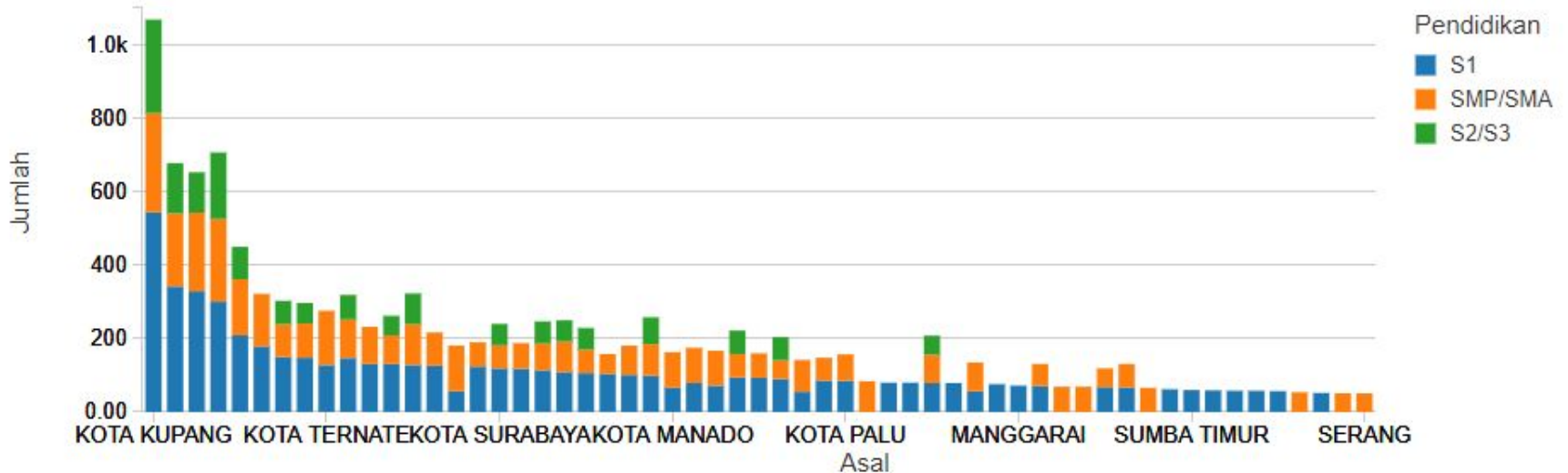
**SEMUA** partai bisa dikatakan **HETEROGENEOUS**

\* pengecualian hanya untuk partai lokal Aceh \*

# Daerah Asal Caleg

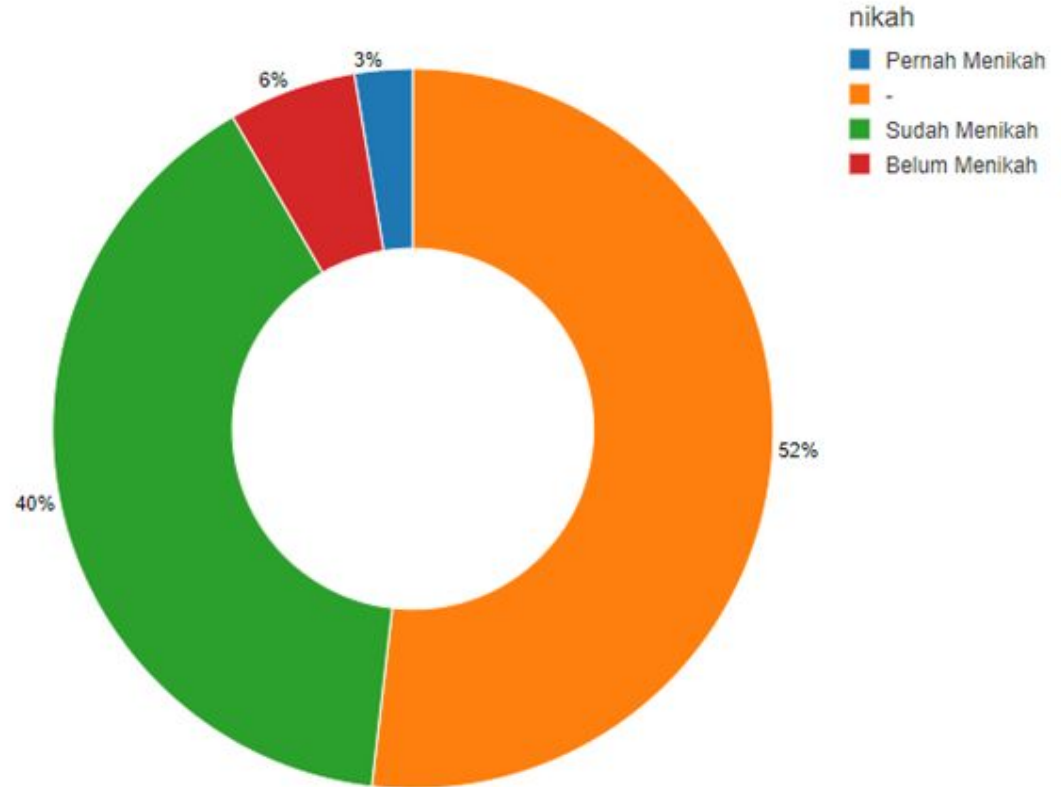
# KOTA KUPANG

adalah daerah penghasil caleg terbanyak tahun 2019

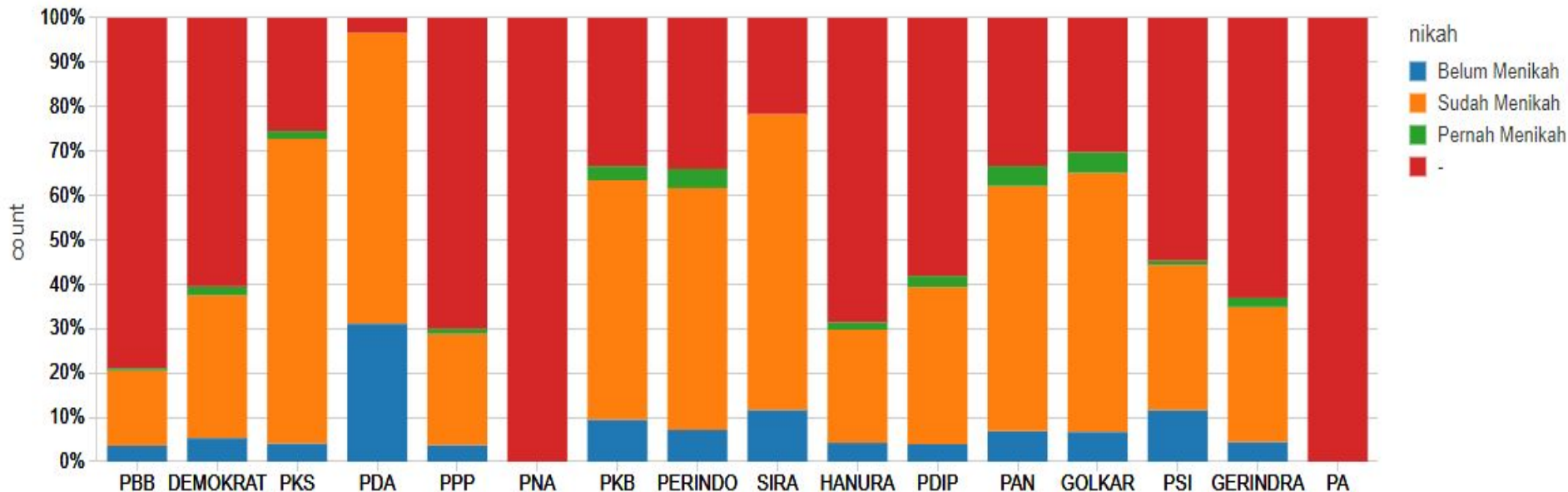


\* grafik hanya menampilkan kategori pendidikan dengan jumlah di atas 50 orang \*

6% atau 1390  
Caleg masih  
Single alias  
**JOMBLO**





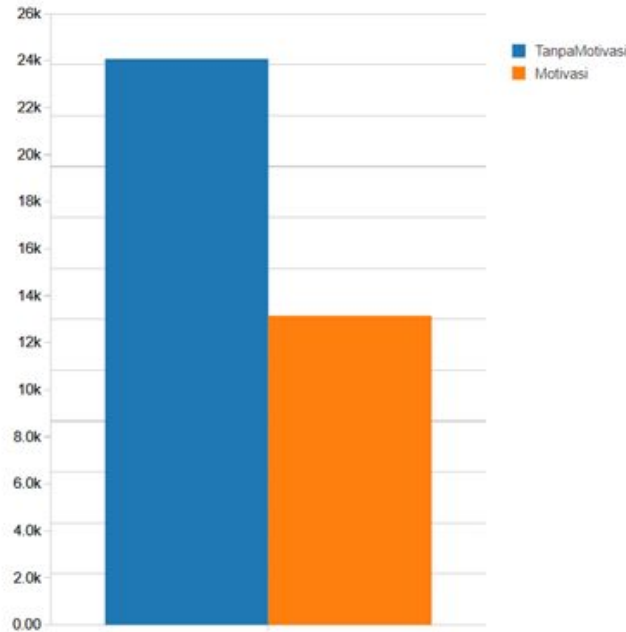


Partai Nangroe Aceh, Partai Aceh, PBB dan PPP merupakan 4 partai terbesar dengan rasio Caleg yang belum mengisi status nikah

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/8387562705802125/1631563187894747/751175171559615/latest.html>

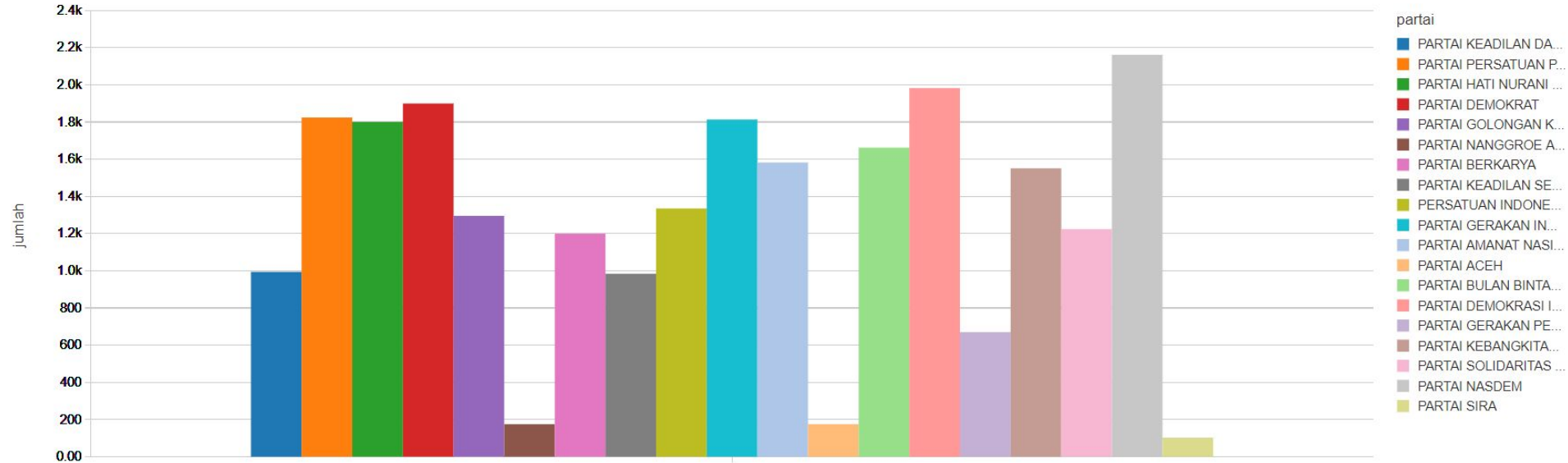
# Motivasi Caleg

Jumlah caleg yang menampilkan motivasi dan tidak menampilkan motivasi

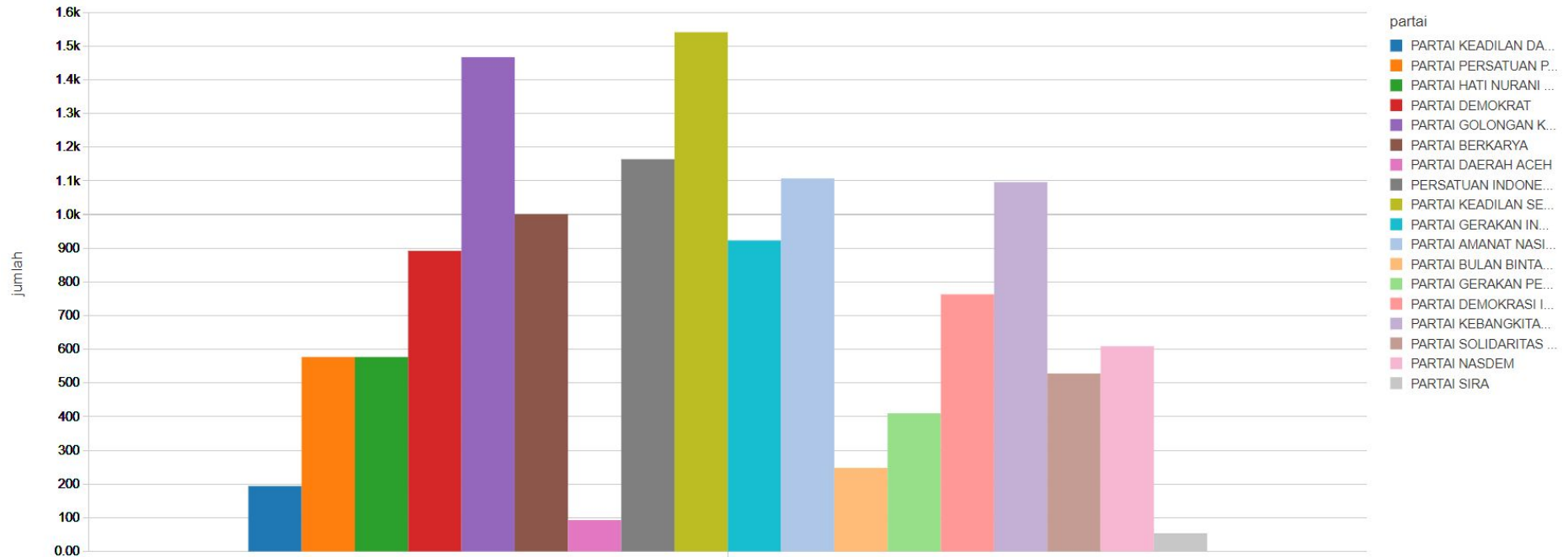


**54%** caleg yang menampilkan motivasinya

# Jumlah caleg yang tidak menampilkan motivasi berdasarkan partai



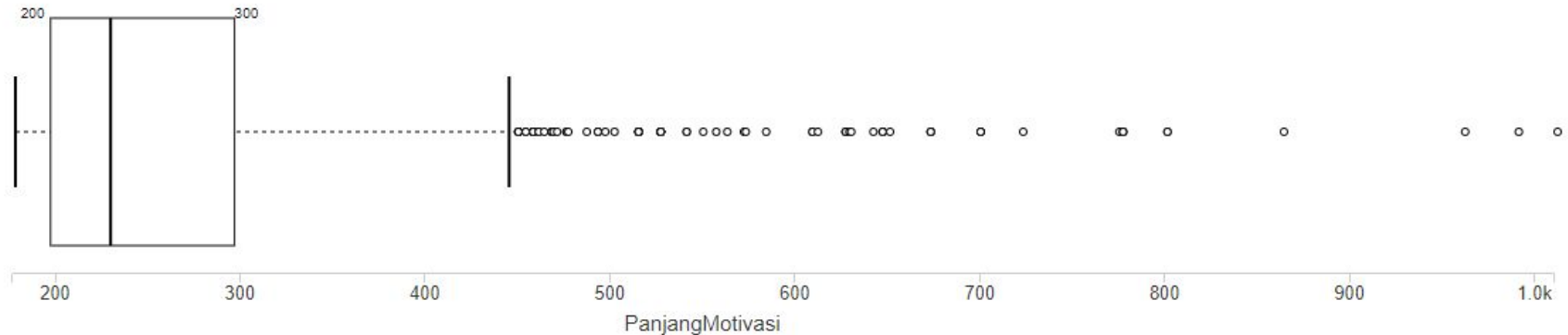
# Jumlah caleg yang menampilkan motivasi berdasarkan partai



# Boxplot Motivasi Keseluruhan

```
4 SELECT COALESCE(length('motivasi'), 0 ) AS PanjangMotivasi FROM `hasil_new_csv` WHERE `partai` IS NOT NULL AND length(`partai`) > 2 ORDER BY length('motivasi') DESC
```

▶ (1) Spark Jobs



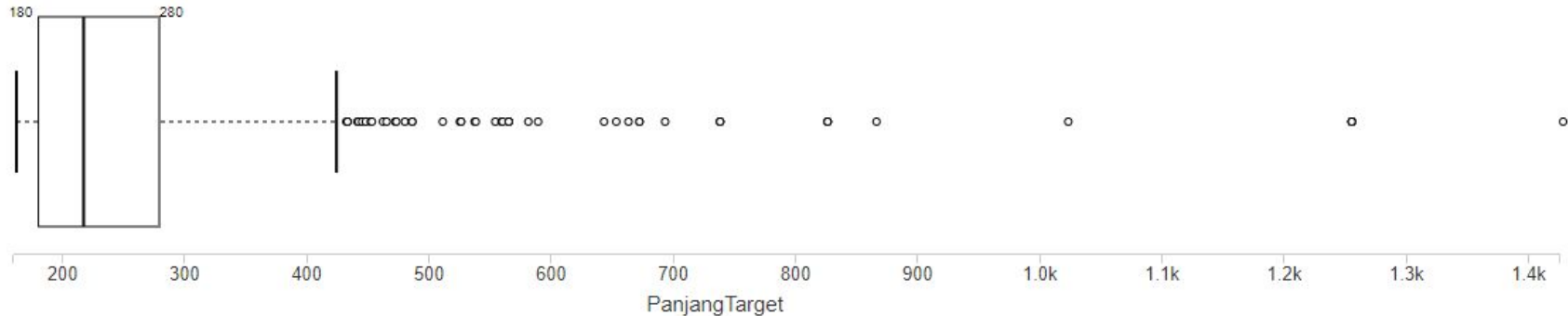
Showing sample based on the first 1000 rows.

Berdasarkan 1000 motivasi terpanjang, 75% caleg menuliskan motivasi di bawah 300 karakter.

# Boxplot Target Keseluruhan

```
4 SELECT COALESCE(length(`target`), 0 ) AS PanjangTarget FROM `hasil_new_csv` WHERE `partai` IS NOT NULL AND length(`partai`) > 2 ORDER BY  
length(`target`) DESC
```

► (1) Spark Jobs



Showing sample based on the first 1000 rows.

Berdasarkan 1000 motivasi terpanjang, 75% caleg menuliskan target di bawah 280 karakter.

# Pengaruh Panjang Motivasi & Target terhadap Pemahaman Caleg

Padahal, Menurut Steven O'Brien dan Vernenda P. Hamm Baugh, banyaknya tulisan yang ada akan mempengaruhi kemampuan seseorang dalam mendiskusikan topik secara panjang lebar.

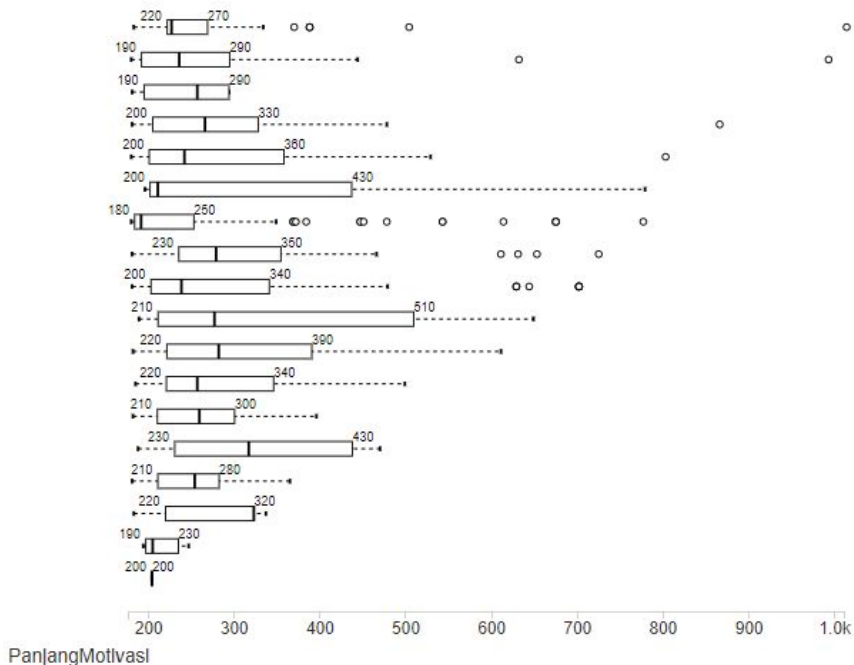
S. O'Brien dan V. P. H. Baugh, "Effects of minimum word counts on writing tasks," *Modern Psychological Studies*, vol. 19, no. 1, pp. 11-15, 2013.

# Panjang Motivasi Per Partai

```
6 SELECT COALESCE(length(`motivasi`), 0 ) AS PanjangMotivasi, `partai` FROM `hasil_new_csv` WHERE length(`partai`) > 2 ORDER BY length(`motivasi`) DESC
```

► (1) Spark Jobs

partai	PARTAI BERKARYA
	PARTAI DEMOKRASI INDONESIA PERJUANGAN
	PARTAI KEBANGKITAN BANGSA
	PARTAI DEMOKRAT
	PARTAI AMANAT NASIONAL
	PARTAI DAERAH ACEH
	PERSATUAN INDONESIA
	PARTAI SOLIDARITAS INDONESIA
	PARTAI GOLONGAN KARYA
	PARTAI GERAKAN PERUBAHAN INDONESIA
	PARTAI NASDEM
	PARTAI KEADILAN SEJAHTERA
	PARTAI HATI NURANI RAKYAT
	PARTAI KEADILAN DAN PERSATUAN INDONESIA
	PARTAI GERAKAN INDONESIA RAYA
	PARTAI PERSATUAN PEMBANGUNAN
	PARTAI BULAN BINTANG
	PARTAI SIRA





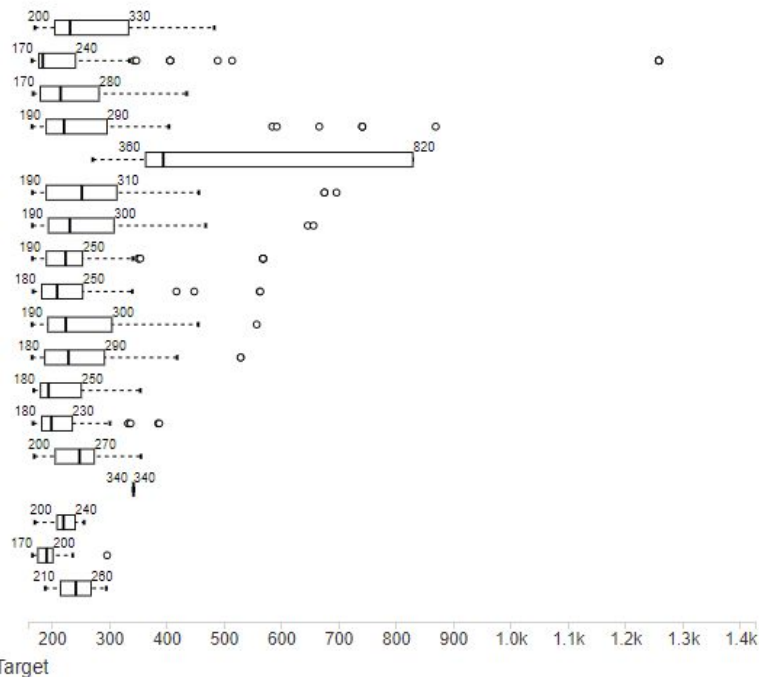
# Panjang Target Per Partai

```
6 SELECT COALESCE(length(`target`), 0) AS PanjangTarget, `partai` FROM `hasil_new_csv` WHERE length(`partai`) > 2 ORDER BY length(`target`) DESC
```

► (1) Spark Jobs

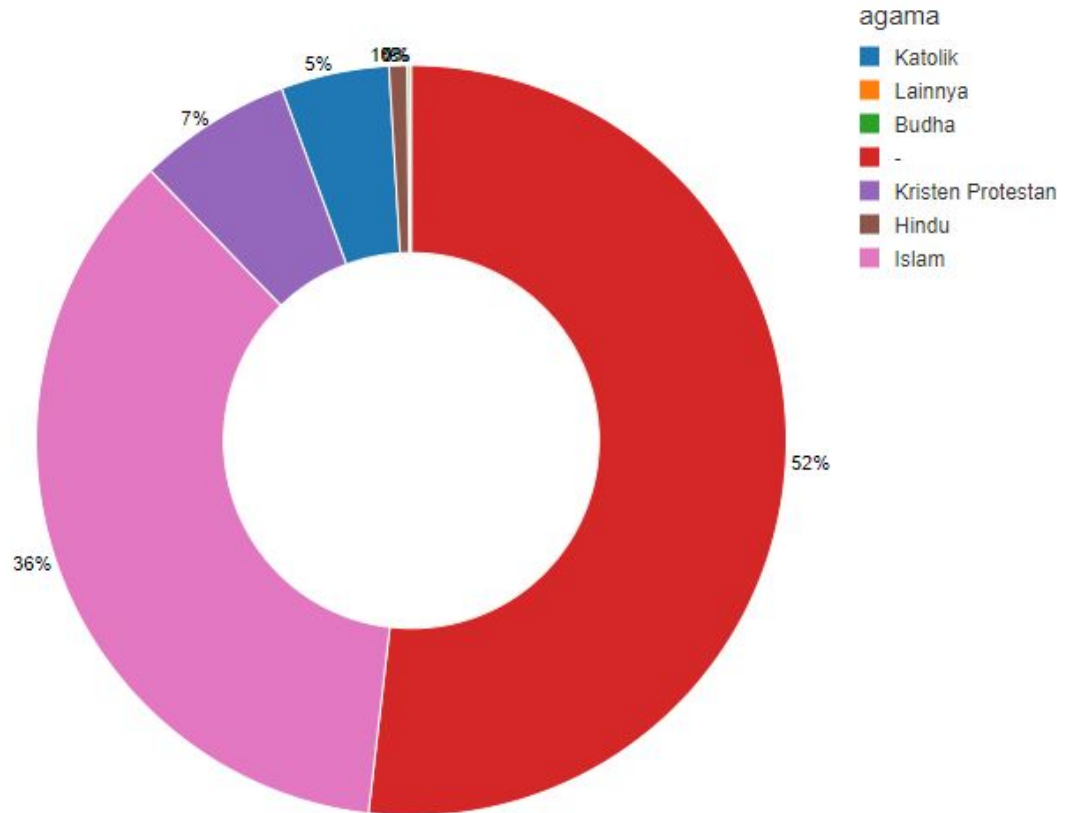
partai

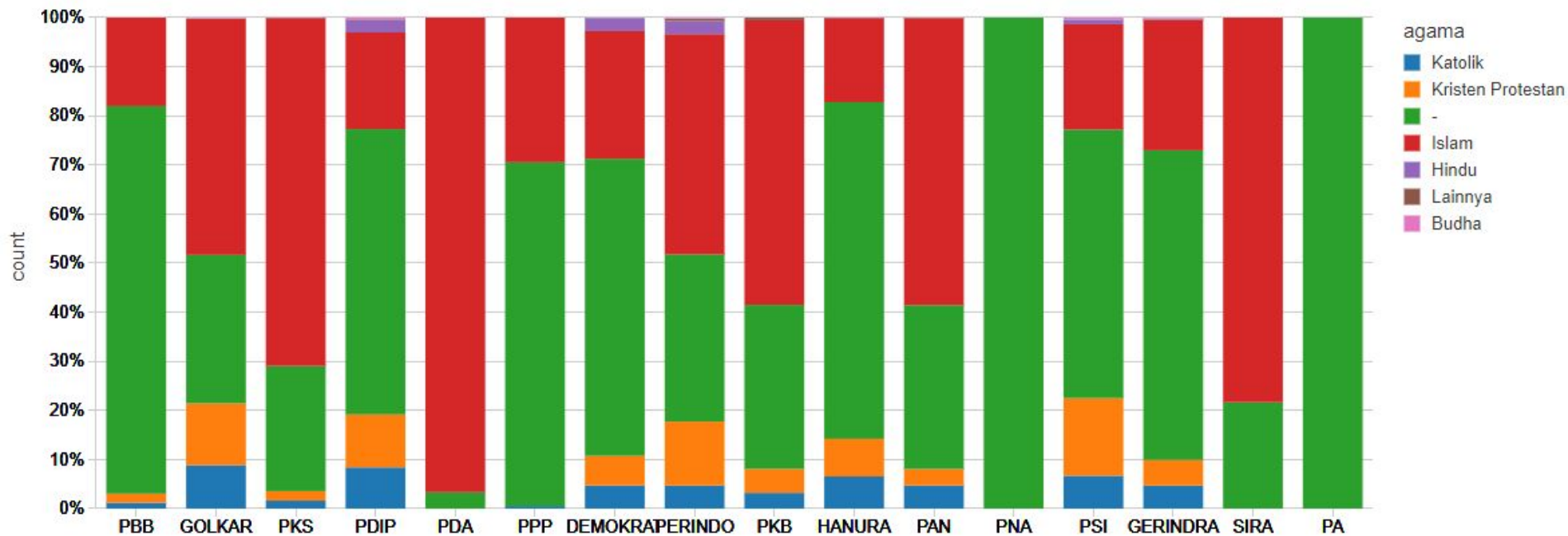
PARTAI GERAKAN PERUBAHAN INDONESIA
PARTAI DEMOKRAT
PARTAI GERAKAN INDONESIA RAYA
PARTAI DEMOKRASI INDONESIA PERJUANGAN
PARTAI DAERAH ACEH
PARTAI SOLIDARITAS INDONESIA
PERSATUAN INDONESIA
PARTAI NASDEM
PARTAI GOLONGAN KARYA
PARTAI AMANAT NASIONAL
PARTAI BERKARYA
PARTAI HATI NURANI RAKYAT
PARTAI Keadilan Sejahtera
PARTAI PERSATUAN PEMBANGUNAN
PARTAI SIRA
PARTAI Keadilan dan Persatuan Indonesia
PARTAI KEBANGKITAN BANGSA
PARTAI BULAN BINTANG



[illegible]

Agama merupakan salah satu field yang terindikasi cukup sensitif, sehingga sebanyak 52% Caleg tidak mengisi field Agama





Partai yang identik dengan Partai Islam tidak semua Calegnya beragama Islam. Partai PKS, PAN, PBB dan PKB yang identik dengan Partai Islam memiliki beberapa persen dari keseluruhan calegnya yang merupakan non-muslim.

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/8387562705802125/1631563187894747/751175171559615/latest.html>



Motivasi Caleg dari  
Semua Partai berbasis  
pendukung Islam

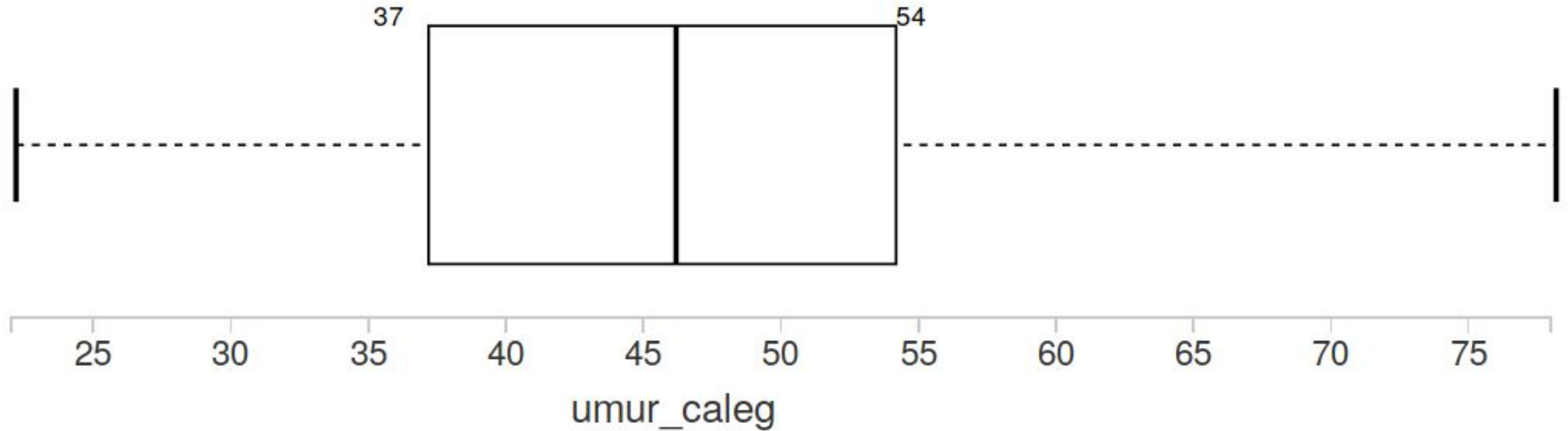


## Target Caleg dari Semua Partai berbasis pendukung Islam



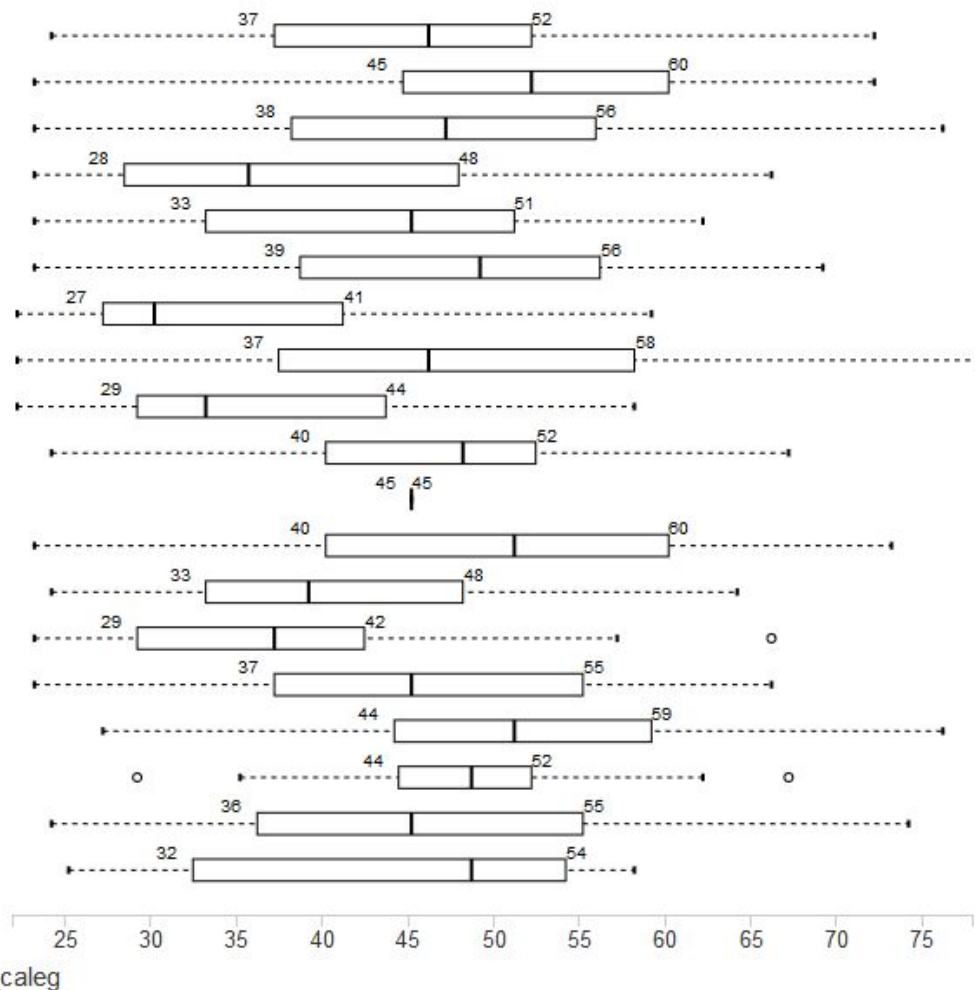
<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173b9c8/8387562705802125/1249899926353975/751175171559615/latest.html>

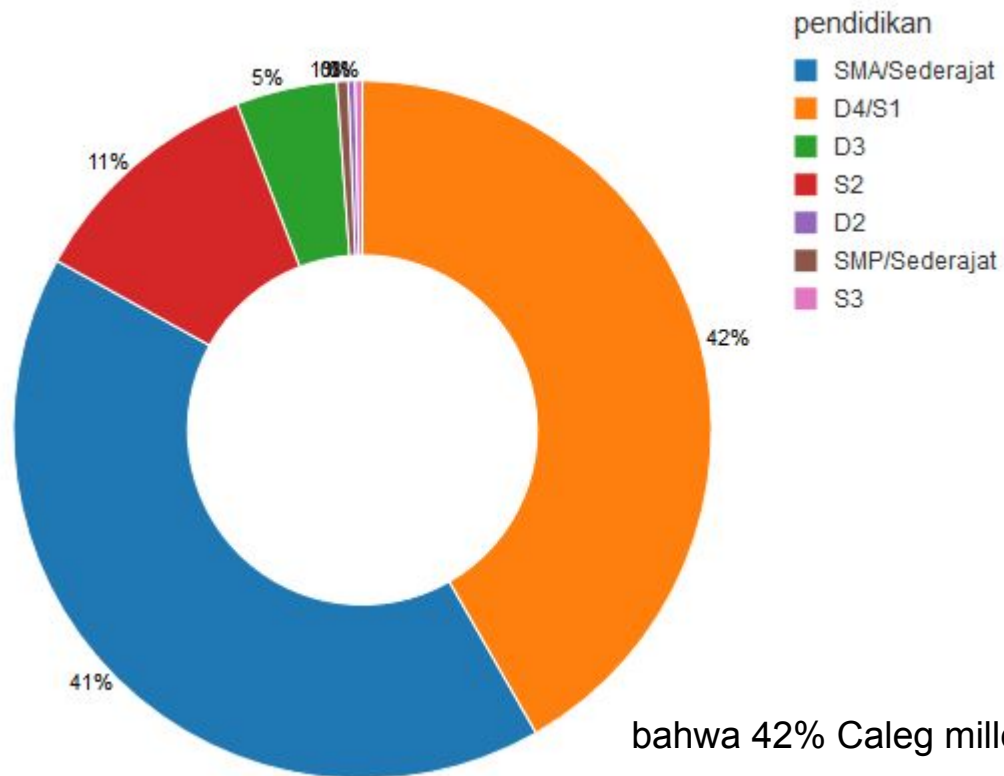
# Usia Caleg 2019



keseluruhan para caleg didominasi umur antara **37-54 tahun**

PARTAI AMANAT NASIONAL  
 PARTAI BERKARYA  
 PARTAI BULAN BINTANG  
 PARTAI DAERAH ACEH  
 PARTAI DEMOKRAT  
 PARTAI GERAKAN INDONESIA RAYA  
 PARTAI GERAKAN PERUBAHAN INDONESIA  
 PARTAI GOLONGAN KARYA  
 PARTAI Keadilan dan Persatuan Indonesia  
 PARTAI KEBANGKITAN BANGSA  
 PARTAI NANGGROE ACEH  
 PARTAI NASDEM  
 PARTAI SIRA  
 PARTAI SOLIDARITAS INDONESIA  
 PERSATUAN INDONESIA  
 PARTAI DEMOKRASI INDONESIA PERJUANGAN  
 PARTAI Keadilan Sejahtera  
 PARTAI PERSATUAN PEMBANGUNAN  
 PARTAI HATI NURANI RAKYAT





bahwa 42% Caleg millenials memiliki tingkat pendidikan D4/S1



# Motivasi Caley Millenials



# Koding dapat dilihat di:

1. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/6829307718073260/2350732941474656/3544324064342174/latest.html>
2. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/8387562705802125/1631563187894747/751175171559615/latest.html>
3. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/4623731963464209/4158765976397395/3906008156204768/latest.html>
4. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/1899304840311955/3935103582901914/6404188321450132/latest.html>
5. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/6259840299075061/3213044115266772/3320013622639166/latest.html>
6. <https://github.com/nurainir/caleg2019>