



Politus Analytics
www.politusanalytics.com

POLITUS TÜRKİYE PANORAMASI

Seçimlere Giderken Hangi Seçmen Ne Hissediyor ve Neyi Konuşuyor

2023 CUMHURBAŞKANLIĞI SEÇİMLERİ

Giriş

Bu rapor Cumhurbaşkanlığı Seçimi sürecine dair seçmen eğilimleri ve özellikleri hakkında 2023 yılı verilerden derlenen zamansal ve coğrafi analizler sunmaktadır.

Bulgular Twitter'dan elde edilen içeriklerin yapay zeka yöntemleri ile işlenmesi sonucunda elde edilmiştir. Analizlerin zaman aralığı hafta seviyesi, coğrafi birimi ise il seviyesidir.

Raporda yer alan seçmen parametreleri şu şekildedir:

- Duygu dünyası
- Siyasi konular
- İdeolojik eğilimler

Hakkımızda

Politus Analytics, Avrupa Araştırma Konseyi tarafından desteklenen ve Koç Üniversitesi bünyesinde yürütülen Politus Projesi kapsamında faaliyet yürütmektedir. Proje kamuoyundaki trendleri otomatik olarak ölçmek için yapay zeka ve büyük veri yöntemlerini kullanan bir veri platformu oluşturmaktadır.

Platform, Türkiye'deki temel siyaset, toplum ve piyasa eğilimleri hakkında temsili, yüksek frekanslı, gerçek zamanlı ve coğrafi detaylı panel veriler sağlamaktadır. Politus Analytics, sosyal medyadaki dijital izleri doğal dil işleme, ağ analizi ve derin öğrenme yöntemleri ile otomatik olarak analiz ederek tüketici ve seçmen eğilimlerinin, ideolojilerin, inançların ve değerlerin, toplumsal gündemde önem verilen konuların değişimi gibi kamuoyu dinamikleri hakkında ölçümler yapmaktadır. Bu ölçümler, coğrafi olarak konumlandırılmıştır. Böylece veritabanı bölgesel ve ulusal düzeyde analizlere izin verirken yaş, cinsiyet gibi temel demografik değişkenlere göre de ayrıştırma olanağı sağlar.

Politus Analytics, sosyal medya platformlarından topladığı verileri etik ve veri mahremiyeti açısından uyumlu yapay zeka yöntemleri ile incelemektedir.



Yöntem

Rapordaki bulgular Twitter API yardımıyla toplanan kullanıcı ve içerik bilgilerinin uygun Doğal Dil İşleme ve Derin Öğrenme teknikleriyle işlenmesi sonucu elde edilmiştir. Politus Analytics'in veritabanında an itibarıyla toplamda 310 milyon tweet yer almakta ve bu sayı her geçen gün artmaktadır. İlgili veriler GDPR ve KVKK kapsamındaki etik ve hukuki sınırlar çerçevesinde muhafaza edilmektedir.

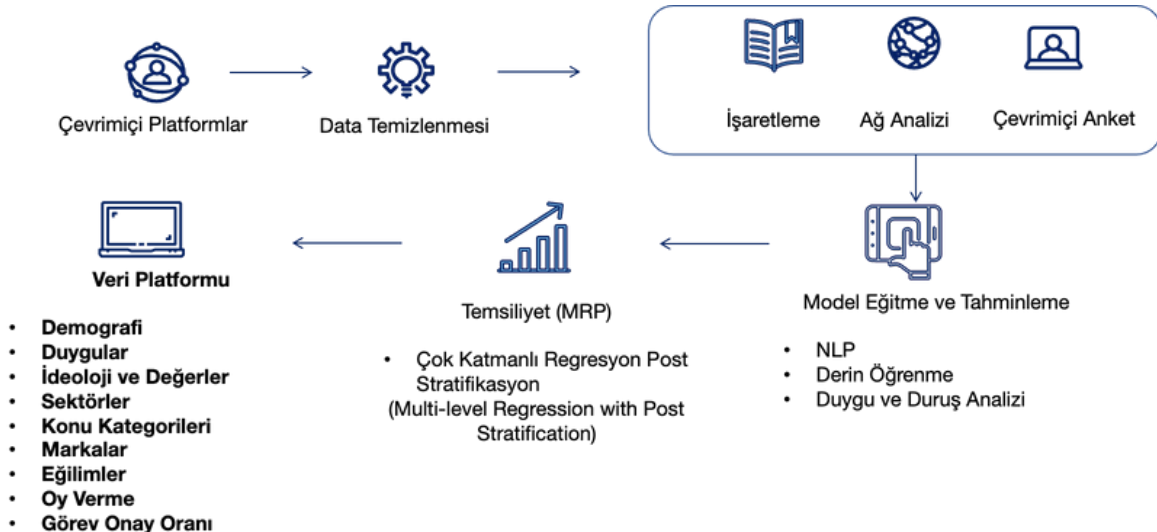
Makine öğrenmesi modellerinin girdisi olan veriler, alanlarında uzman sosyal bilimcilerin desteği ile oluşturulmaktadır. Uzmanlar, kavramlar ile ilgili detaylı tanım rehberleri oluşturduktan sonra, tweet'ler bu rehberlere göre en az iki işaretleyici tarafından işaretlenir. İşaretleyicilerin ortak ya da farklı olarak işaretlediği tüm tweet'ler tekrar uzmanlar tarafından geliştirilen kavramsallaştırmalar kapsamında değerlendirilerek son haline getirilir. Bu kavramsallaştırmalar üzerine çok detaylıca çalışılması sayesinde yapay zeka modelleri nadir rastlanan zor kavramları dahi ayırt edebilecek şekilde eğitilmiştir.

Ayrıca Politus Analytics sonuçları, resmi veriler veya konu bazında yapılmış diğer anketler ile korelasyon değerlendirilmesi yapıldıktan sonra yayınlanmaktadır.

Duygu endeksleri, bir metnin genel duygusal tonunu yansıtan ölçütlerdir. Öncelikle, duygu analizi araçları ile tweet'lerin içerdiği duygu belirlenerek her hafta için bu tweet'lerin ortalama duygu puanı hesaplanır. Bu vesileyle, yalnızca Twitter kamuoyunun duygusal değişimini değil, aynı zamanda konuşulan konuların duygular ile bağlantılı olarak değerlendirilebildiği endeksler haftalık olarak takip edilebilmektedir.

Veri oluşturulması ile ilgili araştırma süreci ve iş akışı grafiği aşağıda yer almaktadır.

Araştırma Süreci ve İş Akışı





Seçmen Duyguları

2023 CUMHURBAŞKANLIĞI SEÇİMLERİ

Seçmen duyguları arasında adaydan bağımsız olarak en baskın duygu öfke olmakla birlikte, Twitter platformunun yapısı da öfke duygusunun daha fazla yer almasında etkilidir; öfkeli kullanıcılar daha fazla tweet atma eğilimindedir. Bu nedenle duygular ayrı ayrı değerlendirilmiş, duygular arasında bir karşılaştırma yapılmamıştır. Öfke duygusu konu bazında değerlendirildiğinde ülke gündemini meşgul eden önemli hukuki davaların öfke duygusunu yukarı yönlü etkilediği gözlemlenmektedir.

6 Şubat Kahramanmaraş depremi ve sonrasında öfke ve üzüntü duyguları en yüksek seviyede iken mutlulukta da beklendiği üzere ciddi bir düşüş olmuştur.

Konu özelinde ve haftalık bazda umut ve umutsuzluk duyguları iniş ve çıkışlar içermekle birlikte, 2023 başından itibaren umut artarken umutsuzluk düşüş trendindedir.

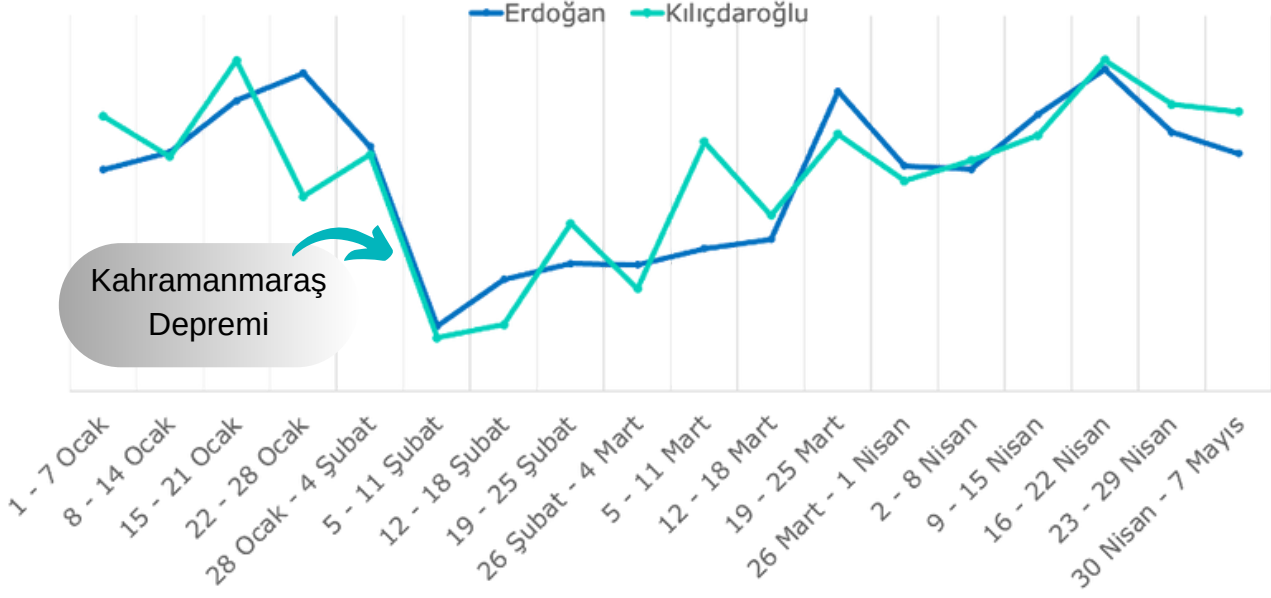
Umut

Umut duygusu Erdoğan destekçileri arasında istikrarlı bir trend izlerken Kılıçdaroğlu destekçileri arasında yükselme eğilimindedir. Umut duygusundaki en belirgin çıkış, Kılıçdaroğlu'nun adaylığının açıklandığı hafta gerçekleşmiştir.



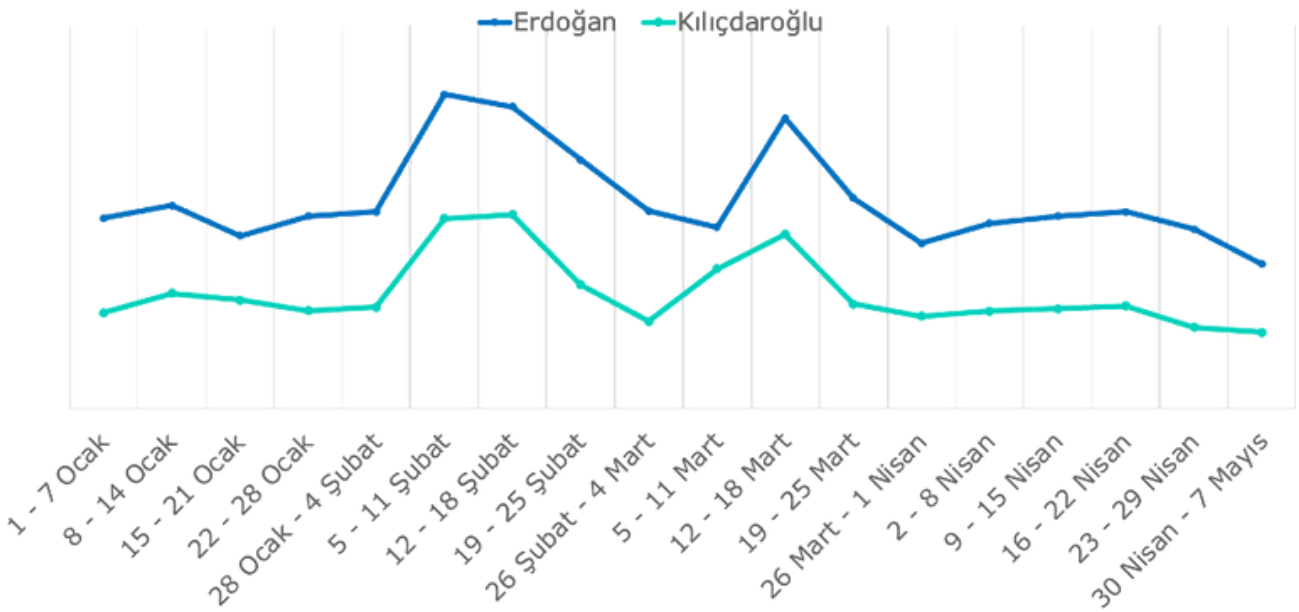
Mutluluk

Seçmenin mutluluk durumunu belirleyen en önemli olay, kuşkusuz Kahramanmaraş merkezli depremdir. Deprem felaketi, her iki seçmen grubunu da benzer şekilde etkilemiştir.



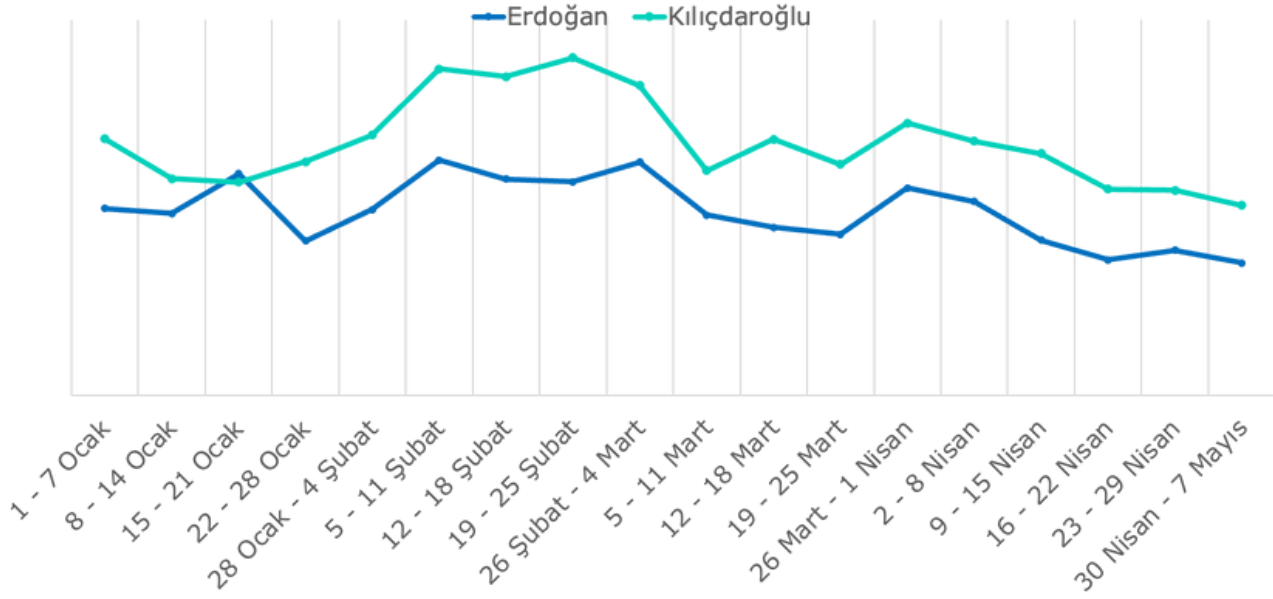
Minnet

Erdoğan destekçilerinin minnet duygusu Kılıçdaroğlu destekçilerine göre daha yüksek görünmektedir.



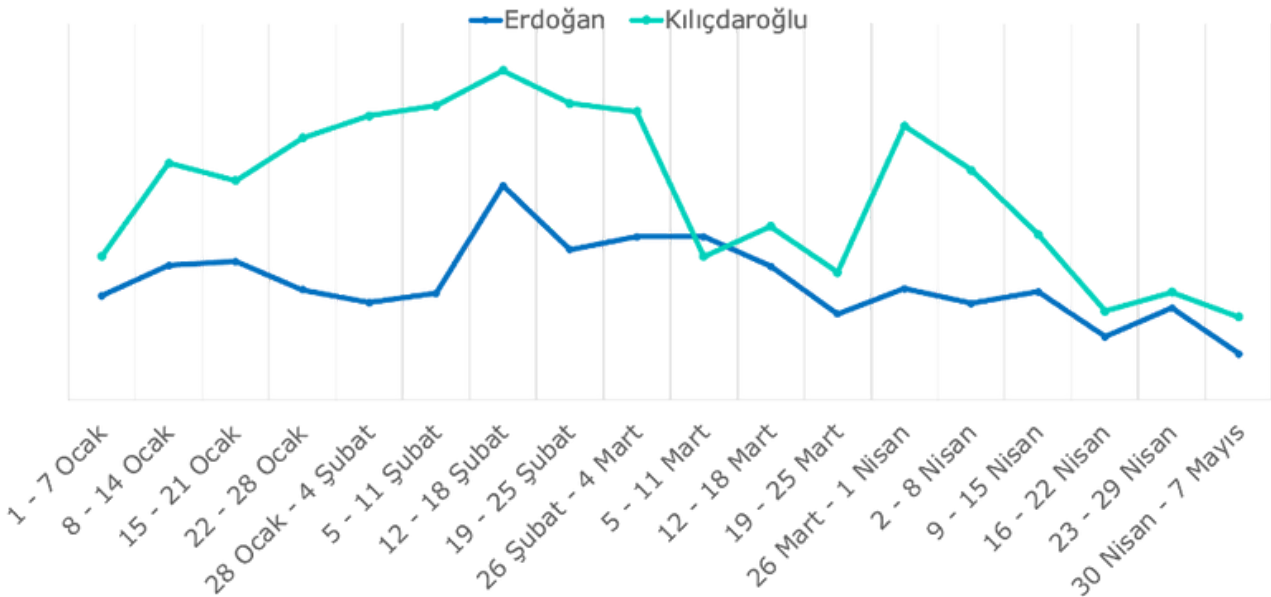
Öfke

Depremle birlikte (minnet duygusuna paralel olarak) tüm seçmenler arasında öfke duygusunda bir yükseliş yaşadığı gözlenmiştir. Ancak tüm bu süreç boyunca Kılıçdaroğlu destekçilerinin öfkesi Erdoğan destekçilerinin üzerinde seyretmekle beraber genel olarak bir azalma trendinden söz edilebilir.



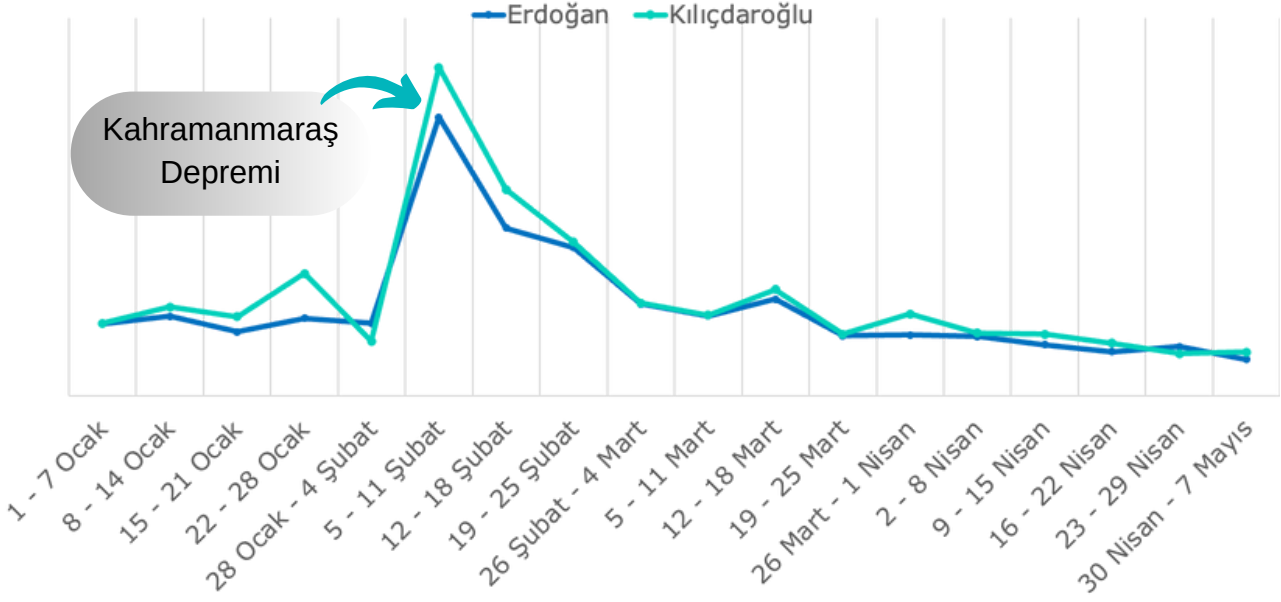
Umutsuzluk

Kılıçdaroğlu'nun adaylığının açıklandığı hafta ve sonrasında Kılıçdaroğlu destekçileri arasında oldukça yüksek olan umutsuzluk duygusunun anlamlı ölçüde azaldığı görülmektedir. Son haftada Kılıçdaroğlu destekçileri arasındaki umutsuzluk en düşük seviyesine gerilemiştir. Erdoğan destekçileri arasında da umutsuzluk açısından aşağı yönlü bir trend görülmektedir.



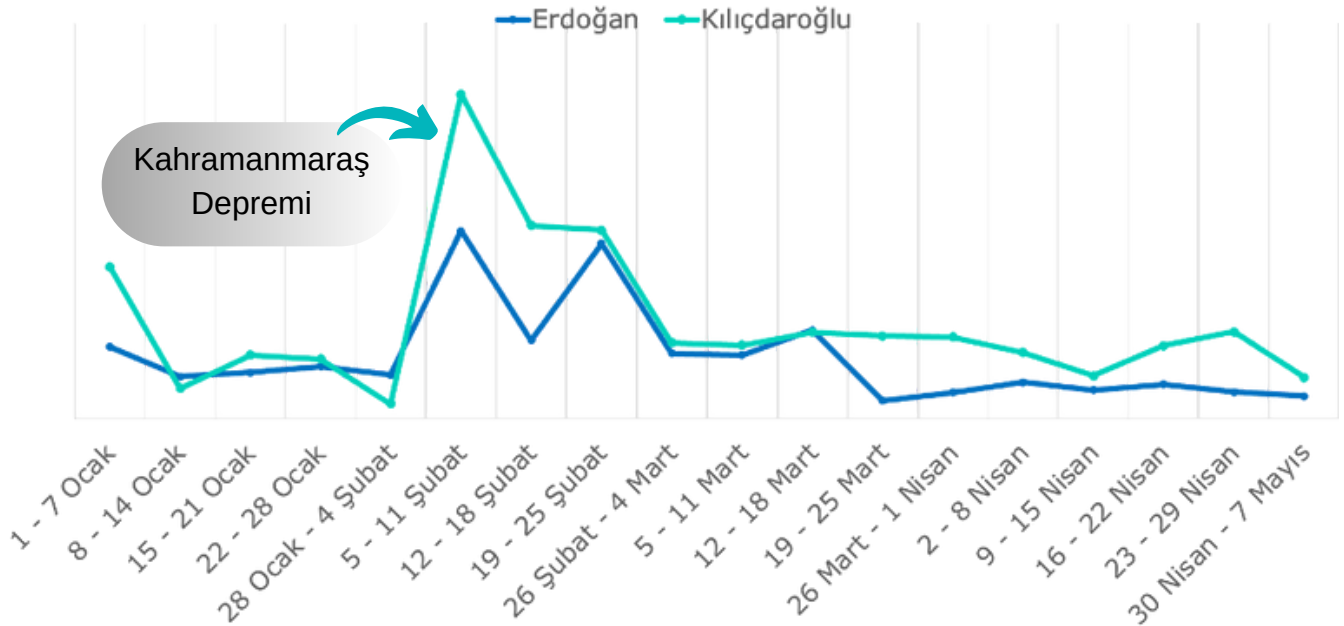
Üzüntü

Depremın tüm seçmenlerde derin bir üzüntüye yol açtığı, Politus verilerine açıkça yansımıştır. Üzüntü duygusu açısından Erdoğan ve Kılıçdaroğlu destekçileri arasında ciddi bir fark gözlemlenmemektedir.



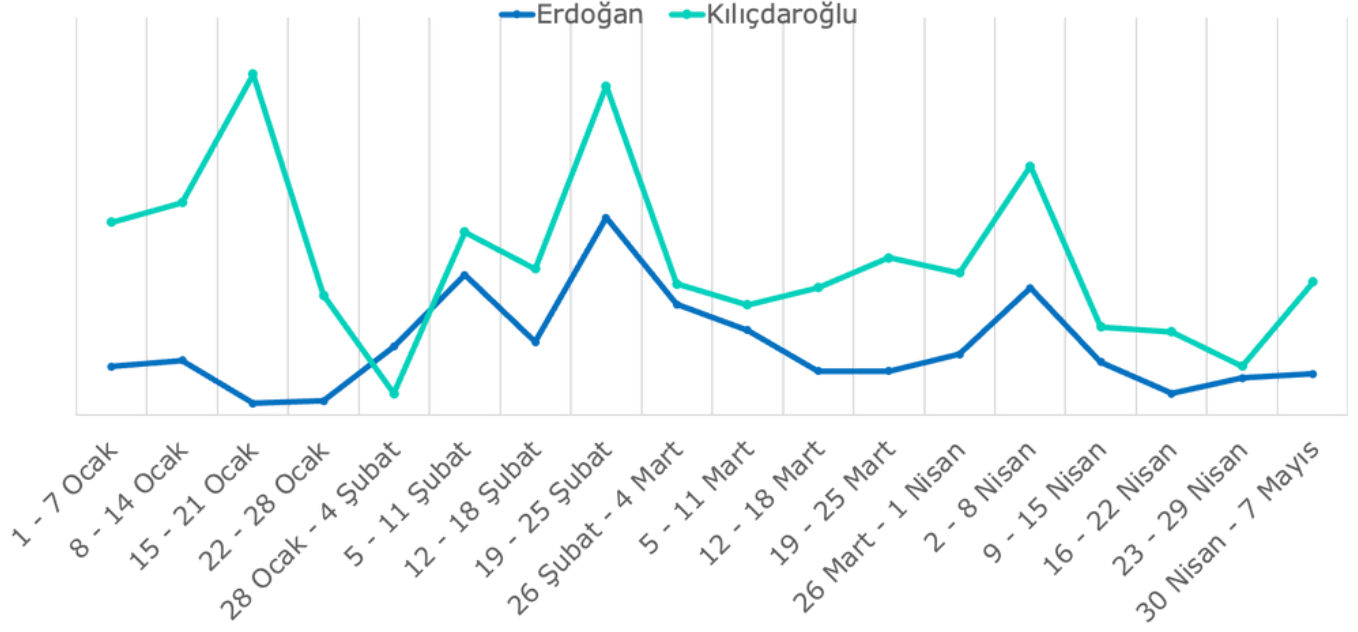
Kaygı

Ancak, kaygı hissiyatının, özellikle deprem döneminde, Kılıçdaroğlu destekçileri arasında daha yüksek olduğunu görmekteyiz.



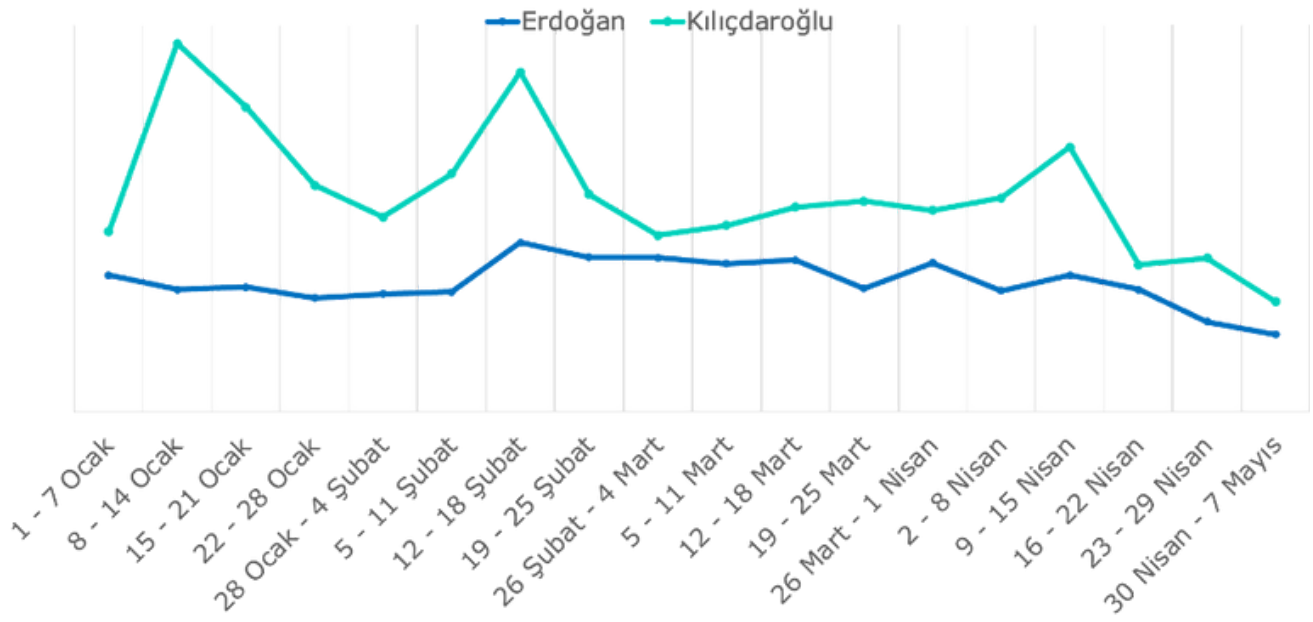
Korku

Kılıçdaroğlu destekçileri arasında korku hissiyatının, Erdoğan destekçilerine göre çok daha inişli çıkışlı bir trend izlediği ve görece yüksek olduğu görülmektedir.



Şaşkınlık

Benzer bir şekilde, Erdoğan destekçileri arasında şaşkınlık düzeyinin görece istikrarlı seyrettiğini, Kılıçdaroğlu destekçilerinin şaşkınlık duygusunun ise oldukça iniş çıkışlı olduğu görülmektedir.

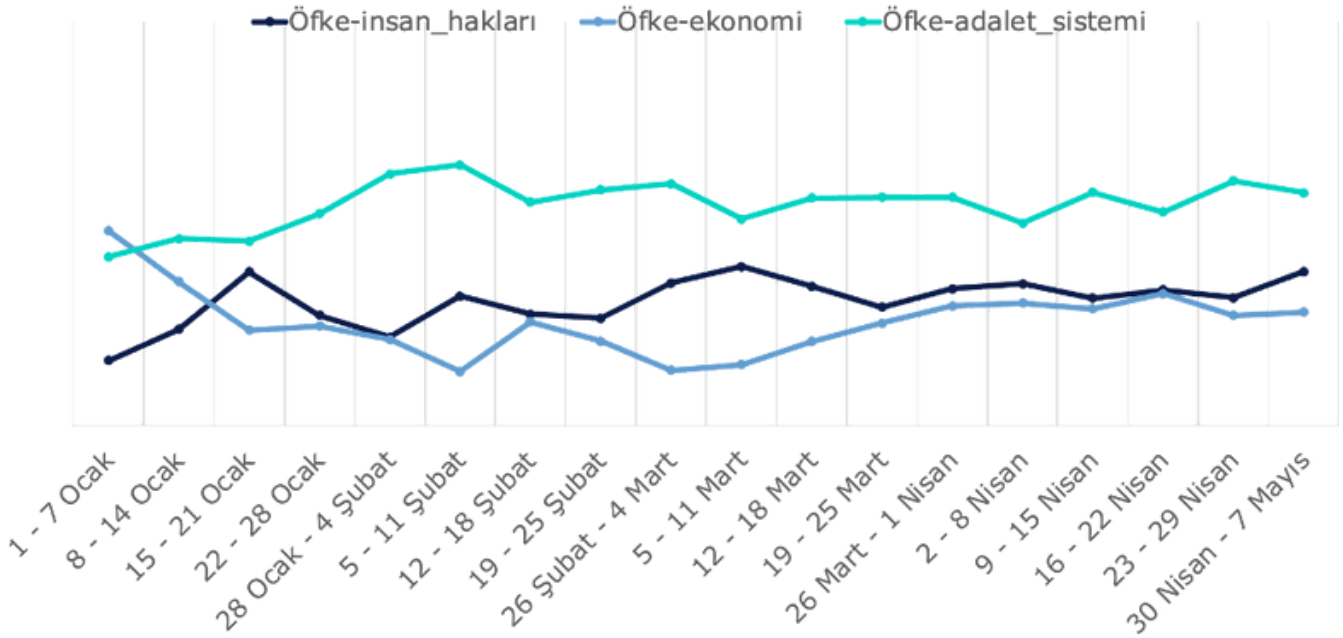


Konu Bazında *Analizler*

2023 CUMHURBAŞKANLIĞI SEÇİMLERİ

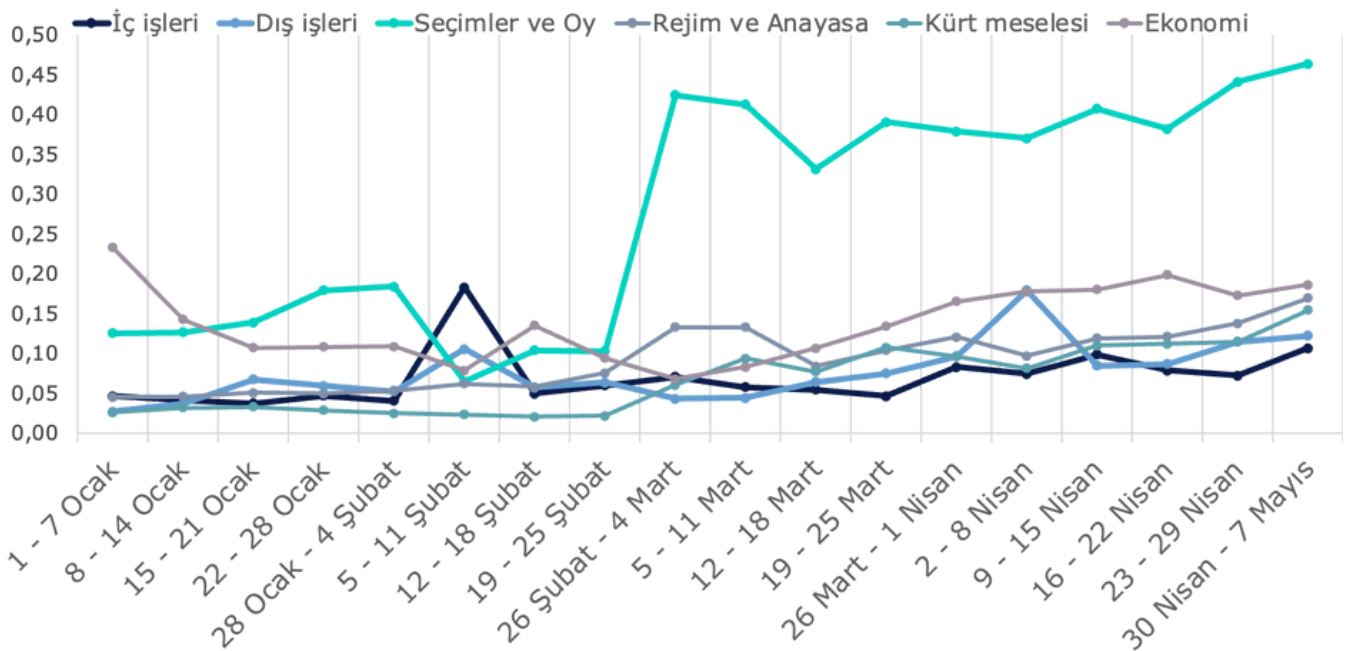
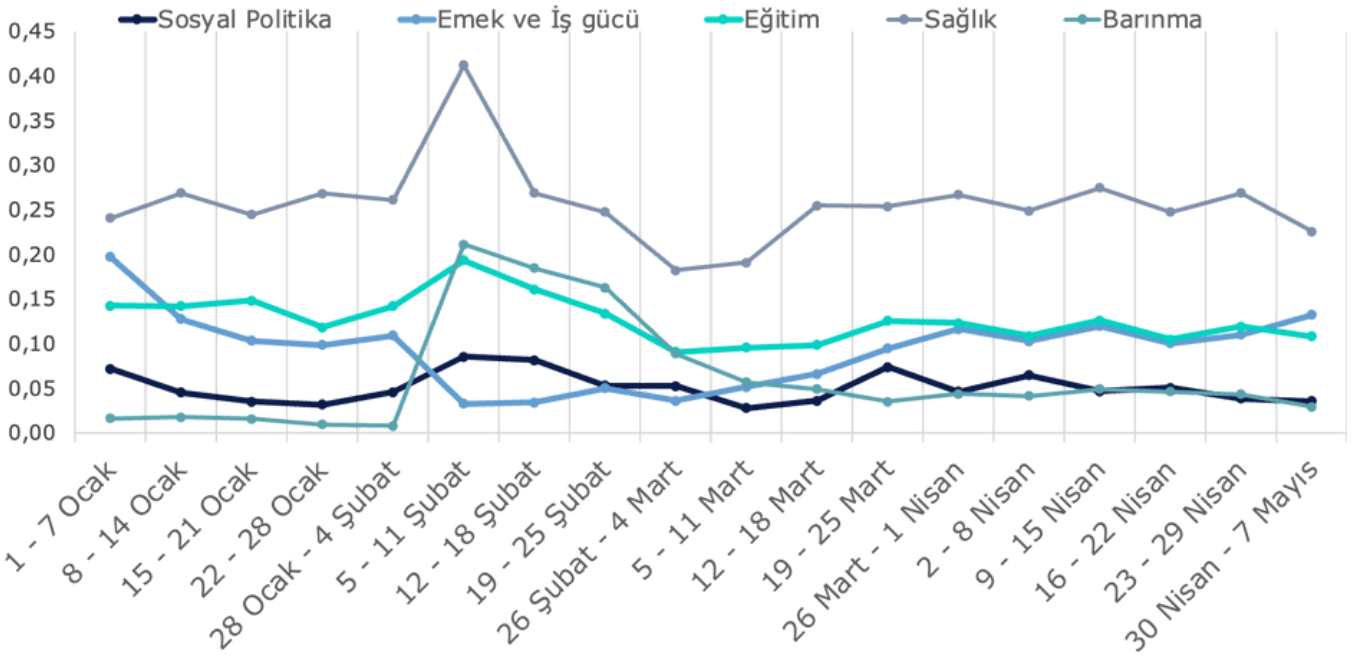
Öfke & İnsan Hakları, Adalet ve Ekonomi

Öfke, seçmenler arasında en yoğun olarak tespit ettiğimiz duygudur. Özel olarak, adalet sistemi konusunda duyulan öfke insan hakları ve ekonomiye göre anlamlı derecede yüksektir. Ekonomi konusunda duyulan öfke, depremle beraber en düşük seviyesine inmişse de, 4 Mart haftasından itibaren doğrusal bir artış göstermektedir. Son haftada ekonomi konusundaki öfke insan hakları konusundaki öfkenin üzerine çıkmıştır. İnsan hakları konusunda duyulan öfkenin İmamoğlu, Kaftancıoğlu, H.K.G. davası gibi dava tarihlerine göre artış eğiliminde olduğu gözlemlenmektedir.



Gündem Konularının Zamansal Değişimi

Aşağıdaki iki grafik, çeşitli konuların ne yoğunlukla konuşulduğunu ve bunun zaman içerisindeki değişimini göstermektedir. Barınma ve sağlık konularının ağırlığı depremle birlikte ciddi bir artış gösterirken Mart ayının başından itibaren oy tercihi ve seçim konularına yönelik ilgide yüksek bir artış yaşandığı görülebilir.

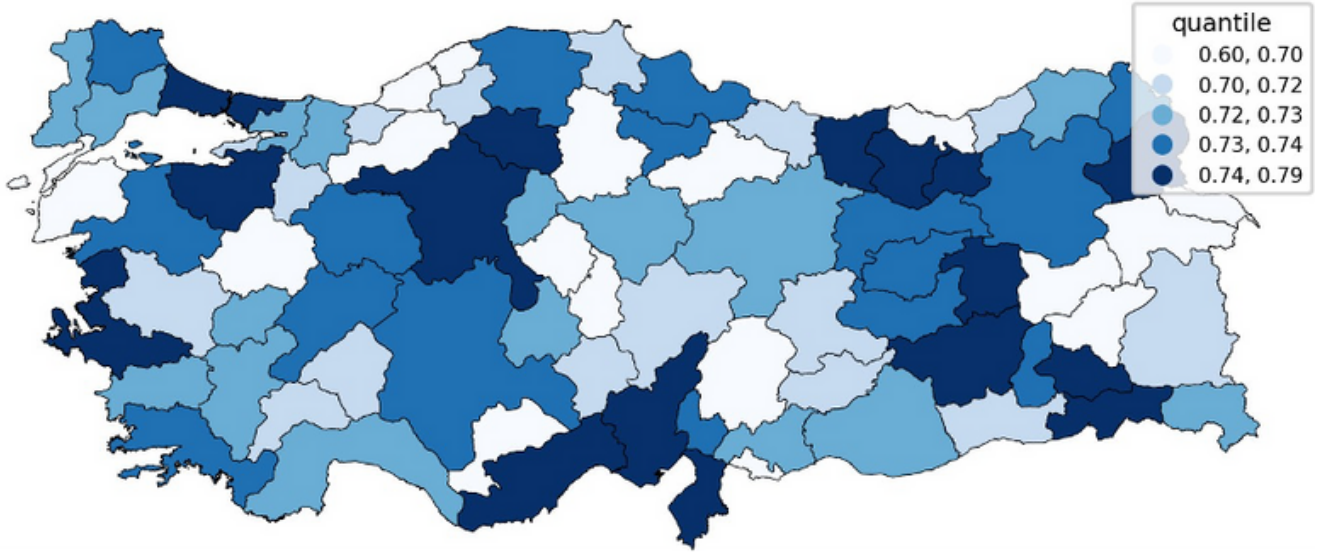


Hangi İlde Hangi Konular Konuşuluyor?

Adalet

Adalet konusunun bahsi, tüm bölgelerde benzer bir dağılım göstermektedir. Büyükşehirlerin adalet konusundan daha çok bahsettiği söylenebilir.

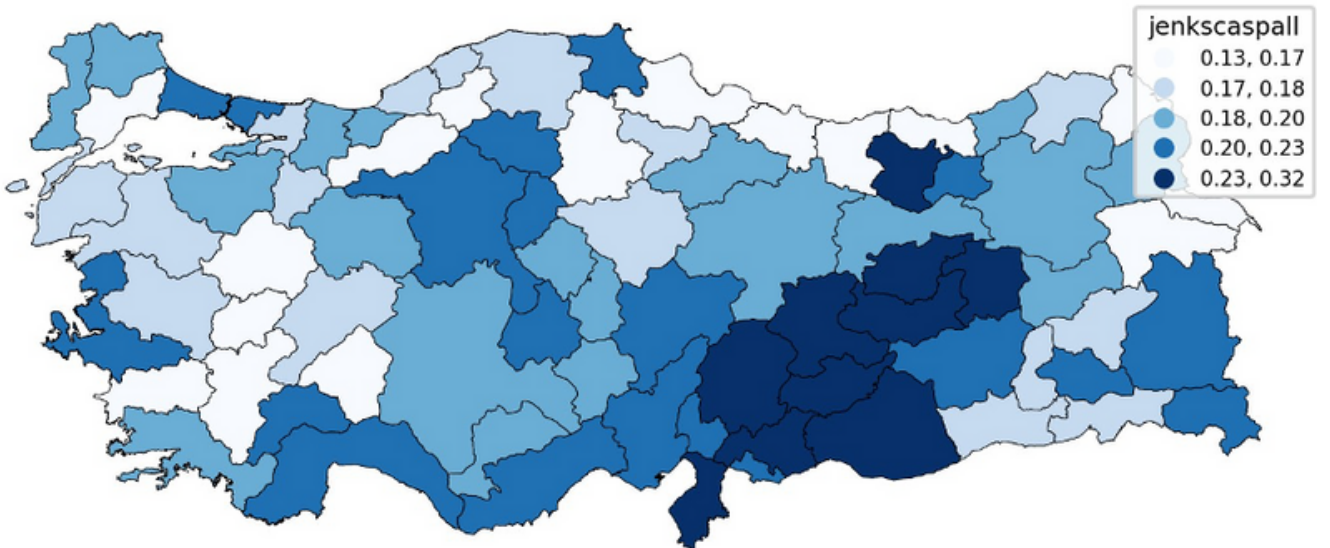
Adalet hakkında tweet atan kullanıcıların illere göre yoğunluğu



Barınma

Barınma konusu tahmin edilebileceği gibi deprem bölgesinde ön plana çıkmaktadır. Renk yoğunluğunun en yüksek olduğu iller deprem bölgesi olarak görünmektedir.

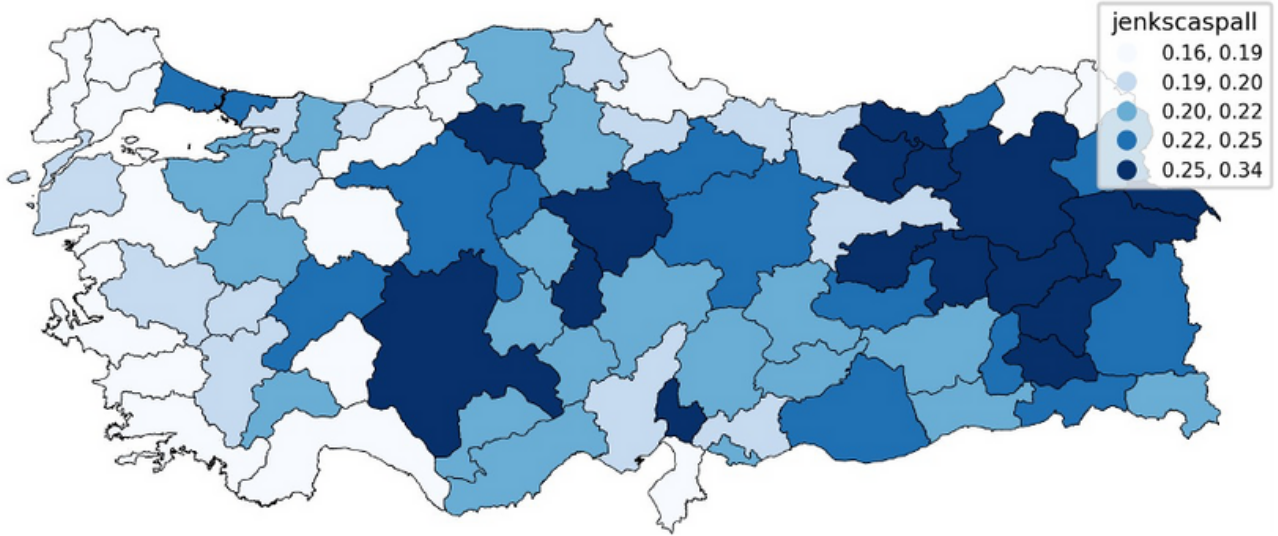
Barınma hakkında tweet atan kullanıcıların illere göre yoğunluğu



Dış işleri

Dış işleri ile alakalı konular ülkenin batısından ziyade orta ve doğu kesimlerinde daha çok konuşulmaktadır.

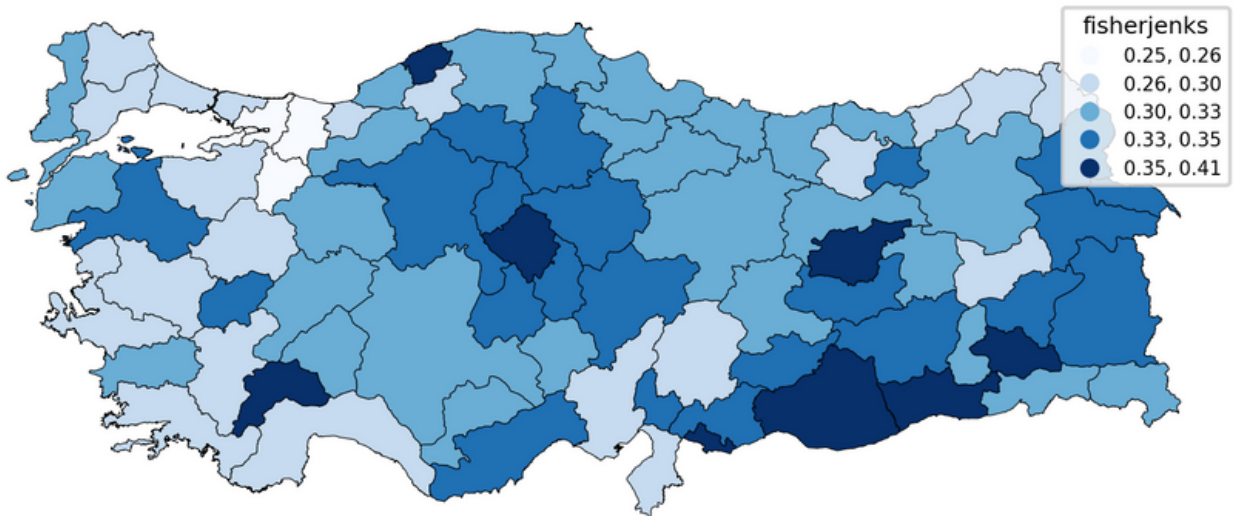
Dış işleri hakkında tweet atan kullanıcıların illere göre yoğunluğu



Eğitim

Eğitim, tüm bölgelerimiz için yoğun bir biçimde ilgi gösterilen bir konu olarak karşımıza çıkmaktadır.

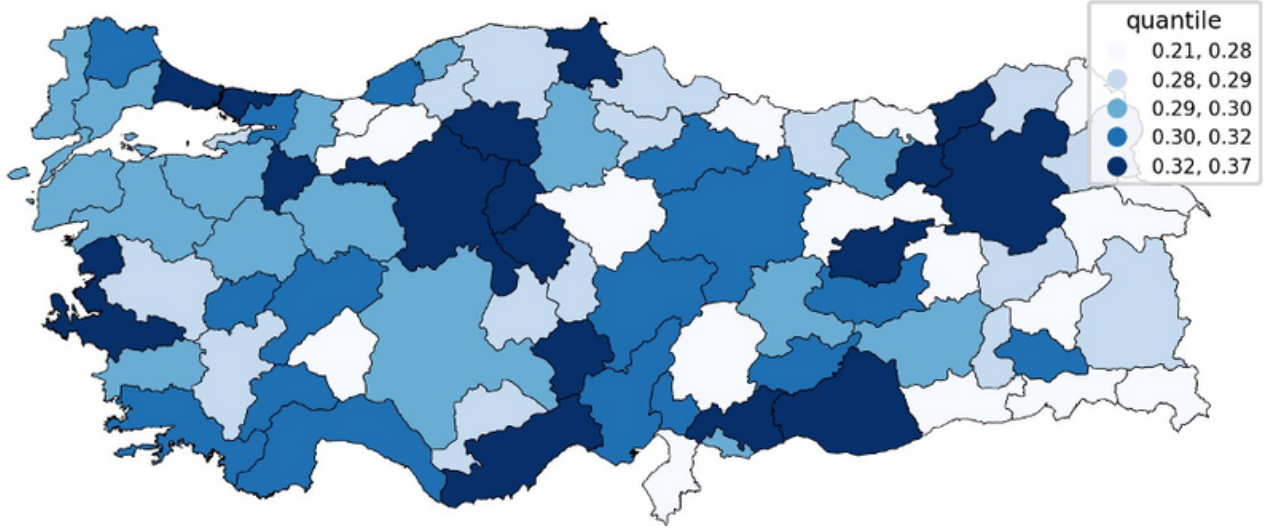
Eğitim hakkında tweet atan kullanıcıların illere göre yoğunluğu



Ekonomi

Ekonomi de, adalet ve eğitim gibi, tüm bölgelerimizi benzer şekilde meşgul etmektedir. Yine büyükşehirlerde Ekonomi konusunda daha çok konuşulduğu gözlemlenebilir.

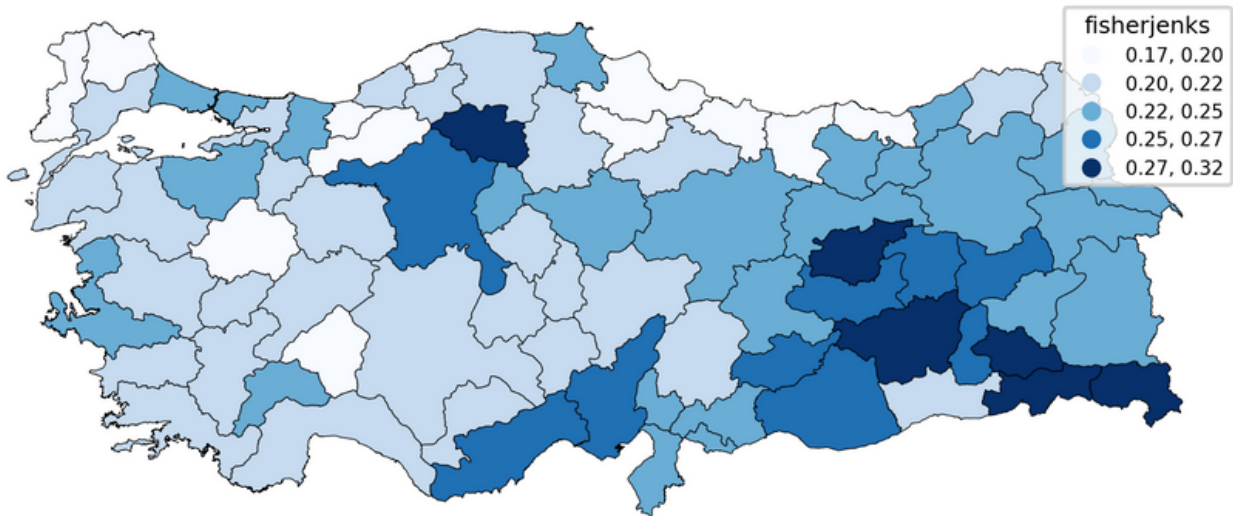
Ekonomi hakkında tweet atan kullanıcıların illere göre yoğunluğu



İç işleri

İç işleri ve güvenlik konularının ülkenin doğu ve güneydoğusunda daha çok konuşulduğu söylenebilir.

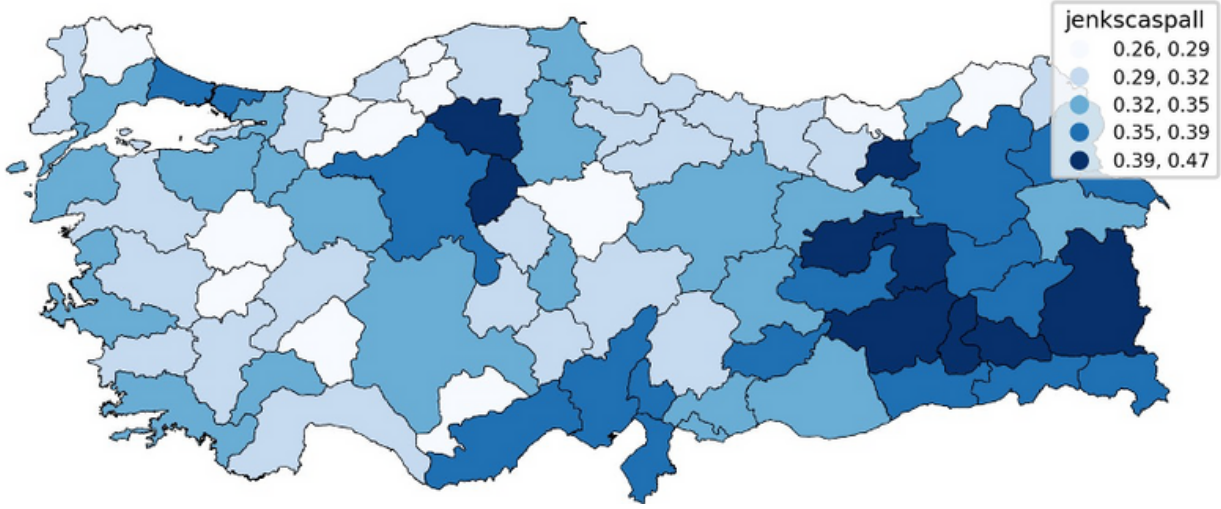
İç işleri hakkında tweet atan kullanıcıların illere göre yoğunluğu



İnsan Hakları

İnsan hakları konusu, daha çok Doğu ve Güneydoğu Anadolu bölgelerinde gündem olmaktadır.

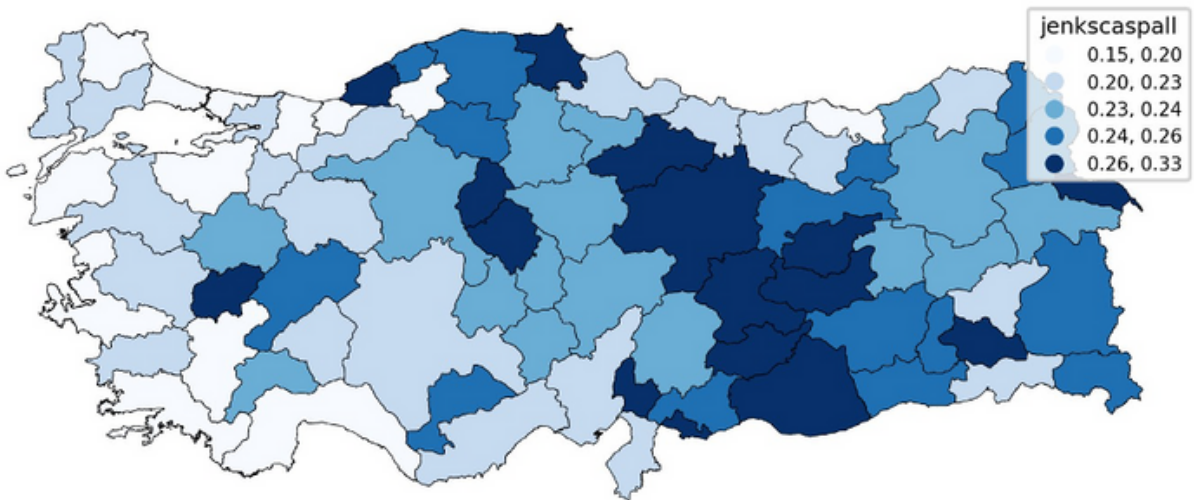
İnsan hakları hakkında tweet atan kullanıcıların illere göre yoğunluğu



Emek ve İş gücü

İş ve iş gücü konularının Orta Anadolu'nun doğu kısmında daha çok gündem olması kayda değerdir.

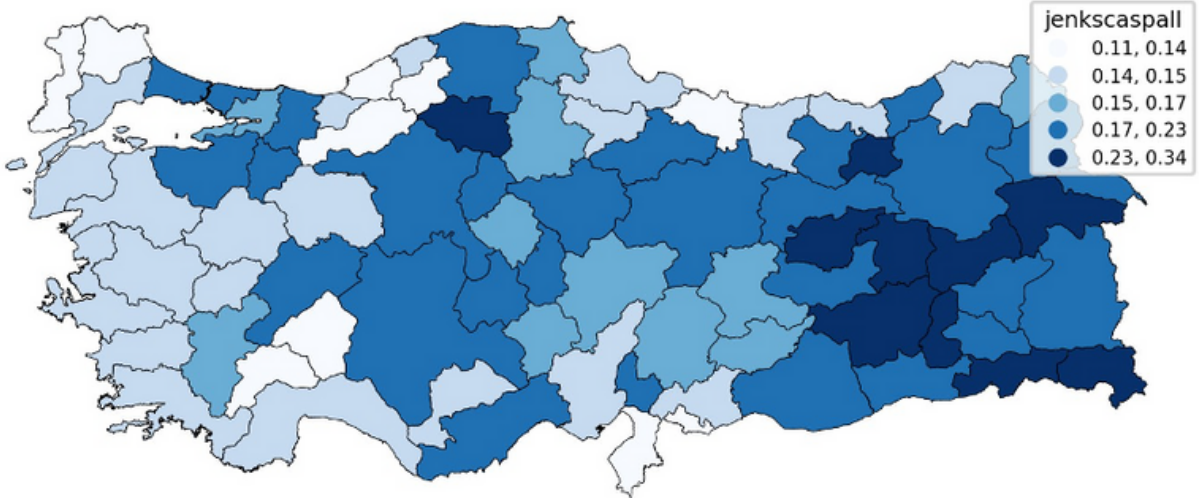
İş ve iş gücü hakkında tweet atan kullanıcıların illere göre yoğunluğu



Kürt Meselesi

Kürt meselesi, Marmara'nın doğusu, Orta Anadolu, Doğu ve Güneydoğu Anadolu bölgelerinin gündeminde daha çok yer tutmaktadır.

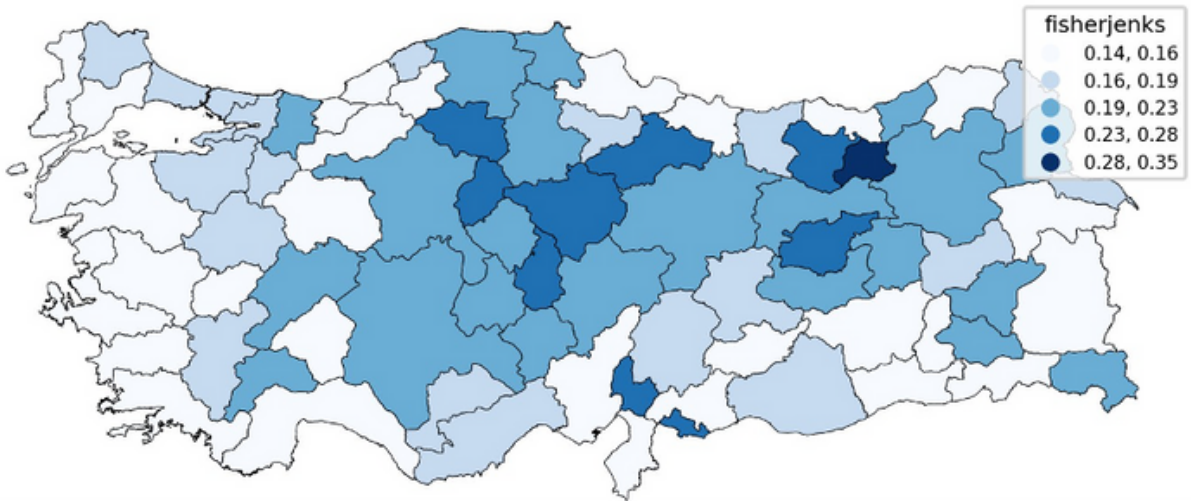
Kürt meselesi hakkında tweet atan kullanıcıların illere göre yoğunluğu



Milli Savunma

Milli savunma ile alakalı konular, milliyetçi seçmenin yoğun olduğu Orta Anadolu illerinde en çok konuşulmaktadır.

Milli savunma hakkında tweet atan kullanıcıların illere göre yoğunluğu

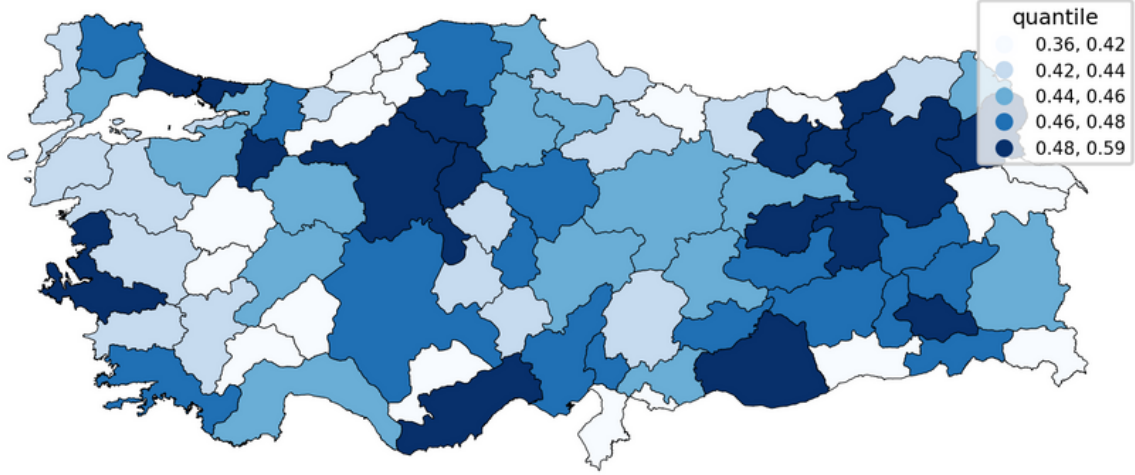




Oy Verme ve Seçimler

Seçim konusu da, ekonomi, adalet ve eğitim gibi, tüm ülkeyi benzer şekilde ilgilendiren bir konu olmakla birlikte büyükşehirlerde görece daha fazla konuşulmaktadır.

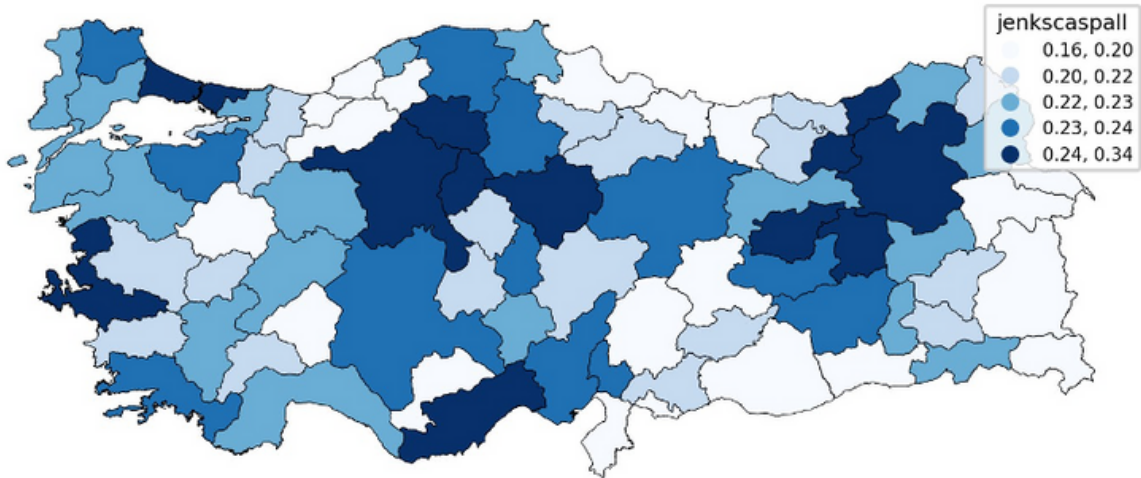
Oy ve seçimler hakkında tweet atan kullanıcıların illere göre yoğunluğu



Rejim ve Anayasa

Rejim ve anayasa konuları, seçim sistemine olan ilgiye paralel bir görünümde dir.

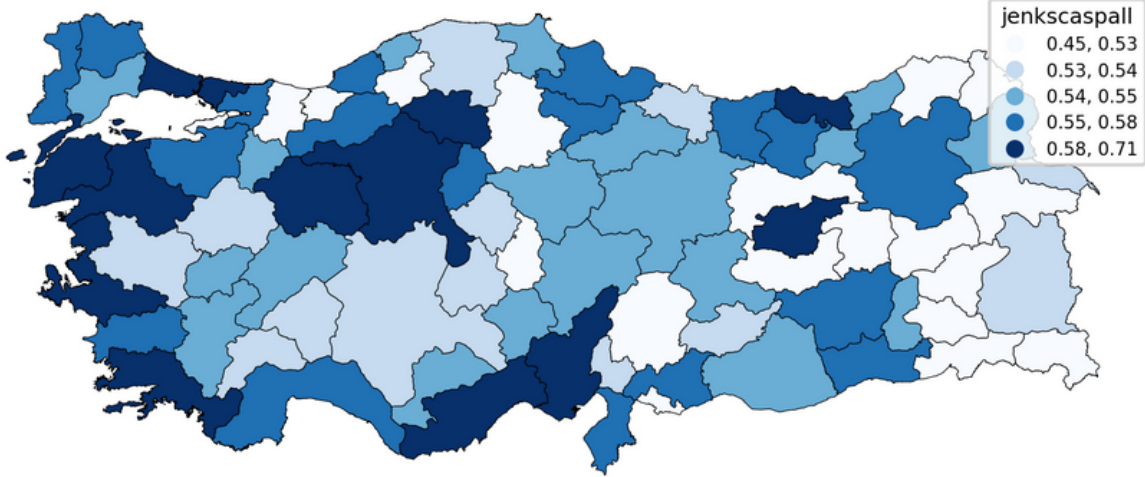
Rejim ve anayasa hakkında tweet atan kullanıcıların illere göre yoğunluğu



Sağlık

Sağlık, Ege ve Akdeniz kıyılarında daha çok gündem olmaktadır.

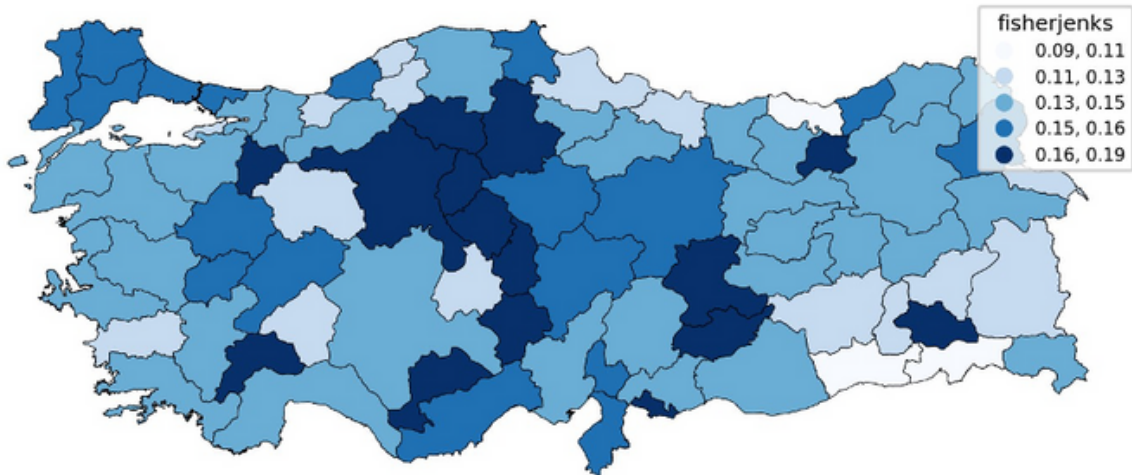
Sağlık hakkında tweet atan kullanıcıların illere göre yoğunluğu



Sosyal Politika

Sosyal politikalar konusundan Orta Anadolu bölgesinde daha fazla bahsedildiği görülebilir.

Sosyal Politikalar hakkında tweet atan kullanıcıların illere göre yoğunluğu

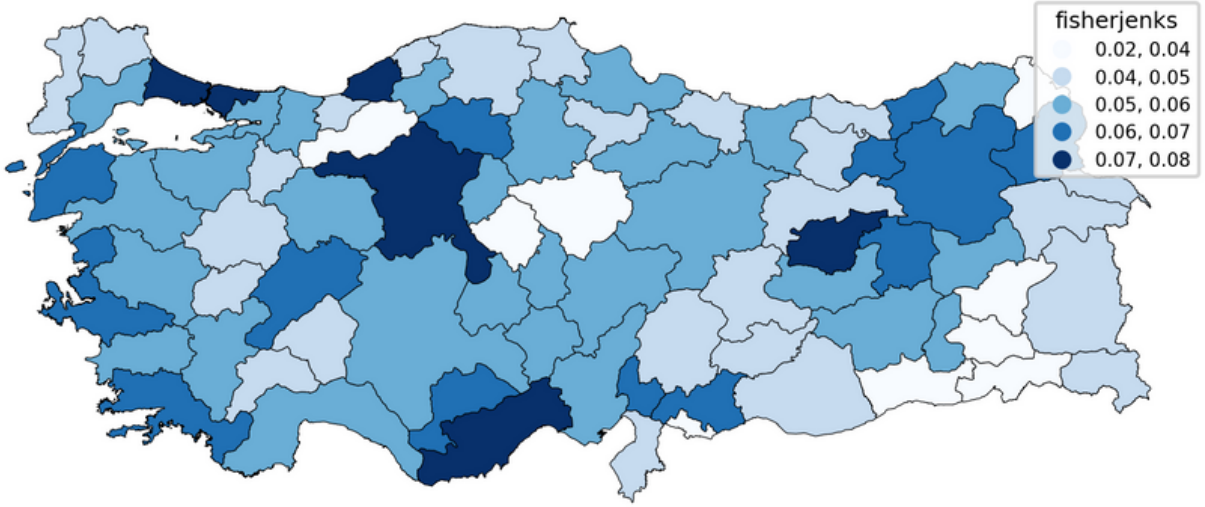




Yolsuzluk

Yolsuzluk konusuna dair ilginin ülkenin batısına doğru bir artış gösterdiği gözlemlenmekle beraber, konu hakkında en çok konuşulan iller farklı bölgelere dağılmış haldedir.

Yolsuzluk hakkında tweet atan kullanıcıların illere göre yoğunluğu

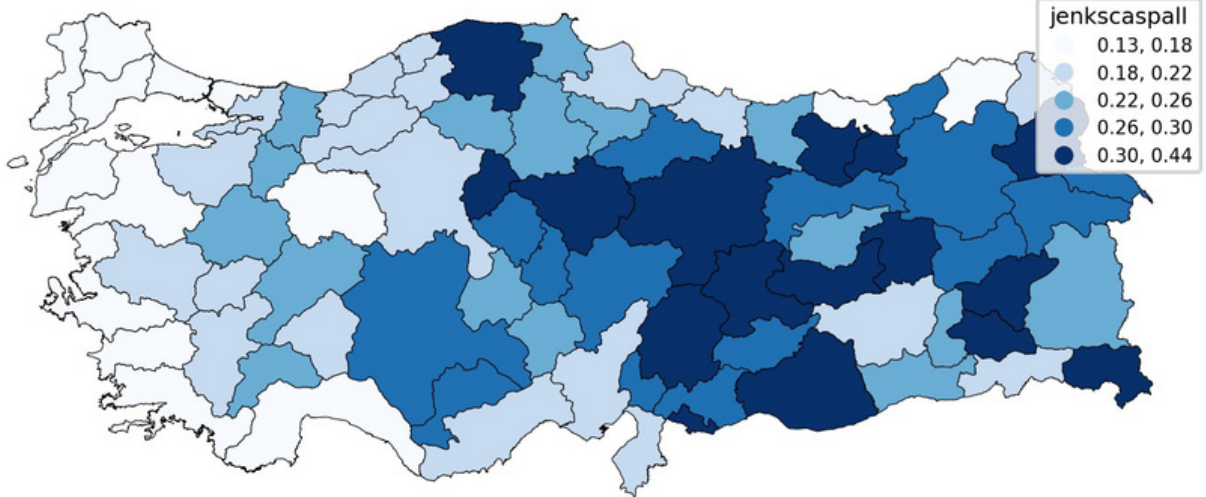




Aşağıdaki haritalar, çeşitli ideolojik görüşlerden seçmenlerin hangi şehirlerde daha yoğun olduğunu göstermektedir. Söz konusu ideolojiler sırasıyla muhafazakarlık, liberalizm, Türk milliyetçiliği, kemalizm, islamcılık, feminizm, ve çevreciliktir. Bu ideolojiler arasında en yoğun şekilde tespit ettiğimiz ideoloji Türk milliyetçiliği olmuştur.

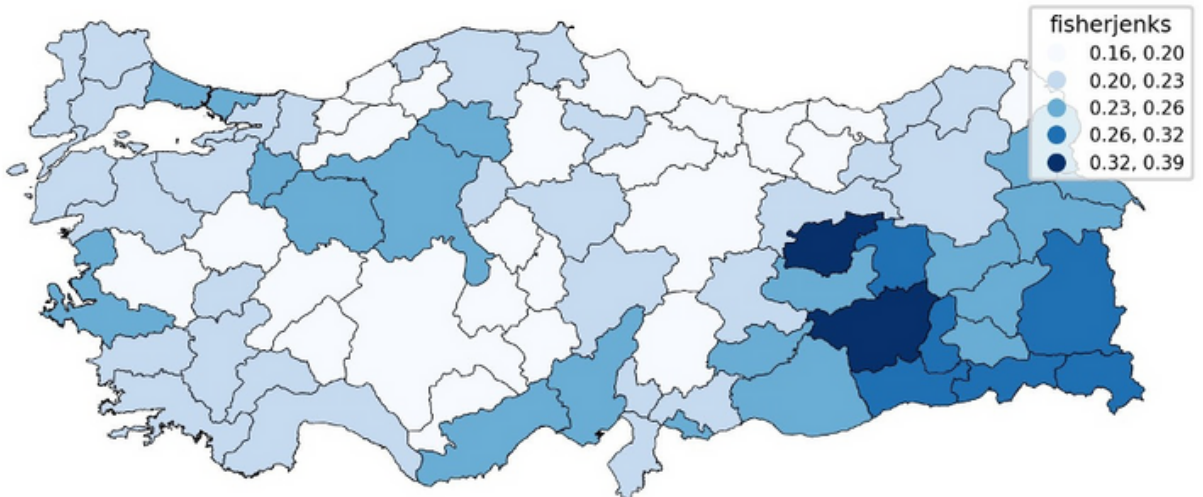
Muhafazakarlık

Muhafazakarların illere göre yoğunluğu



Liberalizm

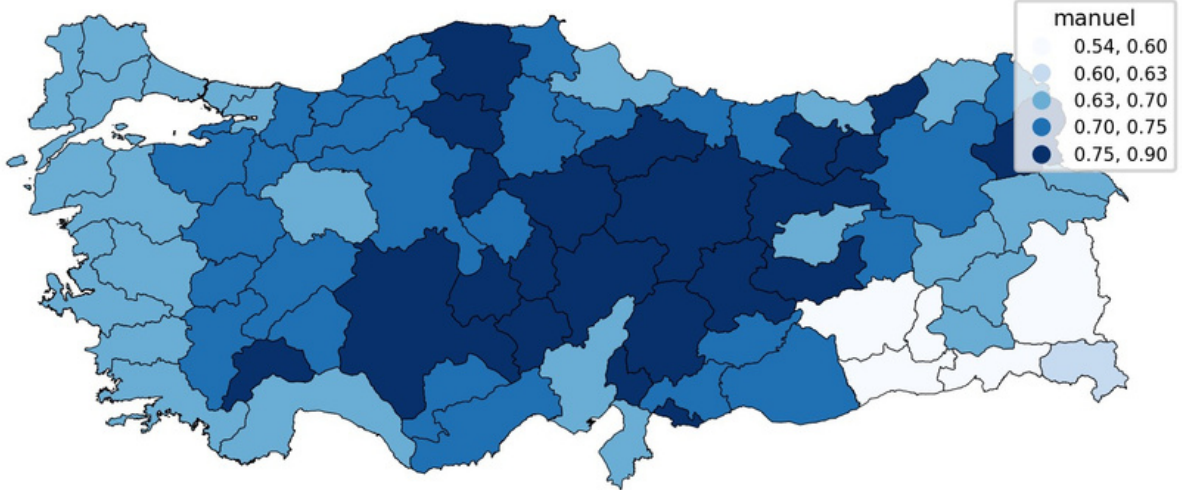
Liberallerin illere göre yoğunluğu





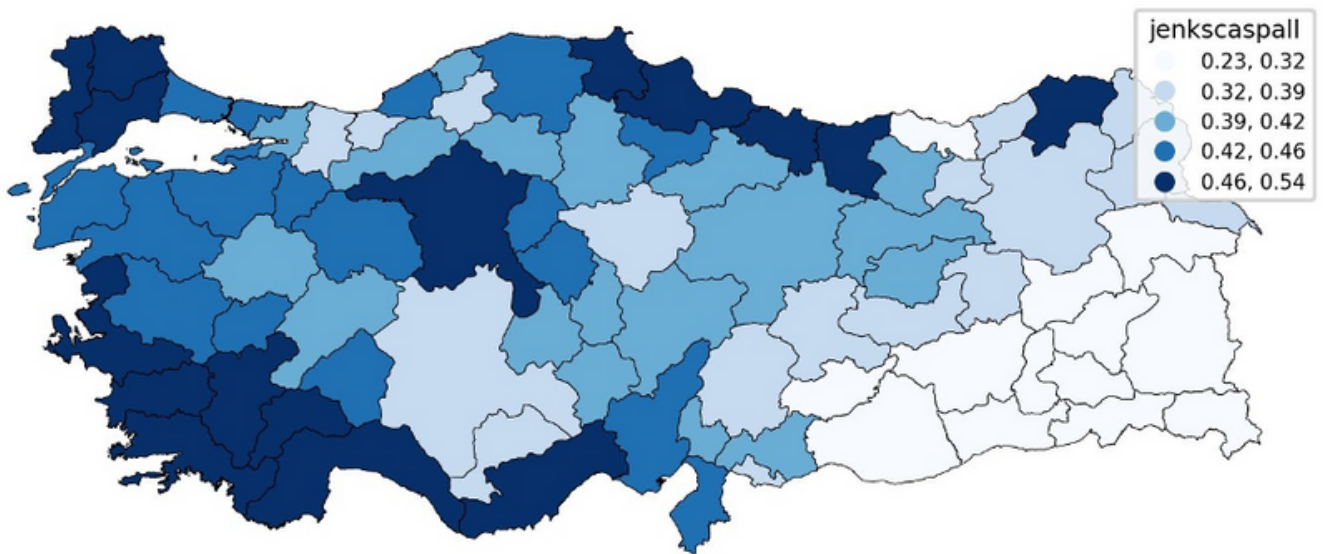
Türk Milliyetçiliği

Türk milliyetçilerinin illere göre yoğunluğu



Kemalizm

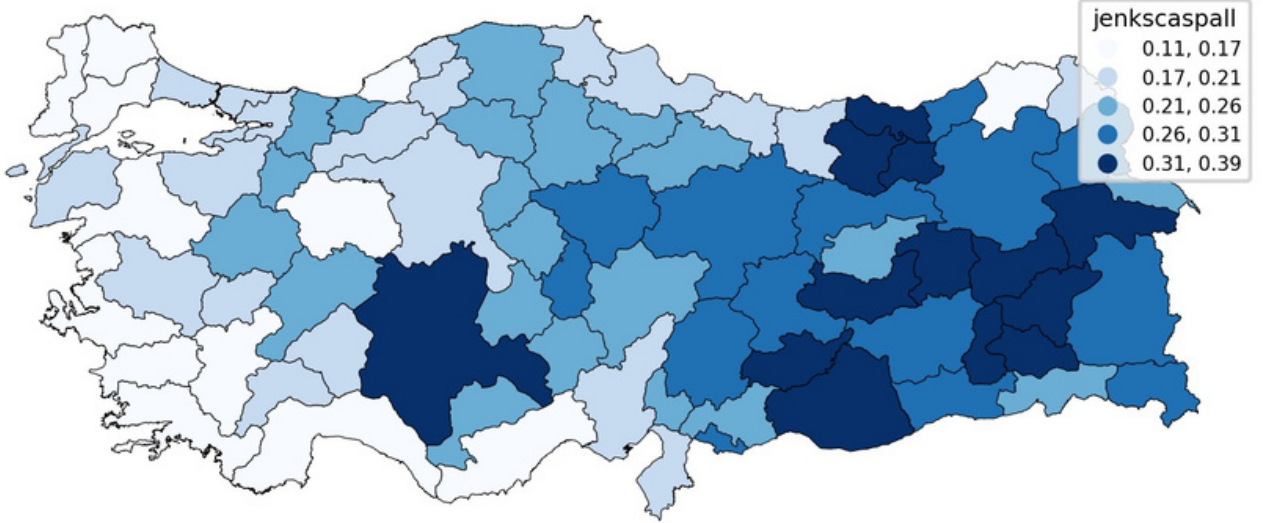
Kemalistlerin illere göre yoğunluğu





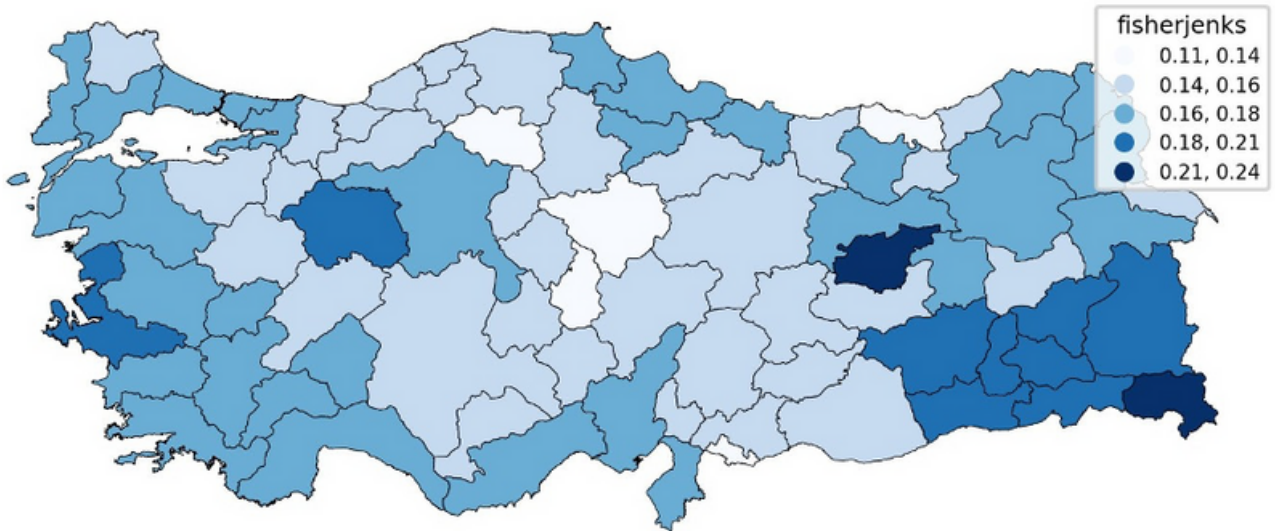
İslamcılık

İslamcıların illere göre yoğunluğu



Feminizm

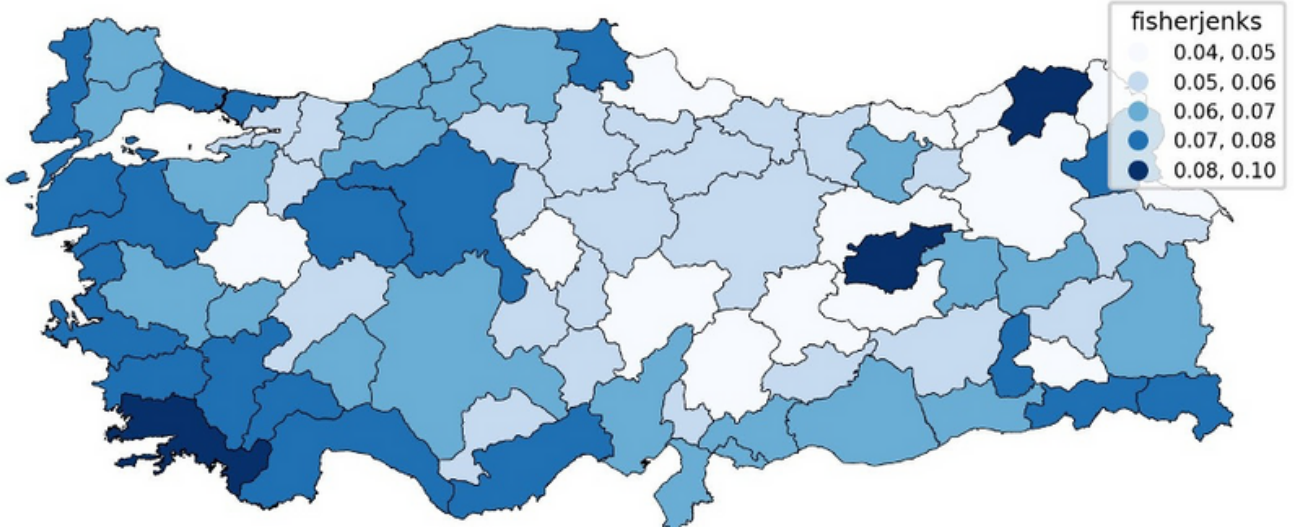
Feministlerin illere göre yoğunluğu





Çevrecilik

Çevrecilerin illere göre yoğunluğu



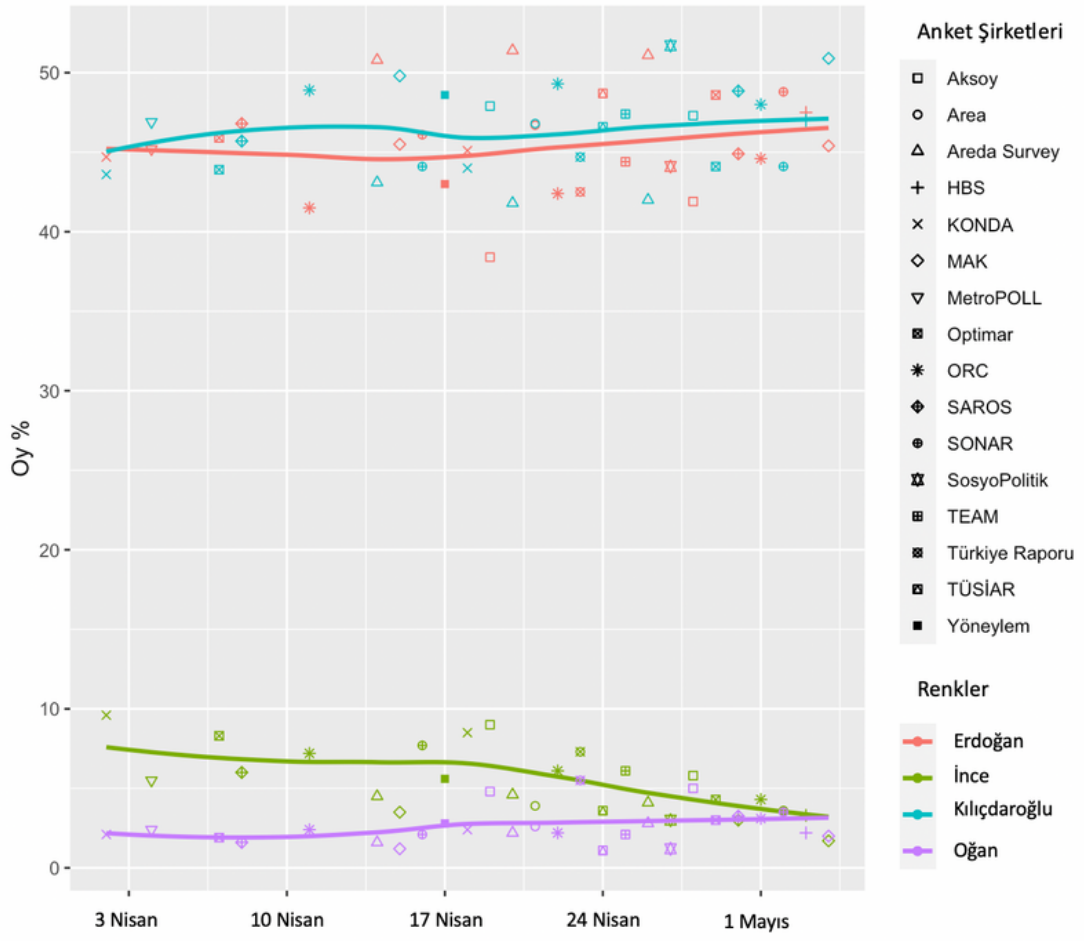


Hazırlayanlar: OĐuz Gurerk, Erdem Yoruk, Mehmet Fuat Kına, Őukru Atsızelti

Kritik 2023 Cumhurbaşkanlığı Seçimleri yaklaşırken kamuoyu anketlerinin sıklığı da gitgide artıyor. Siyasi partiler tarafından Cumhurbaşkanlığı adayını belirlemede dikkate alınan kamuoyu anketleri, karar alma süreçlerini olumlu ya da olumsuz şekilde etkileyebiliyor. Buradaki riskler, anketi düzenleyen kurumların yöntemsel yaklaşımlarındaki hatalardan ve/veya siyasi eğilimlerinden ve taraflılığın kaynaklanıyor. Türkiye’de seçim anketlerine yönelik bir anket birleřtirmesi [poll aggregation] uygulanagelen bir alıřma olmamakla beraber, bu tarzda var olan birkaç alıřma, anketlerin basit aritmetik ortalamasından ibaret bir hesaplama yöntemi kullanıyor. Ancak, ařaĐıda bahsedilen sorunlardan oturu, aritmetik ortalama, yanıltıcı sonuçlara yol aabiliyor.

Politus Analytics olarak 2023 Cumhurbaşkanlığı Seçimleri’ne yönelik yapılan anketleri en isabetli şekilde birleřtirebilmek iin istatistiksel bir yöntem uyguladık. Bu yöntem, anket řirketlerinin gemiř seçimlerdeki hata oranlarını aĐırlıklandırma sürecine dahil eden bir yaklařımı temel alıyor. Gemiř seçimlerdeki hata payları, yukarıda bahsedilen iki tipteki hata payının (yöntemsel ve siyasi) kumulatif olarak olumune olanak saĐlıyor. Bir anketin tahmini, seçim sonucuyla ne kadar ortuřuyorsa, o anketi yapanın birleřtirmedeki goreli aĐırlılıĐının o kadar fazla olması gerektiĐi onculune dayanan bir anket birleřtirme yöntemi oneriyoruz. ote yandan, anket řirketlerinin 2018 ve 2019 seçimlerindeki hata oranlarını hesaplamaya dahil ederek her bir řirketin uzun bir surte genel olarak ne kadar “isabetli” olduĐunun bir endeksini de ıkarmıř oluyoruz. Bu endeksi de her řirketin 2023 Cumhurbaşkanlığı Seçimleri ile ilgili tahminlerini aĐırlıklandırmada kullanıyoruz.

Buna gore nisan ayının bařından bu yana Muharrem İnce'nin oy oranın gittike azaldığı goruluyor. KılıdaroĐlu ve ErdoĐan arasındaki fark ise gorece istikrarlı bir trend takip etmekle beraber, ikisinin de oy oranı artış gosteriyor. ote yandan elde ettiĐimiz sonu, ilk turda seçimi bitirmek iin gereken yuzde elli sınırına iki adayın da eriřemediĐini iřaret ediyor.



1 Methods

1.1 Introduction to poll aggregation

Polls serve as accounting devices to monitor vote intentions at a point in time, and ideally, all survey data should accurately reflect population preferences. Although the central roles of surveys and sampling are facing some challenges in the twenty-first century with the threats to their accuracy from increasing (and differential) nonresponse and major cost pressures, still, they continue to have essential roles that other sources and methods of information gathering cannot easily replace. However, we know that polls suffer not only from sampling error¹, but also from “house effects”: pollster-specific bias introduced by particular sampling frames, non-transparent weighting formulas, or political allegiances of the firms involved. Thus, differentiating a potential source of bias in a measurement of opinion from genuine swings in the political mood is not an easy task.

In this report, we evaluate pre-election polls in Turkey, which are often the subject of heated debates, from a perspective of survey methodology and introduce a new method of poll aggregation, which is based on past performance assessments of the pollsters with respect to an accuracy measure first introduced by Martin et al. (2005). For various reasons (to be briefly explored next), averaging estimates across surveys can be a useful strategy for mitigating particular types of error. *But caution is merited, as a single unbiased, high-quality survey will frequently outperform an aggregation of biased, lower-quality results, [9].*

1.1.1 Choices in aggregation and common errors in polling

Aggregating polls involves combining the results of multiple polls on a particular topic or race to get a more accurate prediction of the target, which can be done by taking a simple average of the polls, or by using sophisticated methods such as Bayesian models, various smoothing algorithms, or statistical modeling of underlying time-series process when combining different surveys. However, “*attığın taş ürküttüğün kurbağaya değiyor mu*” is another topic. Poll aggregation can help reduce the margin of error and increase the accuracy of predictions given that differences in survey design and sample size are accounted, which are – as we will see – not possible or accessible in most cases in Turkey.

In general, large sample surveys will produce more precise estimates than small sample surveys given all else being equal since a larger sample size allows for a smaller margin of error, which means that the results are more likely to be representative of the population being studied. However, unlike sampling error, nonsampling error cannot be reduced simply by conducting more and larger surveys. Note that margin of error is a statistical concept that is used to indicate the level of uncertainty associated with a sample estimate, which is calculated based on the assumption that the sample was selected using random sampling. If the sample was not selected based on a method of random sampling, then the margin of error is not a valid measure of the uncertainty associated with the estimate, and in that case, it would be misleading to report a margin of error. So, it is important to report the limitations of the sample and the method used to select it.

¹as well as additional types of error (other than nonresponse) that are not reflected in the usually reported margins of error: frame, measurement, and specification, etc. as observed in the extensive literature on total survey error. See [10, 4, 7] for a comprehensive and systematic account of potential sources of bias.

Traditionally, “pollsters are used to minimizing nonsampling error by adjusting for differences between sample and population. Depending on the goals of the adjustment, various techniques are used, where the simplest approach is weighting – for instance, people in groups who are less likely to be included in the set of respondents (because of sample design or nonresponse) are given higher weights, etc.”, [6]. However, “survey weighting is a mess. It is not always clear how to use weights in estimating anything more complicated than a simple mean or ratios, and standard errors are tricky even with simple weighted means”, [5]. What is more, unfortunately, is in fact that survey reporting practices in Turkey are often opaque in terms of methodology², [2]. Aydaş (2020) documents that the share of polls with sampling methods unreported makes up about 69% of all publicly available pre-election polls in Turkey³.

Therefore, any method of poll aggregation (or a statistical modeling of repeated polling data) based on unreported polls lacks the scientific foundation on which it can be brought forward. However, scrutinizing a large number of pre-election polls on the basis of their “predictive accuracy” with respect to a measure we are to make use of enables us to argue whether pollster-specific biases and/or political allegiances of the polling firms involved are in fact in effect. Afterwards, we propose a method of poll aggregation based on the premise that the more predictive accuracy a poll has, the more its pollster deserve relative weight in the aggregation.

1.2 A measure of accuracy

Here, we rely on a methodology formalised by Martin et al. (2005); and via its generalization by Arzheimer and Evans (2014), we empirically and systematically study error in election polling, taking advantage of the fact that polls are typically conducted for each election, and that the election outcome can be taken to be the ground truth. Accordingly, we then move on to estimating the magnitude of the polls’ errors, or biases, in the 2018 Turkish presidential election cycle and the March 2019 municipality elections⁴.

First, we analyse the measure in its original form. In [8], the magnitude of a poll’s bias is defined relative to an election outcome, given by the following measure A , which measures the accuracy of a poll as a predictor of a two-party election result. Specifically,

$$A = \log \left(\frac{r/d}{R/D} \right), \quad (1)$$

where R/D is the ratio of actual votes for the Republicans and Democrats, respectively, and r/d is the ratio of support for the two parties in a given pre-election poll.

“It is easily computed and summarized, and it can be used as a dependent variable in multivariate statistical analyses of the nature and extent of biases that affect election forecasts and to identify their potential sources”, [8]. It is comparable across elections with different outcomes and among polls that vary in their treatment or numbers of undecided voters because looking at the odds of the two major parties takes undecided respondents and nonvoters out of the equation consistently,

²In the master’s thesis, Aydaş (2020) presents an extensive examination of the reporting practices of pre-election polls in Turkey for all elections conducted between 2011 and 2019, where the dataset includes 374 polls for 11 elections.

³What may be more surprising is that this is the highest reporting rate, which is from the November 2015 General Elections, as Aydaş (2020) finds, “the polls for the 2019 Istanbul Mayoral Elections have the lowest rate of reporting sampling method where only 8% of the polls did so”, [2].

⁴Ankara, Istanbul, and Istanbul (repeated).

while calculating the ratio of the odds focuses on whether the advantage of the winning party was adequately reflected in the survey [8, 1].

In particular, apart from the expected values of odds ratio other than 1, due to known sampling errors, values that exceed sampling errors can be regarded as real bias affecting the poll. If the odds ratio is greater than 1, this indicates a poll that favoured the *Republican* candidate compared to the actual result of the same candidate: the poll referred to an overestimated proportion of votes to the r -candidate in relation to the actual result. Conversely, an odds ratio less than 1 implies a poll that favoured the *Democratic* candidate compared to the election votes for that candidate. Finally, the transformation of the odds ratio by taking the natural logarithm⁵ makes the measure symmetric around zero (no bias). Thus, the farther the value of the measure is from 0, the worse the predictive ability of the poll is. On the other hand, when the odds ratio is close (or exactly equal) to 0, it can be said that the pre-election poll accurately predicts the official results [8, 1].

Martin et al. (2005) further shows A 's superiority to earlier measures of poll accuracy [8]. In its original form, however, A is restricted to two-party systems, and is therefore inapplicable to the majority of democratic systems [1]. To generalize the approach for a choice between k parties, Arzheimer and Evans (2014) define \mathbf{p} as a vector of proportions p_1, p_2, \dots, p_k of respondents who support party i in a given poll, and \mathbf{v} as a vector of proportions v_1, v_2, \dots, v_k of citizens who actually vote for the respective party, [1]. Depending on the question at hand, (self-declared) nonvoters (or the undecided) can be either excluded from the analysis following the lead of Martin et al. (2005) or coded separately as a pseudoparty [1]. Applying this generalized terminology to a two-party race, eq. (1) becomes

$$A = \log \left(\frac{\frac{p_1}{p_2}}{\frac{v_1}{v_2}} \right) = \log \left(\frac{\frac{p_1}{1-p_1}}{\frac{v_1}{1-v_1}} \right) \quad (2)$$

From eq. (2), Arzheimer and Evans (2014) gives a straightforward definition of a “party-specific” measure A'_i of polling accuracy as follows. For the i th of k parties, A'_i is

$$A'_i = \log \left(\frac{\frac{p_i}{1-p_i}}{\frac{v_i}{1-v_i}} \right) = \log \left(\frac{\frac{p_i}{\sum_{j=1}^k p_j}}{\frac{v_i}{\sum_{j=1}^k v_j}} \right) \text{ for } j \neq i. \quad (3)$$

Then, let \mathbf{a} be the vector of k party-specific measures of bias A'_1, A'_2, \dots, A'_k for a pre-election poll conducted by a pollster. A'_i retains the interpretation of A , and for the two-party case, the absolute values of A'_i and A are identical. Positive values indicate that a poll overestimates support for party i , whereas negative numbers show that the poll is biased against i . If the poll is in perfect agreement with the result of the actual election, all A'_i s are zero.

In most cases, however, “a survey sample is treated as a single realization of a random process that under essentially identical conditions could have produced an infinite number of similar but slightly different samples” so that it can be considered as a measure of public opinion at a point in time, revealing vote intentions. Thus, “ A and A'_i are just estimates for the true systematic bias that results from house effects, social desirability, or real changes in the population after the poll was taken. As a consequence, the question of the precision (standard error) of these estimates is crucial”, [1].

⁵The natural logarithm is denoted by “log” instead of “ln”.

Take, for example, a poll by KONDA⁶ for the Presidential election on June 24, 2018. Then, the corresponding row vector⁷ $\mathbf{a}_{KONDA} = (A'_{Erd.} \ A'_{Ince} \ A'_{Aks.} \ A'_{Dem.} \ A'_{Kar.} \ A'_{Per.})$ becomes

$$(-0.028 \quad -0.13 \quad 0.37 \quad 0.076 \quad -0.24 \quad -0.69) ,$$

where each coordinate shows a specific bias with respect to the candidates (Erdoğan İnce Akşener Demirtaş Karamollaoğlu Perinçek). For comparison, the traditional poll error for KONDA is found by subtracting the election outcome from the announced poll numbers, which is

$$(-0,69 \quad -2,64 \quad 2,91 \quad 0,6 \quad -0,19 \quad -0,1) .$$

Whereas A'_i captures party-specific bias, applied researchers will also be interested in overall measures of (in)accuracy, so Arzheimer and Evans (2014) therefore propose a composite measure B , which is simply the average of the absolute values of the individual A'_i 's, as well as a weighed alternative measure B_w , which additionally considers the parties' respective electoral shares v_i , [1].

$$B = \frac{\sum_{i=1}^k |A'_i|}{k}$$

$$B_w = \sum_{i=1}^k v_i \times |A'_i|$$

Taking absolute values before averaging is necessary because positive and negative bias components would otherwise cancel each other out. As it can also be used to compare the performance of individual polling firms, we are to analyze a number of individual-level survey outcomes from a number of pre-election polls that had been conducted by several different polling firms throughout various election cycles by making use of different combinations of measures of bias as introduced above to provide a performance ranking for the pollsters.

Table 1 comprises all the publicly available pre-election polls announced between May 15th and June 24th, 2018⁸ – ranked according to the B_w scores, from better to worse. Table 2 comprises the poll numbers from which the party-specific biases with respect to measure A'_i are calculated - in order of date⁹.

⁶conducted between June 9-10, 2018.

⁷rounding to two significant figures in the scientific notation.

⁸As there are pollsters that conducted several polls during the period, the table depicts the last poll of those particular pollsters. The dataset is compiled from [3].

⁹Calculating similar tables of measures of bias for every contesting party in each election cycle is certainly informative; however, one may want to quantify the extent a pollster is biased, and whether this bias is statistically significant. Assessing the statistical significance of any differences between sample and population is straightforward using a goodness of fit test when the sample is in fact a random sample. When it is not, there are no suitable test statistics for summarizing a poll's total bias with clear interpretation.

Pollster	N	Date	Erdogan	Ince	Aksener	Demirtas	Other	B_w
ORC	4250	18.06.2018	0,013	-0,076	0,160	0,023	0,314	0,047
A&G	5054	22.06.2018	0,045	-0,081	0,095	0,061	-0,795	0,069
Foresight	500	9.06.2018	-0,063	-0,033	0,082	0,257	-0,611	0,077
GENAR	5000	24.06.2018	-0,035	-0,115	0,026	0,257	0,607	0,084
Argetus	2360	16.06.2018	-0,043	-0,115	0,095	0,191	0,750	0,089
Optimar	2476	17.06.2018	-0,031	-0,135	0,148	0,235	0,380	0,092
SONAR	3000	31.05.2018	-0,163	0,028	0,270	-0,016	0,875	0,125
SAROS	No	23.06.2018	-0,019	-0,220	0,392	0,278	-0,795	0,139
Mediar	2401	9.06.2018	-0,177	-0,043	0,228	0,222	0,963	0,152
Gezici	1812	17.06.2018	-0,167	-0,081	0,473	0,202	0,169	0,166
Vera	2419	8.06.2018	-0,159	-0,095	0,360	0,378	-0,321	0,174
MAK	5400	4.06.2018	-0,035	-0,321	0,578	0,061	0,952	0,176
TÜSIAR	5561	12.06.2018	-0,067	-0,294	0,503	0,169	1,023	0,188
KONDA	No	13.06.2018	0,045	-0,376	0,443	0,278	0,169	0,197
Metropoll	2888	10.06.2018	-0,127	-0,228	0,306	0,413	0,943	0,204
Konsensus	1880	15.05.2018	-0,179	-0,195	0,774	0,169	-0,321	0,229
REMRES	5674	10.06.2018	-0,353	-0,061	0,560	0,349	1,089	0,286
AKAM	No	22.06.2018	-0,316	-0,120	0,666	0,359	0,704	0,290
Piar	2464	20.06.2018	-0,398	-0,180	0,917	0,309	0,914	0,367

Table 1: Publicly available pre-election polls announced between May 15th and June 24th, 2018 – in the order of B_w ’s.

Pollster	N	Date	Erdogan	Ince	Aksener	Demirtas	Other	B_w
GENAR	5000	24.06.2018	51,50	28,40	7,60	10,50	2,00	0,084
SAROS	No	23.06.2018	51,90	26,30	10,60	10,70	0,50	0,139
A&G	5054	22.06.2018	53,50	29,10	8,10	8,80	0,50	0,069
AKAM	No	22.06.2018	44,50	28,30	13,50	11,50	2,20	0,290
Piar	2464	20.06.2018	42,50	27,10	16,70	11,00	2,70	0,367
ORC	4250	18.06.2018	52,70	29,20	8,60	8,50	1,50	0,047
Optimar	2476	17.06.2018	51,60	28,00	8,50	10,30	1,60	0,092
Gezici	1812	17.06.2018	48,20	29,10	11,40	10,00	1,30	0,166
Argetus	2360	16.06.2018	51,30	28,40	8,10	9,90	2,30	0,089
KONDA	No	13.06.2018	53,50	23,40	11,10	10,70	1,30	0,197
TÜSIAR	5561	12.06.2018	50,70	24,90	11,70	9,70	3,00	0,188
Metropoll	2888	10.06.2018	49,20	26,15	9,82	12,06	2,78	0,204
REMRES	5674	10.06.2018	43,60	29,50	12,30	11,40	3,20	0,286
Mediar	2401	09.06.2018	47,96	29,88	9,15	10,18	2,83	0,077
Foresight	500	09.06.2018	50,80	30,10	8,00	10,50	0,60	0,152
Vera	2419	08.06.2018	48,40	28,80	10,30	11,70	0,80	0,174
MAK	5400	04.06.2018	51,50	24,40	12,50	8,80	2,80	0,176
SONAR	3000	31.05.2018	48,30	31,40	9,50	8,20	2,60	0,125
Konsensus	1880	15.05.2018	47,90	26,80	14,80	9,70	0,80	0,229
Election Result		24.06.2018	52,59	30,64	7,29	8,4	1,08	-

Table 2: Poll numbers as originally reported (other than the B_w column) - in order of date.

1.3 Distribution of B_w and candidate-specific biases

Recall that the smaller the realization of the B_w measure, the better the predictive accuracy of the pollster. What we see below is a presentation for the empirical distribution of the B_w -scores given in Table 1.

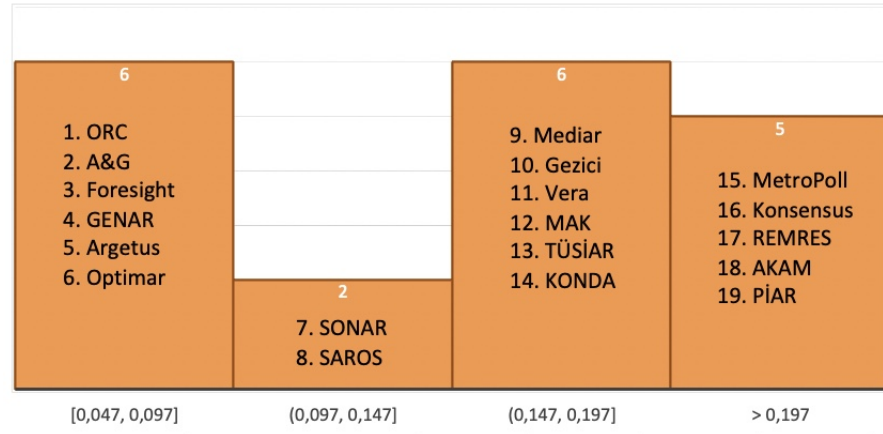


Figure 1: Histogram for the B_w -scores in the 2018 Presidential election.

What we have in Figure 2 is the distribution of candidate-specific biases in 2018 Presidential election. The thin green column in the middle of the figure stands for the perfect agreement between a poll's measurement of a candidate's vote and the election outcome.

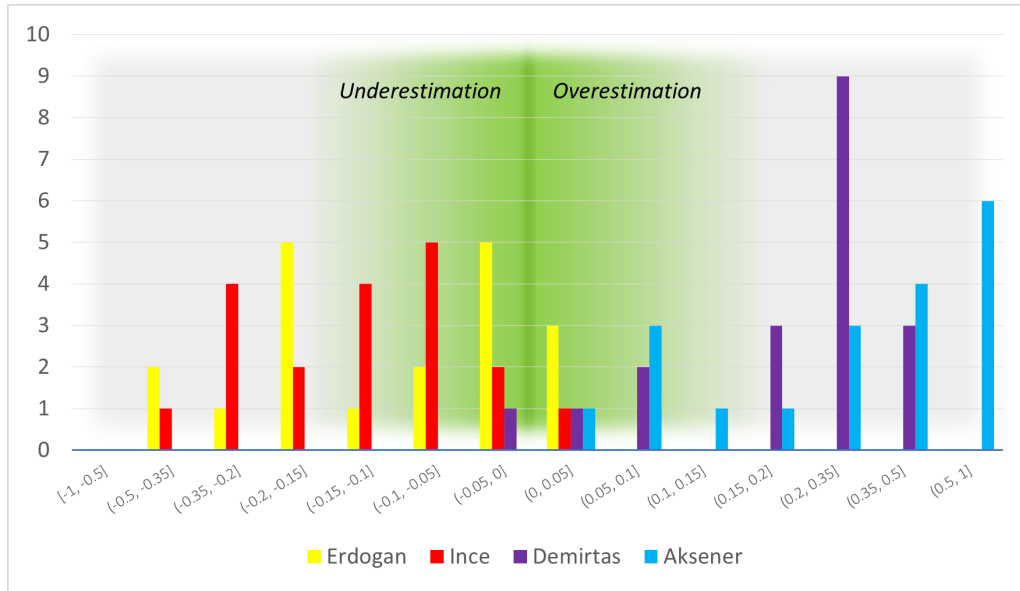


Figure 2: Histogram for the candidate-specific biases.

For instance, we observe a positive bias towards Akşener so consistent as one might call it a systemic overestimation since an overestimation of this size (> 0.35 , of which 10 out of 19 pollsters measured) affects the (predicted) vote distribution of the rest considerably. Additionally, we see that the polling firms tended to underestimate vote shares of Erdoğan and İnce. One might also be interested in biases for a particular subset of candidates (namely, for example, Erdoğan and İnce) or for all candidates via the composite measure of bias B_w (as stated previously).

1.4 Related to the graphs: Erdoğan vs (?)

We combine performances of the polling firms in 2019 Municipality elections (based on the usual measure B_w) with the scores acquired from the 2018 Presidential election in order to calculate the average weights for the polling firms announcing numbers today so that we can draw a poll aggregation graph, weighted based on the previous successes of the firms. If a firm is new, that is, it has never published a measurement for any previous election, then we give that pollster a weight as high as the lowest one in our sub-sample of pollsters, so that the firm is not punished for only being new¹⁰.

Municipality	Pollster	N	Date	AK Party	CHP	Other
Istanbul	SAROS	no	25.03.2019	-0,141	0,193	-0,714
	AREA	3320	14.03.2019	-0,024	0,061	-0,442
	ORC	no	24.03.2019	0,040	-0,063	0,206
Istanbul (repeat)	AREDA	9317	1.06.2019	0,113	-0,129	0,410
	KONDA	3498	16.06.2019	0,004	-0,004	0
	MetroPoll	1613	15.06.2019	-0,044	0,012	0,701
	ORC	4368	19.06.2019	0,193	-0,193	0
	SAROS	3000	14.06.2019	0,033	-0,045	0,321
Ankara	AREA	2060	14.03.2019	-0,045	0,035	0,123
	SAROS	no	25.03.2019	-0,081	0,107	-0,412

Table 3: Candidate-specific biases from the 2019 Municipality elections.

Pollster	Avg. B_w
KONDA	0,101
MetroPoll	0,118
AREA	0,1
SAROS	0,116
AREDA	0,164
ORC	0,098
Yöneylem	0,204
Sosyo Politik	0,204

Table 4: The average B_w 's for the polling firms. Do not forget that a score closer to zero is exponentially better.

¹⁰In this case, the pollsters Yöneylem and Sosyo Politik (as well as other polling organizations unlisted) are given the lowest weights, which is based on the score of MetroPoll in the 2018 Presidential election since it is MetroPoll that is the least succesful (amongst our subgroup of pollsters given in Table 4) with respect to the measure B_w .

References

- [1] Kai Arzheimer and Jocelyn Evans. A new multinomial accuracy measure for polling bias. *Political Analysis*, 22(1):31–44, 2014.
- [2] İrem Aydaş. Pre-election polls in Turkey. Unpublished master’s thesis, 2020.
- [3] İrem Aydaş. Replication data for: Pre-election polls in Turkey, 2021.
- [4] Paul P. Biemer. Total survey error: Design, implementation, and evaluation. *The Public Opinion Quarterly*, 74(5):817–848, 2010.
- [5] Andrew Gelman. Struggles with Survey Weighting and Regression Modeling. *Statistical Science*, 22(2):153 – 164, 2007.
- [6] Andrew Gelman. Failure and success in political polling and election forecasting. *Statistics and Public Policy*, 8(1):67–72, 2021.
- [7] Robert M. Groves and Lars Lyberg. Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*, 74(5):849–879, 01 2010.
- [8] Elizabeth A. Martin, Michael W. Traugott, and Courtney Kennedy. A review and proposal for a new measure of poll accuracy. *The Public Opinion Quarterly*, 69(3):342–369, 2005.
- [9] Josh Pasek. Predicting Elections: Considering Tools to Pool the Polls. *Public Opinion Quarterly*, 79(2):594–619, 04 2015.
- [10] Herbert F. Weisberg. *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press, 2009.