TUNISIAN ELECTION DATA FILES

This dataset was created by Julia Clark, Alexandra Blackman, and Aytuğ Şaşmaz as part of a project on the 2018 municipal elections. It contains results from nearly all of Tunisia's post-revolution nation-wide elections until the 2021 suspension of parliament, including the 2011 National Constituent Assembly Elections, 2014 Presidential Elections (rounds 1 and 2), 2014 Parliamentary Elections, 2018 Municipal Elections, and 2019 Parliamentary Elections. Not included are the 2019 Presidential results, which had not yet been released in early 2020 when the dataset was constructed.

The dataset includes cleaned elections results compiled from the Tunisian Electoral Commission (ISIE) and other sources, that are reconciled across years and polling stations, aggregated to the municipality (municipalities from 2018 to 2023), and with additional coded and calculated variables. This file describes the contents of the dataset, as well as the data sources used to create it, along with specific information on missingness and other limitations for specific years or locations.

The dataset and related files are provided here for public use with attribution, under the Creative Commons Attribution 4.0 International License. For any questions about the data (and particularly if you find any issues that are not documented below), please contact Julia Clark.

Suggested citation: Clark, Julia M., Blackman, Alexandra D., and Şaşmaz, Aytuğ. 2024. "What Men Want: Politicians' Strategic Engagement With Gender Quotas." ***Comparative Political Studies*****.**

VARIABLES

Variables in: mun_new_list_all.xlsx (see attached codebook for more details)

'year' = year of election
'dis_en' = electoral district **[not included in municipal election data]**
'gov_en' = governorate
'mun_uid' = unique ID for 2018 municipalities
'mun_en_2018' = name of 2018 municipality
'registered_voters' = number of registered voters
'valid_ballots' = sum of list, blank, and cancelled ballots (should be the same as 'total_voters')
'total_voters' = total number of voters who cast a valid ballot
'list_ballots' = ballots cast for a list or candidate (i.e., not blank or cancelled) **[2011 and 2014 only, but calculated for 2018 and 2019]**
'cancelled_ballots' = ballots cancelled by voters **[2011, 2014 parl only]**
'blank_ballots' = ballots cast for no list/candidate (protest vote) **[2011, 2014 parl only]**
'list_type' = type of list/candidate (party, coalition, or independent)
'list_name_en' = list name in English; for 2011, this appears to have been Google translated by Democracy International (DI), and we have made some corrections; for 2014-2019, we have translated or transliterated names for party and coalition list, but not independent lists
'list_name_ar' = list name in Arabic
'list_short' = abbreviation for the list

'list_order' = order in which the list appeared on the ballot; varies by electoral district/municipality [missing for 2019 parl]
'votes' = number of votes a list/candidate received
'cand_name_en' = Candidate name for presidential elections **[2014 prez only]**

'reg_men' = number of registered voters (male) **[2018 and 2019 (calculated) elections only]**
'reg_women' = number of registered voters (male) **[2018 and 2019 (calculated) elections only]**
'reg_18to35' = number of registered voters (between 18 and 35) **[2018 elections only; note age group differs from 2019]**
'reg_36to45' = number of registered voters (between 36 and 45) **[2018 elections only; note age group differs from 2019]**
'reg_46to60' = number of registered voters (between 46 and 60) **[2018 and 2019 (calculated) elections only]**
'reg_over60' = number of registered voters (over 60) **[2018 and 2019 (calculated) elections only]**

'reg_18to25_f' = number of registered voters (between 18 and 25 & female) **[2019 elections only]**
'reg_18to25_m' = number of registered voters (between 18 and 25 & male) **[2019 elections only]**
'reg_26to45_f' = number of registered voters (between 26 and 45 & female)* **[2019 elections only]**
'reg_26to45_m' = number of registered voters (between 26 and 45 & male)* **[2019 elections only]**
'reg_45to60_f' = number of registered voters (between 45 and 60 & female)* **[2019 elections only]**
'reg_45to60_m' = number of registered voters (between 45 and 60 & male)* **[2019 elections only]**
'reg_over60_f' = number of registered voters (over 60 & female)
'reg_over60_m' = number of registered voters (over 60 & male)

'turn_men' = percent of voters who were men **[2014 parl and prez only]**
'turn_women' = percent of voters who were women **[2014 parl and prez only]**
'turn_18to21' = percent of voters who were 18 to 21 **[2014 parl and prez only]**
'turn_22to30' = percent of voters who were 22 to 30 **[2014 parl and prez only]**
'turn_31to40' = percent of voters who were 31 to 40 **[2014 parl and prez only]**
'turn_41to50' = percent of voters who were 41 to 50 **[2014 parl and prez only]**
'turn_51to60' = percent of voters who were 51 to 60 **[2014 parl and prez only]**
'turn_61to70' = percent of voters who were 61 to 70 **[2014 parl and prez only]**
'turn_over70' = percent of voters who were over 70 **[2014 parl and prez only]**

*Note that for 2019, the ISIE voter registration data included the categories "26to45" and "45to60"; we do not know which category includes those age 45.

Variables in: parties_rep.xlsx

'list_short' = abbreviation for the list
'list_name_en_master' = list name in English; for 2011, this appears to have been Google translated by DI and we have made some corrections; for 2014-2019, we have translated or transliterated names for party and coalition list, but not independent lists
'list_name_ar_master' = list name in Arabic
'rcd' = indicator for whether party is RCD successor party [author coded; see file for included parties]
'islamist' = indicator for whether party is an Islamist party [author coded; see file for included parties]

'group_fp' = indicator for whether a party was ever a member of the Popular Front (FP) coalition [author coded; see file for included parties]
'group_destourian' = indicator for whether party was a member of a destourian coalition and/or merged into other destourian parties [author coded; see file for included parties]
'group_cpr' = indicator for whether part was a CPR splinter group or coalition [author coded; see file for included parties]

'mun_2018' = indicator for whether the party ran in the 2018 municipal elections
'nca_2011_seats' = number of seats won by the party in 2011
'parl_2014_seats' = number of seats won by the party in 2014
'parl_2019_seats' = number of seats won by the party in 2019
'total_seats_2011_2019' = total national seats held between 2011 and 2019

Variables in: population_mun_2014.xlsx

'mun_uid' = unique ID for 2018 municipalities
'tot_pop' = total population (based on 2014 census estimates)
'tot_pop_m' = total male population (based on 2014 census estimates)
'tot_pop_f' = total female population (based on 2014 census estimates)

Variables in: mun_id.xlsx
'u_id' = unique ID for 2018 municipalities
'seats' = number of municipal council seats in a given municipality (2018)

DETAILS OF DATA COLLECTION BY ELECTION YEAR

This section provides more information on data sources, as well as known issues with specific years, variables, and observations.

2011 National Constituent Assembly Elections

The original returns from the 2011 elections are no longer publicly available from the ISIE website (http://www.isie.tn/resultats/election-2011/). Instead, we use a dataset collected by Democracy International (DI) and Mourakiboun (`nca2011.csv`), described here: https://github.com/fwelections/tunisiaelectiondata.

Missing sector information
Overall, there are 168 out of 2084 sectors missing election data for 2011, 68 of which have results from other years.

The `nca2011.csv` data does not include a separate field for sector and has 305 centers without sector indicators in their names that cannot be matched to other centers in the master polling center list (i.e., they were not used in subsequent elections). These are nearly all secondary schools or colleges that also have asterisks next to their names (... اعدادية م*) and have higher "center numbers" (e.g., 91, 92, ... instead of 1, 2, ...), so we speculate that they were extra or "at-large" voting centers within each district.

Because they cannot be identified in the list of known polling centers, we cannot map them to specific sectors, and the 2011 results by sector (e.g., `2011_sec_list.xlsx`) are missing these results. However, to include these results at the 2018 municipal level, we have aggregated the totals for these centers to the delegation level, and then either:

1. Used these as the municipal totals, when delegation boundaries are equal to the municipal boundaries, OR
2. For delegations that are split across municipalities, divided the delegation total by the number of municipalities and added these to the municipal totals. This assumes that votes within a "split" delegation were equally distributed across all municipalities within the delegation boundaries.

Because we round the totals after dividing delegation totals among municipalities (many of which would have been less than one), this means that the totals summed across municipalities are slightly higher than the totals summed across centers (e.g., `sum(ctr$registered_voters)` = 8291028, while `sum(mun$registered_voters)` = 8292048, a difference of 1020 of 0.01%).

2014 Parliamentary Elections (Author data entry)

Unlike the 2018 and 2019 results, ISIE did not make the 2014 parliamentary results available in excel files; rather the website (http://www.isie.tn/resultats/proces-verbaux-legislatives/) contains PDF images of the handwritten reports from each polling center desk. We have therefore constructed the results by hand from these PDFs with the help of research assistants Houda Ould Khal, Karim Boudlal, Khaled Ferjani, and Manaa Lajnef.

Missing center-desks

A total of 347 desks out of 10,567 are missing from the data, which translates into missing data from 184 out of 4534 polling centers, assuming the same number of centers as the round 1 presidential elections. PDFs for these desks are not available on the ISIE website. A majority of these desks are in:

- *Gafsa*: 153 missing, including all of Redeyef, Sned, El Ksar, El Guetar, and Mdhila delegations
- *Ben Arous*: 56 missing, including some centers and selected desks in La Nouvelle Medina, Bou Mhel El Bassatine, Hammam Lif, Mohamedia, Mornag, and Megrine
- *Kebili*: 53 missing, including all of Douz Nord and Faouar delegations
- *Bizerte*: 31 missing, including some centers and selected desks in Utique, Bizerte Sud, Bizerte Nord, Tinja, Djoumine, Sedjnane, Mateur, and Menzel Bourguiba

Note that in a few cases, the PDFs appear to exist, but are mislabeled (e.g., in Ariana Ville, Ennasr 2, there is a PDF for desk 04010801902, which labeled as desk 2 but is actually desk 3).

For 58 of the missing desks (mostly of those missing in Kebili), we are able to use the DI data described below, although this is limited to results from the top 5 parties, so results for smaller parties and independent lists are not available.

2014 Parliamentary Elections (DI)

DI also aggregated 2014 parliamentary results from ISIE (see https://github.com/fwelections/tunisiaelectiondata). These were used for this dataset to provide a cross-check, and to fill in totals for the major parties at 223 of the polling center desks where the ISIE PDFs are now missing (see above).

However, the DI dataset has certain limitations that prompted the author's collection effort described above:

1. Data was only compiled for those parties that got more than ~3% of the vote (Ennahdha, Nidaa Tounes, Afek Tounes, Popular Front, and UPL). Together, these parties captured approximately 75% percent of the votes cast, but do not allow for more detailed examination of the number of votes cast for independent lists or smaller parties.
2. There are 1434 missing polling-center-desks, resulting in missing data for 29 municipalities in Sfax and Kairouan
3. 21 center-desks are missing vote totals for Nidaa (and it appears based on total list votes that these would have been significant)
4. At center-desk 17100916901, totals are missing for all parties
5. Totals for the Afek Tounes list were reported in the Kebili district, where the party did not run

Note that there are 10 center-desks where the calculated sum of votes for each list is greater than the list_ballots total variable in the dataset; four of these are likely to be transcription errors and have been corrected in our 2014 data:

- 12010502006: should be 453 instead of 253
- 09040204502: should be 392 instead of 192
- 01060506205: should be 375 instead of 275
- 11010200701 should be 291 instead of 201

Six others do not appear to be simple transcription errors and have been left as-is: 12030103801, 08100115601, 01110710301, 03100109706, 01110509806, 01010200503.

2014 Presidential Elections

Presidential round 1 (R1) and round 2 (R2) data are imported directly from delegation-level ISIE excel files (http://www.isie.tn/resultats/resultats-presidentielles) with the exception of Kelibia and Sbikha, which the authors compiled from ISIE's posted PDFs.

List ballots
The ISIE excel files have a column called "مجموع الأصوات حسب المكتب" or "Total votes by office". Because this is always equal to the sum of candidate votes it has been coded as `list_ballots`.

Valid ballots
We did not have total numbers of blank or cancelled ballots for the presidential elections at the time of data collection, so where needed (e.g., when calculating effective number of lists, or ENL), list ballots is used as a denominator instead.

2014 Turnout Data

Demographic turnout variables
These come from `Stats_formatted.csv`, which is DI's cleaned version of ISIE's `statistiqueparticipation.xlsx`). Note that these are NOT the percent turnout of the specific group (e.g., how many registered women voted), but rather the percent of all voters in each demographic category (e.g., percent of all voters who were women). Because these numbers are reported as percentages at the polling center level, they are averaged when aggregating by sector and municipality.

Calculated turnout
For presidential elections R2, turnout as a percent of registered voters (`turnout_per`) is calculated using registered voters from the parliamentary elections as a denominator, as this is not separately available for the presidential data but should be the same.

2018 Municipal Elections

This data comes directly from the ISIE excel files (called "preliminary results"), available at http://www.isie.tn/elections/elections-municipales-2018/resultats/, with the exception of Mdhila (see below).

Mdhila
Mdhila municipality in Gafsa was missing from the ISIE preliminary results at the time of data collection. We created a separate excel file for Mdhila based on a PDF of the final results: https://www.isie.tn/wp-content/uploads/2018/06/%D8%A7%D9%84%D9%85%D8%B8%D9%8A%D9%84%D8%A9.pdf, which is merged in in `1_clean_2018.R`. Because these results are at the municipal level, data for Mdhila is missing in center and sector-level files, as well as those for old municipalities (which require sectors to calculate).

Registered voters

This is from the 2018 turnout file (`registered_2018.xlsx`) posted by ISIE and is available at the municipal level only.

Extra list
There is one list (مشروعنا) in Agareb that appears in the election data but does not have an ISIE list number. For this, we have assigned an arbitrary number for list_number (999999999).

Missing results
There is one list in Carthage (الفينيق) with list_id (20602179) that appeared in the ISIE file of approved lists but has no results. This list is dropped.

2019 Parliamentary Elections

Data for 2019 comes from the ISIE website (http://isie.tn/ar/2019-التشريعية-الانتخابات-نتائج/) and was downloaded in PDF and converted to excel using Tabula, an open-source software.

Registration data
The number of registered voters for the presidential elections and the demographic breakdowns by age/gender (should be the same for parliamentary) is available only at the delegation level. We divide them across sectors based on the percent of population in the delegation to have this variable at the sector and municipal level.

There are a few cases where the number of registered voters in a sector exceeds the total population, or where the number of list ballots cast exceeds the number of registered voters. This is likely due to (a) changes in the distribution of the population across the delegation since the 2014 census (or errors in the census data), and/or (b) people being assigned/registered at polling stations other than those in the sectors where they actually live. Calculated fields based on registration numbers (e.g., `turn_per`) are set to NA where this is the case.

Note the ISIE voter registration data included the categories "26to45" and "45to60"; we do not know which category includes those age 45.

Missing votes for certain lists
There are a few lists that are missing votes because they were not in the ISIE PDF results at the time of data collection, including:

- El Ksour (Kef), desk 07040606401, votes for Ennahda
- Hammam el Onf (Ben Arous), desk 03070507802, votes for independent list "من أجل تأسّس جدّد"
- Sousse Jawhoura (Sousse), desk 20100210506, votes for independent list "كائمت حسب جيّار المحبت"
- Sousse Jawhoura (Sousse), desk 20100511305, votes for independent list "كائمت إذا الشعب يوما أراد الحياة"
- Kesra (Siliana), desk 8100716601, votes for independent list "كائمت الامتياس"
- Kelibia (Nabeul), desk 11070511401, votes for Hezb el Amal
- Kelibia (Nabeul), desk 11070310801, votes for independent list "فلوكة"

Missing Sfax 2 centers
There are 12 polling centers in Sfax 2 that are listed in the ISIE polling center list but do not have results (unclear whether these are missing from ISIE PDFs or whether these centers were simply not used, although likely the former as these are in the highly populated area of Sfax municipality):

- Sfax South: center 62 (م إبتدائية مركز والي - الأفران الشمالية)
- Sfax Medina: centers 103, 107, 109, 111, 112, 114, 115, 116, 117, 120, 121, 122

Turnout variable inconsistencies
Due to errors in either the numerator and/or denominator, there are some cases where turnout variables (`turnout_per`, `turnout_per_pop`, `turnout_per_list`, `turnout_per_pop_list`) are greater than 100% in the sector and municipal-level data. These values are set to `NA`.

At the sector level, there is also some missing turnout data due to errors:

- *Ariana:*
    - L'Ariana Ville, Ariana Ville
    - El Menzah VI, Ariana Ville
    - El Yasmina, Ariana Ville
    - Ennahli, Kalaât El Andalous
    - El Ghazala, Raoued
- *Beja:*
    - Medjez El Bab Ville, Medjez El Bab
- *Tunis:*
    - El Taoufik, La Medina

In addition, one sector in the southern-most part of **Tatouine** (Borj El Khadhra, Remada, `sec_uid` = 53571912) in the Saharan desert with a census population of 78 had no ballots cast in 2018 and 2019 and a turnout of 0 for these years.

HOW WERE POLLING STATIONS MATCHED ACROSS ELECTIONS?

We created a master file for all polling stations to help reconcile and merge election data across years and datasets. Because the file generates a unique center code that is used in other files, it can only be updated manually.

Center numbers
For the 2014 elections onwards, ISIE published an 11-digit syntax to identify each polling center and desk that uses the following structure (from left to right):

- Electoral district (دائرة انتخابية, *circonscription*): 2 digits
- Delegation (معتمدية, *delegation*): 2 digits
- Sector (عمادة, *secteur*): 2 digits

- Polling center (مركز الاقتراع, *centre de vote*): 3 digits
- Desk (مكتب الاقتراع, *bureau de vote*): 2 digits

These codes are used in the PDF vote tabulations completed for each desk (e.g., http://www.isie.tn/wp-content/uploads/filebases/pv-elections-legislatives/Tunisie/Tunis1/Sijoumi/E.P.2%20MARS%2034%20MALLASSINE%20-ALI%20BELHOUANE/01030102901.pdf) and reported in the excel sheets available from ISIE.

For example, center 01010100101 is:

- Electoral district: 01 (Tunis)
- Delegation (معتمدية, delegation): 01 (Hraira)
- Sector: 01 (Hraira)
- Polling center: 001 (Primary school Zouhour 4 Prolongee)
- Desk: 01

These codes appear in the 2014 files, and in some of the 2018 and 2019 ISIE files, although not always consistently. However, they are not in the 2011 data. To merge data on polling stations across sources and years, we developed a master list of polling stations based on station name in Arabic and English and sector/delegation location.

Reconciling 2011-2014 center names

Most differences in polling station names between 2011 and 2014 are due to the sector-name suffixes that follow the name of the polling station (which in most cases, is a school). These were double-checked for matching errors. In most cases, the name of the sector embedded in the polling station name from the 2014 data is the same as the 2018 and 2019 data, so these have been used at the unique name, overriding any differences in 2011 polling station names.

For example, centers 15010200401-4 in Kasserine's Ezzouhour Ouest have the name "E.P.OULED BOU ALLEG EZZOUHOUR OUEST2) إبتدائية أولاد بوعلاق - الزهور الغربي" in the 2011 data, while the 2014 data give the name as "م إبتدائية أولاد بوعلاق – الزهور الشرقي" (Ouled Bou Alleg Ezzouhour Est).

Incorrect districts in 2018 ISIE polling centers

There are 49 centers in the 2018 ISIE center list that have incorrect districts (Tunis 1 instead of Tunis 2, and Sfax 2 instead of Sfax 1):

- Centers in Tunis 1 in 2018, should be Tunis 2:
  - Cité El Khadhra
  - El Menzah
  - El Omrane
  - El Omrane supérieur
  - Ettahrir
- Centers in Sfax 2 in 2018, should be in Sfax 1:
  - Sakiet Eddaïer