This dataset contains text transcripts of 2,005 TV campaign advertisements (part of the Free Electoral Advertisement Period -- *Horário Gratuito de Propaganda Eleitoral*) of every presidential candidate in Brazil from 1989 to 2022.

We combined the recordings that were collected and shared by DOXA (Laboratório de Estudos Eleitorais, em Comunicação Política e Opinião Pública. IESP/UERJ) for the 1989, 1994, 1998, 2002, 2006, and 2010 elections and those made available on YouTube by a variety of channels for the 2014, 2018, and 2022 elections (see references in `YouTube_Sources`).

We transcribed these recordings using the automatic transcription software Trint and then reviewed and edited these transcriptions to best match the original recordings. All transcripts are labeled and included in the `master_dataset` spreadsheet.

Files/Folders:

- `master_dataset`: dataset that lists all 2,005 advertisements labeled by candidate, round, and year. It includes each advertisement's full transcription. It is available in *.csv* (tab) and *.rds*. Column names are as follows:
  - `file_name`: id of each advertisement by candidate and year.
  - `year`: election year, either 1989, 1994, 1998, 2002, 2006, 2010, 2014, 2018, or 2022.
  - `round`: round in which advertisement was aired, either 1 or 2.
  - `candidate`: candidate name (either full name or first/last name only).
  - `text`: transcribed text from the advertisement.
- `Candidate_List.xlsx`: spreadsheet that lists all presidential candidates from 1989 to 2022. Column names are as follows:
  - `candidate`: candidate name (either full name or first/last name only).
  - `year`: election year, either 1989, 1994, 1998, 2002, 2006, 2010, 2014, 2018, or 2022.
  - `round`: round in which advertisement was aired, either 1 or 2.
- `YouTube_Sources.xlsx` list of all YouTube links used to collect the advertisements from 2014 to 2022. Column names are as follows:
  - `candidate`: candidate name (either full name or first/last name only).
  - `party`: candidate's party acronym.
  - `year`: election year, either 1989, 1994, 1998, 2002, 2006, 2010, 2014, 2018, or 2022.
  - `round`: round in which advertisement was aired, either 1 or 2.
  - `type`: type of advertisement, either "insercao" (30 sec ad aired throughout the day) or "programa" (longer ad section aired during a special advertisement twice a day).
  - `length`: duration of advertisement.

- `url`: YouTube link to video.