

TSE - Public Data

The folder `tse_public_data` contains instructions on how to get the public data available through the TSE website : <http://resultados.tse.jus.br>

- TSE stands for: "Tribunal Superior Eleitoral" ("Superior Electoral Court").

`cities-zones-json/*`

The directory `cities-zones-json/` contains all the information of cities and respective election zones and sections for all the states in Brazil.

These JSON files contain the needed information to request the election data of each voting machine, public available through the TSE website (<http://resultados.tse.jus.br>).

- `%UF` : stands for a Brazil state code in this document.
- `ufs.json` : list of all the states in Brazil.
- `zonas-%UF.json` : all the election zones of state `%UF` in Brazil.

Example:

For the file `zonas-ac.json` (state `AC`):

```
{
  "ac": [
    {
      "n": "01066",
      "m": "PORTO WALTER",
      "zs": [
        {
          "0004": [
            "0077",
            "0078",
            "0079",
            "0080",
            "0115",
            "0132",
            "0144",
            "0145",
            "0146",
            "0164",
            "0166",
            "0245",
            "0311",
            "0313",
            "0314",
            "0316",
            "0320",
            "0321",
            "0325",
            "0338",
            "0361",
            "0372",
            "0373",
            "0429",
            "0430",
            "0442",
            "0447"
          ]
        }
      ]
    }
  ],
  ...
}
```

- State: "ac"
- City code: "01066"
- City name: "PORTO WALTER"
- City election zones: ["0004"]
- Zone "0004" sections: ["0077", "0078" ... "0447"]

Requesting the Voting Machine public files

Each city/zone/section has a Voting Machine that generated a file with the results of the poll for the section. The result file is called "Boletim de Urna" (Ballot Box) and has the extension ".bu".

To request the ".bu" file of a section you need a hash number, that can be request through an auxiliar JSON:

```
// Dart code:
var uf = "ac"; // state code.
var cityCode = "01066"; // city code.
var zoneCode = "0004"; // zone code.
var sectionCode = "0077"; // section code.

var urlAux = "https://resultados.tse.jus.br/oficial/ele2022/arquivo-urna/406/dados/$uf/$cityCode/$zoneCode/$sectionCode/p000406-$uf-m$cityCode-z$zoneCode-
```

The data at `$urlAux` above is a JSON. Inside it, you can find a "hash" entry in HEX format that can be used to request the ".bu" file of the section.

```
// Dart code:
var uf = "ac"; // state code.
var cityCode = "01066"; // city code.
var zoneCode = "0004"; // zone code.
var sectionCode = "0077"; // section code.
var hash = "10111213141516171819a0a1a2a3a4a5a6a7a8a9b0b1b2b3b4b5b6b7b8b9c0c1c2c3c4c5c6c7c8c9d0d1d2d3"; // The BU hash code (not real example).

var urlBu = "https://resultados.tse.jus.br/oficial/ele2022/arquivo-urna/406/dados/$uf/$cityCode/$zoneCode/$sectionCode/$hash/o00406-$cityCode$zoneCode$sec-
```

With the correct URL it's possible to download the ".bu" file of a Voting Machine, containing the poll results of the election of the respective section.

This process is the same performed by the TSE website:<http://resultados.tse.jus.br>

BU Format

The TSE website provides a Zip file (`formato-arquivos-bu-rdv-ass-digital.zip`) that explains the format of a ".bu" file specified in ASN.1 (Abstract Syntax Notation One) and encoded in BER (Basic Encoding Rules).

Inside it there's some `python/*.py` scripts to read and validate a ".bu" file.

To export the data inside a ".bu" file you can use the script `bu_dump.py`.

`bu-files/` directory

To use this project a `bu-files/` directory (at `/tse_public_data/bu-files/`) is needed. It will contain all the ".bu" files to be analysed.

The files should be downloaded and saved in the structure below:

- `bu-files/$uf/$cityCode/$uf-$cityCode-$zoneCode-$sectionCode.bu`
- **File path example:**
 - `bu-files/ac/01066/ac-01066-0004-0077.bu`
- `$uf` : state code in Brazil.
- `$cityCode` : city code.
- `$zoneCode` : city electoral zone code.
- `$sectionCode` : zone section code.
- **NOTE:** the `bu-files/` directory will be mapped by the `bu-docker` container at `/bu-files`.

Compressed BU files

To facilitate statistics analysis it's provided a compressed version of the BU files JSON at `bu-files/bus-%UF.comp`. Each compressed file have all the ".bu" files information (in JSON format) for each UF (state in Brazil).

This compressed files can be generated/read using the tool `../bu_statistics/bin/bu_compressor.dart`.

BU Docker Image

To facilitate the execution of the `python/*.py` scripts a Docker image is provided at `bu-docker/Dockerfile` (based on the `python` image), that uses inside it the files at `bu-docker/docker-files/`.

A modified version of the `bu_dump.py` that can export the BU data to JSON is provided at `bu-docker/docker-files/python/bu_dump_json.py`.

To build the `bu-docker` image in your Docker environment:

```
cd bu-docker/
./build-bu-docker-image.sh
```

Bash scripts to use the `bu-docker` image:

After build the `bu-docker` image and have the `bu-files/` directory populated with ".bu" files you can use the `run-bu-dump-*.sh` bash scripts.

- `run-bu-dump-json.sh` :
 - Runs the `bu-docker` image and processes a city directory.
 - Mounts `./bu-files/` to `/bu-files` in the container.
 - Usage:
 - `./run-bu-dump-json.sh "/bu-files/$uf/$cityCode"`
 - Note that the directory argument above is the path inside the docker image (`/bu-files/$uf`).
 - Example:
 - `./run-bu-dump-json.sh "/bu-files/ac/01066"`
 - The execution will create a `.json.gz` file (JSON + Gzip) for each ".bu" file in the city directory.
 - Generated JSON file path example:
 - `bu-files/ac/01066/ac-01066-0004-0077.bu.json.gz`
 - NOTE:
 - After the execution the Docker container is automatically removed due the `docker run --rm` parameter.
- `run-bu-dump-uf.sh` :
 - Will process an entire UF (state) directory.
 - Example:
 - `./run-bu-dump-uf.sh ac`
 - `./run-bu-dump-uf.sh sp`
 - `./run-bu-dump-uf.sh rj`
- `run-bu-dump-uf-all.sh` :
 - Will automatically call `run-bu-dump-uf.sh` for all the states (UFs) in Brazil.

Public Data

These ".bu" files are public available data, provided to anyone that accesses the TSE website starting from 2022/10/02.

This document was written in 2022/10/07. It's possible that the access URL has changed after the release date of this project.