Analysing the results from the South African Elections in 2024

Table of Contents

[Table of Figures 3](#_Toc170416195)

[1. Find and critique a dataset 4](#_Toc170416196)

[1.1. Dataset Information 4](#_Toc170416197)

[1.2. Dataset Assessment 4](#_Toc170416198)

[1.2.1. Quality 4](#_Toc170416199)

[1.2.2. Detail 4](#_Toc170416200)

[1.2.3. Documentation 4](#_Toc170416201)

[1.2.4. Interrelation 5](#_Toc170416202)

[1.2.5. Use 5](#_Toc170416203)

[1.2.6. Discoverability 5](#_Toc170416204)

[1.2.7. Terms of Use 5](#_Toc170416205)

[1.3. Interest in the Dataset 5](#_Toc170416206)

[1.4. Research Questions 5](#_Toc170416207)

[2. Model the data 7](#_Toc170416208)

[2.1. E/R Model of the data 7](#_Toc170416209)

[2.2. Cardinality 9](#_Toc170416210)

[2.3. Database tables and fields 9](#_Toc170416211)

[2.4. Normalisation 9](#_Toc170416212)

[3. Create the database 10](#_Toc170416213)

[3.1. Database Structure 10](#_Toc170416214)

[3.2. Enter instance data 10](#_Toc170416215)

[3.3. Reflection 10](#_Toc170416216)

[3.4. List SQL commands 10](#_Toc170416217)

[4. Create a simple Web Application 11](#_Toc170416218)

[4.1. Node.js Module 11](#_Toc170416219)

[4.2. Web Application Screenshots 11](#_Toc170416220)

[References 12](#_Toc170416221)

# Table of Figures

[Figure 1: E/R diagram for the elections. 8](#_Toc170416166)

# Find and critique a dataset

There was a National election in South Africa recently. The media and political parties have made big predictions about what might happen after the election. The government decided to hold the elections on the 29th of May 2024. In South Africa the largest political party is the ANC and they are also the country's majority party for the last 30 years after Apartheid. The largest opposition party in South Africa is the DA and they are the only party to take control of the Western Cape province from the ANC.

Among the predictions that were made was that the ANC would lose their majority for the first time in history and that they would have to look to other parties to form a government. The DA is considered by the media and some as a white party and that the new black parties for the 2024 election such as ActionSA, Rise Mzanzi, and BOSA can steal more votes from the ANC and also from the DA. South Africa also has radical parties such as the EFF who want to nationalize everything and they are also considered a popular option. Just before the election, the former South African president founded his own party, the MK, and they are also considered a radical party that wants to remove the country's constitution.

## Dataset Information

In South Africa the Independent Electoral Commission (IEC) is an independent constitutional body that is responsible for all elections in the country. The IEC usually makes the election results available on their website where you can view the processed results, but they also make the raw data available which can be downloaded as a CSV file from their website.

The IEC collects election data for the National, Provincial, and Municipal elections. This data analysis project is only going to look at the National election data.

Dataset URL: <https://results.elections.org.za/home/downloads/npe-results>

This dataset contains the election data for South Africa's National election, it has the results for each party by municipal district, voting district, and the province. The dataset contains information such as how many voters are registered per voting district, spoiled votes, and total votes.

## Dataset Assessment

### Quality

The quality of the data is of a high standard as it comes directly from the IEC's website, which means that the data has been verified by the IEC.

### Detail

The dataset contains the basic information such as the registered voters, spoiled votes and how many votes each party received. The data is by each province, municipal district, and voting district.

### Documentation

There is not any documentation available about the data. The IEC only describes the following when selecting the data for download:

P = Province M = Municipality W = Ward B = Ballot VD = Voting District

### Interrelation

It may be possible to link this data with demographic data from StatSA to see how the demographics of the country voted for the different parties.

### Use

The data should be enough to see how the voters voted across the country and how the voters voted in each district. Are there any differences between the districts, certain areas have more spoiled votes than other areas. What does the political parties' support look like across the country.

The data does not contain any demographic data and will not be able to confirm how the demographic voted. It also does not contain any data on the unregistered voters, and it is unclear why this part of the population is not registered to vote and why they stay away.

### Discoverability

The IEC in South Africa is the only entity that has the election information and any other platform that has this information gets it from the IEC. The data is freely available on the IEC website.

### Terms of Use

According to the IEC website, their data is available for personal, educational, and non-commercial purposes. Any commercial use must request the IEC for permission and the IEC owns all intellectual property rights related to the data and content on their website.

## Interest in the Dataset

Elections in South Africa have always been an important aspect for South Africans because it is a time when every person in the country can vote after the end of Apartheid. It is 30 years later, and the political landscape is changing, making it important to analyze this data.

## Research Questions

Voter Turnout and Participation

What is the voter turnout in each voting district?

Calculation: Total Valid Votes / Registered Population.

How does voter turnout compare between urban and rural areas?

Urban vs. rural categorization based on municipality or voting station names.

What is the distribution of spoilt votes across different voting districts?

Analysis of the number of spoilt votes in each district.

Party Performance and Trends

Which party received the most votes in each municipality and province?

Aggregating votes by party within each municipality and province.

How did smaller parties perform compared to major parties in each region?

Analysis of vote shares for smaller parties vs. major parties.

What are the top three parties in each voting district based on the number of votes received?

Ranking parties within each district by votes received.

Regional and Temporal Analysis

What are the voting patterns across different provinces?

Comparative analysis of party votes across provinces.

How do voting results vary between different election years (if historical data is available)?

Trend analysis over multiple election cycles.

Are there any notable differences in voting behavior between weekdays and weekends (if election dates are available)?

Analysis based on the generated datetime.

# Model the data

## E/R Model of the data

Before an E/R diagram can be drawn, the data contained in the CSV file must first be looked at. It is extremely important to understand the content and what it means before we can ask questions and whether the content is good enough. The CSV file I downloaded from the IEC website contains the results for the National election and it contains all the provinces as a category in South Africa and there are a total of nine provinces in the country. South Africa’s nine provinces are the Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, Northern Cape, North West, and the Western Cape. This makes it possible to compare the results of each province with each other.

There is also a category for all the municipalities in the country and these municipalities can only belong to one province of the country. It is also possible that there is more than one province that has a similar municipal name, but this should not be a problem because the municipal names are already listed with the correct province name. Each municipality has a summary of different neighborhoods and these are classified as a voting district with a unique voting district number. We can use the municipal and voting districts to see where most of the votes came from in each province and also identify which areas have the most voters. The data also identifies how many voters registered in each voting district, how many of those votes were counted as spoiled votes, and how many of those voters actually went to vote in their voting district. This voting district data can help us determine what the voter turnout was and whether certain voting districts had a lower turnout. It will also be interesting whether spoiled votes are more in a voting district with a lower turnout.

The results also include the different parties for each voting district and how many of the votes they received. The party result should indicate to us how well each party did in the different districts. The last column of the CSV contains the final date that all the results were finalized by the IEC. We can use the date to indicate that these are indeed 2024 results and the database can then be used in the future to compare with previous elections or any future elections.

This data can be classified into different categories which can be combined to form the E/R diagram. Such a diagram contains entities, attributes, and relationships.

This E/R diagram contains the following:

* Entities
  + Provinces
  + Municipalities
  + Voting districts
  + Results
  + Parties
* Attributes
  + ProvinceID
  + Province\_name
  + MunicipalityID
  + Municipality
  + VotingID
  + VD\_number
  + VS\_name
  + Registered\_voters
  + Spoilt\_votes
  + Total\_votes
  + ResultsID
  + Party\_votes
  + PartyID
  + Party\_name
* Relationships
  + A province can have multiple municipalities, but a municipality can only have one province.
  + Municipalities can have multiple voting districts, but voting district can only have one municipality.
  + Voting districts can have multiple election results and election results can belong to one voting district.
  + A party can have many election results across different voting stations.

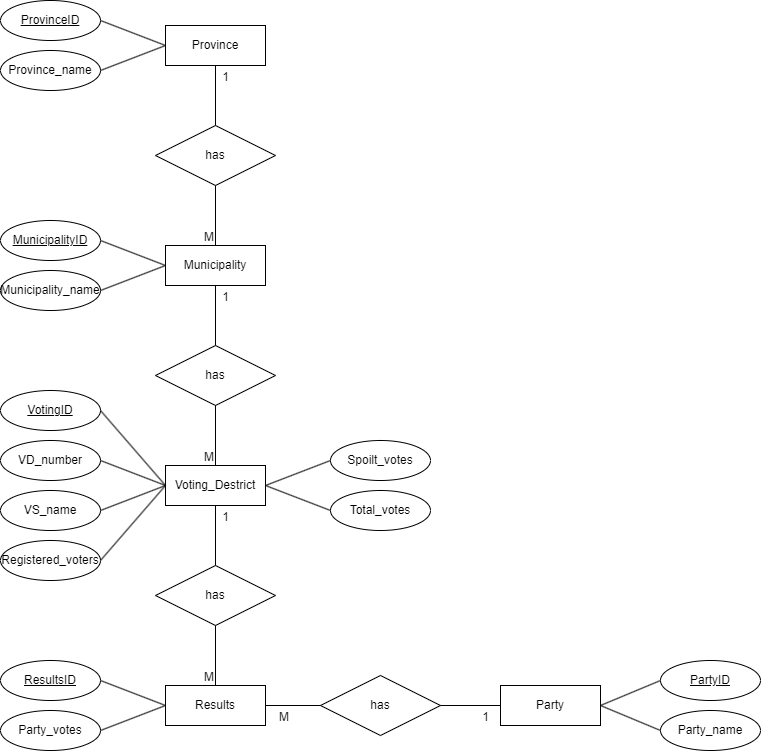


Figure : E/R diagram for the elections.

## Cardinality

## Database tables and fields

## Normalisation

# Create the database

## Database Structure

## Enter instance data

## Reflection

## List SQL commands

# Create a simple Web Application

## Node.js Module

## Web Application Screenshots

# References