MIS781 Business Intelligence and Database

**Example - Election Database Design and Report**

Group Assignment

This report has been prepared by:

GROUP Number

x, y, z Name of all students

TABLE OF CONTENTS

INTRODUCTION2

General Introduction2

Target Audiences2

Database Design Objectives2

Database Design Benefits2

Assumptions3

Attributes and Business Rules3

DATABASE DESIGNS6

Database 1: Voter Turnouts and Voting Results6

Master Table6

1NF Table6

2NF Table7

3NF Table9

Database 2: Candidate Details with Voting Updates11

Master Table11

1NF Table12

2NF Table13

3NF Table14

Database 3: Public Opinion Poll16

Master Table16

1NF Table16

2NF Table17

3NF Table18

REFERENCES19

# **INTRODUCTION**

# Introduction (2 marks)

## -Overall introduction of your report (Database for simulated phishing campaigns dashboard)

## -Mention your three main Target Audience

Example:

1. Phishing Campaign Manager
2. Individual User or Department Manager
3. HR Manager or CEO

# **Example:**

# **General Introduction**

The database designs presented in this report made use of presidential election data in the United States of America across different years. With the country’s electoral college process being unique, the election system can be deemed complicated and perplexing for most people. Hence, to assist different stakeholders in understanding the process, evaluating results, and making decisions on various aspects of the election setting, these database designs have been created.

# **Target Audiences**

* **Database 1:** Civil society
* **Database 2:** Political parties
* **Database 3:** Election campaign teams

# **Database Design Objectives**

* **Database 1:** In consideration of the civil society’s role in voter education and election monitoring (International Foundation for Electoral Systems et. al, 2003), the *Voter Turnouts and Voting Results* database design aims to provide easily accessible, perceptive, and reliable information on both voter participation of various segments of the electorate and official popular vote results in order to assist with the beneficiary’s voter engagement initiatives.
* **Database 2:** ….
* **Database 3:** …

# **Database Design Benefits**

* **Database 1:** With the *Voter Turnouts and Voting Results* database design, the civil society would have a clear picture of the turnout results of their segments of interest by finding all information needed in one place. This would accelerate preparations for voter education training sessions for recipients that are at risk of not voting due to lack of information, apathy, or sense of alienation, or those susceptible to manipulation (International Foundation for Electoral Systems et. al, 2003). Having the popular vote result together in the same database would also allow them to easily interpret the impact of voter turnout to actual voting results. Also, the presence of the details in each state’s voting-eligible population and registered voters would assist them in strategy and decision formulations. In addition, information on the political parties and their respective ideologies would provide a helpful tool in voter education. In this case, target setting, and information dissemination would be accurate, therefore making initiatives on increasing voting turnouts efficient and effective.
* **Database 2:** …
* **Database 3:** …

# **Assumptions**

**DATABASE 1**

* *Voting-eligible population* is composed of people who are above 18 years old and have the right and capacity to vote regardless of voter registration status.
* *Registered voters* are people who have officially registered themselves to vote.
* Data on *voting turnouts* and *popular votes* are official presidential election results data collected from valid resources.
* Data for *poll turnout* are collected from an official surveying body which conducts survey manually and electronically.
* *New voters* are people voting for the first time regardless of their registration date.
* *Ethnic minorities* are ethnic groups not belonging to the following groups: White/European American, Black/African American, Alaska Native/American Indian, Hispanic/Latino, Asian American, Hawaiian Native
* *Remote communities* are towns which populations are less than 15,000.
* *Candidates* are official political party presidential representatives that have won the preliminary elections.
* *Ideology* refers to the political party’s current ideology and can change over time.
* *……*

**DATABASE 2**

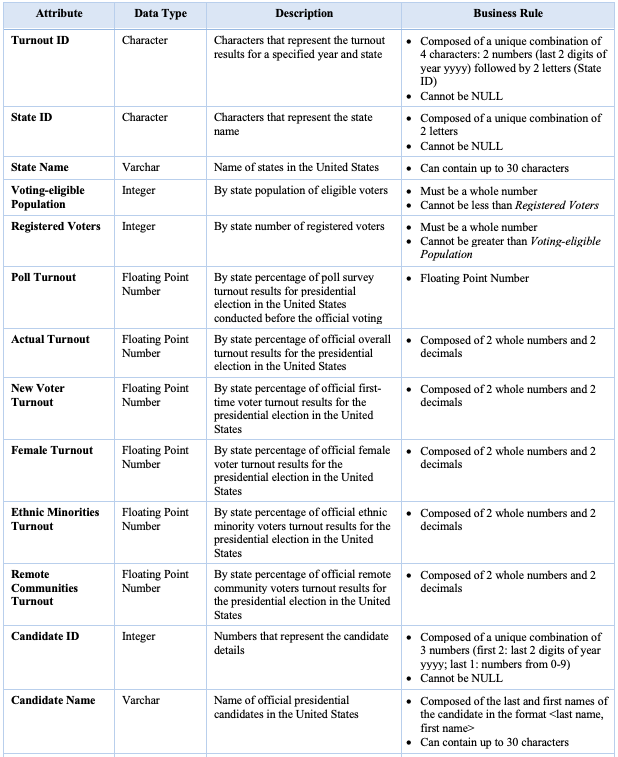
* The data represented in the database is for a particular day in between the voting days for the different regions of states in US.
* The *candidate* must be a natural-born citizen of the United States, be at least 35 years old and above.
* …

**DATABASE 3**

* …

…

# **Attributes and Business Rules**



# **2-DATABASE DESIGNS**

## (4 marks for each database and commentary: 1 mark for 1NF, 1m for 2NF, 1m for 3NF, 1 m for comments)

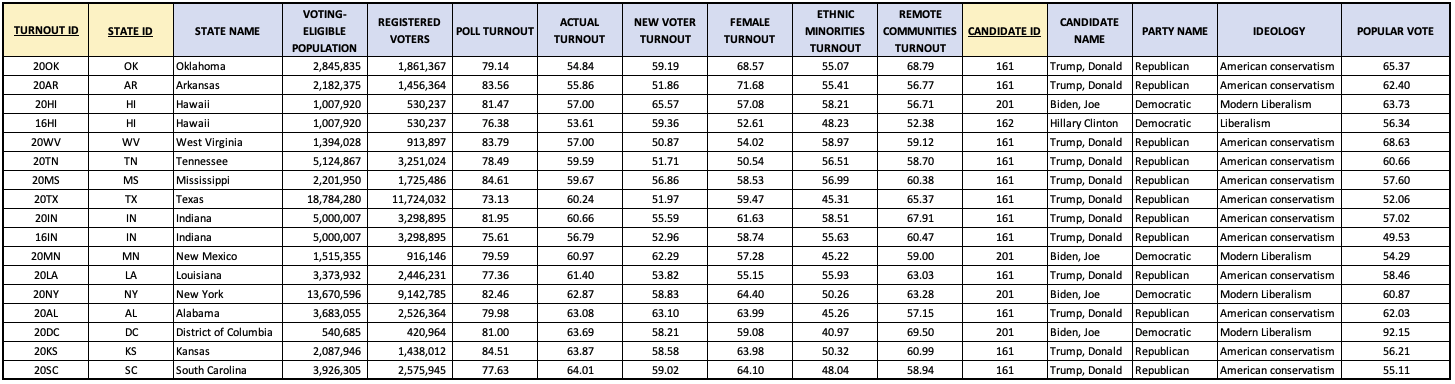
**Example:**

# **Database 1: Voter Turnouts and Voting Results**

Example:

*The Voter Turnouts and Voting Results Database Design* includes by state historical data of general presidential election popular vote results and voter turnouts of various categories, state information, and details of each political party’s presidential candidate.

# 1NF

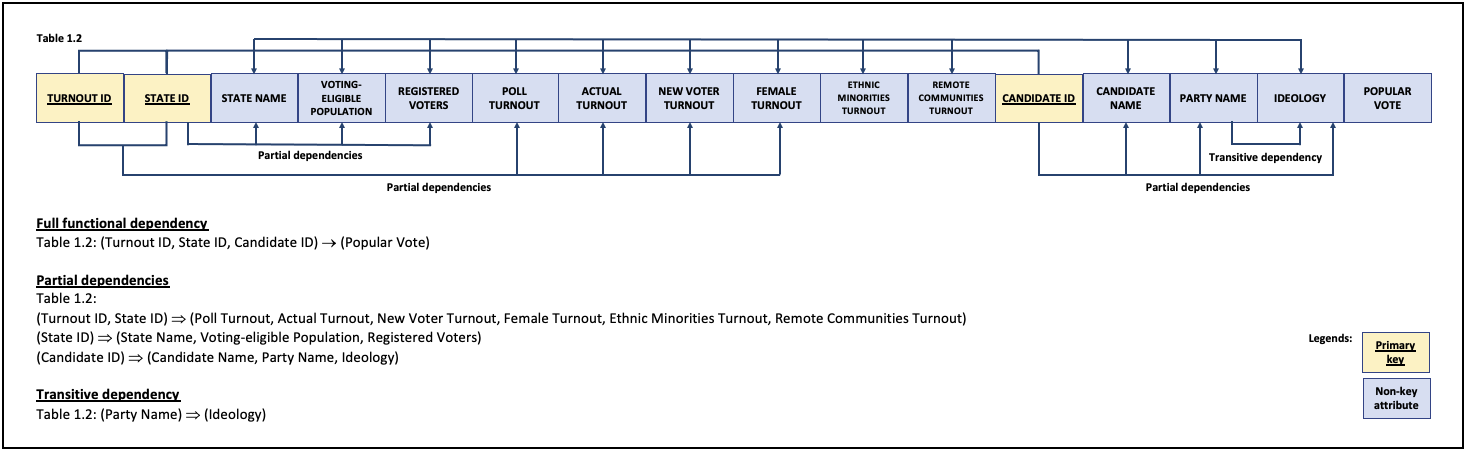


*Table 1.2. Voter turnouts and popular vote results per state*

*Table 1.2* has been created in order to meet the requirements of the first normal form, where all cells are single valued, entries in a column are of the same type, and rows are uniquely identified:

* *Turnout ID, State ID, and Candidate ID* columns were created to act as primary keys in order to distinctly determine the records.

**Dependencies and Key Attributes**



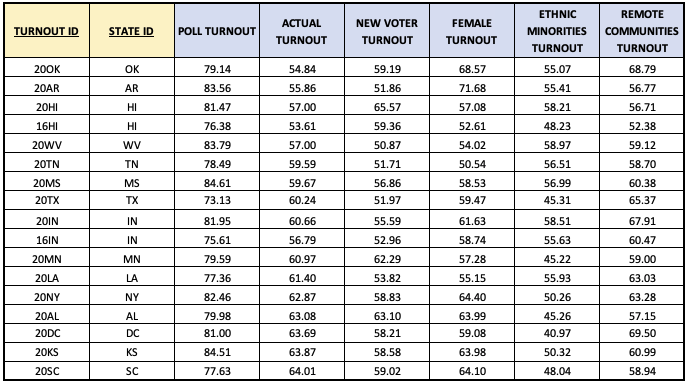
*Figure 1.1. 1NF table dependency diagram*

Even in its first normal form, the following anomalies still exist in *Table 1.2*:

* Insertion: Candidates who have not yet won the popular votes cannot be inserted
* Insertion: Turnouts for states that does not have results for the popular votes cannot be entered
* Deletion: If a candidate is deleted from the table, the corresponding records for the *Popular vote results*, *Voter turnouts*, *Voting-eligible population, and Registered voters* tables will not be lost
* Deletion: Given the relationship of party name and ideology, changing one item would remove the other information from the database
* Update: If a presidential candidate was replaced before the general election, update will be done for all tuples with the candidate’s information

To get rid of these anomalies, we further normalize the table into the second normal form.

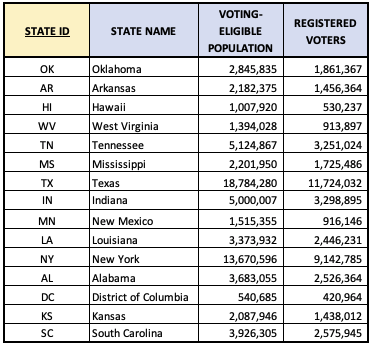
# 2NF



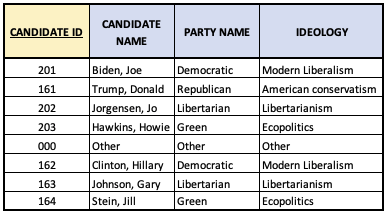
*Table 1.3. Voter turnouts by state broken down into categories*



*Table 1.4. Popular vote results by state*



*Table 1.5. State details*

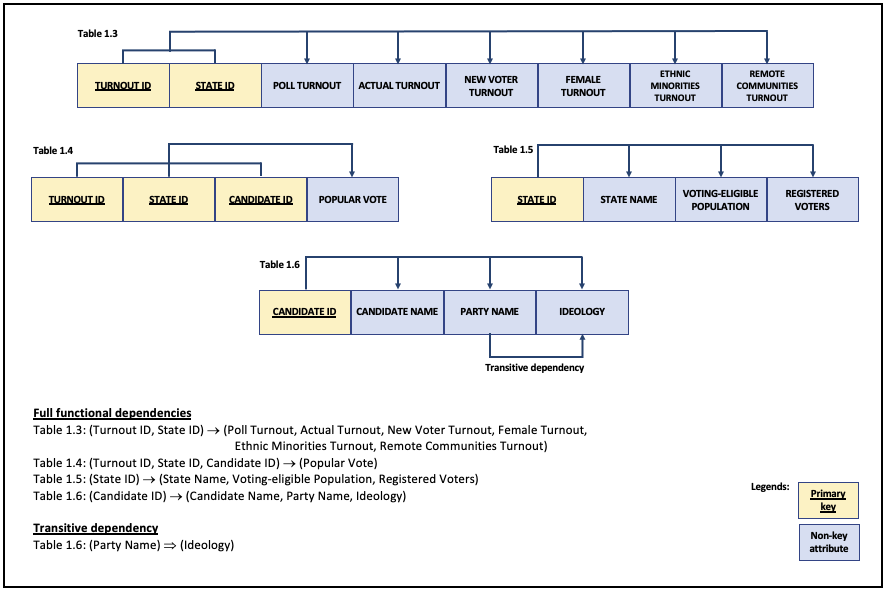


*Table 1.6. Candidate details*

By creating *Tables* *1.3, 1.4, 1.5, and* *1.6*, partial dependencies have been removed and the tables were reduced to the second normal form, where all non-key attributes in the table are fully functionally dependent on the primary key. With this, the following anomalies were eliminated:

* Insertion: Can now enter candidates who have not yet won the popular votes
* Insertion: Can now enter turnouts for states that does not have results for the popular votes
* Deletion: If a candidate is deleted from the table, the corresponding records for the *Popular vote results*, *Vote turnouts*, *Voting-eligible population, and Registered voters* tables will not be lost
* Update: If a presidential candidate was replaced before the general election, update will only be done on the *Candidate details* table

**Dependencies and Key Attributes**



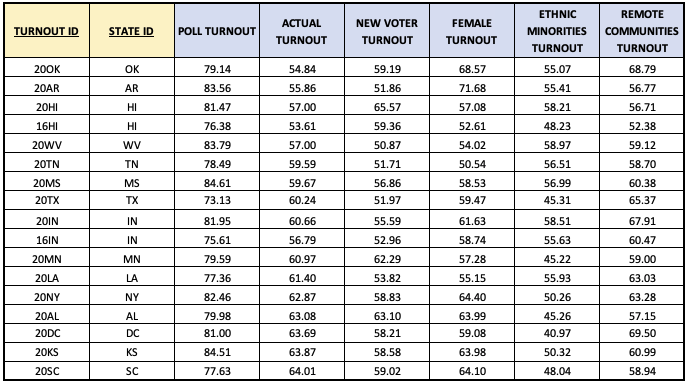
*Figure 1.2. 2NF tables dependency diagram*

Even in its second normal form, the following anomalies still exist in *Table 1.6*:

* Insertion: Cannot enter the fact that a particular political party has an ideology unless *Party Name* is recorded
* Deletion: If a particular candidate switches party before the general election, the information on the previous party’s ideology is lost
* Update: If the ideology of the political party changes, all candidate records need to be updated

To get rid of these anomalies, we further normalise the table into the third normal form.

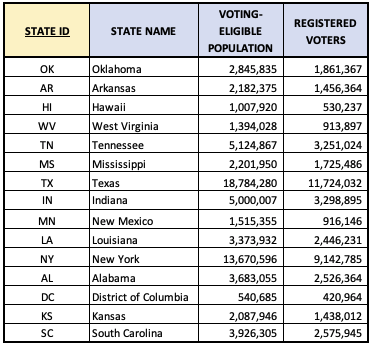
# 3NF



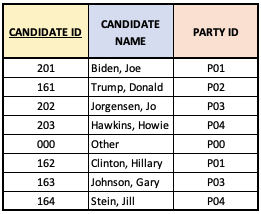
*Table 1.3. Voter turnouts by state broken down into categories*



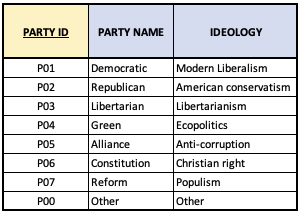
*Table 1.4. Popular vote results by state*



*Table 1.5. State details*



*Table 1.7. Candidate details (version 2)*

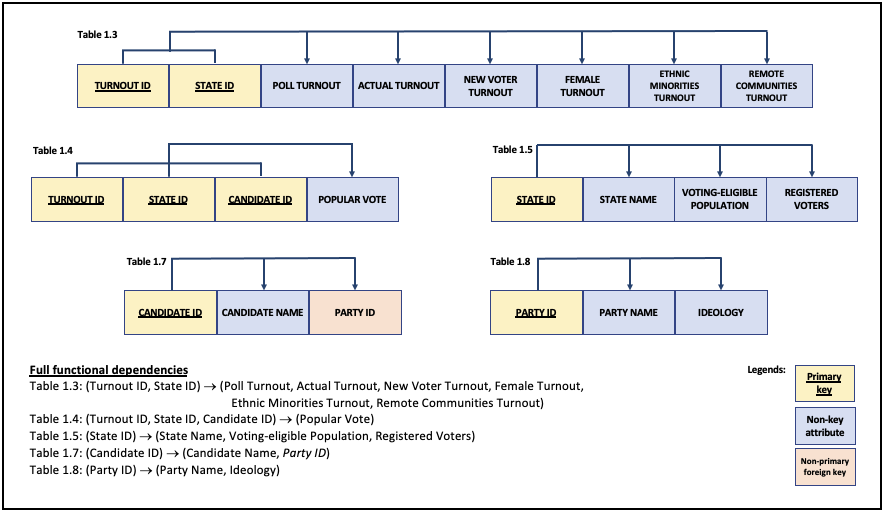


*Table 1.8. Political parties and ideologies*

By creating the *Political parties and ideologies* table, the transitive dependency between *Party Name* and *Ideology* has been removed and the *Candidate details* table was reduced to the third normal form, where all non-key fields in a table are determined solely by its primary key. With this, the following anomalies were eliminated:

* Insertion: Can now enter ideologies for new/other parties that take part in the election
* Deletion: If a particular candidate switches party before the general election, the information on the previous party’s ideology is maintained
* Update: The ideology of each political party appears only once

**Dependencies and Key Attributes**



*Figure 1.3. 3NF tables dependency diagram*

With all the anomalies removed, we are now confident that the relations in *Database 1* are satisfactory and are more effective in inserting, updating, and deleting data.

# **Database 2:…**

# **Database 3:…**

# **REFERENCES**

# (use APA ref 1 mark) and report presentation (1 mark)

See: <https://www.deakin.edu.au/students/studying/study-support/referencing/apa>

Brancati, D. (n.d.). *Global Elections Database.* The Project*.* Global Elections Database. <http://www.globalelectionsdatabase.com/>

Harewood, W. G. (2016). *Comparing the advantages & disadvantages of first past the post, alternative voting [AV] and proportional voting representation [PR] electoral systems*. [Excerpt]. Parliament of Canada.

<https://www.ourcommons.ca/content/Committee/421/ERRE/Brief/BR8555618/br-external/HarewoodWGlenn-e.pdf>

International Foundation for Electoral Systems, International IDEA, UN Department of Economic and Social Affairs (2003). *Voter Education*. ACE Project. [*https://aceproject.org/main/english/ve/index.htm*](https://aceproject.org/main/english/ve/index.htm)

McDonald, M. P. (2018, September 5). *2016 November General Election Turnout Rates.* United States Elections Project. <http://www.electproject.org/2016g>

McDonald, M. P. (2020, December 7). *2020 November General Election Turnout Rates.* United States Elections Project. <http://www.electproject.org/2020g>

Open Knowledge Foundation and Open Data for Development Network. (2016). *Election Results*. Global Open Data Index. <https://index.okfn.org/dataset/elections/>

Pew Research Center (2008). *Inside Obama’s Sweeping Victory.* Pew Research.

<https://www.pewresearch.org/2008/11/05/inside-obamas-sweeping-victory/>

Sharda, R. et. al. (2016). *Business Intelligence, Analytics, and Data Science: A Managerial Perspective (4th edition).* Pearson.

**MIS781 – Business Intelligence and Database**

**Trimester 1, 2024**

**Group Assignment**

**GROUP MEMBER CONTRIBUTION FORM**

Group Number:

|  |  |  |  |
| --- | --- | --- | --- |
| **Name (Print)** | **Student ID** | **%Contribution** | **Signature\*** |
| **1.** |  | **33.3%** |  |
| **2.** |  | **33.3%** |  |
| **3.** |  | **33.3%** |  |

**Note**: \* By signing here, I hereby declare that this is my original work.

If every member of the group contributes equally, the figure entered in the ‘%Contribution’ column should be 33.3% for a 3-member group. This page should be signed by each member of the group and attached to the end of the report.

Note: Where there is significant variation in group contributions, marks may be adjusted individually among group members.