Group 14 - Positive Point System

Contact Info:
Dom Lamastra - lamastd1@tcnj.edu
Nithya Nalluri - nallurn1@tcnj.edu
Andrew Paige - paigea1@tcnj.edu
Joni Van Sanders - vansadj2@tcnj.edu

Phase 1: Proposal

1. Inception: Team Identification

Project name: Positive Point System
Team members: Nithya Nalluri, Dom Lamastra, Andrew Paige, & Joni Van Saders

2. Inception: Intro Video

PHASE 1: Link to Original Intro Video**:**

https://drive.google.com/file/d/1QHV-qS62qMyEYpS7r9Q4YKWKHHc4vKF6/view

3. Revisions of all other Phases

Revisions are highlighted

Phase 2: Proposal

1. Inception: Team Identification

Project name: Positive Point System
Team members: Nithya Nalluri, Dom Lamastra, Andrew Paige, & Joni Van Saders

2. Inception: Objectives and Overview

Our plan for the project:

We plan on creating a relational database using three datasets on EV car ownership, population (census), and gas emission data. By connecting them using the attribute county, we will be able to display to the user of the website the comparison between counties. We will do this by creating an interacting map of NJ by county and the user of our website will be able to click on one or multiple counties and compare their sustainability.

We did not complete a map because that was not required as per what Dr. Degood told us however we did focus a lot of time on making sure we were able to deliver on the back end functionality of our website. We were able to accomplish what we said we would do.

We also added a feature that allows the user to see data that compares EV cars of different car brands. Therefore, you get to see all owned EV car brands in a particular region.

Why this particular project was chosen:

We chose this project because the Sustainable NJ website does not have a clear way of comparing different NJ counties to one another. And having this feature on their website would help improve the local governments and citizen understanding of how sustainable their communities really are. And we will add recommendations to counties with high energy usage by showing the possible actions (from the energy part of the Sustainable NJ actions) that counties can do to get points. The points can be then used for getting awards and certifications.

<u>What we hoped to accomplish and learn as a team from the project:</u>

As a team, we hope to learn more about database systems. The modern world has datasets to describe almost all areas of life. Our job is to translate this information to a readable format, which is in our case a web application. The data we are using is sustainably data for municipalities in New Jersey, census data, and gas emission data. After this class, we may go on to work at any large company and that company will have databases implemented for their records. We hope this project will increase our comfort level with databases so in the future if a manager or employer asks us to use a database we will be confident doing so.

In addition, this is a large scale project that we are working on as a team. Teamwork is often overlooked as a necessary skill for the workplace. This project will hopefully strengthen our ability to work as a team and trust each other to achieve a common goal. We can use team efficient methods learned in software engineering to have a solid foundation and understanding for how our team will be organized. It is important that as deadlines get closer and work needs to be done we stay organized and complete all of the tasks required in a timely manner.

The topic of sustainable energy solutions in NJ municipalities was given to us, as all groups in the class are doing something revolving around this topic. However, it is an interesting topic to perform a study on. With the addition of electric cars into the modern world, gas emissions have gone down but energy usage has gone up. As we are all college students, we are hoping this project will teach us more about the world around us. There is only one earth and we have to make sure we are educated so we can do our part to take care of it.

Our group will implement a database that joins tables between gas emission, population, and sustainable energy data. We want to first establish how densely populated the cities that we are looking at are. From this, we want to explore car usage. The more densely populated an area is, the less likely people are to use cars because in densely populated cities a lot of people can use public transport or they can walk to arrive where they need to go. However, in all areas of New Jersey a part of that population will have cars. We want to see what percentage of the population have their own cars and what percent of that use gas cars. Cars are a big source of gas emission and pollution. Seeing this data can show where cars make the biggest impact on the air. Places with high car usage may benefit from advocating for electric vehicles.

Our data is going to be shown to users across the state as a web application. Our web application will have 2 types of users, a developer user and a citizen user. A developer user has high level access because they can manage the database and respond to any concerns the rest of the users have. The rest of the users will be county citizens and even county officials, and they can look at the data and make energy decisions based on the data that they see. The developers can create a web based map of all the counties in NJ and display municipality data, then the other users can view the map and analyze it seamlessly.

Overall Team member's skills, interests, background, and experience:

Despite the teams being randomly chosen, our group got a wide variety of talents, skills, and experience in our group. Nithya's skills, interests, and experiences demonstrate creativity, and her experiences with MUSE and REU research will only enhance our efforts to make an application that can serve the needs of many moving forward. Nithya is responsible for originating our point system idea and we will hope to have more of that out of the box thinking moving forward.

Joni's research abilities can help us find the most relevant and unique data sets we can find to ensure our study is being done thoroughly. In addition, Joni's traveling interest can provide motivation for this project as we are looking to highlight which municipalities in New Jersey are most energy efficient, and traveling to a place that overuses their energy may not be worth the trip.

Dom's python coding experience will make him a very useful back end developer. Dom is also a good human. Dom has had much experience in the past working with reading csv and excel files into python, as Dom had to do this in his internship on multiple occasions but he knew nothing about databases then. Making this data transition from spreadsheet to python as smooth as possible will help the project stay organized in its entirety so we can finish it in a timely manner.

Andrew's experience with flask and reactJS is very beneficial to the outcome of the project. These tools will help us make the project's usefulness. The data that we find will be hard to read and follow stored in these giant csv files. Andrew's background will help us make the user interface easy on the eyes and convenient for the user to find what they are looking for without being overwhelmed with tables and tables of data.

3. Elaboration: Requirements Modeling and Analysis

Our overall project breakdown:

1) EV Car and Gas Emissions data:

a) EV Car data includes information on the number of EV vehicles in each county and their share of the overall number of vehicles in each county, and the Gas Emissions data includes information about overall gas emissions in each county.

b) Gas cars vs. EV car: We will be comparing the proportion of gas cars to EV cars in each county and examining trends in the transition to electric vehicles.

   i) The points system (10–15 points), known as the actions by sustainable NJ, will be used to evaluate the energy efficiency of each county and will take into account things like the utilization of renewable energy, waste reduction, and transportation efficiency.

2) <u>Population and census data:</u>

   a) Each municipality in New Jersey will have their own respective population and population densities. Only calculating the total emissions of gas cars and energy usage of EV cars is not enough. This guarantees that the data would be biased, Newark NJ will be higher than Fanwood NJ in gas emissions easily because the population of Newark is way higher. The data would be less biased if it was normalized in some way, and we will normalize the data with population and population density.

   b) Our project will gather census datasets, and datasets that have the total amount of land per county/municipality and use that to find the population density of each region. This data will be joined with other tables in our database to get a non biased description of gas emissions and EV car data.

3) <u>REVISED REVISIONS:</u>

4) EV Car Data

EV Car Data includes municipality name, county name, the year, the % of ev cars compared to all total cars, the # of ev cars in a municipality, and the # of personal cars in a municipality.

5) Emissions Data

Emissions data is municipality name, county name, the year, and the total number of carbon emissions in a given municipality.

6) Population Data

Population data is municipality name, county name, the year, the population of the municipality, the square miles of the municipality, and the municipality score of a given municipality

7) Car Brand Data

The brand of the car, the avg_accel time of the car per sec, the average topspeed of the car, the average range in km, the average efficiency of the car, the average fastcharge of the car, and the avg priceeuro, and the avg price in dollars of the car.

8) NJ EV Registration Data

The county, the car brand, and the number of cars bought in that county

<u>Some of the questions we will explore with the chosen datasets:</u>

1) How does population density affect EV car usage and gas consumption in different ~~counties~~ (municipalities)?
2) How efficiently are ~~counties~~ (municipalities) using energy, and which areas need to improve their energy efficiency?
3) How can ~~counties~~ (municipalities) gain more points in the sustainability NJ actions system, and what measures can they take to improve their energy efficiency and reduce waste?
4) What brands of cars are more prevalent in certain municipalities in NJ?

How could the data help them identify sustainability problems, and opportunities to propose positive change?

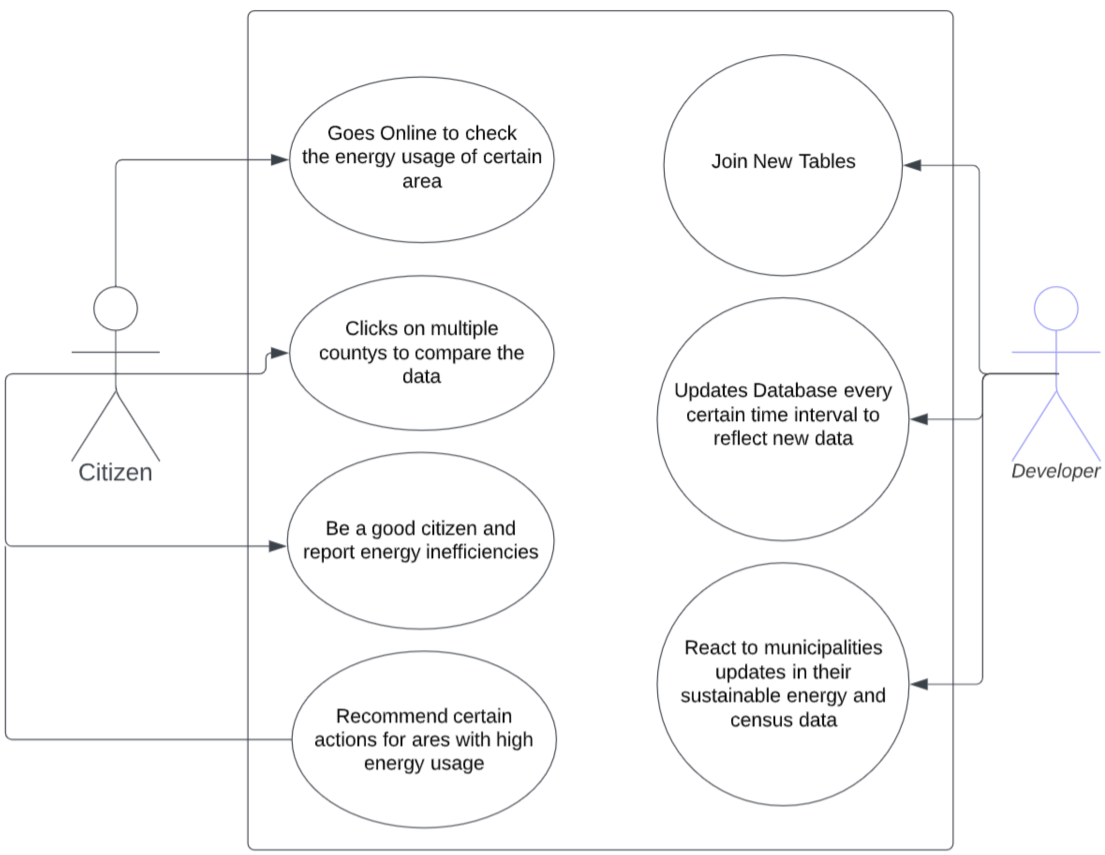
The current point system used by the Sustainable NJ website to compare energy efficiency across ~~counties~~ (municipalities) has limitations in accurately representing population density differences. Lower population density areas receive fewer points than highly populated ~~counties~~ (municipalities), despite potentially using less energy. Our group's goal is to improve the point system by taking population density into account as a confounding variable. This will provide a more accurate picture of energy usage across ~~counties~~ (municipalities), assisting in determining which areas truly require energy efficiency improvements. By removing this bias, we can encourage positive change and encourage energy wasters to reduce their consumption.

What is an overall sustainability issue you will be exploring?

The sustainability issue we are exploring is the transition from gas powered vehicles to electric vehicles, it is a complex issue with various sustainability and ethical complications. Gasoline and diesel vehicles contribute to air pollution, which has a significant impact on the environment and public health. In addition, the production and consumption of gasoline and diesel lead to greenhouse gas emissions, which can contribute to climate change. The stakeholders affected by the problem is everyone that lives on the planet, due to the planet being affected by environmental issues. Consumers also play a role in the adoption of electric vehicles by choosing to purchase them. Ethical issues presented by the problem include being able to better show the local governments and citizens how they can understand the amount of gas emissions they produce compared to other ~~counties~~ (municipalities) in New Jersey.

4. Use Case Diagram and Descriptions

Use case diagram:



1) Choose two representative user interactions with your completed project. Write a detailed textual use case for each.

Use Case #1: User clicks on multiple ~~counties~~ (municipalities) to compare the data

Primary Actor: Citizen

Goal in context: To compare energy data between ~~counties~~ (municipalities)

Preconditions: The user has to go to our website on a web browser

Trigger: User selects the county that they want to see data for

Scenario:

1. User clicks on county they want data for (User selects municipality from dropdown)
2. User clicks on other ~~counties~~ (municipalities) they want to compare (user submits a range of similar municipalities they want to see)
3. User can view the data and observe trends

Exceptions:

1. The width of the users screen can only show a limited number of ~~counties~~ (municipalities) at a time

Priority: Essential, must be implemented

When available: Project's completion

Frequency of use: Multiple times per visit

Channel to actor: Via User Interface

Secondary Actors: None

Channels to secondary actors: None

Open Issues:

1. How does the user know the website is up to date? How can the user trust the website?

Use Case #2: Recommend certain actions for areas with high energy usage

Use Case: Recommend certain actions for areas with high energy usage

Primary Actor: Citizen

Goal in context: Reduce the energy used in ~~counties~~ (municipalities) with high energy usage

Preconditions: None, data is on the site

Trigger: User clicks on county on map and will display data and recommendations

Scenario:

1. User clicks on county they want data for
2. User observes data about selected county
3. User observes recommendations for how the county can reduce their energy usage

Exceptions:

1. Not enough data on county to make recommendations

Priority: Essential, must be implemented

When available: Project's completion

Frequency of use: Multiple times per visit

Channel to actor: Via User Interface

Secondary Actors: None

Channels to secondary actors: None

Open Issues:

1. How can we measure if our recommendations are effective?
2. How do we know that the people will follow the recommendations?

5. Works Cited

"Participating Municipalities & Approved Actions." *Home - Sustainable Jersey*, https://www.sustainablejersey.com/.

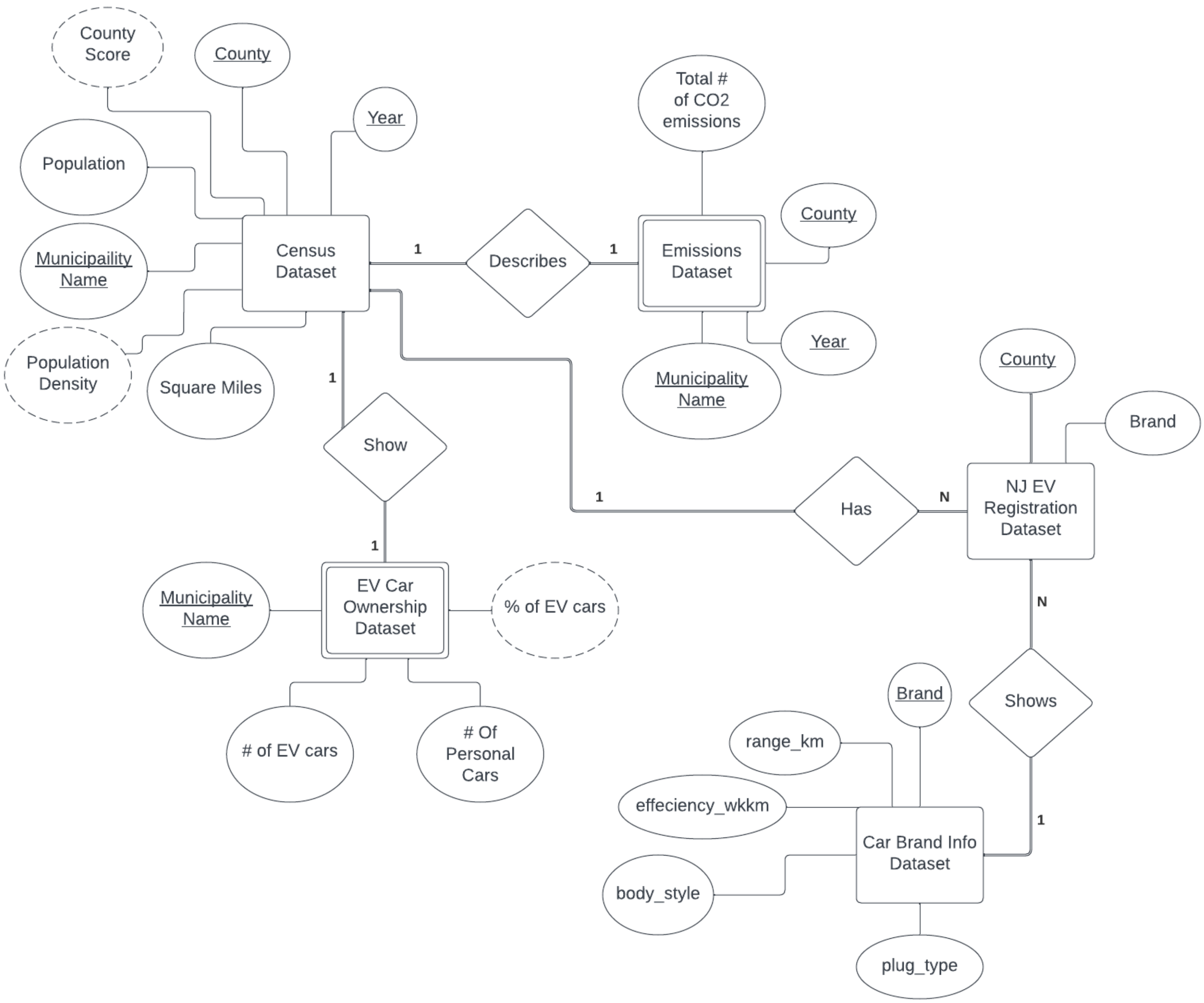
Phase 3: Proposal

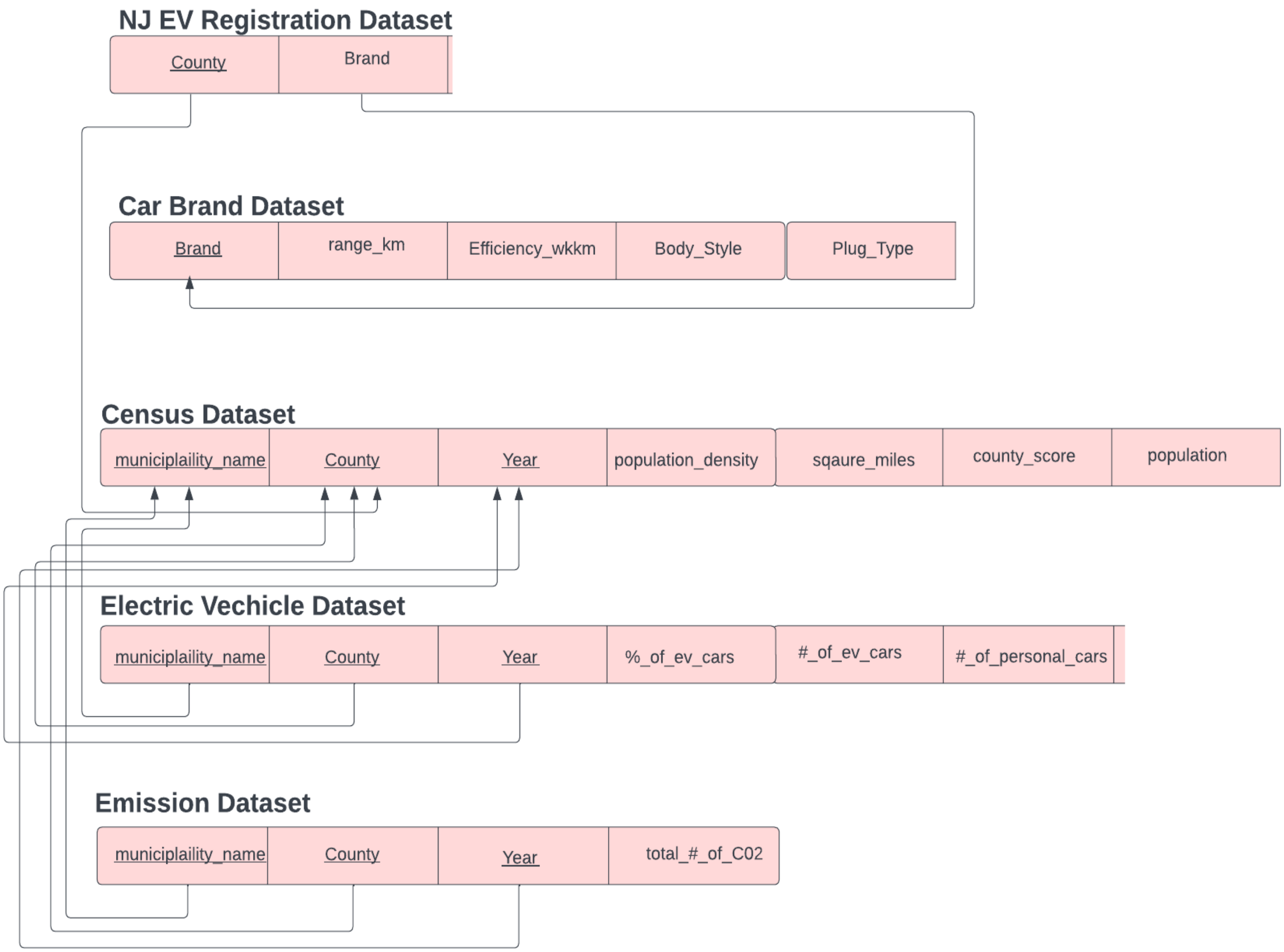
1. Team Identification

Project name: Positive Point System

Team members: Nithya Nalluri, Dom Lamastra, Andrew Paige, & Joni Van Saders

2. Entity Relationship (ER) Diagram And A Relational Scheme

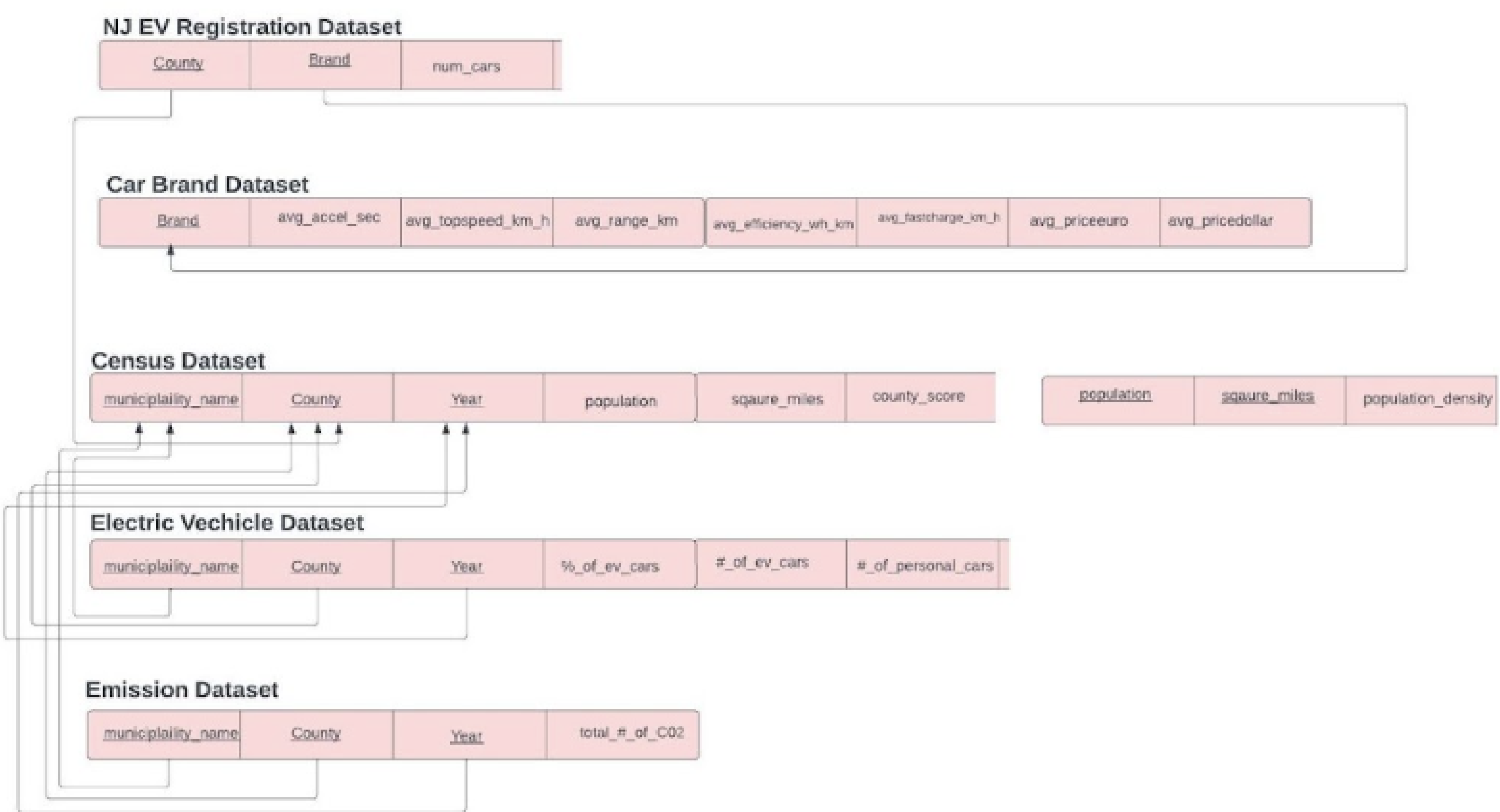




ER Diagram Explained:

In our diagram, there are multiple keys per table. Every table is about the municipality, county, and year or car brand and can be sorted by either, therefore we are making the municipality, county, and year name the key for 3 of our tables, and the county will be a foreign key with the car brand that connects the car brand to the county. One table will just be county to car brand and the last dataset will have statistics about each car brand. The municipality name, county name, and year will be a primary key for Census Data and a foreign key in the other two tables. We are not using any specialization/generalization in our project. Our tables are connected by relationships. The show relationship connects the census data to EV dataset and is 1:1 because every 1 Municipality has 1 set of EV data. The show relationship is total on both sides because every municipality has EV data. The Describes relationship connects the census data to Emission dataset and is 1:1 because every 1 municipality has 1 set of Emission data. The Describes relationship is total on both sides because every municipality has Emission data. The Has relationship connects the car brand to the county, and the shows relationship connects the car brand to the car brand data.

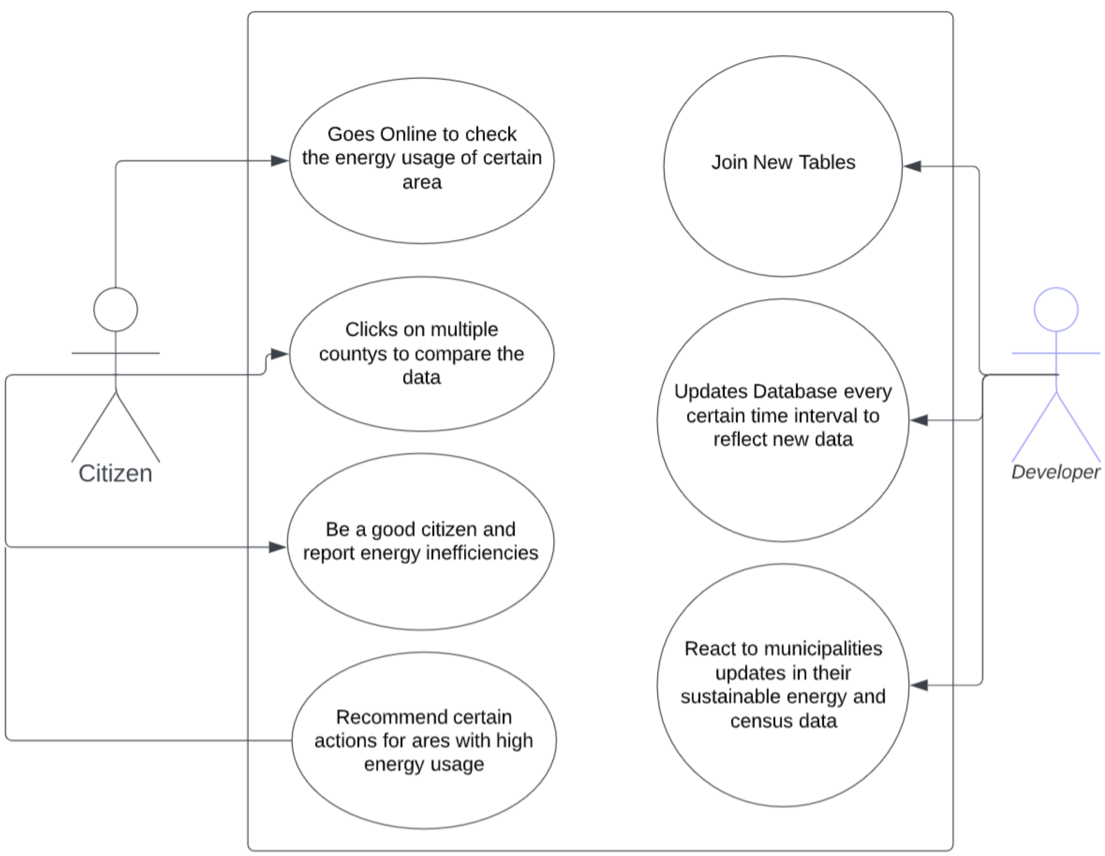
We had a transitive functional dependency arise where population density is dependent on population and square miles. Attached is our revised schema.



Our database has 5 relations. The first one, NJ_EV_Dataset is uniquely identified by the composite key of county and brand. Next we have a car brand dataset that is uniquely identified by the primary key of the car brand. This contains information about the average statistics for all car brands such as average acceleration. Next, Census Data relation is used to calculate the population density, and also has a key of municipality and county and year. This relation has info for each municipality. Next, Electric Vehicle and Emissions datasets use municipality and county and year to uniquely identify ev car data and emissions data as previously described in phase two. We use these datasets to calculate emissions from cars and counties then factor in population density to create our scores for each municipality.

3. Use Case Diagram and Descriptions

Use case diagram:



Our system has two types of users, the developers and the citizens. The developers are necessary to create the application. They can also update the product as newer data is generated as time goes by. The other actor types are citizens. A citizen could be anyone that is curious about which ~~counties~~ (municipalities) are more eco friendly than others. But the purpose of this project is really for the politicians in the ~~counties~~ (municipalities) to make decisions based on the data that they see. If their municipality is not eco friendly then they can take steps to fix it. The goal of our project is to see politicians make changes based on visiting their website, being told to be more eco friendly, and then doing so. The scope of our system is nothing more than entering the url and looking at data. But we hope the story that our data tells is impactful and convincing enough for ~~counties~~ (municipalities) to clean up their act, if necessary.

4. Narrative

Our database can be thought of as a room and in that room there are entities, or smaller boxes in the room. And then in our boxes there are different items in each different box, which are the attributes. The attributes go into a box based on the features of the attribute. But, some of those items can be connected to multiple boxes based on similarities they share.

So in our case, our room is our database. And we have three boxes in our room. We have a box for census data. We have a box for electric vehicle data and we have for emission data.

We have 3 main entities, we have the electric car entity, we have the gas emissions entity, and we have a census data entity. Each entity comes from a separate dataset that we have gathered through research. The datasets themselves are not connected to each other. It is our job to connect all the datasets to each other in order to show the trends that we are trying to show.

The census data dataset is called community profile data by municipalities. The ev car dataset [ev] is ownership data. and the last dataset is the community scale greenhouse gas (GHG) emissions. The job of the developers is to take this data and reformat it to show trends about the outside world.

I stated there are multiple attributes which are represented as boxes. We set up the primary key as a municipality name for each unique municipality. Within the census data, the other mini boxes in it would be population, and the number of square miles that the municipality is. There is a 1:1 relation between census and emissions because it shows the relationship between census data and gas emissions per municipality. It is 1:1 because for every municipality there is only one row about gas emissions. The gas emission one has only two attributes, the same identification number and the total number of carbon emissions. This is a weak relation because both the entities are dependent on another entity. The emission entity is dependent on the census and the EV dataset is dependent on the census dataset as well. Without the census data everything crumbles.

Once all of the data is connected, the next step is to interpret the connections we have made. Our project is intended to show ~~counties~~ (municipalities) if their car emissions are high enough to be considered wasteful or not. All of the datasets represent a piece of how we solve this, and together they will show an unbiased view of all the data. We will use statistics to normalize the total number of carbon emissions rates per municipality while taking into account population density of every municipality and the % of EV car users in the area.

As developers, our job is simple as described in our use case diagram. We join new tables as new data arrives, we update existing tables with current data, and we stay attentive to changed policy to look out for new information.

In addition, if the user's interest is grabbed by different car brands, we show which counties correlate to each car brand. Therefore, you will be able to see how all the car brands are used.

Our project will also be available to the people. People can log on and view the current eco friendliness data we are providing. Multiple counties can be compared at the click of a button. If someone suspects energy inefficiency that can be reported too. Counties that are not energy inefficient will be recommended ways to change their behaviors.

5. Database Size and Type

Emissions Dataset:
1130 row * 4 col = 4520 cells

Census Dataset:
1130 rows * 7 cols = 7910 cells
1130 rows * 6 cols = 6780 cells
1130 rows * 3 cols = 3390 cells
Electric Vehicle Dataset:
1130 rows * 6 cols = 6780 cells

Car Brand Dataset:
103 rows * 5 cols = 515 cells

33 rows * 8 cols = 264 cells

NJ EV Registration Dataset:
161202 rows * 2 cols = 322,404
1021 rows * 3 cols = 3063 cells
Dataset Type: The main search type of our database will be a select, select a score from a municipality of choice. However, the score will be calculated through complex SQL queries in the back end covering many of the concepts covered in relational algebra like aggregate functions and joining tables. Every 1 search will require the calculation of population density, the calculation of EV car%, and a normalization of these two categories + a normalization of off road vehicle data. Therefore every search will use 6 searches for our algorithm.

6. Works Cited

"Participating Municipalities & Approved Actions." *Home - Sustainable Jersey*, https://www.sustainablejersey.com/.

Phase 4: Proposal

1. Team Identification

Project name: Positive Point System
Team members: Nithya Nalluri, Dom Lamastra, Andrew Paige, & Joni Van Saders

2. Boyce–Codd normal form (BCNF)

Our Database passes the conditions for 1NF

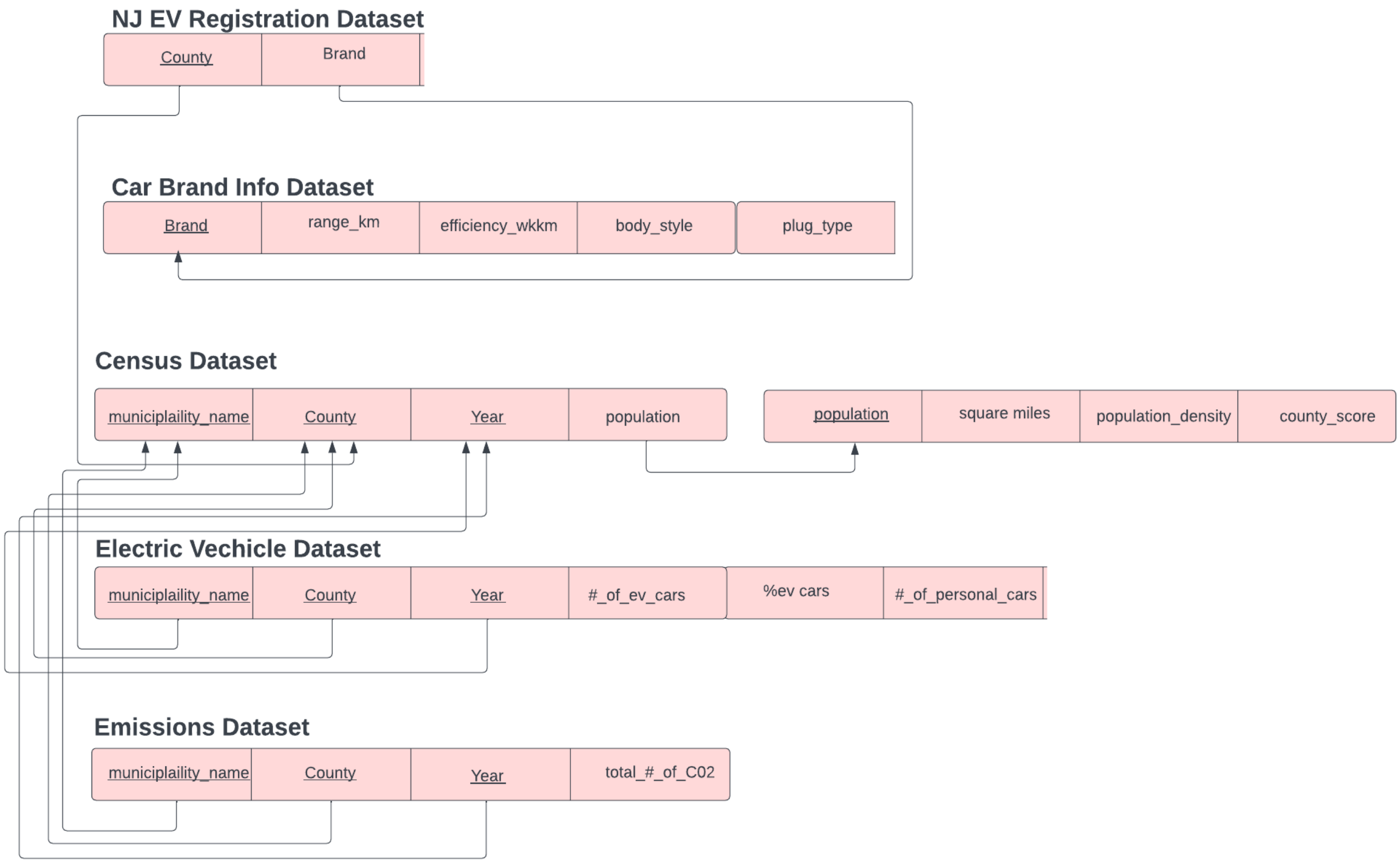
Our Database do the conditions for 2NF

Our Database passes the conditions for 3NF after making the following change:

Because we had to break down a transitive dependency. Population density is dependent on population, and population is dependent on {year, county, municipality_name}.

Our Database passes the conditions for BCNF because we satisfy all the following conditions:

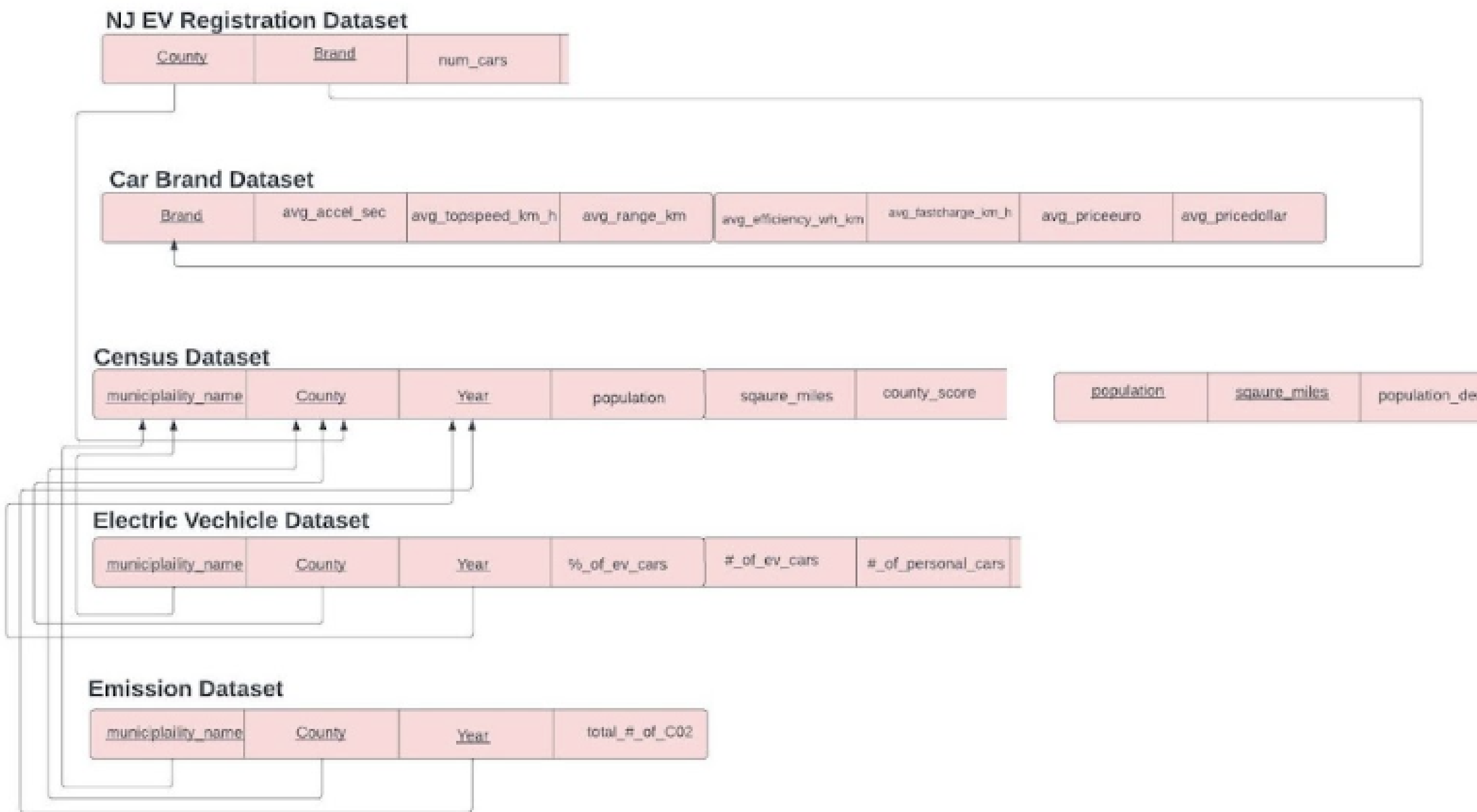
1. The key of the municipality name we are using is a primary key.



2. We only have 1 primary key, so there are no primary keys dependent on other primary keys.
3. No other non prime attributes are functionally dependent on the primary key
4. All attributes are dependent on the primary key (no transitive dependencies).

Our Database satisfies these conditions.

Updated Database Schema



3. Define the different views (virtual tables)

A view of emission vs car data:
CREATE VIEW EV_Car_Data_Vs_Emission AS
SELECT * FROM Emission
INNER JOIN Electric_Vehicle ON municipality_name = municipality_name, county = county, year = year

A view of EV Car Data vs Population Data
CREATE VIEW Car_Data_Vs_Population AS
SELECT * FROM Emission FROM EV_Car_Data_Vs_Emission INNER JOIN Census_Data ON municipality_name = municipality_name, county = county, year = year

A view of All Brand Information
CREATE VIEW Brand AS
SELECT * FROM NJ_EV_Registration
INNER JOIN Electric_Vehicle_Dataset ON Brand = Brand

A View on Brand vs Population (data for each brand)
CREATE VIEW Brand_Vs_Population AS
SELECT Brand FROM NJ_EV_Registration

INNER JOIN Census_Data ON county = county

A view of Brand vs Emission
CREATE VIEW Brand_Vs_Emission AS
SELECT Brand FROM NJ_EV_Registration
INNER JOIN Emission ON county = county

(throughout the project, we may add or remove more views as we see fit and we will have no transaction requirements)

```
\c group_test5
CREATE TABLE census_data (
Municipality VARCHAR (50) NOT NULL,
County VARCHAR(50) NOT NULL,
Muni_and_County VARCHAR(50),
Year INTEGER NOT NULL,
Square_Miles FLOAT NOT NULL,
Population INTEGER NOT NULL,
PRIMARY KEY(Municipality, County, Year)
);

\copy census_data FROM 'cleaned_cd.csv' WITH (FORMAT CSV, HEADER true, DELIMITER ',');

ALTER TABLE census_data DROP COLUMN Muni_and_County;

--CREATE TABLE pop_density_data AS SELECT population, square_miles FROM census_data;

--ALTER TABLE pop_density_data ADD pop_density FLOAT;

--UPDATE pop_density_data SET pop_density = population / square_miles;

ALTER TABLE census_data ADD pop_density FLOAT;

UPDATE census_data SET pop_density = population / square_miles;

ALTER TABLE census_data ADD score FLOAT;

--CREATE TABLE for_score (Municipality VARCHAR (50) NOT NULL, County VARCHAR(50) NOT NULL, YEAR INTEGER NOT NULL, percent_of_ev FLOAT, total_mtco2e INTEGER, pop_density FLOAT, norm_ev FLOAT, norm_co2 FLOAT, norm_pd FLOAT, score FLOAT);
```

```
--INSERT INTO for_score (municipality, county, year) SELECT census_data.municipality,
census_data.county, census_data.year FROM census_data;

--INSERT INTO for_score (municipality, county, year, percent_of_ev, total_mtco2e,
pop_density) SELECT census_data.municipality, census_data.county,
census_data.year, electric_vehicle.percent_of_evs, emissions.total_mtco2e,
census_data.year FROM census_data INNER JOIN electric_vehicle ON
census_data.municipality = electric_vehicle.municipality AND census_data.county =
electric_vehicle.county AND census_data.year = electric_vehicle.year JOIN emissions
ON census_data.municipality = emissions.municipality AND census_data.county =
emissions.county AND census_data.year = emissions.year;


--ALTER TABLE pop_density_data ADD PRIMARY KEY (Population, square_miles);

CREATE TABLE electric_vehicle (
Municipality VARCHAR (50) NOT NULL,
County VARCHAR(50) NOT NULL,
Year INTEGER NOT NULL,
Total_Personal_Vehicles INTEGER NOT NULL,
number_of_EVs INTEGER NOT NULL,
percent_of_EVs FLOAT NOT NULL,
PRIMARY KEY(Municipality, County, Year)
);

\copy electric_vehicle FROM 'cleaned_ev.csv' WITH (FORMAT CSV, HEADER true,
DELIMITER ',');

CREATE TABLE emissions (
Municipality VARCHAR (50) NOT NULL,
County VARCHAR(50) NOT NULL,
Muni_and_County VARCHAR(50),
Year INTEGER,
Total_MTCO2e FLOAT NOT NULL,
PRIMARY KEY(Municipality, County, Year)
);

\copy emissions FROM 'cleaned_em.csv' WITH (FORMAT CSV, HEADER true,
DELIMITER ',');

ALTER TABLE emissions DROP COLUMN Muni_and_County;

CREATE TABLE car_brand_info (
```

```
Brand VARCHAR(20) NOT NULL,
Model VARCHAR(40) UNIQUE NOT NULL,
Accel_sec FLOAT,
TopSpeed_km_h INTEGER,
Range_Km INTEGER,
Efficiency_Wh_km INTEGER,
FastCharge_km_h INTEGER,
PriceEuro INTEGER,
PriceDollar INTEGER,
PRIMARY KEY(Model)
);

SELECT census_data.municipality, census_data.county, census_data.year,
census_data.pop_density, emissions.total_mtco2e, electric_vehicle.percent_of_evs
INTO for_score FROM census_data LEFT JOIN electric_vehicle ON
census_data.municipality = electric_vehicle.municipality AND census_data.county =
electric_vehicle.county AND census_data.year = electric_vehicle.year LEFT JOIN
emissions ON census_data.municipality = emissions.municipality AND
census_data.county = emissions.county AND census_data.year = emissions.year;

ALTER TABLE for_score ADD COLUMN norm_pd FLOAT;
ALTER TABLE for_score ADD COLUMN norm_ev FLOAT;
ALTER TABLE for_score ADD COLUMN norm_co2 FLOAT;
ALTER TABLE for_score ADD COLUMN score FLOAT;

\copy car_brand_info FROM 'Car_brands_output.csv' WITH (FORMAT CSV, HEADER
true, DELIMITER ',');

CREATE TABLE revised_car_brand_info AS SELECT brand, AVG(Accel_sec)
avg_accel_sec, ROUND(AVG(topspeed_km_h), 0) avg_topspeed_km_h,
ROUND(AVG(range_km), 0) avg_range_km, ROUND(AVG(efficiency_wh_km),0)
avg_efficiency_wh_km, ROUND(AVG(fastcharge_km_h),0) avg_fastcharge_km_h,
ROUND(AVG(PriceEuro),0) avg_priceEuro, ROUND(AVG(PriceDollar),0)
avg_priceDollar FROM car_brand_info GROUP BY brand ORDER BY brand;

ALTER TABLE revised_car_brand_info ADD PRIMARY KEY (Brand);

DROP TABLE car_brand_info;

ALTER TABLE revised_car_brand_info RENAME TO car_brand_info;

CREATE TABLE nj_ev_reg (
County VARCHAR(30) NOT NULL,
Brand VARCHAR(40) NOT NULL,
```

```
num_cars INTEGER,
PRIMARY KEY(County, Brand)
);

\copy nj_ev_reg FROM 'NJ_EV_output_grouped.csv' WITH (FORMAT CSV, HEADER true, DELIMITER ',');
```

Our tables will be broken down this way. All users will have access to the same tables. Overall, the virtual tables/views would allow users to retrieve specific information related to the municipality, EV ownership, car brand data, and greenhouse gas emissions in an easier way. The transaction requirements would ensure that the data is kept up-to-date and accurate, enabling users to make informed decisions based on the latest information available.

4. Complete set of SQL queries to satisfy the transaction requirements

Assuming the views are named CensusData, EVData, and EmissionData

1) Query for retrieving population of a specific municipality:
    a) SELECT population FROM CensusData WHERE municipality_name = 'Municipality X';

2) Query for retrieving the top 5 municipalities with the highest number of EVs:
    a) SELECT municipality_name, COUNT(*) AS num_evs
       FROM EVData
       GROUP BY municipality_name
       ORDER BY num_evs DESC
       LIMIT 5;
        i) Notes: DESC is short for "descending" and is used to sort query results in descending order

3) Query for retrieving the total greenhouse gas emissions for a specific municipality:
    a) SELECT total_emissions FROM EmissionData WHERE municipality_name = 'Municipality A';

4) Query for retrieving the top 10 municipalities with the lowest greenhouse gas emissions:
    a) SELECT municipality_name, total_emissions
       FROM EmissionData
       ORDER BY total_emissions ASC
       LIMIT 10;
        i) Notes: ASC is short for "ascending" and is used to sort query results in ascending order

5) Query for receiving the top 10 Brands of cars driven in counties with the highest total_#_of_Co2 rate
   a) SELECT Brand
      FROM Brand_Vs_Emission
      ORDER BY total_#_of_CO2 ASC
      LIMIT 10;
      i) Notes: ASC is short for "ascending" and is used to sort query results in ascending order
6) Query for seeing Census Data only in the year 2020
   a) SELECT * FROM CensusData WHERE year='2020'

SELECT municipality, population, year FROM census_data WHERE municipality='Aberdeen township';

SELECT municipality, square_miles, year FROM census_data WHERE county= 'Union';

SELECT nj_ev_reg.brand, nj_ev_reg.num_cars, avg_topspeed_km_h FROM car_brand_info INNER JOIN nj_ev_reg ON car_brand_info.brand = nj_ev_reg.brand;

6. Works Cited

"Participating Municipalities & Approved Actions." *Home - Sustainable Jersey*, https://www.sustainablejersey.com/.

Final Phase: Additional Information

1. Your proposal to improve municipal sustainability based on the data

Our project will improve municipalities sustainability based on the data because the data will generate a score that municipalities can use to make better decisions about going green. The sustainable jersey website gives out some scores but they do not take into account the confounding variables that come up when answering the questions of how green-conscious this particular municipality is. Things like population, population density, and what car brands people drive are not taken into account by the folks at sustainable jersey. Giving the sustainable jersey access to this database will allow them to strengthen the quality of the scores the municipalities receive.

Our group has implemented an additional feature where the users can select a range of other municipalities with similar scores as them. Therefore, municipalities can work together that may have similar problems to find common solutions. If one municipality does not know how to fix an issue they can bounce ideas off with municipalities in a similar boat.

Knowing what brands people like to drive will help the municipality go green immensely. If one municipality has a popular EV car, they can add dealerships of that car to the municipality in order to try and get even more people to buy it. Therefore people will be more enticed to switch to a specific electric vehicle and decrease the size of their carbon footprint. Maybe the advertisements for going green became old and tiring, but advertising a specific car brand may help the municipalities relationship with that car brand and it also may get more people to help the earth indirectly. The information used from this green technology will provide us with a better future. With our project up and running, we can all live on a greener planet and save the environment.

2. How our proposal can be implemented in detail

Our proposal could be implemented in detail using the same database and front end approach that our prototype of the project has. The major difference is that we would need an interface to easily update the values. Things like population, population density, carbon emission, and percent of EV cars are important in creating the scores for each municipality. These attributes need to be updated constantly, and we can't rely on the person who is updating the tables to have the ability to write SQL queries. Therefore we need a user interface pre programmed with the UPDATE and SET commands in the back end.

In addition, the score categories can be altered. Our group created a database and determined the ethical implications. We are not environmentalists. If there are more confounding variables that can affect how green a municipality is, our database can be modified to support those confounding variables. Population density, co2 emission, and amount of owned electric vehicles are three important variables in determining a municipality's interest in protecting the environment but it is not the only one. If an environmental scientist determines we should add more, our database can handle more tables with information that can be used to generate more accurate scores. Not only can we have different categories, but the weights of the categories that we do have can be altered. We have population density set at 20% currently, but it can be altered to be a different percentage that would give a better representation of what will best help the environment.

3. Ethical implications of our proposal (positive and negative)

Our solution will affect multiple stakeholders. For example, a citizen can look at our project and think about things he or she can do differently to better help the planet. Our website can make a person want to switch to an electric vehicle, or spend less time driving to lower their carbon emission rate. Our website shows the top 5 electric vehicles per county which can make a person want to buy more cars.

Our solution can also affect government officials and legislators. They can look at our project and make legislative changes to help the community prosper. For example, they can advertise or even require people to lower their carbon footprint. They can help fund electric car companies and add electric car dealerships to the area. They can look at the most popular electric vehicles and advertise those car brands more. Most importantly, mayors of municipalities can contact mayors of other municipalities within the same score range as them. The score range is referring to other municipalities that have a similar score to a single municipality. This can be either the 5% range, the 10% range, or the 15% range that is available on the website. They can work together to try and figure out how to reduce their carbon footprint and help the environment prepare for a better tomorrow.

Some unintended risks that may consequence from our project is that the data may not be up to date and misleading. The scores are based on values that change frequently. If a towns carbon emission rate changes dramatically from one year to the next, this change should be visible to the public eye as soon as possible, but this piece of data only needs to be updated yearly. The cars that are listed have prices attached, which would need to be updated much more frequently. Lastly, to find out which 5 cars are most popular would have to be queried very frequently. This should not be an issue because if this project was put into use the data should be updated frequently and someone should be hired to make sure that this happens as fast as possible. In addition, if someone cannot afford an electric vehicle, even though our website is advertising electric vehicles, people should not buy things that they can not afford. I don't think this risk is a big concern however because people should be trusted to be more aware of what they can and cannot afford.