Tampere University

Erick Rico Esparza, 152698072

COMPARATIVE ANALYSIS OF URBAN RECYCLING RATES IN NYC

Case

TAU.280 Contemporary Circular Economy Challenges and Solutions
Supervisor: Marika Kokko
Intended credits: 5
29.03.2025

CONTENTS

1. PROBLEM 2
   - 1.1 Overview of Recycling Rate Challenges 2
   - 1.2 Recycling as a Pillar of the Circular Economy 2
   - 1.3 Objectives 3
2. BACKGROUND 4
   - 2.1 Recycling and Circular Economy: Literature Insights 4
   - 2.2 Comparative Recycling Strategies in Other Regions 5
   - 2.3 Recycling in New York: Context and Challenges 6
   - 2.4 Key Research Question 7
3. METHODOLOGY 8
   - 3.1 Dataset Description and Geographic Context 8
   - 3.2 Data Dictionary 10
   - 3.3 Software and Tools 11
   - 3.4 Data Import, Exploration, and Preprocessing 11
   - 3.5 Analytical Approach 12
   - 3.6 Study Limitations 13
4. RESULTS AND DISCUSSION 14
   - 4.1 Data Analysis and Results 14
     - 4.1.1 Identifying Influential Factors 14
     - 4.1.2 Evaluating Policy Impacts 16
     - 4.1.3 Detecting Structural Barriers 17
     - 4.1.4 Analyzing Seasonal Patterns 19
     - 4.1.5 Proposed Data-Driven Strategies 20
     - 4.1.6 Statistical Significance of Yearly Variation 22
   - 4.2 Discussion 23
   - 4.3 Conclusion 24

REFERENCES 25

APPENDIX A: DSNY ZONES 26

APPENDIX B: RESULTS IN TABLE FORMAT 27

APPENDIX C: PYTHON CODES AND OUTPUTS 29

.

1. PROBLEM

1.1 Overview of Recycling Rate Challenges

Urban recycling systems face many challenges that can affect the overall performance of waste management. In New York City, recycling rates can vary widely between different districts and over time. For example, differences in population density, local infrastructure, and community awareness often lead to significant variability among neighborhoods. Moreover, recycling performance can change across different periods—such as monthly or fiscal year intervals—due to seasonal trends or shifts in local policies. External factors, including economic conditions, educational programs, and the implementation of local initiatives, also influence these rates. As noted in Todd's 'Technical and Economic Analysis of New York City's Recycling System', understanding these challenges is crucial for developing effective recycling strategies [1]. This study focuses on examining these variations by comparing the recycling diversion and capture rates reported in NYC.

1.2 Recycling as a Pillar of the Circular Economy

Recycling is a central component of the Circular Economy (CE) since it basically reduces waste, conserves resources, and minimizes environmental impacts by returning materials to the production process [2]. When materials are efficiently recycled, less new raw material is needed, which in turn lowers energy consumption and greenhouse gas emissions. This efficient reuse of resources not only helps decrease landfill use but also supports sustainable growth by reducing our reliance on virgin resources.

Moreover, effective recycling systems foster innovation in product design and business models. For instance, companies may redesign products to make them easier to recycle or to incorporate recycled materials, thereby creating closed-loop systems that offer both economic and environmental benefits [3]. These improvements can inevitably lead not only to cost saving and job creation, but also to better resource recovery, which ultimately would make recycling a true pillar of the CE.

1.3 Objectives

This analysis focuses exclusively on New York City's recycling diversion and capture rates, and its principal objective is to recognize the factors that influence recycling performance in the city to ultimately provide data-driven insights based on an open dataset that covers the period from 2016 to 2019 [4]. This way, recommendations could be made in order to improve recycling efficiency.

However, more objectives were set to be accomplished by the end of this study. Specifically, these are:

- **Identify Influential Factors:** Determine which variables have the strongest correlation with changes in recycling rates across different districts and time periods.
- **Evaluate Policy Impacts:** Explore whether variations in recycling performance suggest the influence of local recycling initiatives or policy changes.
- **Detect Structural Barriers:** Identify obstacles that may be limiting the efficiency of recycling efforts, such as infrastructure issues or community engagement challenges.
- **Analyze Seasonal Patterns:** Examine how recycling rates vary by month and year to identify potential seasonal trends.
- **Propose Data-Driven Strategies:** Based on the analysis, suggest actionable recommendations to enhance recycling performance and support the circular economy.

2. BACKGROUND

2.1 Recycling and Circular Economy: Literature Insights

Recycling rates are often considered a cornerstone of CE strategies because they measure how effectively resources return to productive use instead of being discarded. Several studies highlight that low recycling rates in urban areas may slow down the development of a CE, since valuable materials end up in landfills rather than going back into production. In a linear system—where products follow a “take, make, dispose” path—cities face increased waste management costs and higher resource demand. Researchers, conversely, argue that shifting to a circular model saves materials, lowers emissions, and reduces the dependence on virgin resources [2] [3].

CE concepts have evolved over time. Early work focused for instance on reducing pollution and resource consumption. Recent research, however, proposes a more comprehensive and resource-efficient model where products and materials remain in use across multiple life cycles and thus, minimizing waste [2]. This approach goes beyond basic recycling by also encouraging design innovations, new business models, and better policy frameworks to keep materials circulating for as long as possible. Furthermore, a full CE depends not only on recycling but also on redesigning products for durability and reuse, and on developing efficient recovery systems [3].

Moreover, in the context of urban environments, “diversion rate” and “capture rate” are two important metrics used to measure the effectiveness of recycling programs. *Diversion rate* typically shows how much waste is diverted from disposal through recycling or composting, while *capture rate* usually refers to how much of a specific recyclable material is actually recovered, compared to what could have been collected [4] [6].

Even with these metrics, it is well-known that urban areas still face obstacles even with the implementation of different kinds of strategies [1]. Some regions may lack efficient collection infrastructure, or others may struggle with less consistent data. So, in the end, tracking these rates could not only help local authorities but also waste management organizations to see where gaps exist and how to improve program performance.

2.2 Comparative Recycling Strategies in Other Regions

In addition to the general literature, several case studies from other regions offer concrete examples of how enhanced recycling strategies can be implemented. For example, research conducted in Minnesota has demonstrated that municipalities with a well-funded recycling infrastructure and clear regulatory guidelines achieve noticeably higher recovery rates. In these areas, local governments invest in both technology and public outreach programs, which not only improve collection efficiency but also reduce contamination levels. Such investments have led to measurable improvements in material recovery, highlighting the potential for adapting these strategies in different urban settings. Moreover, comparative studies have shown that integrating economic incentives—such as reduced disposal fees and bonuses for high collection rates—can further boost recycling performance [7].

Furthermore, some studies suggest that relying only on extrinsic incentives—such as financial rewards or social influence—or solely on intrinsic motivations—like personal satisfaction or psychological attachment to a cause—may not be sufficient on their own to drive a substantial increase in recycling rates. Research indicates that when policies focus exclusively on one type of incentive, the behavioral change observed is often limited and short-lived [8]. In contrast, a more effective approach appears to be one that blends both extrinsic and intrinsic motivators. For example, while monetary incentives can provide immediate benefits and encourage quick action, integrating these with educational programs and community engagement efforts can foster a deeper sense of environmental responsibility.

Ultimately, these findings suggest that a combination of incentives—including a modernization of infrastructure, a rigorous application of policies, and better public-private partnerships—can create a more effective recycling system. The combination of stimuli can create a more sustainable shift in behavior, leading to longer-term improvements in recycling performance [8]. So, by analyzing these successful models, cities could better understand the mix of factors that lead to higher recovery rates and potentially tailor similar approaches to their own unique contexts [1]. For instance, while a city might have some effective recycling programs, it could still experience issues with contamination of maybe recyclable materials, which could potentially lower the overall quality and, thus, profitability.

2.3 Recycling in New York: Context and Challenges

New York City presents a particularly complex recycling landscape due to its high population density and diverse urban environment. Detailed analyses, such as Todd's technical and economic evaluation of NYC's recycling system, reveal that while the city has made significant strides in establishing recycling programs, there are still persistent challenges. These include issues with material contamination, inconsistent performance across districts, and the difficulty of maintaining high recovery rates in densely populated areas [1]. Other studies further emphasizes that local socio-economic factors and the design of current recycling policies contribute to these disparities, making it difficult to achieve uniform results citywide. For example, some local governments may have strong regulations but lack adequate funding for widespread education campaigns. Others may even rely on voluntary programs, leading to irregular results [9] [10].

Adopting a circular approach often requires rethinking financial incentives, consumer habits, and product design. For instance, deposit systems for beverage containers or electronics can encourage people to return items, which raises capture rates. Research showed that building a big and unified recycling processing center called materials recovery facility (MRF) could significantly increase recovery rates and yield net economic benefits compared to the existing approach [1]. This type of facility would have lowered overall disposal costs and better capture valuable materials. Of course, certain waste streams would still have needed unique outreach, more accessible drop-off points, or tailored processing methods. However, although it is not quite clear whether this sort of facility was actually built by the government or they kept relying on private contractors, research suggests that a mix of policy tools, public awareness efforts, and modernized recovery infrastructure can help expand urban recycling and advance the CE.

Furthermore, the dynamic nature of NYC's urban environment—with its rapid demographic shifts, varying local infrastructure, and evolving public policies—compounds the challenge of improving recycling rates. For instance, while some districts may benefit from strong community engagement and robust local initiatives, others could lag behind due to inadequate investment in recycling infrastructure. This uneven performance underscores the importance of adapting recycling programs to meet the specific needs of each of the different zones and districts, and in order to accomplish this, it is quite crucial that governments review successful programs in other places, to implement the best possible option in their own city, in this case, in New York City [1].

Such localized insights are quite necessary for developing strategies that not only improve general recovery rates but also advance the broader goals of a CE in the city. In the end, it is clear that recycling strategies have evolved, and discussions about improving recovery through enhanced infrastructure continue to surface.

Figure 1. Boroughs of New York City (the five major governmental districts that make up NYC) [11].

2.4 Key Research Question

Based on literature, and to achieve the goals, this work aims to answer a question:

> ***Which factors correlate most strongly with changes in urban recycling rates, and how do these factors shape the progress of the circular economy in a city like New York?***

Many academic sources explore how policy frameworks, infrastructure quality, education efforts, and local demographics can influence recycling performance. However, in terms of NYC, not all zones nor districts face the same challenges, so by drawing on what researchers have found regarding New York State, this analysis will focus solely on comparing different zones and districts within the City of New York over a defined time period. The goal is to ultimately identify which factors are most closely related to higher diversion and capture rates. This work could then guide recommendations for improving recycling programs and supporting the transition to a more CE.

3. METHODOLOGY

3.1 Dataset Description and Geographic Context

As mentioned, this analysis utilizes the NYC recycling dataset available from the city's open data portal [4]. This dataset covers the period from 2016 to 2019, a timeframe that is already defined in it; but selected for its completeness and relevance to current recycling practices. So, by focusing on this period, I aim to capture recent trends in recycling performance and assess the effectiveness of ongoing recycling strategies.

In addition, the dataset covers various geographic levels within New York City, including its 59 so-called "Community Districts," which are distributed among the city's five main governmental districts. This wide coverage makes it easier to compare data from multiple areas, as each Community District can have different waste collection patterns or local policies. However, although the city government has recently launched a plan to split the city into 20 different *Commercial Waste Zones* (in other words, 20 *DSNY Zones*, as shown in *Figure 2*) established by the *Department of Sanitation of New York*, that plan has only just begun, and its impact remains limited so far [12]. Because of this, the dataset's original timeframe (2016–2019) precedes the practical use of these new zones, so at that point, no one was yet operating under the full 20-zone design. As a result, the dataset groups New York City into only 7 zones, creating a simpler division than the new Commercial Waste Zones.

In the case of this study, the 7-zone approach is more practical and it remains the primary way these records are organized. Consequently, any comparisons or analyses in this study rely on those 7 zones, reflecting the city's waste management situation before the new DSNY plan could be put in place. Nevertheless, to better understand how the old 7 zones are grouped in the dataset, and to make sense of these new DSNY zones, refer to *Table A1* in *Appendix A*. For example, all 8 new zones that are planned to be in Manhattan are combined into a single zone labeled 'Manhattan' in the 2016-2019 dataset. This table summarizes each 'New DSNY Zone" with its corresponding 'Old DSNY Collection Zone' along with their district codes and key identifiers, which help link the dataset to the geographic breakdown shown both in *Figures 2* and *3*. This grouping ensures that all the necessary data is understandable to perform a thorough analysis.

Figure 2. NYC Commercial Waste Zones (DSNY) [12].

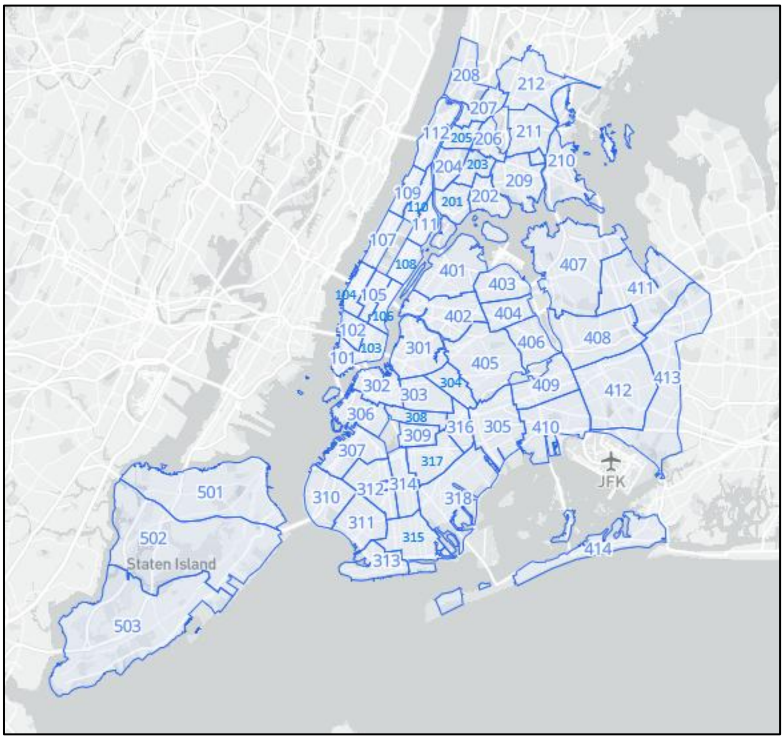

Figure 3. Community district codes in New York City [13].

3.2 Data Dictionary

To ensure a clear understanding of the dataset, this section provides a detailed data dictionary, which was directly obtained from the file downloaded from the NYC's open data site and placed in its entirety down here [4]. The dataset includes several columns such as "Zone," "District," "Fiscal Month," "Fiscal Year," "Diversion Rate," and "Capture Rate." Each of these fields is defined as follows: the "Zone" indicates the old DSNY zone, the "District" provides a more specific location within NYC, "Fiscal Year" and "Fiscal Month" denote the time period for the reported data, and the "Diversion Rate" and "Capture Rate" are key metrics used to evaluate recycling performance.

Table 1. Dataset Dictionary (Duplicated from NYC Open Data Source) [4].

Column Name	Column Description	Term, Acronym, Code Definitions or Additional Notes
Zone	Old DSNY Collection Zone	Old DSNY Zones include the following: Manhattan, Bronx, Brooklyn North, Brooklyn South, Queens West, Queens East, Staten Island
District	One of NYC's 59 community districts which correspond to Sanitation districts.	Community Districts are divided into their associated DSNY Zones. DSNY districts are coterminous with NYC's community districts.
Fiscal Month Number	The number of the fiscal month.	The NYC fiscal year starts in July so July is fiscal month 1.
Fiscal Year	Year that corresponds to NYC fiscal year.	The NYC fiscal year starts in July and ends in June.
Month Name	Name of the month	
Diversion Rate-Total	Total Recycling / Total Waste	Diversion rate is tonnage diverted divided by the sum of tonnage diverted and disposed. Disposed materials are sent via transfer stations to landfills or waste-to-energy facilities outside of NYC. Diverted materials are sent to reuse or recycling facilities inside or outside of NYC.
Capture Rate-Paper	Total Recyclable Paper / Max Recyclable Paper in Waste Stream	Capture rate is the amount of materials set out for residential recycling collection as a percentage of designated recyclable materials in both recycling and refuse streams. This ratio measures how much of the targeted materials are actually being recycled, which is a measure of how successfully such materials are recycled.
Capture Rate MGP	Total Recyclable Metal, Glass, Plastic & Beverage Cartons / Max Recyclable Metal, Glass, Plastic & Beverage Cartons in Waste Stream	
Capture Rate-Total	(Total Recycling - Leaves) / (Max Recyclable Paper + Max Recyclable Metal, Glass, Plastic & Cartons) x 100	

Key variables such as "Diversion Rate" and "Capture Rate" are central to this analysis, as they directly reflect how much waste is redirected from landfills and how much recyclable material is actually recovered. Any derived metrics or calculated fields will be explained in the analysis section, ensuring that readers have a complete understanding of the data used in this study.

3.3 Software and Tools

For this analysis, I used Python 3 running on Google Colaboratory [14]. The environment provided an interactive workspace with a user-friendly interface and the ability to easily share code and results. The main libraries used for data manipulation and visualization were *pandas*, *matplotlib.pyplot*, and *seaborn*. These tools allowed for efficient handling of the dataset and the creation of clear and informative graphs that illustrate trends and relationships within the recycling data.

Figure 4 below shows a code snippet that demonstrates how the libraries were imported in the Google Colab environment. In this snippet, we can see that I used Python 3 language along with the primary libraries for this analysis

Figure 4. Libraries importing code snippet.

```
import pandas as pd             # For handling DataFrames and data manipulation
import matplotlib.pyplot as plt  # For creating various types of graphs
import seaborn as sns           # For advanced visualization such as heatmaps
```

We can observe that each library has been imported with an alias—*pandas* as *pd*, *matplotlib.pyplot* as *plt*, and seaborn as *sns*. Using aliases helps simplify the code, making it easier to call functions from these libraries without having to write out their full names each time. This practice not only improves code readability but also enhances efficiency during the analysis process.

3.4 Data Import, Exploration, and Preprocessing

In this section, we describe the process of importing and exploring the dataset. The recycling data file obtained from the New York City open data site [4], and named "*NYC-recycling-diversion-and-capture-rates.csv*," was uploaded to Google Colab using the file uploader feature. Once the file was in the environment, it was imported into a pandas DataFrame using the *pd.read_csv()* function. *Figure 5* shows the code used.

Figure 5. File import into Google Colab.

```
from google.colab import files

df = pd.read_csv("/content/NYC-recycling-diversion-and-capture-rates.csv")
```

Here, *df* represents the DataFrame containing the dataset. This dataset covers the abovementioned period from 2016 to 2019, and includes data for all 59 districts, which are grouped into 7 zones.

After importing the dataset, I used the command *df.head()* to display the first few rows (refer to *Figure C1* in *Appendix C*). This step was performed to verify that the data was correctly loaded and that all columns—such as those described in the Data Dictionary (*Section 3.3*)—appear as expected. Furthermore, to ensure data integrity, I checked for missing or null values using the command *df.isnull()*. The output confirmed that the dataset was clean, as all cells returned "False," indicating that there were no missing values (refer to *Figure C2* in *Appendix C*).

3.5 Analytical Approach

After data importation and validation of information within the dataset, an analytical approach was implemented using the abovementioned libraries (refer to *Appendix C* to see Python codes used). In this case, descriptive statistics were employed to understand the variation in recycling rates across New York City. The mean was calculated to obtain the average recycling rates for each zone and district, providing a clear picture of overall performance. For instance, the average recycling rates per zone and district were calculated to identify areas that recycle more effectively versus those that lag behind. This step provided a quantitative basis for comparing performance across the city. Although the median was not directly used in this analysis, it is noted as a potentially useful measure for future studies to identify central tendencies that remain unaffected by extreme values. Likewise, while the standard deviation was not explicitly computed for the graphs, its use would be valuable to assess the variability in recycling rates between different zones and across years.

In addition, visualization techniques were applied to explore trends over time and differences between geographic areas. Bar charts were created to compare recycling performance among various zones and districts, while line graphs were used to display trends over fiscal months and years. A heatmap was also generated to display the correlation matrix between key recycling metrics, helping to examine the relationships between key recycling metrics. Pearson correlation analysis was then applied to quantify these relationships, which was chosen because it effectively measures the linear association between two numerical variables. After this, to assess the robustness of the patterns observed, a basic statistical significance test was applied to validate whether the observed differences between zones or years were statistically significant. This combination of methods allowed for a comprehensive analysis of the recycling data, ultimately supporting data-driven recommendations for improving recycling performance in NYC.

Finally, it is important to note that geographic data or shapefiles (SHP) were not available in the source file, so libraries such as *geopandas* or *folium* were not required, emphasizing that only *pandas*, *matplotlib.plotpy* and *seaborn* were used in this data analysis. So instead of working with SHP files, online GIS resources from NYC open data sites were used to interpret the results, enrich the discussion, and ultimately support the final conclusions.

3.6 Study Limitations

This study is exposed to many limitations that must be taken into account when interpreting the results. First, the temporal coverage of the dataset spans only from 2016 to 2019, which means that any recycling trends or changes that have occurred after 2019 are not captured in this analysis. Second, although the dataset provides comprehensive recycling metrics, it does not include certain external factors—such as detailed socio-economic indicators or community outreach efforts—that may also influence recycling performance. Finally, potential biases may arise from inconsistencies in data collection across different zones, which could affect the comparability of recycling rates among districts.

These limitations, underscore the need for cautious interpretation of the results and suggest areas for further research. In addition, it is crucial to mention that the integration of more contextual variables in future studies, such as population density, income levels, and education programs, would be beneficial to ultimately gain a deeper understanding of recycling performance. Addressing these gaps could inevitably help refine the analysis and provide better recommendations for improving urban recycling systems.

4. RESULTS AND DISCUSSION

4.1 Data Analysis and Results

This section presents the findings from the dataset described in *Chapter 3*, linking each subsection to one of the five objectives defined in *Section 1.4* of this study.

4.1.1 Identifying Influential Factors

Initially, the analysis focused on determining which zones of New York City reported the highest and lowest diversion rates and capture rates, and whether there were distinctive patterns that correlated with other potential variables.

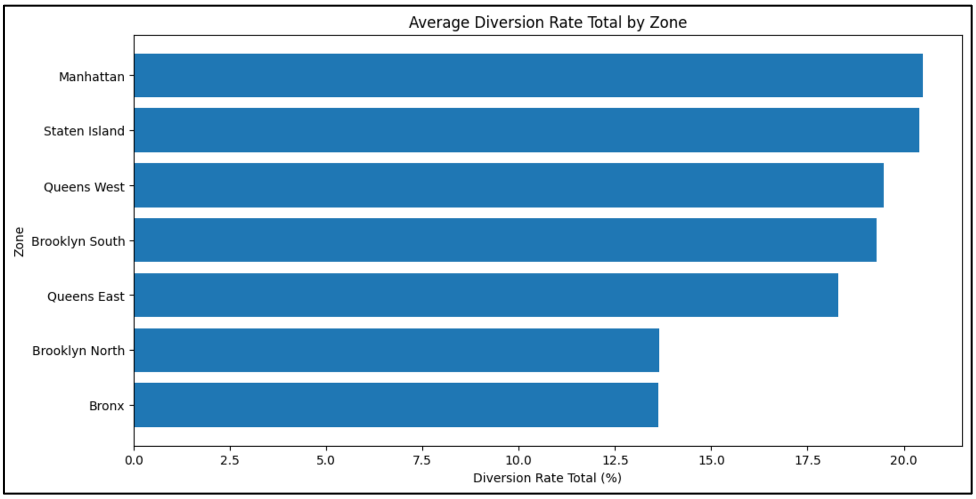

Figure 6. Average Diversion Rate Total by Zone.

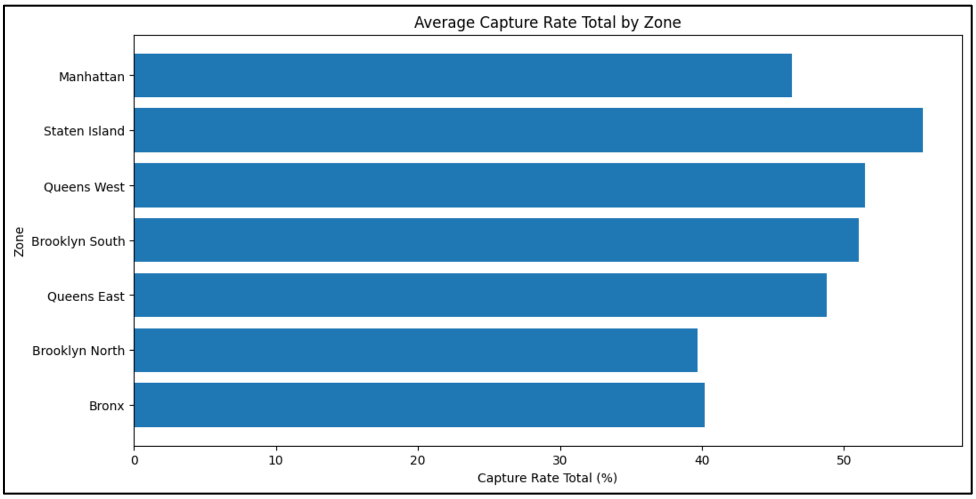

Figure 7. Average Capture Rate Total by Zone.

From these visualizations—and the data in *Table B1* in *Appendix B*—two zones stand out for higher performance: *Manhattan* and *Staten Island* post average diversion rates close to 20.5%. Besides, *Staten Island* also shows the highest overall capture rate (about 55.5%). Meanwhile, *Brooklyn North* and the *Bronx* consistently fall below 14% in their diversion rates and hover near 40% in capture rate.

These findings point to zone-specific factors that could include local infrastructure, socio-economic characteristics, or policy enforcement disparities. As a first indicator, it appears that economic or demographic variables (e.g., lower-income neighborhoods) may be contributing to lower recycling rates in certain areas. The following figure highlights that the Bronx and parts of Brooklyn North overlap with higher poverty concentrations.

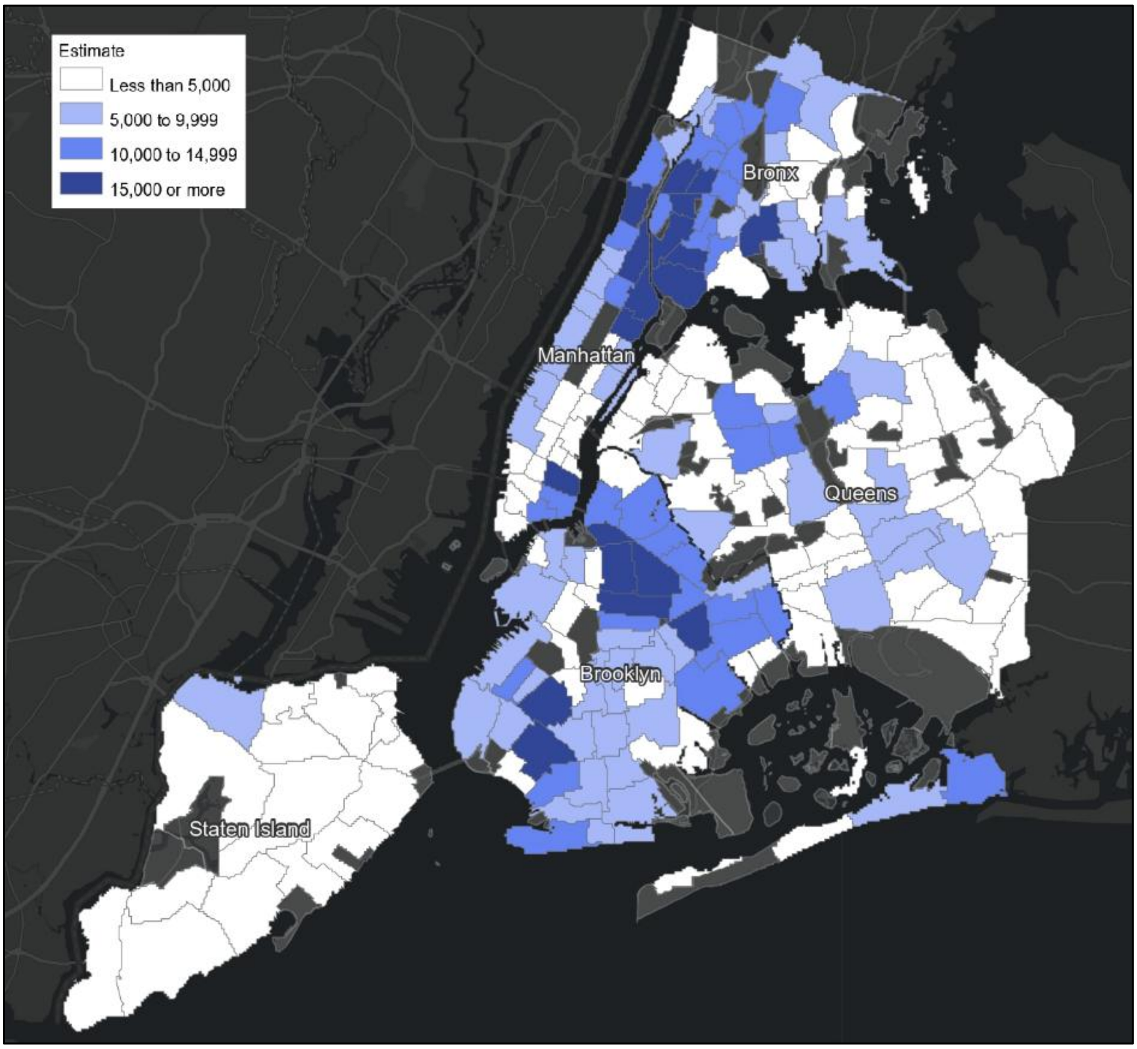

Figure 8. Population with Income Below Poverty Level. Source:
U.S. Census Bureau, 2018-2022 ACS Summary File [15].

4.1.2 Evaluating Policy Impacts

To see if recycling initiatives or policy shifts influenced performance, annual trends for both diversion and capture rates were examined from 2016 to 2019.

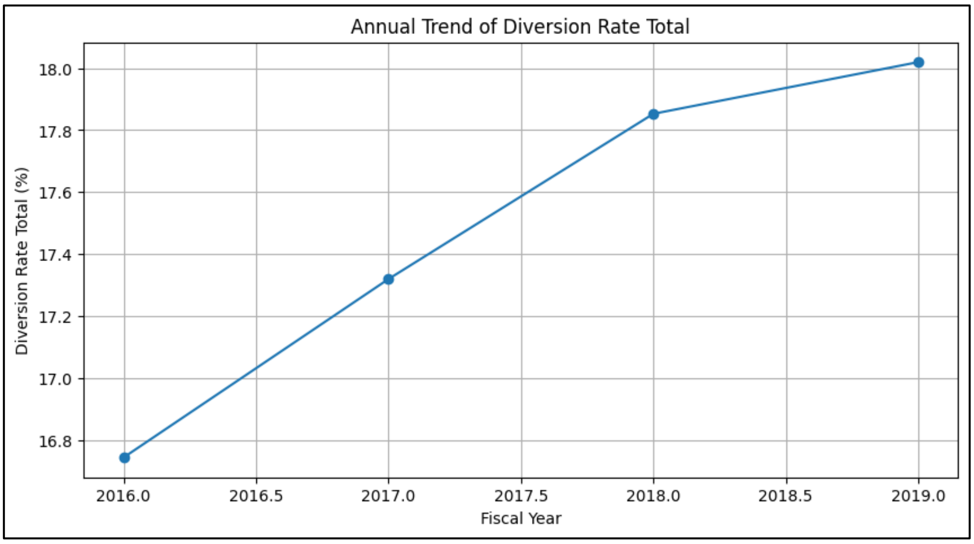

Figure 9. Annual Trend of Diversion Rate Total.

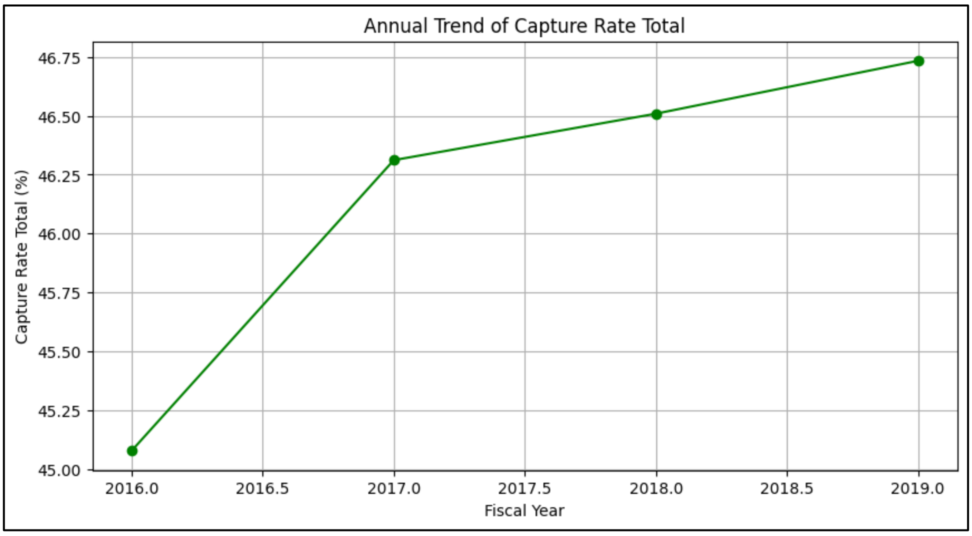

Figure 10. Annual Trend of Capture Rate Total.

The above data in *Figures 9* and *10* show that the *Diversion Rate-Total* field rose from about 16.7% in 2016 to 18.0% by 2019. Similarly, the *Capture Rate-Total* one had its most significant jump between 2016 (around 45%) and 2017 (46.3%), followed by smaller incremental increases in 2018 and 2019 (to see more detailed values refer to *Table B2* in *Appendix B*).

2017 appears to be a transitional year in which changes—possibly local initiatives, expanded public outreach, or updated recycling regulations—helped boost participation. This aligns with more observations that NYC introduced or strengthened various citywide waste management programs around that time [16]. Though this chapter focuses on the main dataset's materials, other sources suggest that multiple facets of recycling policy might have enhanced in 2017 for e-waste collection programs [17], which may also suggest a progress in collection policies for other types of waste.

4.1.3 Detecting Structural Barriers

A separate look at the lowest performing districts provided clues to potential barriers in both *Figures 11* and *12*:

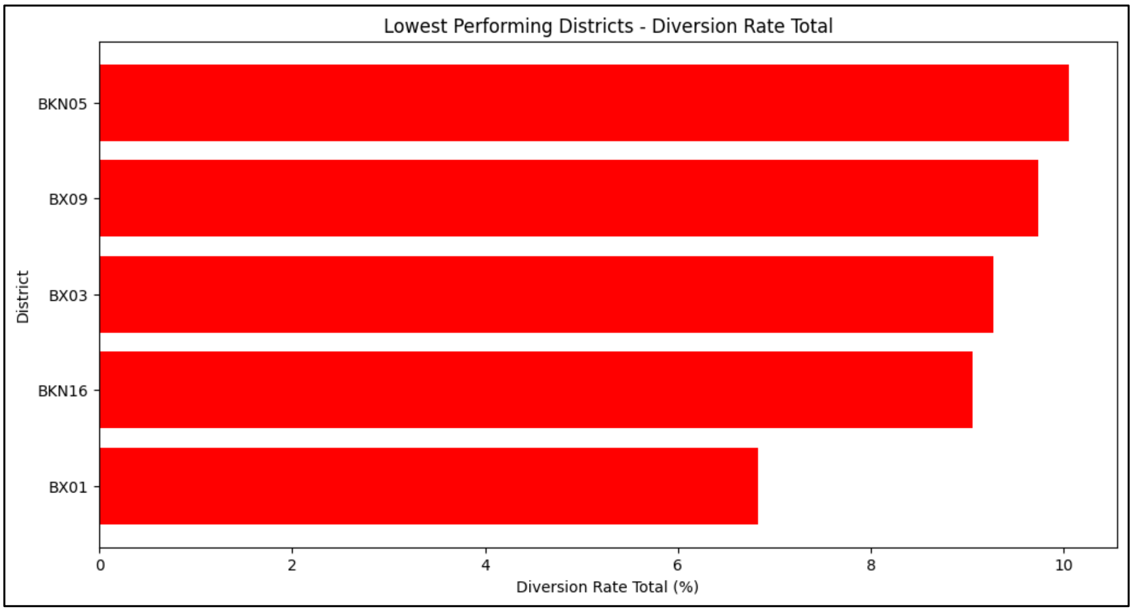

Figure 11. Lowest Performing Districts – Diversion Rate Total. This bar chart shows the five districts with the lowest diversion rates, sorted in ascending order. Each label corresponds to a Community District code: BKN05: Brooklyn North 05, BX09: Bronx 09, BX03: Bronx 03, BKN16: Brooklyn North 16, and BX01: Bronx 01. For more details regarding code districts, please refer to *Table A1* in *Appendix*

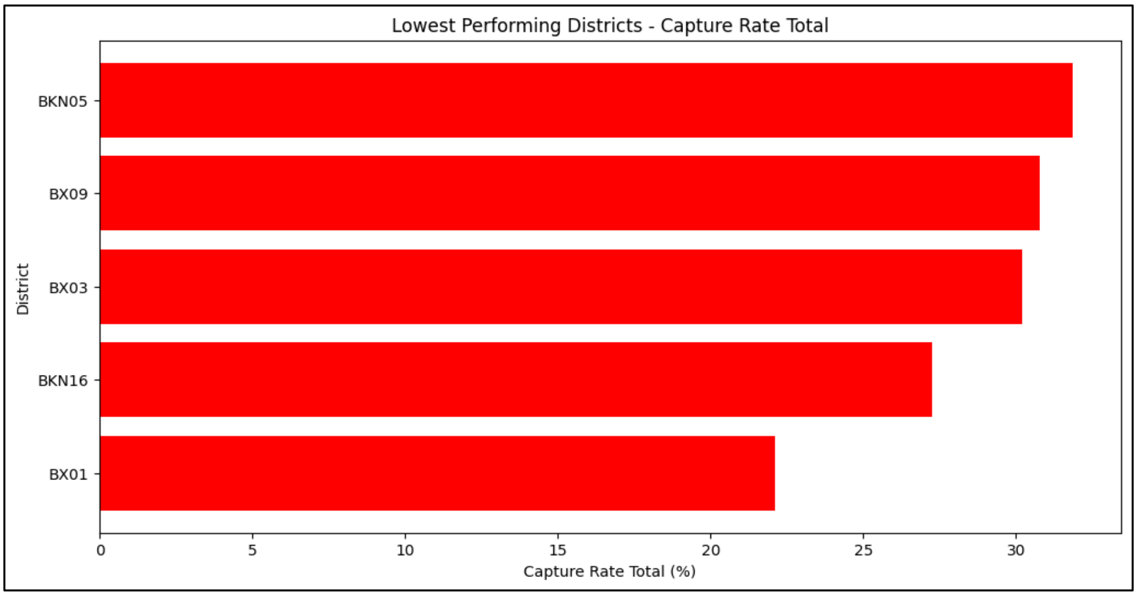

Figure 12. Lowest Performing Districts – Capture Rate Total. This bar chart presents the five districts with the lowest capture rates, also sorted in ascending order. Each label represents a Community District code: BKN05: Brooklyn North 05, BX09: Bronx 09, BX03: Bronx 03, BKN16: Brooklyn North 16, and BX01: Bronx 01. For more details regarding code districts, please refer to *Table A1* in *Appendix A*.

Specific community districts in the *Bronx* (BX01, BX03, BX09) and *Brooklyn North* (BKN16, BKN05) show diversion rates close to or below 10%. Their capture rates are also significantly low—one district barely reach 22% capture. From these graphs (see precise values in *Table B3* in *Appendix B*) some *probable* barriers that come to my mind and that could be happening in certain parts of the city—or at least that maybe happened a few years ago—and *possibly* are the reason for these results, are: insufficient recycling infrastructure (few drop-off points or inadequate container availability), lower public awareness due to fewer educational campaigns, or even some socio-economic barrier such as difficulties in areas with higher poverty rates, where it goes without saying, that limited resources or competing priorities could even lead to less engagement in recycling practices. These patterns could be related to the anecdotal evidence from the Poverty-Level Map (*Figure 8*), where many of these districts overlap with higher-poverty areas. Besides, another thing was pointed out in the *2017 Waste Characterization Study for NYC*, which noted that people who have never been asked to recycle often lag in recycling participation, despite the large public education [16]. In the next figure, we can see an estimate of the inhabitants of NYC owning a bachelor's degree or higher. It is notable that the majority of these individuals are concentrated in *Manhattan*, the richest borough in New York City, but also, we can observe how few people with such degrees are in the *Bronx* and some parts of *Brooklyn*, although the numbers in *Queens* and *Staten Island* are quite similar, so more research and data are needed.

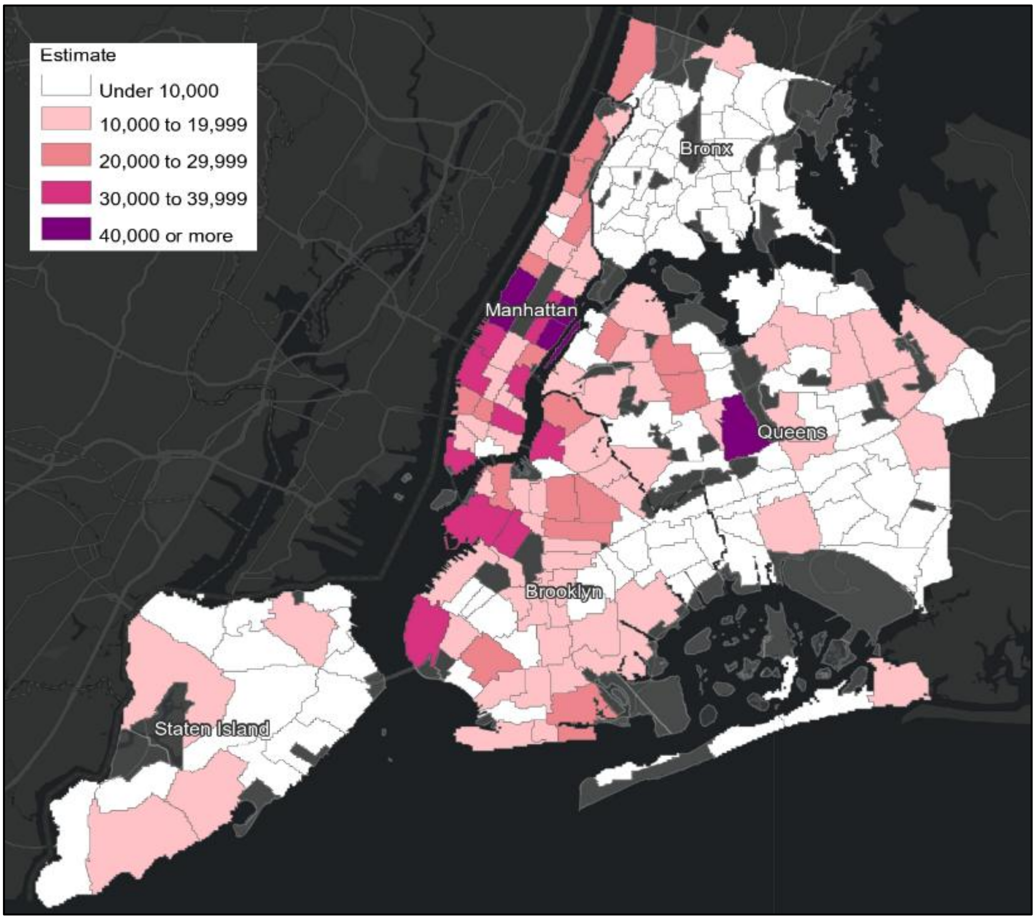

Figure 13. Population with a Bachelor's Degree or Higher. Source: U.S. Census Bureau, 2018-2022 ACS Summary File [18].

4.1.4 Analyzing Seasonal Patterns

Another key focus was determining how recycling rates fluctuated across the fiscal year's monthly breakdown.

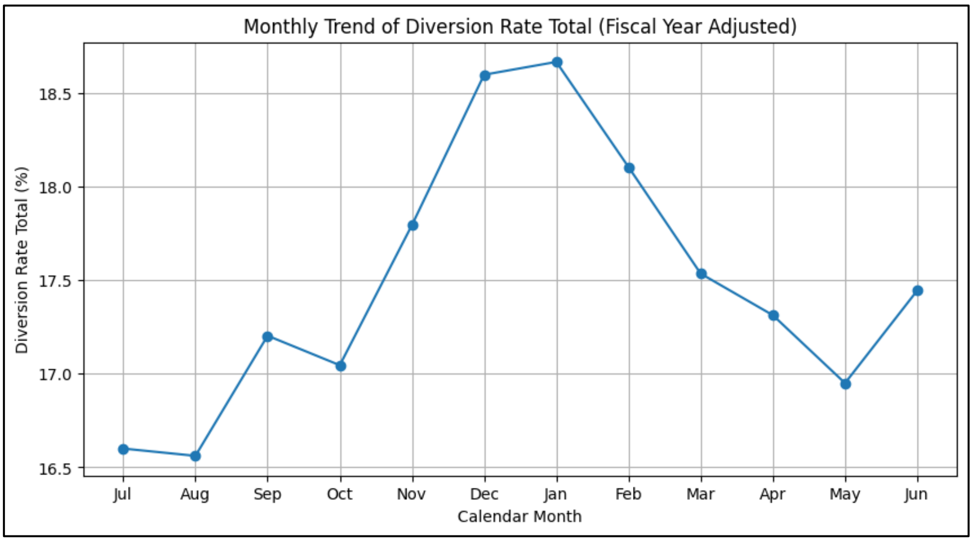

Figure 14. Monthly Trend of Diversion Rate Total (Fiscal Year Adjusted).

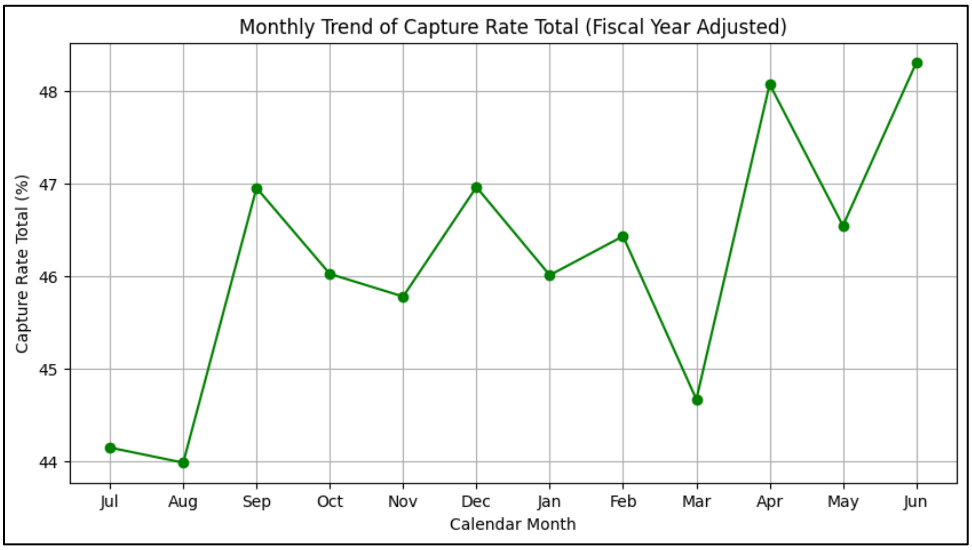

Figure 15. Monthly Trend of Diversion Rate Total (Fiscal Year Adjusted).

From these charts—see specific values in *Table B4* in *Appendix B*— we notice that both December and January typically have the highest diversion rates (around 18.6%). This could be attributed to holiday season waste (e.g., gift packaging) plus active year-end donation drives and city campaigns. Besides, it is quite clear how July and August drop significantly, to about 16.6%, perhaps due to summer travel and less resident participation. And finally, capture rates peak both in April and June (roughly 48%), which I could think that these are *possibly* linked to spring "green" campaigns and community clean-ups, but this is just an *idea* and further research is advisable. Finally, a small drop appears in March (almost 45%), which might be explained by transitional weather conditions or fewer targeted outreach efforts that month. These findings suggest that public information campaigns could be timed to match seasonal shifts, reinforcing or compensating for the months when residents are less likely to sort and recycle effectively.

4.1.5 Proposed Data-Driven Strategies

Finally, a correlation matrix was computed to see which recycling metrics most strongly affect one another. The parameters selected—Diversion Rate Total, Capture Rate Paper, Capture Rate MGP, and Capture Rate Total—were chosen because they represent the core performance indicators within the dataset for evaluating recycling behavior. By analyzing the relationships among these specific metrics, we can understand how improvements in the capture of specific material types (like paper or MGP—metal, glass, plastic) influence general recycling efficiency and diversion efforts. These variables are directly tied to operational results in waste management and are the most relevant for informing pointed strategies in a CE context.

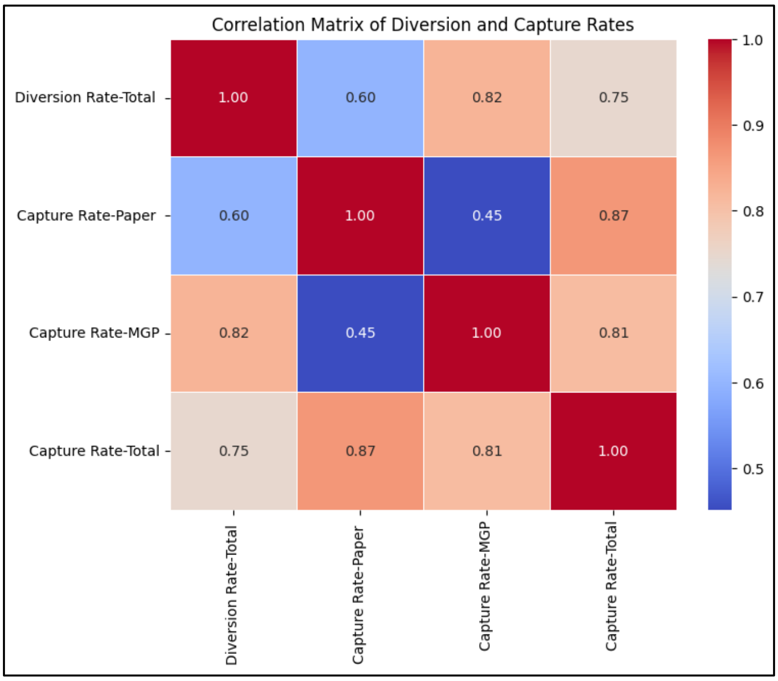

Figure 16. Correlation Matrix of Diversion and Capture Rates.

Now, we can observe some things in the correlation. For instance, there is strong positive correlation (0.87) between *Capture Rate-Paper* and *Capture Rate-Total* which suggests that boosting paper-collection programs can significantly increase total capture rates. While it is true that any capture rate (e.g., for metal, glass, plastic) could logically contribute to a higher total, this result explicitly points to *paper* as the material with the strongest statistical influence on overall capture. Therefore, this insight helps prioritize which material ultimately might offer the highest return on investment when it comes to improving recycling efforts.

Additionally, *Diversion Rate Total* correlates well with *Capture Rate MGP* (metal, glass, plastic), at about 0.82, indicating that improved collection of MGP also lifts overall diversion. Lastly, *Capture Rate Total* and *Diversion Rate Total* are moderately to highly correlated (0.75), affirming that when more designated recyclables are captured, overall landfill diversion also rises. Based on these relationships, some data-driven strategies can emerge. Firstly, it is important to prioritize paper recycling by expanding drop-off sites and targeted education, since paper capture strongly influences total rates. Secondly, a key plan could be simply improving MGP infrastructure, especially in the under-performing districts, using better signage, container availability, and consistent pick-up schedules. And last but not least, it is also crucial to focus on critical months (summer and early spring) by launching or amplifying campaigns to avoid seasonal dips in both diversion and capture rates.

4.1.6 Statistical Significance of Yearly Variation

To complement the annual trends discussed earlier, a basic statistical test was used to determine whether the observed increases in Diversion and Capture Rates between 2016 and 2019 were statistically significant. Specifically, a unilateral ANOVA (Analysis of Variance) was applied to test if the average rates varied significantly across the whole four fiscal years.

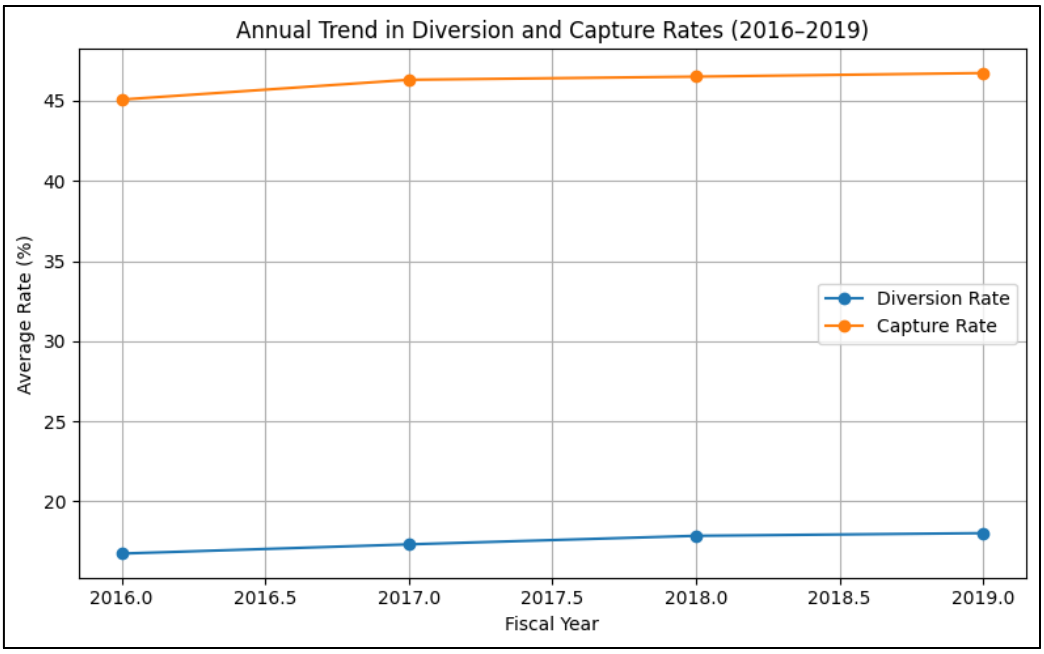

Figure 17. Annual Trend in Diversion and Capture Rates (2016–2019).

The results of the ANOVA confirm that the increases in both Diversion Rate and Capture Rate are statistically significant. For the Diversion Rate, the p-value was less than 0.0001 ($F = 8.56$), and for the Capture Rate, the p-value was approximately 0.0048 ($F = 4.31$). These values are well below the common significance threshold of 0.05, which means that the differences observed between years are unlikely to be due to some random variation.

This definitely adds statistical support to the earlier interpretation that recycling performance improved over the years, especially during and after 2017. The 2017 spike in the Capture Rate might reflect the effect of new or reinforced policy initiatives introduced around that time, as already brought up in *Section 4.1.3*. Although the year-to-year growth in later periods appears moderate, the analysis shows that the general trend is meaningful and not just due to chance.

4.2 Discussion

The findings above align with previous research on urban recycling challenges, which highlights that a combination of infrastructure, policy, and community engagement shapes recycling performance [1] [7]. Socio-economic factors emerge as a key influence: low-performing districts in the Bronx and Brooklyn North coincide with higher poverty levels and low education levels—especially the Bronx—, a pattern that matches the spatial overlap of lower recycling rates with under-resourced areas. *Figures 8* and *13* further support the notion that improving accessibility (for instance, more convenient bins or frequent pickups) and strengthening education could help mitigate structural barriers.

Also, policy changes also appear significant, according to some research and other recent city updates, substantial improvements often stem from mandatory separation laws, extended collection services, or stricter enforcement [1] [16]. Although some sources, including Recycling Today, focus mainly on e-waste, the broader principle is that consistent policy measures—such as extended producer responsibility or deposit return systems—can markedly raise overall diversion [17]. Moreover, seasonal variability likewise plays a role in shaping recycling success. Wintertime peaks in diversion rates coincide with other large cities' experiences, where holiday consumption leads to more recyclable packaging, while summer months often show reduced engagement. Research conducted in different parts of the world supports this finding by noting that local events and public vacations can impact directly on the routine recycling [19]. Adapting to these seasonal dynamics may strengthen program efficacy. Additionally, the correlation analysis highlights the importance of prioritizing the largest waste streams to boost total recycling performance, which aligns with CE models suggesting that once major recyclable fractions are consistently captured, recycling rates improve. Lastly, the ANOVA analysis was crucial to see that the overall curve is quite significant and randomness can be discarded in this case.

Furthermore, recent work suggests that combining extrinsic incentives (for example, deposit refunds) with moral or social motivations (like environmental self-identity or neighbor ties) can sometimes promote sustainable actions. However, it may also weaken people's intrinsic sense of environmental responsibility, reducing the global impact of those incentives [8] [20].

Taken together, these observations point out the value of inter-district collaboration and targeted strategy. For instance, Staten Island's relatively high capture rates could mean better alignment of infrastructure, engagement initiatives, and supportive policy. Transferring these methods or adapting them to the contexts of the Bronx and Brooklyn North may prove more effective than a one-size-fits-all intervention, especially when combined with year-round public education and well-timed campaigns.

4.3 Conclusion

Summing up, the results confirm and expand upon the research question. This analysis indicates that paper and MGP capture rates act as pivotal drivers: districts excelling in these areas generally achieve higher diversion. Socio-economic variables are strongly associated with poor performance, emphasizing the need for tailored infrastructure and more inclusive campaigns. Policy shifts around 2017 likely reinforced upgrades, showing that legislation and more rigorous enforcement can elevate both capture and diversion. Seasonal patterns can either reinforce these gains or undermine them, so planning strategic campaigns during months with historically weaker results may help sustain progress. These combined factors shape the city's trajectory toward a more CE by either accelerating material reuse through good systems or impeding it when social and economic barriers remain unaddressed.

Additionally, while the theory behind diversion and capture is solid, practices in real life could vary from place to place. Some processes might only return materials to the cycle temporarily, use extra resources, or even release unintended emissions. These realities highlight the need for standard rules and systems to improve both diversion and capture rates. Finally, although recycling rates showed slight improvement from 2016 to 2019, applying data-driven strategies in past datasets to focus on areas of opportunity is quite vital for improving public and private strategies in the present and in the long run.

REFERENCES

[1] Todd, C. E. (2002). *TECHNICAL AND ECONOMIC ANALYSIS OF THE NEW YORK CITY RECYCLING SYSTEM*. Retrieved from: https://www.columbia.edu/cu/seas/earth/wtert/newwtert/Research/sofos/Todd_thesis.pdf

[2] Di Maio, F., & Rem, P. C. (2015). A Robust Indicator for Promoting Circular Economy through Recycling. *Journal of Environmental Protection*, *06*(10), 1095–1104. https://doi.org/10.4236/jep.2015.610096

[3] Ellen MacArthur Foundation, (2013). *Towards the circular economy Vol. 1: Economic and business rationale for an accelerated transition*. Chapter 2.

[4] *Recycling diversion and capture Rates | NYC Open Data*. (2019, July 24). Retrieved from https://data.cityofnewyork.us/Environment/Recycling-Diversion-and-Capture-Rates/gaq9-z3hz/about_data

[5] *What is a Diversion Rate?* (n.d.). Plastic Action Centre. Retrieved from: https://plasticactioncentre.ca/directory/what-is-a-diversion-rate/

[6] *What is a Capture Rate?* (n.d.). Plastic Action Centre. Retrieved from: https://plasticactioncentre.ca/directory/what-is-a-capture-rate/

[7] Sidique, S. F., Joshi S. V., & Lupi F. (2010). Factors influencing the rate of recycling: An analysis of Minnesota counties. *Resources, Conservation and Recycling*, *54*(4), 242–249. https://doi.org/10.1016/j.resconrec.2009.08.006

[8] Hornik J., Cherian J., Madansky M., & Lupi F. (1995). Determinants of recycling behavior: A synthesis of research results. *The Journal of Socio-Economics*, *24*(1), 105–127. https://doi.org/10.1016/1053-5357(95)90032-2

[9] Shah, P., & Yang, J. Z. (2025). Using self-affirmation to encourage recycle and reuse behaviors in New York State. *Environmental Development*, 101149. https://doi.org/10.1016/j.envdev.2025.101149

[10] Liu, Z., Yang, J. Z., Bloomfield, A., Clark, S. S., & Shelly, M. A. (2022). Predicting recycling intention in New York state: The impact of cognitive and social factors. *Environmental Development*, *43*, 100712. https://doi.org/10.1016/j.envdev.2022.100712

[11] *File:New York City District Map 2.SVG - Wikimedia Commons*. (2009, July 12). https://commons.wikimedia.org/wiki/File:New_York_City_District_Map_2.svg

[12] *NYC commercial waste zones: What businesses need to know*. Great Forest. (2024, November 18). Retrieved from https://greatforest.com/nyc-commercial-waste-zones/

[13] *NYC Boundaries*. (n.d.). Retrieved from https://boundaries.beta.nyc/?map=cd

[14] *Google Colab*. (n.d.). https://colab.research.google.com/

[15] *Experience*. (n.d.). Retrieved from https://experience.arcgis.com/experience/c625a78991d34ae59deb7a33806ac0d1/page/Poverty-%7C-Below-Poverty?views=Estimate-----------

[16] Garcia, K. (2017). 2017 NYC Residential, School, and NYCHA Waste Characterization Study. In *NYC Department of Sanitation*. https://dsny.cityofnewyork.us/wp-content/uploads/2018/04/2017-Waste-Characterization-Study.pdf

[17] Workman, M. (Ed.). (2018, May 16). *New York City reduces electronics in waste stream by 60 percent*. Recyclingtoday.com. Retrieved from https://www.recyclingtoday.com/news/new-york-ewaste-drop-recycling/

[18] *Experience*. (n.d.). Retrieved from https://experience.arcgis.com/experience/c625a78991d34ae59deb7a33806ac0d1/page/Education-%7C-Bachelor's-Degree-or-Higher?views=Estimate-----------

[19] Yang, J., Jiang P., Zheng M., Zhou J., & Liu X. (2022). Investigating the influencing factors of incentive-based household waste recycling using structural equation modelling. *Waste Management*, *142*, 120–131. https://doi.org/10.1016/j.wasman.2022.02.014

[20] Xu L., Yang H., & Ling M. (2023). Factors moderating the effect of financial incentives on household recycling: A motivation crowding perspective. *Environmental Impact Assessment Review*, *100*, 107078. https://doi.org/10.1016/j.eiar.2023.107078

APPENDIX A: DSNY ZONES

Table A1. Zone Names and District Codes.

Old DSNY Zone (in dataset)	New DSNY Zone (Fig. 2)	Community District	Key identifier (Fig. 3)	Old DSNY Zone (in dataset)	New DSNY Zone (Fig. 2)	Community District	Key identifier (Fig. 3)
Manhattan	Lower Manhattan	MN01	101	Brooklyn South	Brooklyn West	BKS06	306
	Manhattan Southwest	MN02	102		Brooklyn West	BKS07	307
	Manhattan Southeast	MN03	103		Brooklyn Southwest	BKS10	310
	Manhattan West	MN04	104		Brooklyn Southwest	BKS11	311
	Midtown North/South	MN05	105		Brooklyn Southwest	BKS12	312
	Manhattan Southeast	MN06	106		Brooklyn South	BKS13	313
	Manhattan West	MN07	107		Brooklyn South	BKS14	314
	Manhattan Northeast	MN08	108		Brooklyn South	BKS15	315
	Upper Manhattan	MN09	109		Brooklyn East	BKS18	318
	Upper Manhattan	MN10	110	Queens West	Queens West	QW01	401
	Manhattan Northeast	MN11	111		Queens West	QW02	402
	Upper Manhattan	MN12	112		Queens Central	QW03	403
Bronx	Bronx West	BX01	201		Queens Central	QW04	404
	Bronx East	BX02	202		Queens Central	QW05	405
	Bronx West	BX03	203		Queens Central	QW06	406
	Bronx West	BX04	204		Queens Southeast	QW09	409
	Bronx West	BX05	205	Queens East	Queens Northeast	QE07	407
	Bronx West	BX06	206		Queens Northeast	QE08	408
	Bronx West	BX07	207		Queens Southeast	QE10	410
	Bronx West	BX08	208		Queens Northeast	QE11	411
	Bronx East	BX09	209		Queens Southeast	QE12	412
	Bronx East	BX10	210		Queens Southeast	QE13	413
	Bronx East	BX11	211		Queens Southeast	QE14	414
	Bronx East	BX12	212	Staten Island	Staten Island	SI01	501
Brooklyn North	Brooklyn North	BKN01	301		Staten Island	SI02	502
	Brooklyn West	BKN02	302		Staten Island	SI03	503
	Brooklyn North	BKN03	303				
	Brooklyn North	BKN04	304				
	Brooklyn East	BKN05	305				
	Brooklyn East	BKN08	308				
	Brooklyn East	BKN09	309				
	Brooklyn East	BKN16	316				
	Brooklyn East	BKN17	317				

APPENDIX B: RESULTS IN TABLE FORMAT

Table B1. Average Recycling Rates by Zone.

Zone	Diversion Rate-Total	Capture Rate-Total
Bronx	13.614674	40.198558
Brooklyn North	13.658001	39.659
Queens East	18.309736	48.768839
Brooklyn South	19.300317	51.058554
Queens West	19.468078	51.505599
Staten Island	20.410823	55.533794
Manhattan	20.489944	46.331025

Table B2. Annual Recycling Trends.

Fiscal Year	Diversion Rate-Total	Capture Rate-Total
2016	16.743875	45.078814
2017	17.319654	46.311827
2018	17.852903	46.508928
2019	18.019666	46.733341

Table B3. Lowest Performing Districts in Recycling.

District	Diversion Rate-Total	Capture Rate-Total
BX01	6.833435	22.108491
BKN16	9.049374	27.234333
BX03	9.274373	30.208657
BX09	9.737491	30.789084
BKN05	10.049036	31.842562

Table B4. Monthly Recycling Rates (Fiscal Year Adjusted).

Calendar Month	Diversion Rate-Total	Capture Rate-Total
Jul	16.600315	44.150677
Aug	16.561398	43.988231
Sep	17.202731	46.953296
Oct	17.0453	46.025408
Nov	17.79361	45.781062
Dec	18.596262	46.962349
Jan	18.66475	46.011804
Feb	18.103569	46.432204
Mar	17.534393	44.670795
Apr	17.312404	48.072866
May	16.949693	46.545442
Jun	17.443866	48.304598

Table B5. Matrix Correlation.

	Diversion Rate-Total	Capture Rate-Paper	Capture Rate-MGP	Capture Rate-Total
Diversion Rate-Total	1	0.595113	0.823796	0.748865
Capture Rate-Paper	0.595113	1	0.451435	0.866771
Capture Rate-MGP	0.823796	0.451435	1	0.809871
Capture Rate-Total	0.748865	0.866771	0.809871	1

Table B6. ANOVA Test Results.

Metric	F-value	p-value	Significant?
Diversion Rate	8.56	0.000012	Yes
Capture Rate	4.31	0.0048	Yes

APPENDIX C: PYTHON CODES AND OUTPUTS

Figure C1. *df.head() output.*

```
df.head()
```

	Zone	District	Fiscal Month Number	Fiscal Year	Month Name	Diversion Rate-Total	Capture Rate-Paper	Capture Rate-MGP	Capture Rate-Total
0	Brooklyn North	BKN01	10	2019	April	14.687093	44.909160	43.034062	44.146764
1	Brooklyn North	BKN02	10	2019	April	19.950181	34.194020	57.947031	41.213700
2	Brooklyn North	BKN03	10	2019	April	12.164161	33.521557	44.919731	38.155937
3	Brooklyn North	BKN04	10	2019	April	15.541803	35.211361	68.511260	48.750755
4	Brooklyn North	BKN05	10	2019	April	10.051845	22.265430	45.051791	31.530129

Figure C2. *df.isnull() output.*

```
df.isnull()
```

	Zone	District	Fiscal Month Number	Fiscal Year	Month Name	Diversion Rate-Total	Capture Rate-Paper	Capture Rate-MGP	Capture Rate-Total
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
2827	False	False	False	False	False	False	False	False	False
2828	False	False	False	False	False	False	False	False	False
2829	False	False	False	False	False	False	False	False	False
2830	False	False	False	False	False	False	False	False	False
2831	False	False	False	False	False	False	False	False	False

2832 rows × 9 columns

Figure C3. Python code for *Section 4.1.1*

```
# Grouping by zone to analyze average performance
zone_avg = df.groupby("Zone")[["Diversion Rate-Total", "Capture Rate-Total"]].mean().sort_values(by="Diversion Rate-Total")

# Chart: Diversion Rate-Total by Zone
plt.figure(figsize=(12,6))
plt.barh(zone_avg.index, zone_avg["Diversion Rate-Total"])
plt.xlabel("Diversion Rate Total (%)")
plt.ylabel("Zone")
plt.title("Average Diversion Rate Total by Zone")
plt.show()

# Chart: Capture Rate-Total by Zone
plt.figure(figsize=(12,6))
plt.barh(zone_avg.index, zone_avg["Capture Rate-Total"])
plt.xlabel("Capture Rate Total (%)")
plt.ylabel("Zone")
plt.title("Average Capture Rate Total by Zone")
plt.show()

# Showing table
display(zone_avg)
```

Figure C4. Python code for *Section 4.1.2*

```
# Calculating the average annual recycling rates
yearly_avg = df.groupby("Fiscal Year")[["Diversion Rate-Total", "Capture Rate-Total"]].mean()

# Graph: Evolution of the Diversion Rate-Total over time
plt.figure(figsize=(10,5))
plt.plot(yearly_avg.index, yearly_avg["Diversion Rate-Total"], marker='o', linestyle='-')
plt.xlabel("Fiscal Year")
plt.ylabel("Diversion Rate Total (%)")
plt.title("Annual Trend of Diversion Rate Total")
plt.grid(True)
plt.show()

# Graph: Evolution of the Capture Rate-Total over time
plt.figure(figsize=(10,5))
plt.plot(yearly_avg.index, yearly_avg["Capture Rate-Total"], marker='o', linestyle='-', color='green')
plt.xlabel("Fiscal Year")
plt.ylabel("Capture Rate Total (%)")
plt.title("Annual Trend of Capture Rate Total")
plt.grid(True)
plt.show()

# Calculating the average annual recycling rates
yearly_avg = df.groupby("Fiscal Year")[["Diversion Rate-Total", "Capture Rate-Total"]].mean()


# Showing the table with the evolution of recycling rates by year
display(yearly_avg)
```

Figure C5. Python code for *Section 4.1.3*

```python
# Identifying the 5 worst performing districts in Diversion Rate-Total and Capture Rate-Total
lowest_performing_districts = df.groupby("District")[["Diversion Rate-Total", "Capture Rate-Total"]].mean().nsmallest(5, "Diversion Rate-Total")

# Chart: Districts with the lowest Diversion Rate-Total
plt.figure(figsize=(12,6))
plt.barh(lowest_performing_districts.index, lowest_performing_districts["Diversion Rate-Total"], color='red')
plt.xlabel("Diversion Rate Total (%)")
plt.ylabel("District")
plt.title("Lowest Performing Districts - Diversion Rate Total")
plt.show()

# Chart: Districts with the lowest Capture Rate-Total
plt.figure(figsize=(12,6))
plt.barh(lowest_performing_districts.index, lowest_performing_districts["Capture Rate-Total"], color='red')
plt.xlabel("Capture Rate Total (%)")
plt.ylabel("District")
plt.title("Lowest Performing Districts - Capture Rate Total")
plt.show()

# Showing table with low-performing districts
display(lowest_performing_districts)
```

Figure C4. Python code for *Section 4.1.4*

```python
# Correcting fiscal year month names to the actual calendar
fiscal_to_calendar = {
    1: "Jul", 2: "Aug", 3: "Sep", 4: "Oct", 5: "Nov", 6: "Dec",
    7: "Jan", 8: "Feb", 9: "Mar", 10: "Apr", 11: "May", 12: "Jun"
}
df["Calendar Month"] = df["Fiscal Month Number"].map(fiscal_to_calendar)

# Grouping by calendar month
monthly_avg = df.groupby("Calendar Month")[["Diversion Rate-Total", "Capture Rate-Total"]].mean()
months_order = ["Jul", "Aug", "Sep", "Oct", "Nov", "Dec", "Jan", "Feb", "Mar", "Apr", "May", "Jun"]
monthly_avg = monthly_avg.reindex(months_order)

# Graph: Diversion Rate-Total per month
plt.figure(figsize=(10,5))
plt.plot(monthly_avg.index, monthly_avg["Diversion Rate-Total"], marker='o', linestyle='-')
plt.xlabel("Calendar Month")
plt.ylabel("Diversion Rate Total (%)")
plt.title("Monthly Trend of Diversion Rate Total (Fiscal Year Adjusted)")
plt.grid(True)
plt.show()

# Graph: Capture Rate-Total per month
plt.figure(figsize=(10,5))
plt.plot(monthly_avg.index, monthly_avg["Capture Rate-Total"], marker='o', linestyle='-', color='green')
plt.xlabel("Calendar Month")
plt.ylabel("Capture Rate Total (%)")
plt.title("Monthly Trend of Capture Rate Total (Fiscal Year Adjusted)")
plt.grid(True)
plt.show()

# Showing table
display(monthly_avg)
```

Figure C5. Python code for *Section 4.1.5*

```
# Calculating correlation between recycling rates
correlation_matrix = df[["Diversion Rate-Total", "Capture Rate-Paper", "Capture Rate-MGP", "Capture Rate-Total"]].corr()

# Graph: Correlation matrix
plt.figure(figsize=(8,6))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5)
plt.title("Correlation Matrix of Diversion and Capture Rates")
plt.show()

# Showing table
display(correlation_matrix)
```

Figure C6. Python code for *Section 4.1.6*

```
from scipy.stats import f_oneway

# Grouping by year and calculate averages
yearly_stats = df.groupby("Fiscal Year")[["Diversion Rate-Total", "Capture Rate-Total"]].mean().reset_index()

# Preparing data by year
diversion_by_year = [df[df["Fiscal Year"] == year]["Diversion Rate-Total"] for year in range(2016, 2020)]
capture_by_year = [df[df["Fiscal Year"] == year]["Capture Rate-Total"] for year in range(2016, 2020)]

# ANOVA
anova_diversion = f_oneway(*diversion_by_year)
anova_capture = f_oneway(*capture_by_year)

# Showing results
print("ANOVA Diversion Rate:", anova_diversion)
print("ANOVA Capture Rate:", anova_capture)

# Graph
plt.figure(figsize=(8, 5))
plt.plot(yearly_stats["Fiscal Year"], yearly_stats["Diversion Rate-Total"], marker="o", label="Diversion Rate")
plt.plot(yearly_stats["Fiscal Year"], yearly_stats["Capture Rate-Total"], marker="o", label="Capture Rate")
plt.xlabel("Fiscal Year")
plt.ylabel("Average Rate (%)")
plt.title("Annual Trend in Diversion and Capture Rates (2016-2019)")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```