**Recyclables and Organics in Metropolitan Minnesota**

Allen Majd and Gustavo Orozco Nuñez

STAT 301: Analysis of Variance and Multivariate Analysis

**Introduction**

In 1989, the Minnesotan legislature passed the Governor’s Select Committee on Recycling and the Environment (SCORE) laws to monitor and improve waste management efforts throughout the state. Since 1989, all 87 Minnesotan counties have reported various waste management information (funding, costs, waste reduction programs, etc.) to the Minnesota Pollution Control Agency who have compiled the county data to create the SCORE report and SCORE data set.

Examining the SCORE data, analysts constructed two research questions for further analysis: (1) Did the amount of collected recyclables significantly increase between 2009-2014, 2014-2019, or 2009-2019 in Minnesota’s metropolitan counties? (2) Did the amount of collected organics significantly increase between 2009-2014, 2014-2019, or 2009-2019 in Minnesota’s metropolitan counties? Using the SCORE data, a subset data set (metro) was created containing five columns of data: *Year*, *County*, *ID*, *Recycle\_Tons*, and *Organic\_Tons*. *Year* was defined as the data collection year and was selected as the independent categorical variable for research questions 1 and 2. *County* was defined as one of the seven metropolitan counties in Minnesota (Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington). *ID* assigned a unique value to each *County* (Anoka = 1, Carver = 2 , Dakota = 3, Hennepin = 4 , Ramsey = 5 , Scott = 6, and Washington = 7). *Recycle\_Tons* was defined as tons of collected recyclables from each metropolitan Minnesota county and was selected as the dependent numerical variable for research question 1 (RQ1). *Organic\_Tons* was defined as tons of collected organics from each metropolitan Minnesota county and was selected as the dependent numerical variable for research question 2 (RQ2). RQ1 and RQ2 are both represented by the following null and alternative hypotheses.

At least 2 means are different.

**Methods**

R was used to analyze the metro data set. The One-Way Repeated Measures ANOVA was selected as the preferred hypothesis test because the mean amount differences of collected recyclables and mean amount differences of collected organics were compared, only one independent categorical variable (*Year*) and one dependent numerical variable (*Recycle*\_*Tons* for RQ1 or *Organic\_Tons* for RQ2) were selected, and *Year* was a repeated measure with three levels (2009, 2014, and 2019). If an ANOVA result was significant, a Tukey test was performed to obtain more details about the significant ANOVA result. An alpha level of 0.05 was used to evaluate ANOVA and Tukey results.

**Conclusion**

The One-Way Repeated Measures ANOVA results are shown in Table 1. Since the p-value of Recyclables is significant, the null hypothesis is rejected. There is sufficient evidence that the amount of collected recyclables is significantly different between 2009, 2014, and 2019 in Minnesota’s metropolitan counties. Since the p-value of Organics is significant, the null hypothesis is rejected. There is sufficient evidence that the amount of collected organics is significantly different between 2009, 2014, and 2019 in Minnesota’s metropolitan counties.

Table 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **One-Way Repeated Measures ANOVA Results** | | | | |
|  | Recyclables (tons) | | Organics (tons) | |
|  | F-value | p-value | F-value | p-value |
| Year | 4.198398 | **0.0415** | 12.23528 | **0.0013** |

Since both ANOVA results are significant, two Tukey tests are performed (one for Recyclables and one for Organics) to determine which specific year ranges are significant. The Tukey results are displayed in Table 2. For Recyclables, there is a significant difference between 2009-2019. For Organics, there are significant differences between 2014-2019 and between

2009-2019.

Table 2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tukey Results** | | | | | |
|  |  | Recyclables (tons) | | Organics (tons) | |
|  |  | z-value | p-value | z-value | p-value |
| Years | 2009-2014 | 1.271 | 0.4115 | 2.028 | 0.1055 |
| 2014-2019 | 1.619 | 0.2374 | 2.894 | **0.0107** |
| 2009-2019 | 2.891 | **0.0109** | 4.921 | **< 0.001** |

To determine if the significant Tukey results are significant increases and not decreases, bar plots of mean recycle tons and mean organic tons are created (see Figures 1 and 2) and analyzed. Since both bar plots strictly increased from 2009 to 2014 to 2019, the amount of collected recyclables significantly increased between 2009-2019 in Minnesota’s metropolitan counties, and the amount of collected organics significantly increased between 2014-2019 and between

2009-2019 in Minnesota’s metropolitan counties.

Figure 1

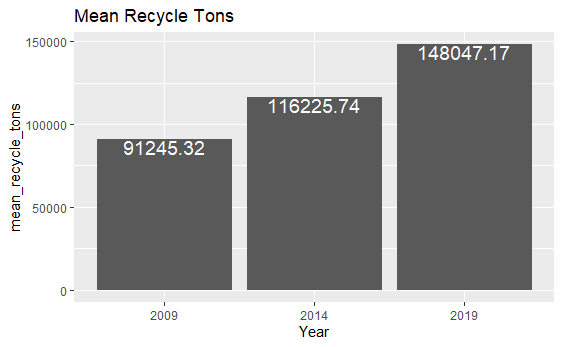
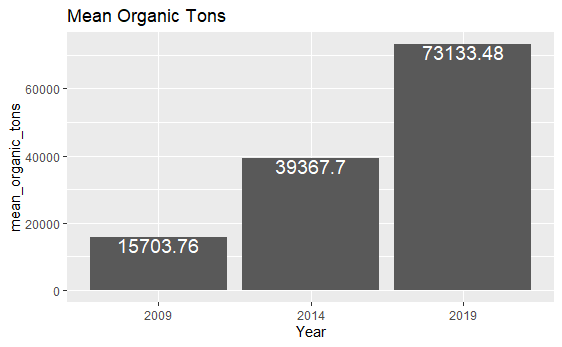


Figure 2



An interesting finding was that mean recycle tons experienced linear growth from 2009 to 2014 to 2019 but mean organic tons experienced exponential growth from 2009 to 2014 to 2019. A likely explanation for the exponential growth of collected organics is that Minnesota’s metropolitan counties have greatly increased their investment in infrastructure that collects and processes organics over the years. A potential limitation of the conducted analysis is that the amount of collected recyclables and organics by the metropolitan counties in Minnesota were not calculated on a per capita basis, which could change the ANOVA and Tukey results and lead to different conclusions. Future studies could be done to consider the per capita values of collected recyclables and organics and to investigate the growth rate discrepancy between recyclables and organics.

References

MPCA Data Services. (2024, February 22). *SCORE Overview*. Tableau Public. https://public.tableau.com/app/profile/mpca.data.services/viz/SCOREOverview/

SCOREOverview?publish=yes

Allen’s GitHub repository: https://github.com/Allen-Majd/ANOVA-Project-Paper

Gustavo’s GitHub repository: https://github.com/G-373/STAT-301\_Project