Abdellah Afellah and Ann Letsinger
STAT 301 Project
Monday, April 24th, 2023

What's in our trash?
An analysis on waste in Minnesota

**Introduction**

In 2017, the Environmental Protection Agency reported 268.7 million tons of waste was generated in the United States, the highest of any country. Knowing what type of waste we produce is one of the first steps in creating better waste management and reduction techniques to improve the health of our planet. The Waste Management Hierarchy, a list of types of waste from most to least preferred, states that source reduction/reuse and recycling/composting, are the most preferred while energy recovery and treatment/disposal are least preferred. Using waste data in Minnesota, we looked at residential recycling waste as well as amounts of waste in certain categories for the entire state.

Our first question was to analyze the main effect of the county on residential recycling waste, the main effect of the category on residential recycling waste, and the interaction between the two. Is there a significant mean residential recycling waste difference between the counties (Hennepin, Ramsey, Anoka)? Is there a significant mean residential recycling waste difference between categories (paper, metal, organic)? Our null hypothesis states that the average tons of recycling waste is equal for all counties ($HO\colon \mu(Hennepin) = \mu(Ramsey) = \mu(Anoka)$ ) and the alternative hypothesis states that at least two counties' averages are different. Concerning the main effect of waste categories, our null hypothesis states that the mean amount of residential recycling waste is the same for all recycled waste categories ($HO\colon \mu(Paper) = \mu(Metal) = \mu(Organic)$ ) while the alternative hypothesis states that the mean amount of recycling waste in at least two categories is different.

We chose our nominal categorical independent variables, Hennepin, Ramsey, and Anoka counties, due to their close proximity to our university. From the other nominal categorical independent variable, category, we selected three types of waste, paper, metal, and organic. The reason for this is paper, metal, and organic waste all belong to the more preferred section of the Waste Management Hierarchy, recycling/composting. Paper and Metal also had a good amount of data reported among the three counties with paper totaling 196 rows and metal totaling 111 rows. While organic waste only had 51 rows, we were interested in including it in our data analysis because some cities located within the counties we chose are implementing organic curbside pick up services.

For our second question, we looked at the effect of the type of waste (landfilled and waste to energy) on timing (1991 and 2017). Is there a significant main effect of timing, significant main effect on type of waste, and is there a significant interaction between timing and type of waste? Our null hypotheses state that the average tons of waste are equal between 1991 and 2017 ($HO{:}\ \mu(1991Waste) = \mu(2017Waste)$ ) and equal between waste to energy and landfilled waste ($HO{:}\ \mu(WTE) = \mu(Landfilled)$ ). The alternative hypotheses state the average tons of waste are not equal between timing

($HA{:}\ \mu(1991Waste) \neq \mu(2017Waste)$ ) and types of waste ($HO{:}\ \mu(WTE) \neq \mu(Landfilled)$ ).

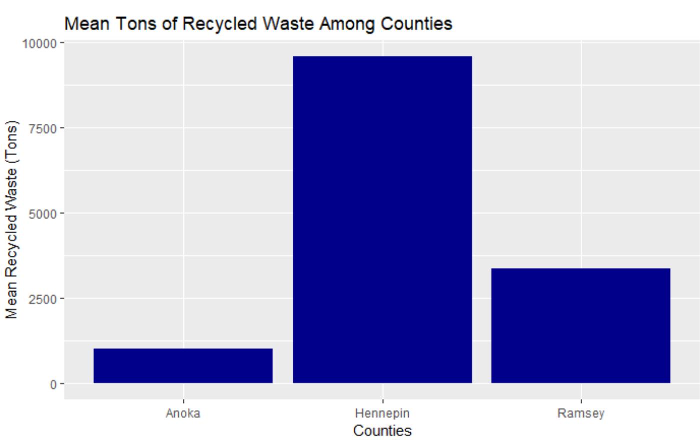
For the independent variables in our second question, we selected landfilled and waste to energy (WTE) types of waste because we were curious to know their relationship since they have different environmental and economic impacts. Landfills can have negative effects on the environment, such as greenhouse gas emissions and groundwater contamination, while WTE can reduce the amount of waste sent to landfills and generate renewable energy.

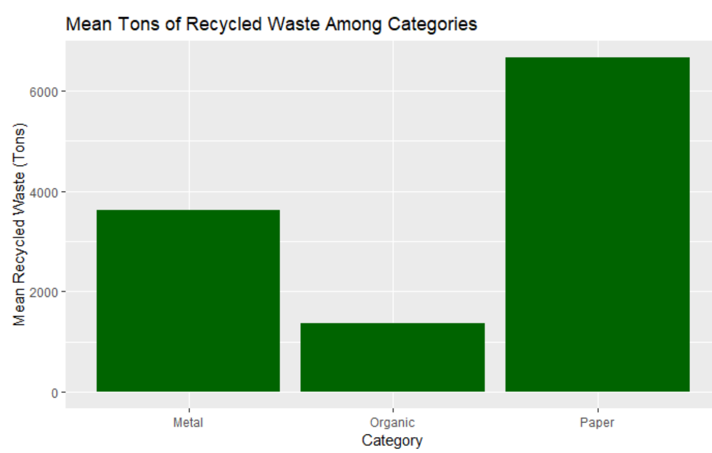
By studying the relationship between these two variables, we can gain insights into the overall effectiveness and sustainability of different waste management strategies. Additionally, understanding the relationship between these variables can help inform policy decisions around waste management, such as incentivizing the development of more WTE facilities or implementing stricter regulations on landfills.

## Methods

For both questions we used a Two-way ANOVA, however for the first question, we used the unbalanced design since the counts for categories were not the same among the counties. The Two-Way ANOVA was the most appropriate for our questions because the independent variables had two or more levels and were not repeated. This assumption can be checked by visual inspection that our dependent variable (residential recycling waste) should be normally distributed within each group defined by the independent variables and its variances are equal across by the independent variables (County and waste Category) .Also our observations within these variables are independent of each other.

## Conclusion




Testing the main effect of the county on residential waste, we rejected the null hypothesis. There is evidence that the mean tons of residential recycling waste are significantly different among the three counties ($p = 0.0000$; $F = 10.66$). Hennepin county had the highest average tons of recycled waste with a total of 9,602.283 tons.

Testing the main effect of the category on residential recycling waste, we rejected the null hypothesis. There is evidence that the mean tons of residential waste are significantly different among the three categories ($p = 0.0402$; $F = 3.304$). Paper had the highest average tons with a total of 6,657.37 tons.

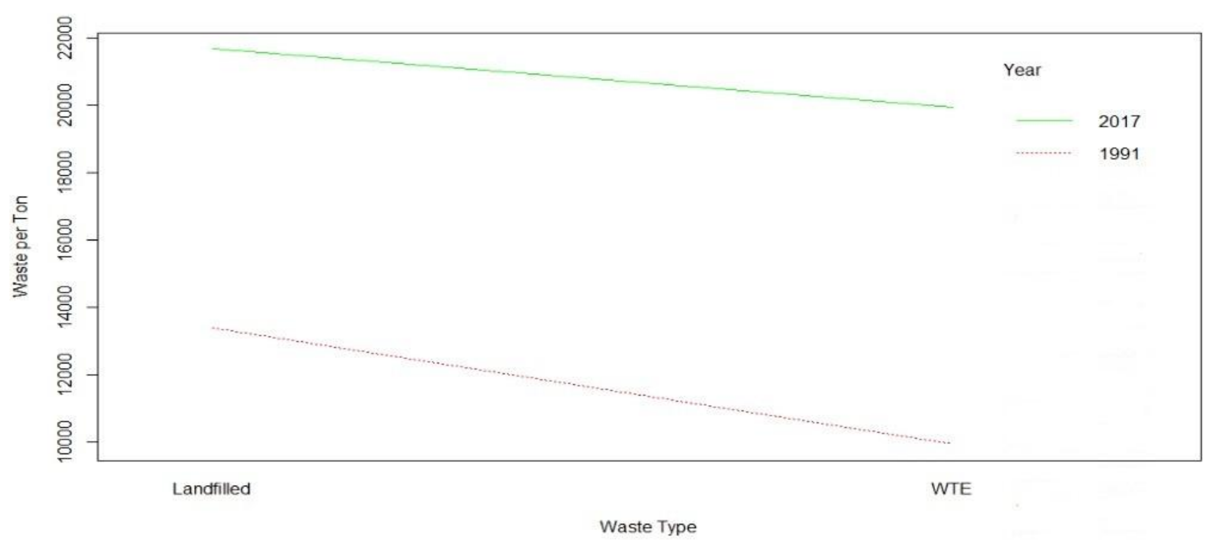
The interaction between county and category also has a significant effect on the amount of waste recycled ($p = 0.00665$, $F = 3.74$). To further understand the differences in means, we filtered the data by each category, and conducted an ANOVA with counties. The results show that a significant difference lies between all the counties and paper. Meanwhile, for the other two categories, there is no significant interaction between all counties and organic and metal categories.

| **County** | **Category** | **P-value** |
|---|---|---|
| Anoka | Paper | ($p = 0.0000$) |
| Hennepin | | |
| Ramsey | | |
| Anoka | Organic | ($p = 0.419$) |
| Hennepin | | |
| Ramsey | | |
| Anoka | Metal | ($p = 0.493$) |
| Hennepin | | |
| Ramsey | | |

Overall, the results show that the most populated county, Hennepin, has the highest amount of recycled waste, which is to be expected. The more people in a county, the more waste produced. Paper was significantly higher than the other two categories tested. One possibility for this is paper is a widely used resource, therefore, more of it exists to be recycled. One thing worth noting is the low amount of recycled organic waste. In 2012, the Natural Resources Defense Council reported that Americans wasted what equates to approximately

400 pounds of food per person per year. With the increased access to organic recycling facilities and methods, maybe this will lead to an increase in organic recycling in the coming years. To improve waste management and reduction techniques, local authorities and waste management companies should consider implementing strategies that target metal and organic waste for recycling and counties (Anoka and Ramsy) that recycle the least waste. Additionally, it would be beneficial to investigate the reasons behind the differences in waste recycled between the counties and categories and develop appropriate interventions to address them.

For our second question, we failed to reject the null hypothesis for timing effect ($p = 0.222$, $F = 1.50$). There is no significance between timing and waste per ton for 1991 and 2017. We also failed to reject the null hypothesis for type of waste effect ($p = 0.729$, $F = 0.12$ ). There is no significance between type of waste and waste per ton for waste to energy and landfilled waste. Testing for the interaction between timing and type waste, there was no significant interaction between the two variables ($p = 0.907$, $F = 0.01$ ).




Based on the results of the Two-way ANOVA, we failed to reject the null hypothesis for both main effects and the interaction for our second question. This means that there is no significant difference in waste per ton between 1991 and 2017, and no significant difference in waste per ton between waste to energy and landfilled waste. Additionally, there is no significant interaction between timing and type of waste. Therefore, we cannot conclude that there is a significant relationship between the type of waste and the timing in terms of waste per ton.

The most interesting finding is that there is no significant difference in waste per ton between 1991 and 2017, despite the increased awareness and efforts towards waste reduction and recycling. This suggests that more work needs to be done to reduce the amount of waste generated in the United States.

In terms of limitations, our analysis only focused on three counties in Minnesota, which may not be representative of the entire state or country. Additionally, our analysis only looked at specific types of waste for each question, which may not capture the full picture of waste generation and management. Finally, there may be other factors beyond the variables we analyzed that affect waste generation and management, such as population growth and economic conditions.