Yaroslav Vechirko

Introduction

My first question is if the average metal recycling in the top 5 counties was the same each year. I selected the top 5 counties based on the total metal recycling between 1991 and 2017. The null hypothesis is 𝐻0: 𝜇(Hennepin)=𝜇(Ramsey)=𝜇(Stearns)=𝜇(Anoka)=𝜇(Benton) and the alternative hypothesis is Ha: at least two means are different. My second question is if the average recycling for each type is the same over time for the top 5 counties. I selected the top 5 counties based on the total amount of recycling minus organic recycling. Organic recycling was not measured each year, so I did not select it. I grouped the years into 3 categories: 1991 to 1999, 2000 to 2008, and 2009 to 2017. The null and alternative hypothesis for time periods is 𝐻0: 𝜇(Glass)=𝜇(Hazardous)=𝜇(Metal)=𝜇(Other)=𝜇(Paper)=𝜇(Plastic) and Ha: at least two means are different. The null and alternative hypothesis for categories is 𝐻0: 𝜇(Period 1)=𝜇(Period 2)=𝜇(Period 3) and Ha: at least two means are different. The null and alternative hypothesis for interaction between period and category is 𝐻0: there is no significant interaction between period and category and Ha: there is a significant interaction between period and category.

Methods

For question 1 I chose one-way ANOVA. I was measuring the average metal recycling for each county. There is only one independent variable, and it has more than two groups. The assumptions are met. The amounts each year are independent, there are equal variances between each year, and error terms are normally and independently distributed. I also performed the TukeyHSD comparison since there was a significant difference among the counties. For question 2 I chose a two-way fixed repeated measure. I chose this because the difference between categories was measured over different time periods. There were two independent variables, and one of them was different time periods. I also performed glht() function since there was difference between the categories. The assumptions for this are that the values follow a normal distribution.

Conclusion

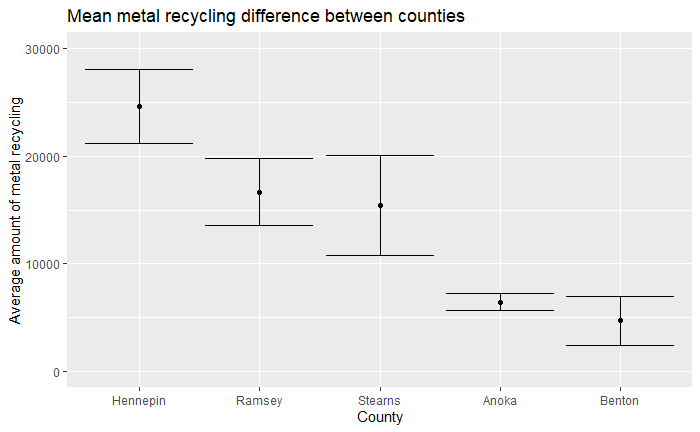
For question 1, the test statistic is 28.94 and the p-value is 0. I can reject 𝐻0, there is evidence that the mean metal recycling each year is different for at least two counties. The most interesting find for this is that some counties can be very similar to each other and very different from others. Anoka and Benton were similar to each other, and Stearns and Ramsey were similar to each other. All the other locations were not similar. Some limitations with this are that not all counties could have accepted metal recycling, some population centers in those counties could have used recycling locations in different counties, or record keeping was not accurate for all locations.

Figure , One-way ANOVA results

Chart

Description automatically generated with medium confidence

Figure



For question 2, the test statistics are 19.45 for category, 0.43 for period, and 0.2 for the interaction. The p-values are 0 for category, 0.65 for period, and 0.99 for the interaction. For category, I reject 𝐻0, there is evidence that the average recycling amount is different for each type of material. For period I fail to reject 𝐻0, there is evidence that the periods are similar. For the interaction, I fail to reject 𝐻0. There is no significant interaction between category and period. The most interesting find is that the average between each period was very similar for every type of material, and that each category was similar to each other except for paper. Paper was significantly different from every other type of recycling except for metal. I expected the amount of recycling to increase over time. The limitations for this question are similar to the limitations from question 1.

Figure , Two-way fixed repeated measure ANOVA results

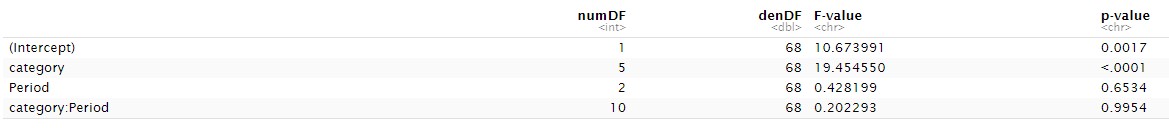


Figure , glht() results

