Combined Group E and Group F

Team Members: Evan Thomas, Sage Markwardt, Vlad O, Hao Fan

1. Do different countries of residence exhibit different trends with carbon emissions, ecological footprint, and ecological footprint composition? For the same question, do different genders exhibit any changes in behavior?

* **Country of Residence (Evan):**   
  When analyzing the average GHA values between survey respondents currently residing in the U.S. or UAE, the overall Ecological Footprint was higher in the UAE by approximately 2.29 GHA.   
    
  While the categorical composition for the U.S. was lower in every individual column, the difference between the two countries varied depending on the specific category.   
  Both Food and Shelter were on the lower end, only having a 0.31 GHA and 0.21 GHA difference between the two countries respectively.  
  On the other hand, the difference between the countries was larger when comparing Mobility, Goods, and Service, coming in at a difference of 0.63, 0.49, and 0.60 GHA respectively.  
    
  Switching over to the CO2 Emissions, the U.S. averaged 14.89 tonnes compared to the UAE’s 18.39 tonnes. For the median emissions, the U.S. had 13.6 tonnes while the UAE had 18.1. Looking at the data a bit closer, while the median and mean emissions are relatively similar in terms of the difference between them, the variance in the data tells a slightly different story.   
    
  When removing two outlier data points from the UAE set, both the U.S. and UAE have a range of data that lies between a maximum of approximately 30 tonnes and a minimum of around 7 tonnes. Even though the range of the set is similar, the data concentrated between the lower and upper quartiles varies more heavily between the countries.  
  For the U.S., the range lies between around 11 tonnes to 19 tonnes. Comparing this to the UAE with a range of around 15 tonnes to 21 tonnes, while the UAE generally trends higher in terms of CO2 emissions, they also have a less “diverse” range (smaller difference between the lower and upper quartile).

1. Is there any connection between different categories in terms of one category directly correlating to another (E.G. if someone has a high Mobility score, does that mean they also have a high Goods or Food score?)

**The table I've provided is a correlation matrix, which shows the correlation coefficients between different categories. The correlation coefficient ranges from -1 to 1, where:**

· 1 indicates a perfect positive correlation.

· -1 indicates a perfect negative correlation.

· 0 indicates no correlation.

**Most Positive Correlation:**- The highest positive correlation is between "Ecological Footprint" and "CO2 Emissions - tonnes" with a coefficient of 0.908672.

**Most Negative Correlation:**- The most negative correlation is between "Age" and "Service - gha" with a coefficient of -0.317333.

**There are some correlations between the categories:**

1. Age and Other Categories:

- Age has negative correlations with most other categories. This suggests that as age increases, there tends to be a decrease in the values of ecological footprint, CO2 emissions, food consumption, shelter, mobility, goods, and services.

2. Number of Earths and Ecological Footprint: - There is a strong positive correlation between the number of Earths and ecological footprint, suggesting that as the ecological footprint increases, the demand on Earth's resources also increases.

3. Ecological Footprint, CO2 Emissions, and Other Categories: - Ecological footprint and CO2 emissions are positively correlated, which is expected since carbon emissions are a significant part of the ecological footprint. There are also positive correlations between ecological footprint and categories such as shelter, mobility, goods, and services.

4. Goods and Other Categories: - Goods consumption is positively correlated with mobility and services, indicating that people who consume more goods also tend to have higher mobility and service-related impacts.

5. Service and Other Categories: - Service consumption is positively correlated with shelter, mobility, and goods, suggesting that people who consume more services also tend to have higher impacts in these other categories.

**It's important to note that correlation does not imply causation, and these relationships are based on statistical associations. Other factors may influence these relationships, and further analysis or experimentation would be needed to establish causal connections. Additionally, the strength of correlations varies, and some correlations in the table are relatively weak.**

1. For the earth overshoot day, do any of the specific categories have more “influence” than others (Does a higher mobility score have more of an impact on the overall overshoot day when compared to something like Goods or Shelter?)

The average correlation for the categories in relation to the overshoot day is 0.19. There are two categories that are a decent amount higher than this average. The category of Goods has a correlation coefficient of 0.33, which means a person with an early overshoot date is reasonably likely to have a higher score in the Goods category as well. Mobility also has a higher than average correlation of 0.29, which will have a noticeable influence on the overshoot date. Both these categories have the most contribution to when a person will hit their overshoot date.

The food category and Services category are close to the average, sitting at .2 and .13 respectively. The lowest category is interesting because it has so little effect on the overshoot date. The correlation coefficient of the Shelter category is 0.011, which is barely above zero and means this category will not really change a person’s overshoot day at all.