STEM Fellowship Big Data Challenge 2024

Investigating the Correlation between Human Activity and Animal Biodiversity in Canada

Favour Atewologun[1] and Oluchukwu Favour Iheanyichukcu[1]

[1]Webber Academy

January 14, 2024

Abstract

Biodiversity remains essential for the stability and security of our ecosystems. However, with the growth of human activity in our current day and age, we must step back and analyze our impact on the biodiversity in our nation Canada. The goal of our project was to determine the human impact on animal biodiversity in Canada. Determinants of human activity in this study were annual human population estimates in Canada and greenhouse gas emissions while the number of animal species was used to measure animal biodiversity. The programming language Python was used as a tool to not only find notable statistical correlations but to also make visual representations of data. The results of this study showed weak and moderate correlation between our explanatory variables, human population and greenhouse gas emissions, and our response variables, the total number of animal species in Canada, and the number of species of birds, fish, and mammals. Overall, this paper draws attention to the overall biodiversity of animals in Canada, and focuses on three specific categories - birds, fish, and mammals - to understand the trends in biodiversity since 1990, suggesting ways to limit negative human impacts and preserve the diversity of animals in Canada.

Keywords: Biodiversity, Human impact, number of species, fish species, mammal species, bird species

1 Introduction

Since the Industrial Revolution, human beings have hurt biodiversity. Where human beings migrate, so do their activities, which include fishing, hunting, the destruction of land through agriculture, and the release of pollutants through agricultural processes[1]. Canada's population currently sits at 40 million, and to be able to meet the demands of the economy, this number is projected to only increase over the next few years[2]. Although several studies have stated the negative impacts of humans on biodiversity, others show that certain animals have been able to adapt to the human stressors around them through processes such as domestication and natural selection.[3] The question remains if our demand for human development will negatively impact the overall animal biodiversity around us, or if they will be able to adapt to the new human-induced challenges around them. Human population and greenhouse gas emissions help us to quantitize aspects of human activity while as we compare it to the number of species in Canada over the last few years. In this research paper, we hope to be able to identify the relationship between human activity and animal biodiversity to be able to improve the state of our natural ecosystems in Canada.

2 Hypothesis

Using our knowledge of the relationship between human activity and nature, we hypothesized that human population and greenhouse gas emissions would have a strong negative correlation with the number of species in Canada.

3 Materials & Methods

3.1 Getting Data

First, we collected data from three public databases: Greenhouse gas emissions from The Government of Canada, Canadian species index from The Government of Canada, and

Population Estimates, quarterly from Statistics Canada. Github, an open source version control software, was used to store the datasets after we downloaded them from their various sources as csv files. We used Google Colab to write code in Python, a language that is famously used to perfom data analysis. More specifically, we used the Pandas module to import our csv files from Github before cleaning, sorting, and visualizing the data. Additionally, Matplotlib was used for extra assistance in data visualization and Scipy was used to analyze other statistical values.

3.2 Sorting the Data

The Canadian species index dataset had information on the annual number of species in Canada, the database related to the human population in Canada showed quarterly estimates of Canada's population, and Greenhouse gas emissions from The Government of Canada had data collected on the yearly emission estimates in Canada. Using the data on human population, we created a dataset that averaged the quarterly population estimates to make annual population estimates. For all three data sets, we specified our year range to 1990 to 2016 and then used the year column to combine them into one.

3.3 Data Analysis

In our study, we compared human population and greenhouse gas emmisions to overall number of species, before investigating impacts on specific species groups such as birds, mammals, and fish using scatter plots. On all our graphs, we used linear regression to estimate the linear relationship between our different variables. Our analysis consisted of several stats values including the P Value to test statistical significance, the slope of our linear regression line, R Value to measure correlation, and standard deviation error to measure confidence.

4 Results

4.1 Human Population

After getting our data, we wanted to compare the population of Canada to the number of species in Canada to understand how the human population impacts animal population. We first compared Canada's population to the total number of animal species in Canada, then chose to study 3 specific categories - birds, fish, and mammals.

In Figure 1, we plotted Canada's population against the total number of animal species in

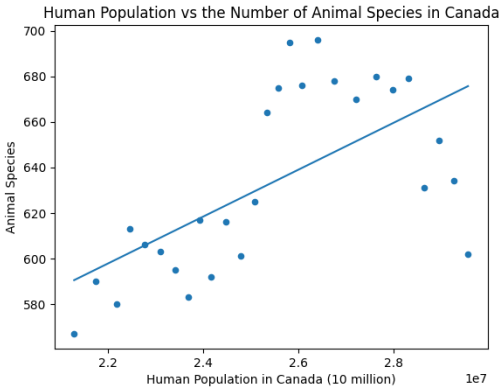

Figure 1: A scatter plot and regression line that compares the population of Canada (x-axis) with the total number of animal species in Canada (y-axis), with data points taken from 1990 - 2016.

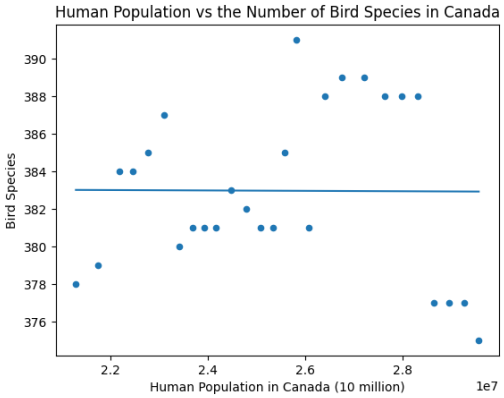

Figure 2: This graph compares the population of Canada (x-axis) with the number of bird species in Canada, taking data points from 1990 - 2016.

Canada. We found that the slope of the line of best fit was 1.028e-05, indicating that an increase of 1 million people in the Canadian population is correlated with an increase of 10 animal species in Canada. The R value of 0.63 indicates a moderate positive correlation, supported by the low p value of 0.0004. The standard error of 2.531e-06 is moderately small.

In Figure 2, we compared Canada's human population with the number of bird species in Canada. The slope of the graph was -1.05e-08, predicting a slight decrease of approximately 0.1 species of birds for every increase of 1 million in the Canadian population. The R value of -0.006 indicates almost no correlation between the two variables. This is supported by the high p value of 0.977, suggesting insignificant results. The standard error is 3.614e-07.

Figure 3 plots Canada's population against the number of fish species in Canada, with the slope of the graph being 1.374e-05. This regres-

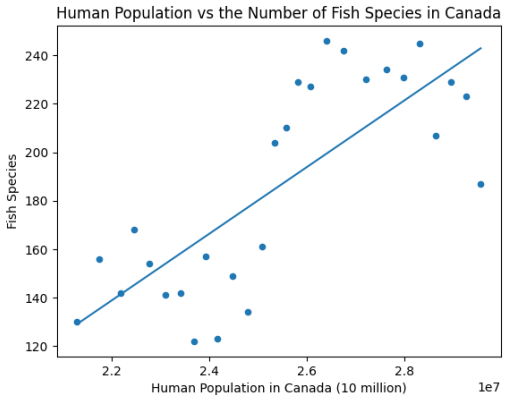

Figure 3: In this graph, there is a regression line that compares the population of Canada (x-axis) with the number of fish species in Canada (y-axis), with data points taken from 1990 - 2016.

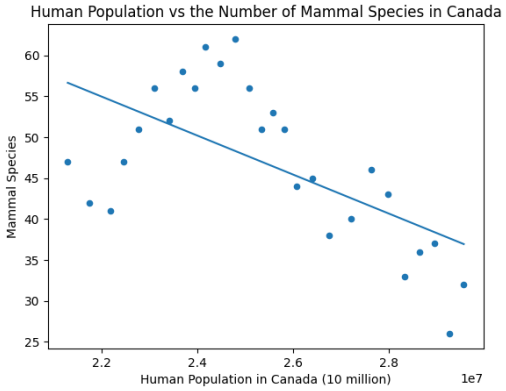

Figure 4: This graph relates the population in Canada (x-axis) to the number of mammal species in Canada (y-axis) using data from 1990 - 2016.

sion line predicts an increase of 13 bird species in Canada for every human population increase of 1 million. A moderate positive correlation is suggested by the R value of 0.77. The p value of 2.505e-06 suggests that the results are significant. The standard error of 2.269e-06 is moderately low.

We compared Canada's population to the number of animal species in Canada in Figure 4. The slope of the graph was -2.379e-06, indicating that a population increase in Canada of 1 million is correlated with a decrease of over 2 mammal species. The R value of -0.614 indicates a moderate negative correlation between the two variables, which is also supported by the p value of 0.0007, suggesting significant results. The standard error of this graph is 6.119e-07.

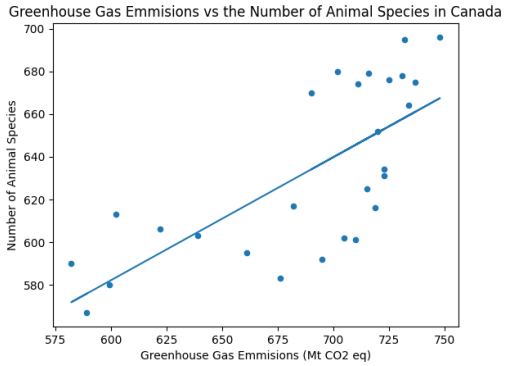

Figure 5: A scatter plot and regression line that compares greenhouse gas emissions in Canada (x-axis) with the total number of animal species in Canada taking data from from 1990 - 2016.

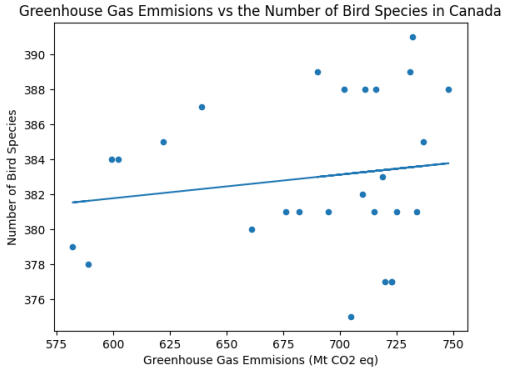

Figure 6: This graph compares greenhouse gas emissions in Canada (x-axis) with the total number of bird species in Canada (y-axis), using data from 1990 - 2016.

4.2 Greenhouse Gas Emissions

We also wanted to compare the impact of the production of greenhouse gases on the number of animal species in Canada, both overall and in the three categories we selected.

In Figure 5, we plotted greenhouse gas emissions against the total number of animal species in Canada. The slope of the line of best fit is 0.576, indicating that for every additional tonne of greenhouse gas emitted, the number of species in Canada increased by that much. The R value of 0.720 indicates a strong positive correlation between greenhouse gas emissions and the number of animal species in Canada. The p-value of 2.292e-05 suggests that the results are significant. The standard error was 0.111.

Figure 6 compares greenhouse gas emissions in Canada with the number of bird species in Canada. The line of best fit has a slope of 0.013, indicating little change in the number of bird species in Canada for every additional tonne of greenhouse gases emitted. The R value of 0.152

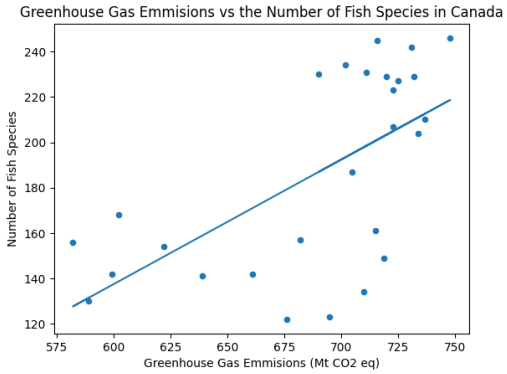

Figure 7: This graph shows the relationship between greenhouse gas emissions in Canada (x-axis) and the number of fish species in Canada (y-axis), using data from 1990 - 2016.

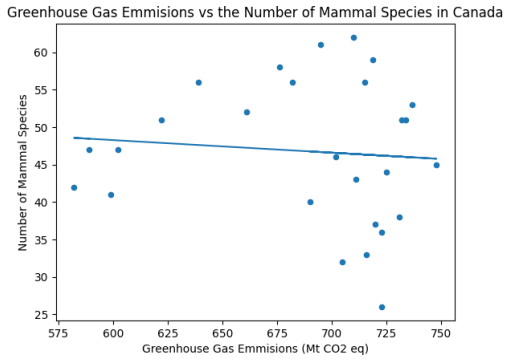

Figure 8: This scatter plot and regression line plots greenhouse gas emissions in Canada (x-axis) against the number of mammal species in Canada (y-axis), taking data points from 1990 - 2016.

suggests a weak correlation between greenhouse gas emissions and the number of bird species in Canada. The p-value of 0.448 indicates the results may not be significant. Considering the small slope of the graph, the standard error of 0.018 indicates that the line of best fit is not a very good predictor of the relationship between the two variables.

Figure 7 plots greenhouse gas emissions against the number of fish species in Canada. The graph had a slope of 0.548, indicating that on average, each additional tonne of greenhouse gases emitted by Canada related to an increase in bird species of just over 0.5. The R value of 0.627 indicates a moderate positive correlation between the two variables, and the p value of 0.0005 suggests the results are significant. The standard error of 0.136 is relatively high.

We compared greenhouse gas emissions with the number of mammal species in Canada in Figure 8. The graph's slope of -0.017 indicates a slight decrease in the number of mammal species in Canada for every tonne of greenhouse gases emitted. The R value of -0.088 indicates a very weak correlation between the two variables, which is confirmed by the high p value of 0.661, suggesting insignificant results. The standard error 0.038 is relatively large, especially when considering the graph's small slope.

5 Discussion

5.1 Analysis of Results

The overall biodiversity in Canada, measured by the number of animal species in Canada, is composed of a number of smaller groups of species, including birds, fish, and mammals. While some populations have seen decreases in number of species due to human activity, others have seen increases. Looking at Figures 1 and 5, the overall correlation between biodiversity and human activity seems to be positive, but the results of the individual categories yields a deeper understanding.

Of the graphs using human population as the explanatory variable, Figures 1 and 3 showed that an increase in population was correlated with an increase in the number of species, suggesting that a higher human population is associated with greater biodiversity in Canada, for animals overall and for fish. This seems counter-intuitive, as there has been an increasing focus on how rapid growth of cities and their populations lead to habitat loss and unstable conditions for animals, endangering animals and even causing some species to go extinct. The increase in species could be due to speciation, where the species in conditions very different from their usual niches are forced to adapt to the rapidly changing environment, forming new species as a result. From these results, we concluded that overall biodiversity and biodiversity of fish are positively correlated with human population in Canada.

As seen in Figure 2, the human population in Canada has very little correlation with the number of species of birds in Canada, indicating that changing human populations have very little to do with bird biodiversity.

Figure 4 has the most intuitive results, where an increase in human population was moderately correlated with a decrease in the number of species of mammals in Canada. This may be due to the fact that mammals are most affected by habitat loss due to humans - because humans and mammals both share land, they are likely to be the most affected by human activities like deforestation and urbanization that increase with

an increase human population. The biodiversity of mammals is most negatively impacted by human populations.

Figures 5 and 7 mimic the results of Figures 1 and 3, this time using greenhouse gas emissions as the explanatory variables. In both figures, an increase in human population is moderately correlated with an increase in the total number of species and number of species of fish in Canada. Higher emissions of greenhouse gases is associated with more biodiversity overall and in fish populations, which, again, seems counter-intuitive. Once again, this may be due to speciation and forced adaptations as a result of more harmful emissions. In addition to this, greenhouse gases would likely have the least impact on fish, as they live in water and the gases tend to affect the ozone and air quality. The result in Figure 7 may be due to the natural rate of speciation in Canada, possibly indicating that the biodiversity of fish in Canada is not very affected by greenhouse gas emissions.

Figure 6, which compares greenhouse gas emissions with the number of species of birds in Canada, has a slope of nearly 0, an R value of 0.152 that indicates a very weak correlation, and a high p value, suggesting that the results may not be significant. From this data, we concluded that the relationship between greenhouse gas emissions and the biodiversity of birds in Canada is not very significant, if present at all.

In Figure 8, we compared greenhouse gas emissions with the number of species of mammals in Canada. Similar to the result in Figure 4, we noticed a negative slope for the line of best fit, but a very weak correlation between the two variables, indicated by the R value of -0.08.

Overall, mammals were the most negatively impacted by human population and activity, likely because their niche most directly overlaps with that of humans. Because resources like shelter are shared with humans, mammals would be most harmed by human impacts like rapidly expanding cities as a result of increasing human populations.

5.2 Application

Governments at the local, provincial, and national level should work together to ensure that our population growth is not destroying the environment and the biodiversity of Canada. By instituting laws that punish excessive emissions from corporations, companies will be discouraged from needlessly expelling more greenhouse gases into the air, contributing to air pollution and ruining the ozone layer[4]. Additionally, conscientious and deliberate considerations should be made when planning to expand cities, towns, or do any major construction, taking note of the animals and habitats that may be affected as a result of these expansions. [5]By remaining dedicated to the goal of protecting our environment, Canada will preserve, and possibly even enhance, its biodiversity. While remaining dedicated to preserving biodiversity is important, protecting sensitive information is crucial as well. While exploring datasets to use for this project, some websites and data were protected by passwords and login keys, barring us from having access to the data. While open access to data is beneficial because it allows for transparency and the gives the public the ability to look for information they want to and complete their own data analysis, careful consideration about how much data is needed for the project is important. As society becomes more reliant on technology, it is important to ensure that the transition does not make valuable and private data vulnerable to attack through different authentication methods. [6]

5.3 Sources of Error

Looking at the shape of all the scatterplots, it seems that many of the graphs did not have a linear shape, preventing us from getting graphs with very strong correlations. If this study were to be repeated, an attempt to linearize the data should have been attempted first. An analysis could have been done on a power model of the data, or a different model like an exponential or polynomial model could have been used. While this could detract from the ease of discovering the general relationship between the explanatory and response variables, using this model could help develop an understanding of the relationship for specific ranges of values of the explanatory variables.

The data used came from the years 1990 - 2016. When looking at the data by year, while the Canadian human population almost consistently increased for the entire period, the number of animal species (both the total number of species and the three selected categories) seemed to peak around the early 2000s before declining until the most recent data point, 2016. This likely affected the shape of the data and its ability to be linearized, as discussed in the paragraph above. Perhaps doing the analysis from 1990 to 2000 and 2000 to 2016 separately may have yielded results with stronger correlations, also giving the opportunity to compare how the biodiversity of animals in Canada before and after the turn of the millennium.

5.4 Future Studies

Future studies could look specifically at populations of endangered animals to see the impact of human activity on species that already at risk of extinction. Additionally, conducting a similar study on the number of species of reptiles, amphibians, and insects would be beneficial, as these animals tend to be much more sensitive to changes in the environment than our selected categories. While our project focused on large scale major effects in the past few decades, this suggested study would likely reflect the small but gradually increasing changes to the environment and its impact on biodiversity in Canada. In future projects related to biodiversity, artificial intelligence could be used to model populations and the number of species in both human-affected and natural environments of Canada, comparing the rate of change in biodiversity in response to differing levels of attempts to limit humans' negative impacts on the environment.

Conclusion

We found that, despite the correlation between overall biodiversity in Canada and human activity possibly being misleading, comparing the number of species in broad categories (birds, fish, and mammals) to human population and greenhouse gas emissions showed a more nuanced result of the humans' impact of biodiversity in Canada. Based on the results, mammals are most intensely negatively harmed by high human populations and large amounts of greenhouse gas emissions. To help protect the biodiversity of mammals and all animals, laws with more strict consequences for companies regarding emissions should be created. By deliberate efforts to protect the environment in Canada, the biodiversity of animals can be protected, slowing the speed of the endangerment and extinction of species. Future studies can further enhance Canada's understanding of biodiversity and emphasize the importance of limiting activities that negatively impact the environment. Cybersecurity will be increasingly more important in protecting the results of these studies from vulnerabilities. Finally, AI can be used to help make predictions and simulations of environments in various conditions, allowing us to understand the impacts of our actions without having to wait years to see long-term results. We hope that our and future studies contribute to protecting the precious biodiversity in Canada.

Acknowledgements

We would like to thank our mentor, Mr. Niven for his support and insightful comments along with Statistics Canada and the Government of Canada for their accessible and expansive datasets. Finally, we would like to appreciate everybody at STEM Fellowship who helped develop a fun and academically stimulating challenge.

References

[1] Philip Hunter. The human impact on biological diversity. how species adapt to urban challenges sheds light on evolution and provides clues about conservation. 2007.

[2] Darren Major. Canada's population surpasses 40 million. 2023.

[3] Understanding Evolution. Natural selection at work. 2020.

[4] Akiko Hirata Shinichiro Fujimori Kiyoshi Takahashi Ikutaro Tsuyama Katsuhiro Nakao Yuji Kominami Nobuyuki Tanaka Yasuaki Hijioka Tetsuya Matsui Haruka Ohashi, Tomoko Hasegawa. Biodiversity can benefit from climate stabilization despite adverse side effects of land-based mitigation. 2019.

[5] Yingjie Li Zhenbo Wang Siao Sun Sanwei He Wei Qi Chao Bao Haitao Ma Yupeng Fan Yuxue Feng Xiaoping Liu Guangdong Li, Chuanglin Fang. Global impacts of future urban expansion on terrestrial vertebrate diversity. 2022.

[6] The Computer Clinic Bicester. The importance of multi-factor authentication for cyber security. 2023.

[7] Deinet S De Palma A Carranza T Cooper N Loh J Baillie JE. Collen B, McRae L. Predicting how populations decline to extinction. 2011.

[8] Crowther T. W. Zohner C. M. Robmann N. M. Lauber T. van den Hoogen J. Kuebbing S. Liang J. de-Miguel S. Nabuurs G.-J. Reich P. B. Abegg M. Adou Yao Y. C. Alberti G. Almeyda Zambrano A. M. Alvarado B. V. Alvarez-Dávila E. Alvarez-Loayza P. Alves L. F. ... Maynard D. S. Delavaux, C. S. Native diversity buffers against severity of non-native tree invasions. August 23, 2023.

[9] Mandrak N. E. Jackson, D. A. Ecology evolution of freshwater fish. 2002.

[10] Sciulli P. W. Fooce C. D. Waite T. A. McKee, J. K. Forecasting global biodiversity threats associated with human population growth. pages 161 – 164, 2004.

[11] Michael L. Morrison John C. Newnam Andrew J. Campomizzi Shannon L. Farrell Heather A. Mathewson Robert T. Snelgrove Raymond J. Carroll-Robert N. Wilkins Bret A. Collier, Julie E. Groce. Predicting patch occupancy in fragmented landscapes at the rangewide scale for an endangered species: an example of an american warbler. August 25, 2011.

[12] Statistics Canada. Census releases.

[13] Government of Canada. Greenhouse gas emissions.

[14] Government of Canada. Canadian species index.