

|  |  |
| --- | --- |
|  | School of Applied Business **Master of Applied Business**  **Postgraduate Diploma in Applied Business** |

**BSNS8030 Data Anayltics**

**Assignment 3 Semester 1, 2024**

|  |  |  |
| --- | --- | --- |
| **Date Issued:** | 26 February 2024 | |
| **Due date and time:** | Research Project Report – Week 14, 12 June 2024 by 08.00 am | |
| **Delivery:** | Turnitin via Moodle | |
| **Total Marks:** | **100** | |
| **Weighting:** | **30%** | |
| **Expected number of hours to spend on this assessment:** | **None** | |
| **Word limit:** | e.g.Maximum 2,500 word limit | |
| **Instructions:** | Complete this **cover sheet** and attach itto your assignment.  Collusion, copying or plagiarism may result in disciplinary action. | |
| **Lecturer:** | **Arfian Zudana** | |
| **Individual declaration:** | We confirm that:   * Our submission is original and entirely our own work. * Where we have used ideas, tables, diagrams etc. of other writers, we have acknowledged the source in every case using APA7 style. * Our submission has not been submitted as assessed work for any other academic course. | |
| **Student name:** | **Student ID:** | **Student Signature** |
| **Rahul Lihare** | **1582566** | **RL** |
| **Jacob G** | **1571326** | **JG** |
| **Kunal Tatiya** | **1568437** | **KT** |
|  |  |  |

Table of Contents

[Introduction 3](#_Toc169571894)

[Executive summary 3](#_Toc169571895)

[Research Objectives 4](#_Toc169571896)

[Data Sources: 4](#_Toc169571897)

[Analysis 5](#_Toc169571898)

[Conclusion 18](#_Toc169571899)

## **Executive summary**

Climate change significantly affects global ecosystems and civilisations, causing extreme weather, rising sea levels and a dramatic decline in biodiversity (“Introduction to Climate Change,” n.d.). To reduce the adverse effects of climate change, practical climate actions like reducing greenhouse gas emissions, promoting renewable energy, and implementing sustainable practices across all sectors are required (Tvaronavičienė, 2021). The report researches and analyses critical factors, including trends in GHG emissions and tourism impact on CO2, and predicts future climate scenarios. This information will play a key role in helping MFAT design policies and strategies for promoting environmental sustainability that align with its strategic goal of finding global solutions to environmental challenges.

The first research objective attempts to analyse the GHG emissions from different sectors of New Zealand, revealing how agriculture and transportation largely influence and contribute to GHG emissions. The descriptive statistics, correlation and regression analysis found that CO2 emissions strongly correlate with greenhouse gas emissions, indicating significant sector-specific impacts.

The second research objective further explores the relationship between tourism and Annual CO2 emissions in New Zealand from 1990 to 2020. The statistical analysis, like hypothesis testing, correlation and regression analysis, suggested that increased tourist arrivals and extended stays significantly impact the annual CO2 emissions in New Zealand. These findings will help MFAT highlight the need for sustainable tourism strategies to reduce environmental impact.

Lastly, the third research objective employs predictive modelling using historical data on some of the factors affecting climate change in New Zealand. Annual surface temperature change, sea level rise, GHG emissions, and climate-related disasters forecast adverse effects on the climate for the next eight years. The findings highlight the pressing need for MFAT to implement sustainable strategies that can help reduce human-related environmental impacts.

This report aims to provide valuable insights and recommendations to MFAT for creating strong environmental policies ensuring a sustainable and resilient future.

## **Research Objectives**

The three research objectives mentioned below are related and framed to align with MFAT’s strategic goals.

1. **Analysing the overall trend of GHG (greenhouse gas) emissions across different sectors of New Zealand:** GHG emissions are the primary factors driving climate change (Ritchie et al., 2023). By analysing the trends in GHG across different sectors in New Zealand, we can find which sectors are highly contributing and how they have changed over time. This valuable insight will embrace MFAT to prioritise the interventions and develop targeted policies for each sector that will help mitigate the overall emissions.
2. **Is there a relationship between tourism and carbon emissions in New Zealand?**: Tourism significantly contributes to GHG emissions worldwide, especially Carbon emissions (Cevik, 2022). By examining the correlation between tourism arrivals and carbon emissions, we try to identify what environmental impact tourism has in New Zealand. If there is any crucial relation, this analysis will help MFAT develop policies to promote sustainable tourism practices to minimise carbon footprints. This further aligns to promote sustainable solutions to global environmental challenges.
3. **Considering the historical surface temperature, rise in sea level, GHG emissions, and other climate-related disasters, how does the prediction look for the next eight years, and how can the insights be used to prioritise environmental policies and resource allocations?** By looking at the historical data of a few factors affecting climate, predictions can be made to forecast what level of climate change will happen in the next eight years. This will help answer the question of how MFAT can align their environmental strategies in the right direction to mitigate future climate changes.

## **Data Sources:**

The datasets used in this report were acquired from various reliable sources mentioned below:

1. **Our world in data**
   1. *Global annual Greenhouse Gas emissions*

<https://ourworldindata.org/greenhouse-gas-emissions>

This dataset contains yearly greenhouse gases emitted by countries across the globe.

The data was filtered with a focus on New Zealand.

* 1. *Global annual CO2 emissions by Countries*

<https://ourworldindata.org/grapher/annual-co2-emissions-per-country?country=~NZL>   
This dataset shows how much carbon dioxide (CO2) each country produces over the years. This report focuses on the CO2 (in a million tonnes) produced mainly by New Zealand.

* 1. *Greenhouse gas emissions by sector in New Zealand*

<https://ourworldindata.org/grapher/ghg-emissions-by-sector?time=latest&country=~NZL>

This dataset shows how much GHG is emitted from various sectors in New Zealand (1990 – 2020).

1. **New Zealand Government Data Catalogue** 
   1. *New Zealand Tourism Forecasts*

<https://catalogue.data.govt.nz/dataset/new-zealand-tourism-forecasts/resource/a3568d41-9642-4f37-b5b0-350d3ae2ffa9>

This dataset from the Ministry of Business Innovation and Employment describes the total number of visitors in a year, the total days spent in New Zealand, and the money spent by them. This forecasted data will be used to explore the relationship between visitors and GHG emissions.

1. **IMF Climate Change Data and Stats NZ**
   1. *Annual surface temperature change*

<https://climatedata.imf.org/datasets/4063314923d74187be9596f10d034914>

This dataset provides yearly global surface temperature change over the years.

* 1. *Change in Mean Sea Levels*

<https://climatedata.imf.org/datasets/b84a7e25159b4c65ba62d3f82c605855>

This dataset provides the rise in mean sea levels around the globe.

* 1. *Climate-related disaster frequency*

<https://climatedata.imf.org/datasets/b13b69ee0dde43a99c811f592af4e821>

This dataset provides the trend in climate-related disasters globally over time.

* 1. *Gross GHG emissions*

<https://www.stats.govt.nz/indicators/new-zealands-greenhouse-gas-emissions/>

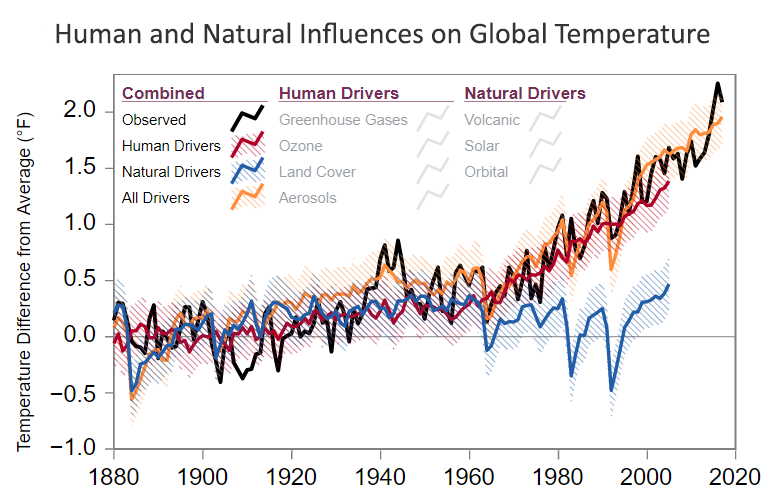
Total GHG emissions in New Zealand

*Note: All the modified datasets are attached to the report.*

## **Analysis**

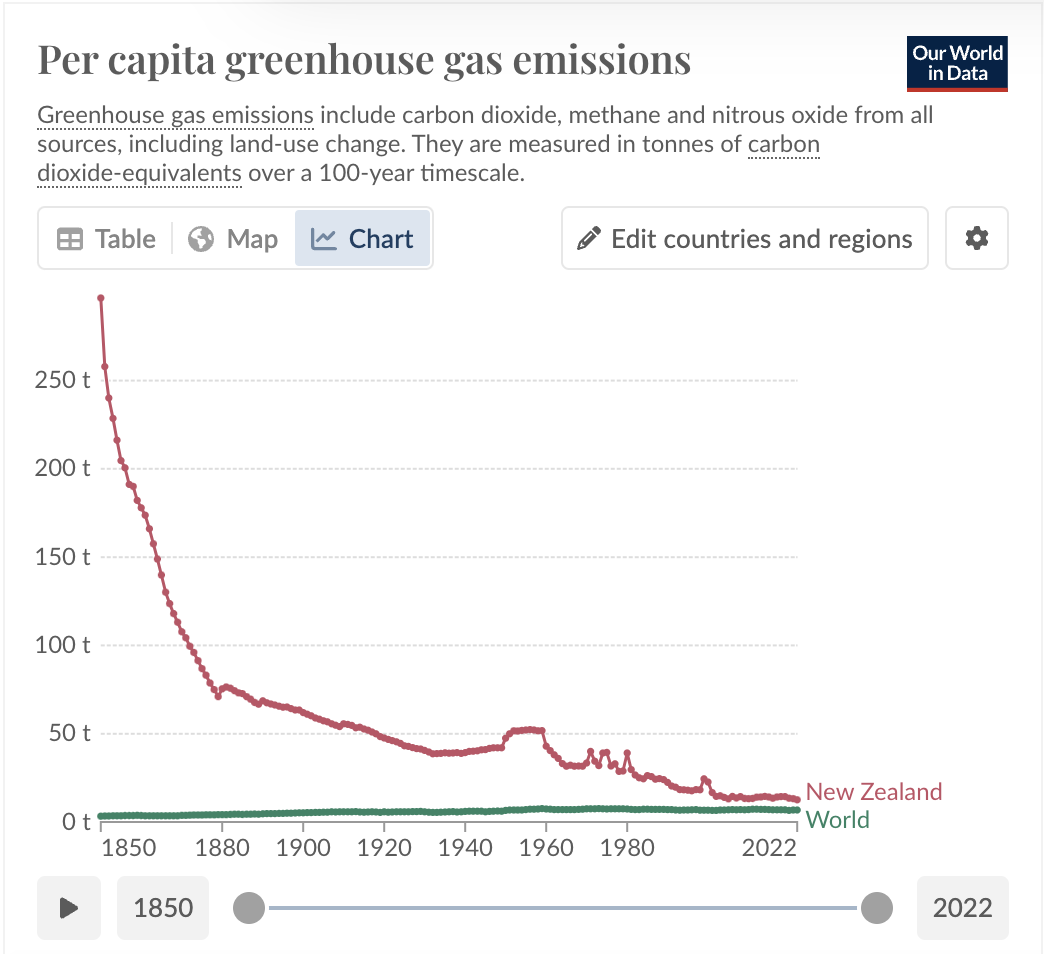
**Research objective 1:**

While we do not have much control over the natural causes of climate change, human-driven causes that significantly impact climate change can be controlled (*What Are the Causes of Climate Change?*, 2024). Activities like burning fossil fuels, deforestation, and livestock farming release harmful gases like CO2 and Methane that trap the heat from the sun, making the planet warmer. Greenhouse gases are at the roots of climate change (Ritchie et al., 2023).



*Human and natural influences on global temperature. Source:* [*Causes of Climate Change | US EPA*](https://www.epa.gov/climatechange-science/causes-climate-change)*. (Causes of Climate Change | US EPA, 2024)*

Being a small country, the greenhouse gases produced per person are prominent in New Zealand (*Climate Change Action*, 2024). Considering CO2, one of the greenhouse gases, in 2021, the average CO2 emissions of OCED countries were 11 tonnes, while New Zealand emitted 15 tonnes of CO2 per capita (Less, n.d.).



*Comparing GHG emission per capita, New Zealand vs World*

*Source:* [*New Zealand: CO2 Country Profile - Our World in Data*](https://ourworldindata.org/co2/country/new-zealand)

Hence, the first objective explores which sectors highly contribute to total GHG emissions in New Zealand. World annual greenhouse gas emission data was acquired from *Our World in Data* and merged with the *GHG emissions by sector in New Zealand*. Additionally, ‘Annual CO2 emissions’ (from 1990 – 2020) of New Zealand were merged as CO2 emission is the highest contributing factor in GHGs.

**Data cleaning and analysis:** Filters were applied for ‘New Zealand’. Unwanted columns like ‘Code’ were dropped. A few missing values were replaced with the median of that particular column.

The datasets were imported into Python (Google Collaboratory), and descriptive statistics were performed.

Below are the details and screenshots:

1. Exploring the top using the **head()** function.

A screenshot of a computer

Description automatically generated

1. Exploring the bottom using the **tail()** function.

A screenshot of a computer

Description automatically generated

1. Exploring the random rows using the **sample()** function.

**A screenshot of a computer

Description automatically generated**

1. Exploring the information about the data frame using **info()**.

**A screenshot of a computer

Description automatically generated**

1. Descriptive statistics and top findings:

**A screenshot of a white sheet

Description automatically generated**

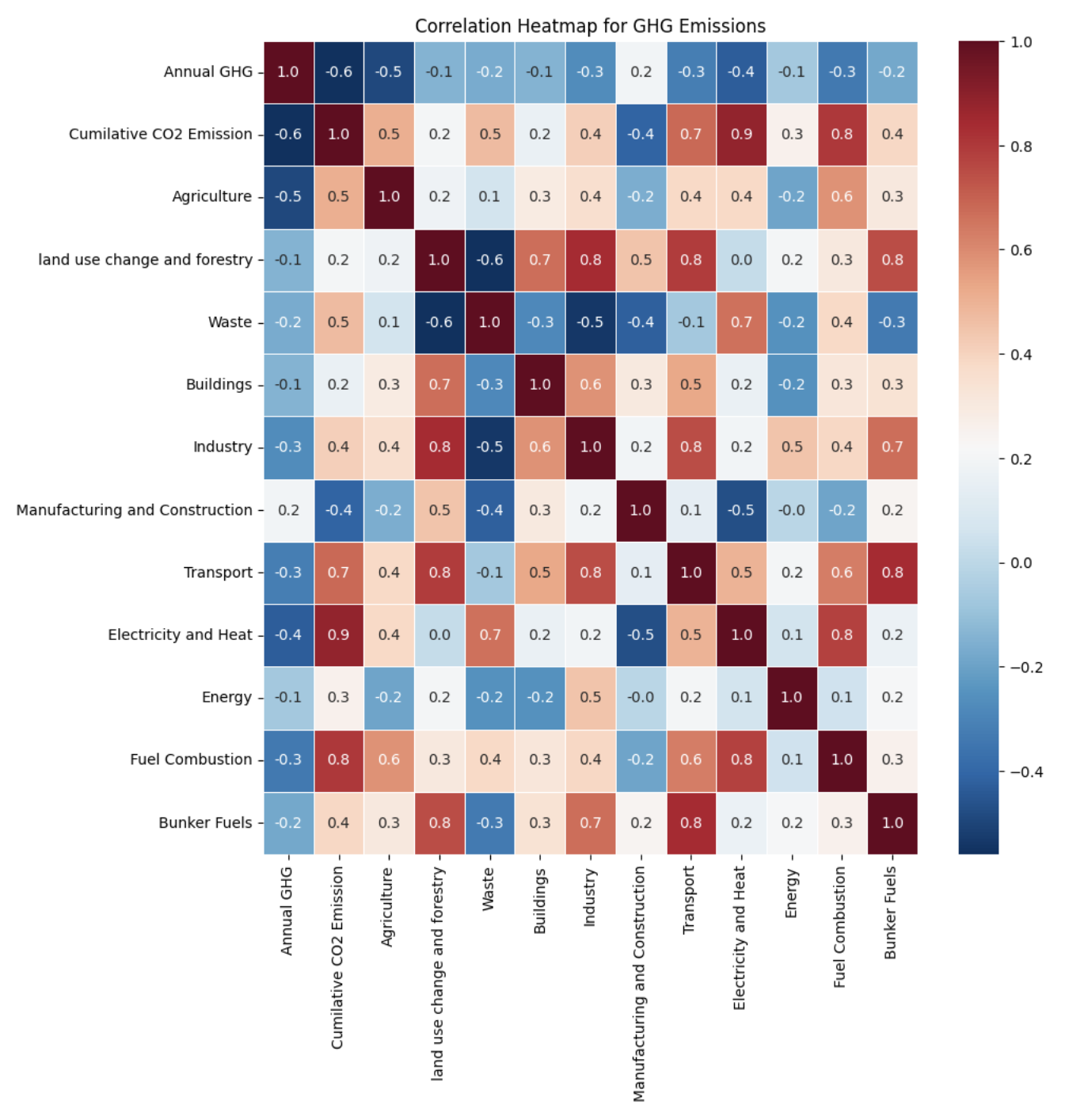
* 1. The average annual GHG emissions from 1990 to 2020 is 65.8 million CO2 equivalent.
  2. The average GHG emissions from Agriculture (42 million CO2 equivalent) are the highest, followed by Transportation (13.0 million CO2 equivalent).
  3. The transportation sector has the highest standard deviation at approximately 2,101,714.66 CO2 equivalents, indicating substantial variability in emissions over the years.
  4. The building sector has the lowest standard deviation, 133,462.57 CO2 equivalents, indicating stable annual emissions.

1. Further, a correlation analysis was conducted to explore which sectors have a high correlation with Annual GHG emissions and CO2 emissions.

**A screenshot of a computer

Description automatically generated**

Correlation table output

****

Correlation heatmap without masking.

Below are the insights from the correlation analysis:

1. The GHG emissions from the transportation sector (0.9) have a high positive correlation with cumulative CO2 emissions, suggesting that the transportation sector significantly contributes to overall CO2 emissions.
2. The transportation sector (0.3) positively correlated with Annual GHG emissions.
3. A strong correlation between industry and manufacturing/construction (0.8) suggests that these sectors' emissions are closely linked, likely due to shared processes and resource use.
4. Buildings, Land Use Change and Forestry, and Agriculture sectors show strong positive correlations with each other (0.7 to 0.8), indicating that emissions in one of these sectors are likely to be accompanied by emissions in the others. This could be due to interconnected activities, such as land management practices affecting forestry and agriculture.

Regression analysis was conducted to analyse how different greenhouse gases contribute to overall Annual GHG emissions.

Here, the dependent variable was **Annual GHG**, and the independent variables were **Annual CO2, Nitrous Oxide and Methane emissions**.



Interpretations assuming *5% significance level*:

1. Adjusted R square is 1, showing a strong relation between Annual GHG, nitrous oxide, methane and CO2 emissions.
2. F-significance is less than the assumed significance of 5%, showing that the model is ‘best fit’.
3. The coefficient is 1 for all the independent variables, showing a perfect positive relation.
4. The P-value is less than 5%, indicating all the variables are statistically significant to Annual GHG emissions.

Furthermore, using regression analysis in Microsoft Excel, Cumulative CO2 emission was analysed in relation to GHG emissions from different sectors. Cumulative CO2 emissions are considered as the dependent variable, and the sectors are independent variables.



Below are the findings, assuming a significance level of 5%.

1. Multiple R is 0.98, close to 1, showing a strong relationship between Annual Cumulative CO2 and GHG emissions from different sectors in New Zealand.
2. Adjusted R Square is 0.94, meaning other independent variables can explain 94% variation in Annual Cumulative CO2.
3. The F-significant is lower than the assumed significance level, showing that the model is the best fit.
4. GHG emissions from Agriculture and Transportation have positive coefficients, indicating that an increase in emission from both sectors is associated with an increase in cumulative CO2 emissions.
5. The P-values of the independent variables are less than the assumed significance level, suggesting they are statistically significant to Cumulative CO2 emissions.

Overall, it can be seen that GHG emissions from the transportation and agriculture sectors are highly correlated to cumulative CO2 emissions. Different gases like CO2, methane and nitrous oxide are perfectly related to Annual GHG emissions in New Zealand. Another regression analysis showed a strong relationship between annual cumulative CO2 and GHG emissions from various sectors in New Zealand, including agriculture and transportation. This analysis helps MFAT identify the industries that heavily contribute to high climate change factors.

***Objective 2: Is there a relationship between tourism and cumulative carbon emissions in New Zealand?***

Tourism accounts for 8% of global carbon emissions (*Causes of Climate Change | US EPA*, 2024). In 2021, tourism in New Zealand accounted for 6452-kilo tonnes of carbon emissions (*Greenhouse Gas Emissions (Consumption-based): Year Ended 2021 (Provisional) | Stats NZ*, n.d.). This section analyses whether there is a relationship between the number of tourist visitors and carbon emissions in New Zealand.

**Datasets**: For analysis, tourism forecast database was acquired from New Zealand’s MBIE.

**Data cleaning and analysis**: The Cumulative CO2 emissions data was merged with the tourism data (Tousrism1.xls). Unwanted columns like Code, visitor spends, and the average stay were removed to make the dataset lean.

1. Descriptive analysis using **describe()**.

A screenshot of a graph

Description automatically generated

The following are the observations:

1. The average length of stay is 19 days.
2. The number of visitors arriving in New Zealand varies greatly, ranging from 1 million to over 4.3 million, resulting in lower and higher tourism periods.
3. The cumulative CO2 emission shows a gradual upward trend over the years, considering the average value of 0.112.

Next, the report analyses if the increase in visitor arrivals in New Zealand increases annual CO2 emissions. Hypothesis testing is performed for this analysis.

A ‘Two-sample, Two-tail T-Test assuming unequal variance at a 5% significance level’ is performed to get the following results (TouristvsCO2.xls).

***Ho****: The increase in average arrivals of visitors does not affect the average annual CO2 emissions.*

***H1****: The increase in average arrivals of visitors affects the average annual CO2 emissions.*



Key observations:

1. T-statistics (-46.079) is lower than t-critical (2.04), providing strong evidence against the null hypothesis.
2. The p-value (3.1248E-32) is significantly lower than the assumed significance level of 5%.
3. There is enough evidence to reject the null hypothesis, meaning the increase in average arrivals of visitors affects the average annual CO2 emissions.

Moreover, correlation and regression analysis were performed to check how the number of visitors relates to annual CO2 emissions, assuming a 5% significance level.

**A screenshot of a computer

Description automatically generated**

A graph of a number of visitors

Description automatically generated

Correlation matrix

****

Regression analysis

The key observations are:

1. Multiple R is 0.897, close to 1, suggesting a strong relationship between the dependent variable, annual CO2 emissions in New Zealand and independent variables like total visitor days and total visitor arrivals.
2. Adjusted R square is (0.791), meaning that 79% of the variation in annual CO2 emissions in New Zealand can be explained by visitor days and arrivals.
3. F-significance is less than the assumed significance level (0.05), suggesting the model is the best fit.
4. P-values are less than the 5% significance level, suggesting visitor days and visitor arrivals are statistically significant to Annual CO2 emissions.

Overall, using descriptive statistics, hypothesis testing (two-sample, two-tail t-test, unequal variance), correlation and regression analysis, there is enough evidence suggesting that tourism (total number of visitors and days spent) significantly contributes to annual CO2 emissions in New Zealand.

***Objective 3: Considering the historical temperature, rise in sea level, GHG emissions and other climate-related disasters, how does the prediction look for the next eight years, and how can the insights be used to prioritise environmental policies and resource allocations?***

Predicting future climate scenarios is essential for anticipating potential impacts on different aspects of the environment, including what disasters can increase due to surface temperature change, how rising sea levels and extreme weather events lead to increased coastal level flooding and biodiversity displacement, etc. (United Nations, n.d.). Studying factors leading to climate change helps us identify natural processes and how they are affected by human-related activities.

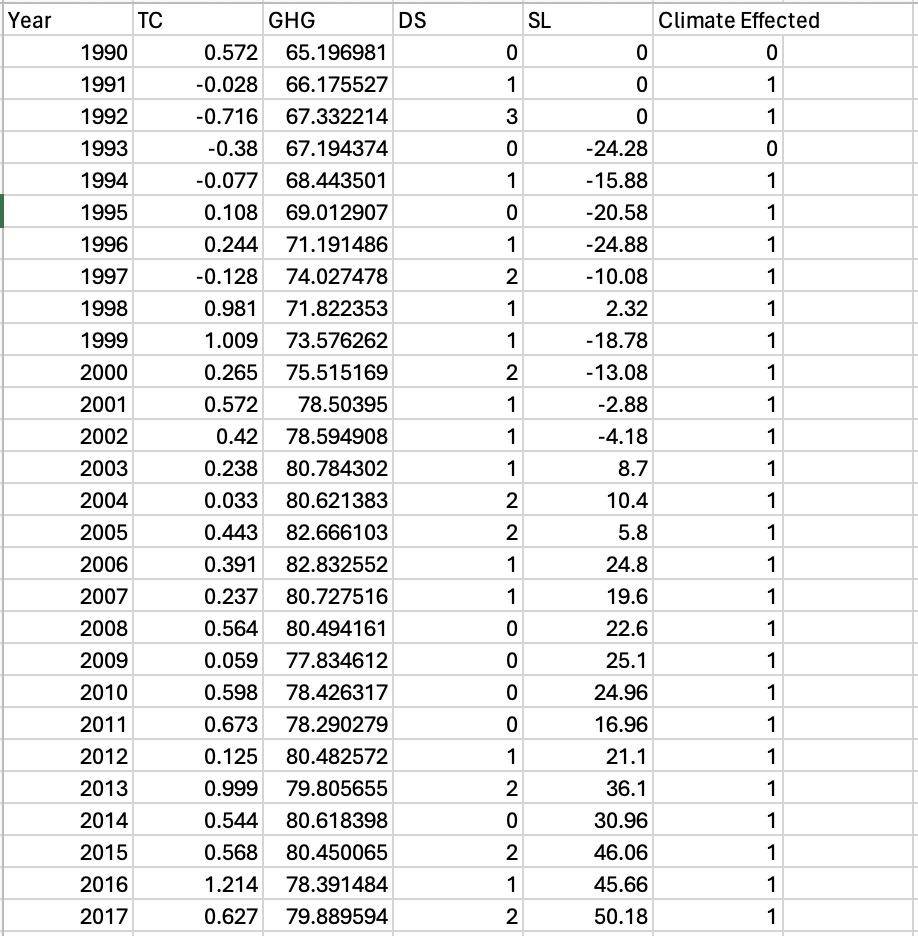
To explore this objective, we used several datasets from IMF climate change, such as annual surface temperature change, change in mean sea levels, GHG emissions, and climate-related disaster frequency. We merged all the datasets in one Excel sheet, filtered them by ‘New Zealand’ and removed unwanted columns and rows (Prediction.csv). These are further named as *‘four factors’* affecting climate change.

Here are the descriptive statistics for the used dataset

**A screenshot of a graph

Description automatically generated**

Further, a new column, ‘Climate Effected’, is added where ‘1’ represents the climate affected, and ‘0’ represents the climate not affected. The values were formulated considering the magnitude of factors from all four columns.

****

The datasets were further segregated into four different CSV files that were used in Python. ARIMA (Autoregressive Integrated Moving Average) model was used to predict the future values of the next eight years based on the historical data from 1990-2022 for all four factors individually. The datasets, Python codes, and results are attached to this report.

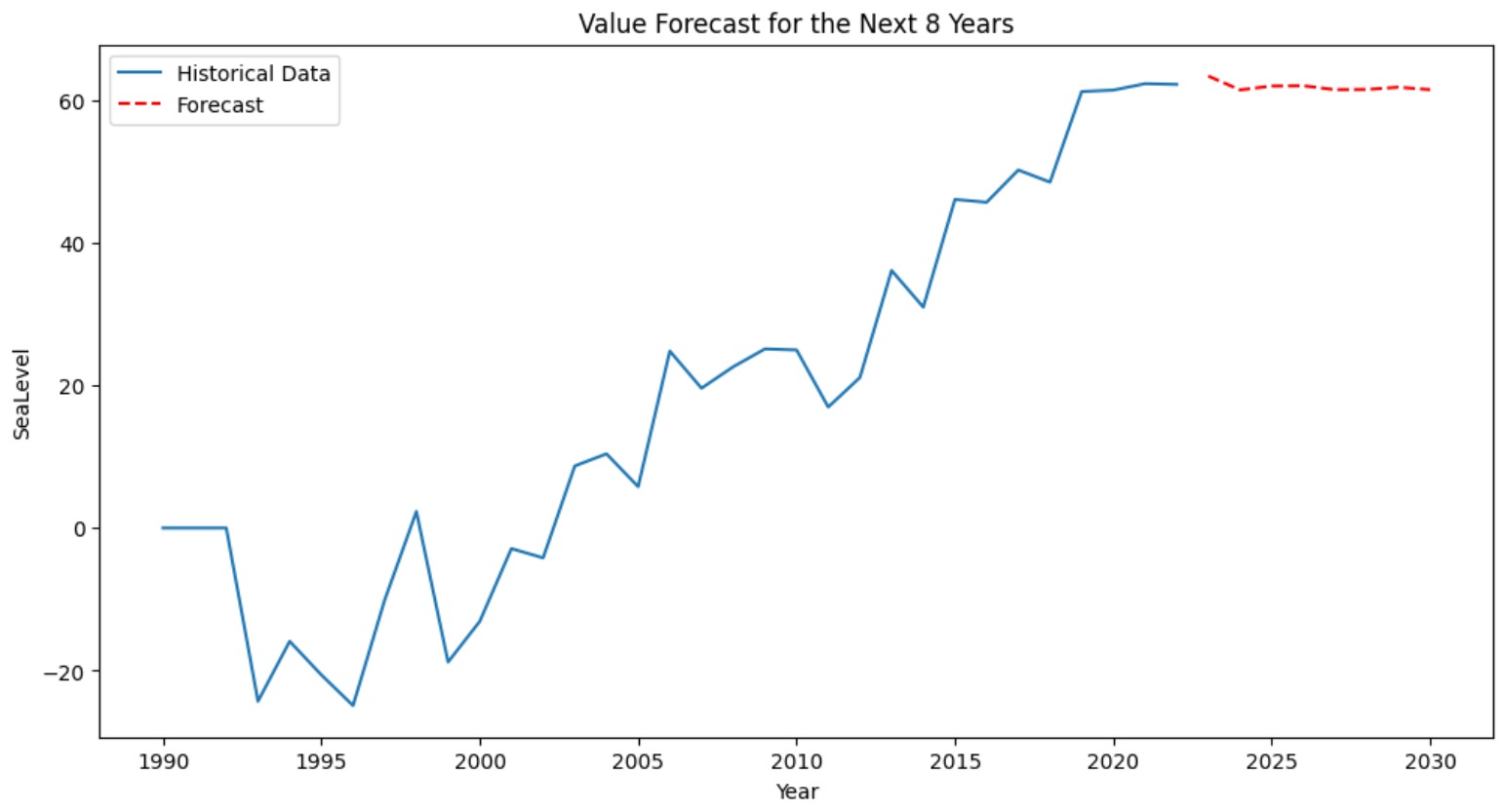
Here’s what the prediction looks like for all four factors.

1. Annual Surface Temperature change prediction

A graph showing the value of a stock market

Description automatically generated

1. Mean sea level change prediction



1. GHG Emissions change prediction

A line graph showing the value of a company

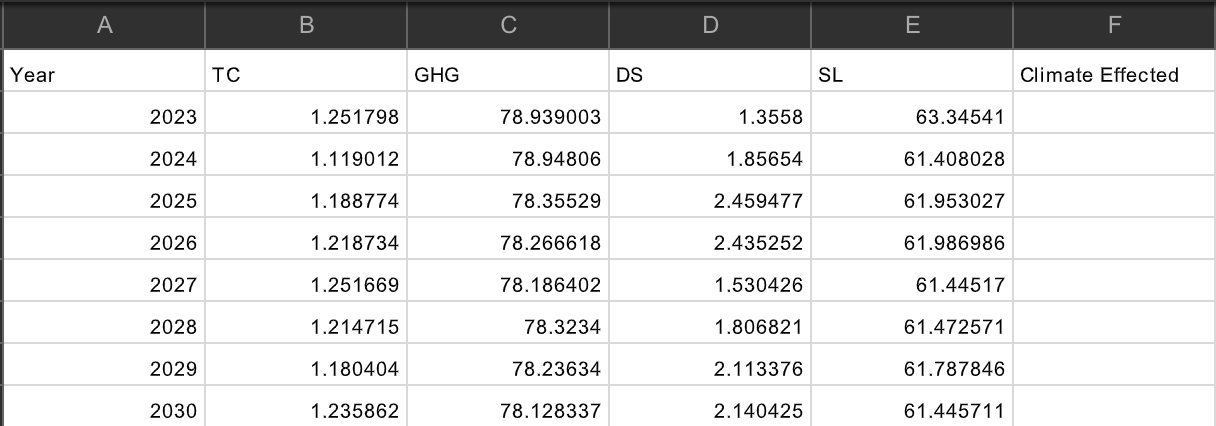
Description automatically generated

1. Climate-related disasters change prediction.

A graph showing the value of a stock market

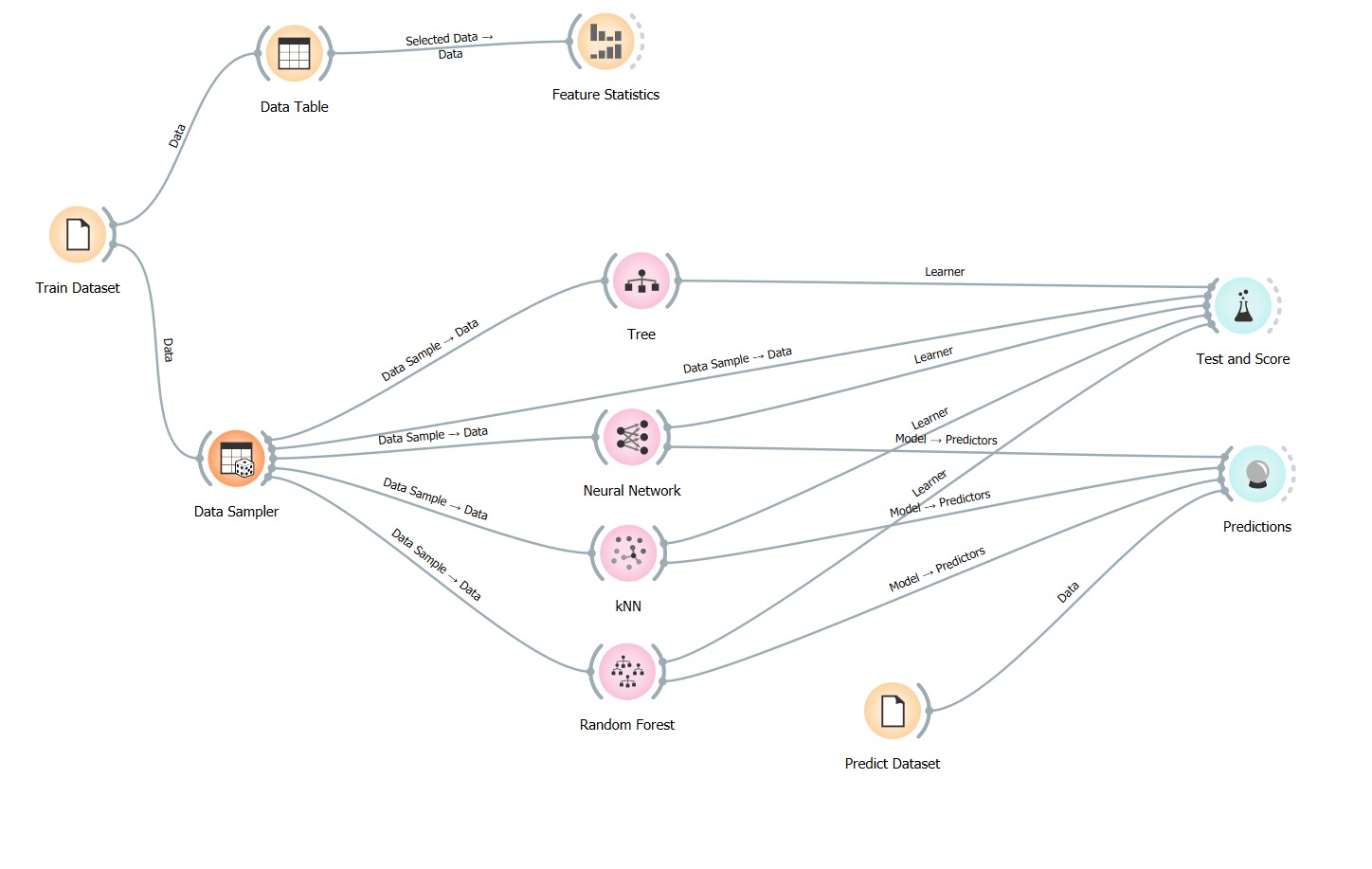
Description automatically generated

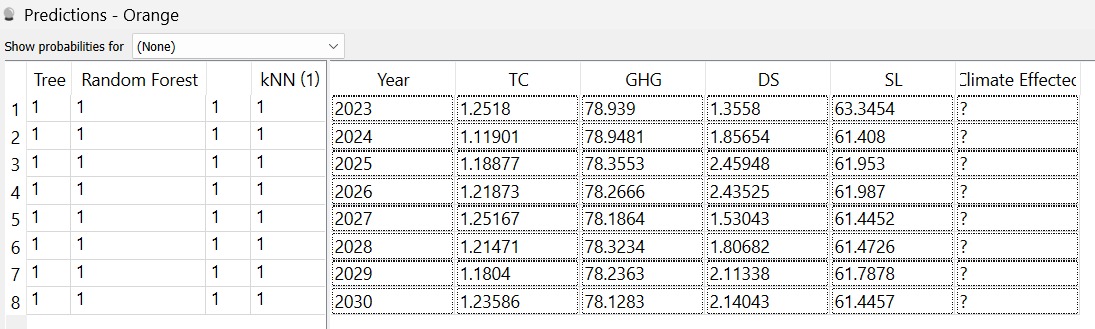
The following are the predicted values for all four factors for the next eight years (from 2023 to 2030) (Predict.csv).

****

Moreover, two datasets (train and predict) were created for data mining in Orange. The training dataset had a column called ‘Climate Effected’ considered the target variable. The predicted dataset had eight years of forecasted values generated from Python’s ARIMA model, and the ‘Climate Effected’ column was left blank for prediction values from Orange.

After training the algorithm using train.csv data using different models like kNN, decision trees, and random forests, the prediction results are below. The accuracy for the three models is 0.988, 0.938, and 0.938 for KNN, decision tree and Random forest, respectively. The predictions are as shown below.





A screenshot of a computer

Description automatically generated

The predictions show that the four factors will continue negatively affecting the climate, disturbing the delicate environmental balance for the next eight years.

Overall, MFAT can focus on employing sustainable strategies to reduce the human impacts that affect these four factors contributing to climate change.

## Conclusion

This research suggests that climate change needs immediate attention in New Zealand. The GHG emissions from agriculture and transportation are significant contributors compared to other sectors. Furthermore, this research established a link between increased tourism and its potential relationship with growing annual CO2 emissions. It further explores the need for sustainable tourism practices. Moreover, alarming predictions demonstrate rising sea levels, increase in surface temperature and climate-related disasters in the next eight years, highlighting the emergency of proactive policies. With the help of the best strategic solutions to mitigate climate actions and strengthen international partnerships, MFAT can push New Zealand to become a sustainable leader globally. Lastly, this report is a blueprint for policymakers to guide them in creating informed strategies to benefit the environment and the next generation.

**References:**

Introduction to Climate Change. (n.d.). In *climateassembly.uk*. Climate Assembly UK. <https://www.climateassembly.uk/documents/5/Introduction_to_climate_change_FINAL_002.pdf>

Tvaronavičienė, M. (2021b). Effects of climate change on environmental sustainability. *E3S Web of Conferences*, *250*, 01005. <https://doi.org/10.1051/e3sconf/202125001005>

New Zealand Ministry of Foreign Affairs and Trade. (2024, June 13). *New Zealand Ministry of Foreign Affairs and Trade*. <https://www.mfat.govt.nz/>

New Zealand Ministry of Foreign Affairs and Trade. (n.d.). *New Zealand Ministry of Foreign Affairs and Trade*. <https://www.mfat.govt.nz/en/about-us/our-strategic-direction>

Cevik, S. (2022, September 9). *Dirty Dance: Tourism and environment*. IMF. <https://www.imf.org/en/Publications/WP/Issues/2022/09/09/Dirty-Dance-Tourism-and-Environment-523271>

*What are the causes of climate change?* (2024, March 26). <https://www.nrdc.org/stories/what-are-causes-climate-change>

*Climate Change action*. (2024, April 16). WWF-New Zealand. <https://wwf.org.nz/climate-change-action>

*Causes of climate change | US EPA*. (2024, April 12). US EPA. <https://www.epa.gov/climatechange-science/causes-climate-change>

*Climate change data*. (n.d.). <https://climatedata.imf.org/pages/climatechange-data>

United Nations. (n.d.). *Causes and effects of climate change | United Nations*. <https://www.un.org/en/climatechange/causes-and-effects-climate-change>