Global Sustainable Development Goals with a focus on: Climate Change

TEAM GLOOM

Tuğba AKIN	Baran DATLAR	Giray ÇELİKKAYA	Sedanur ASLAN
191180008	191180031	19118002	191180011
tubakkin99@gmail.com	barandatlarr@gmail.com	celikkayagiray@gmail.com	sedanur.aslan121@gmail.com

ABSTRACT

In this report, we analyzed climate change and its causes. We visualized the data and conducted research on these visualized data. During this visualization, we used the dataset containing the forest area percentage, glacier main extent, temperature anomalies, global sea mean level and land temperature change by year.

1. INTRODUCTION

One of the top priorities in the Global Sustainable Development Plan is to combat climate change. In the face of increasing temperatures, polluted air, major natural disasters and many factors, the world is no longer able to accommodate living things, whether plant, animal or human. This shows the importance of climate change for life. The fact that countries work together on climate change and prepare joint panels and budgets is an important example of this. In this study, approximately 25 years of data on climate change have been reached, these data have been analyzed, visualized and outputs have been obtained. The changes in the data over time were examined. Various APIs and data sets were used to analyze climate change from different perspectives.

In this study, we used the forestland dataset to show the 30-year change in forestland by country. With yet another dataset, we analyzed glacier area data by year, month, day and hemisphere. The data set, which includes ocean temperature anomaly data between 1880 and 2021, is also among the questions we examined.

2. METHODOLOGY

In this research, we examine various sub-branches on climate change. Among the climate change problems we are working on are; There are subjects such as ocean temperature anomaly, the effect of global warming on land temperature, the percentage of forest areas according to years, the amount of melting of glaciers according to years, the change in sea level according to years. We used many Python libraries (pandas, numpy, scikit-learn, seaborn, plotly, matplotlib) to process and visualize these topics in detail. In order to create a dataframe object and work on the Pandas library in general, the numpy library to put the data in the appropriate form in numerical operations; We used the scikit-learn, seaborn, plotly and matplotlib libraries to visualize the data and graph the ML techniques we used.

We used 2 types of ML techniques. ML techniques are most generally grouped as supervised and unsupervised. The first technique we used, the linear regression technique, is in the subcategory of the supervised group. This technique is used to predict the value of one variable based on the value of another variable. We have seen that the accuracy of this technique can be high because the data in the dataset we use this technique follow a more linear path. Therefore, we taught the algorithm to predict by creating test and train sets, and we got a prediction result with high accuracy. In other words, we have proven that the linear regression technique is suitable for this data set.

The other technique we use is the k-means algorithm, which is the algorithm of the clustering method, which is a subcategory of the unsupervised group. Here k represents the number of clusters. It has K center points. According to these central points, new clusters are formed and continue until the system is stable. The reason for using this clustering technique is that the data in the data set we use is not in a certain linear plane and data groups are formed in the form of clustering. The stability and accuracy level of the clusters formed using this algorithm was high. Considering this, it can be said that we use this algorithm in the right place.

.

3. DATASETS

Our datasets are these:

1. https://data.worldbank.org/indicator/AG.LND.FRST.ZS

```
Data columns (total 2 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Years        31 non-null     int64
 1   Forest Area  31 non-null     float64
```

Figure 1. Forest Area Percentage by Year

2. https://www.kaggle.com/datasets/nsidcorg/daily-sea-ice-extent-data%C2%A0

```
Data columns (total 2 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Years        42 non-null     int64
 1   Mean Extent  42 non-null     float64
```

Figure 2. Glacier Main Extent by Year

3. https://www.statista.com/statistics/736147/ocean-temperature-anomalies-based-on-temperature-departure/

```
Data columns (total 2 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Year     142 non-null    int64
 1   Anomaly  142 non-null    float64
```

Figure 3. Ocean Temperature Anomalies by Year

4. https://www.kaggle.com/datasets/somesh24/sea-level-change/versions/1?resource=download

```
Data columns (total 3 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   gmsl         134 non-null    float64
 1   uncertainty  134 non-null    float64
 2   year         134 non-null    int64
```

Figure 4. Global Sea Mean Level

5. https://global-warming.org/

```
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Time    143 non-null    int64
 1   Land    143 non-null    float64
```

Figure 5. Land Temperature Change by Year

4. RESULTS

4.1.Forest Area Percentage

The data set contains 265 countries and world mean forest area by years from 1990 to 2020. We did not use a few columns as there are no data from 1960 to 1990. We visualized the data set by using sklearn, pandas and pyplot.

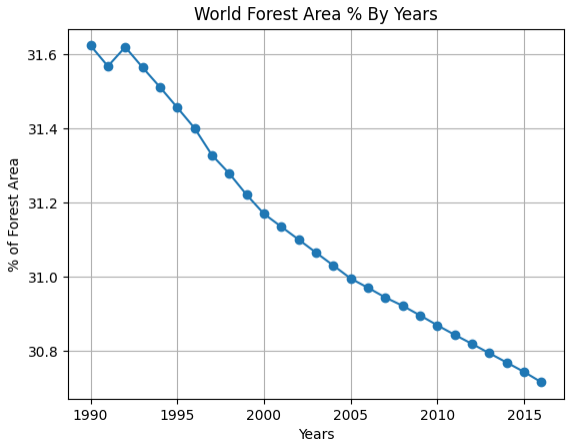

Figure 1. Shows the World Area Percentage Change by Years

First we split the data set to train set and test set. After that we predict with linear regression using the test set. We added prediction as a line with the test set while visualizing.

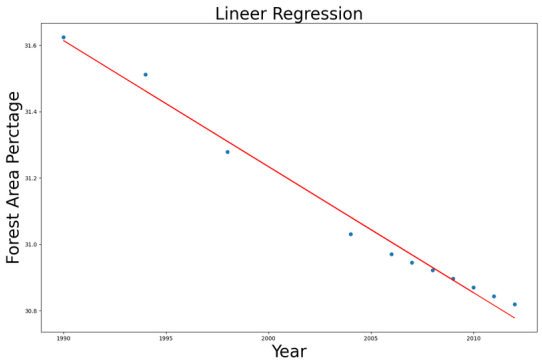

Figure 2. Shows World Forest Area Percentage with Linear Regression Line

4.2.Glacier Main Extent

The dataset consists of six columns. These are year, month, day, extent, missing, source and hemisphere. But we used only two columns. These are year and extent. First we sum the extent data by year and averaged it. Final data set consists of the mean extent by year from 1978 to 2020.

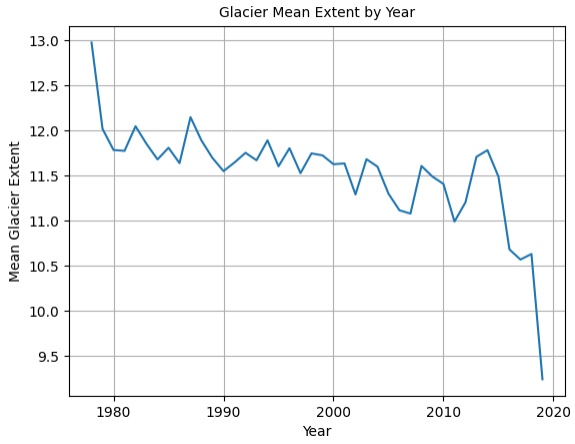

Figure 3. Shows the Glacier Mean Extent by Year

Figure 3. shows glacier mean extent visualized with a line chart.

First we split the data set to train set and test set. After that we predict with linear regression using the test set. We used sklearn for linear regression. We added prediction as a line with the test set while visualizing.

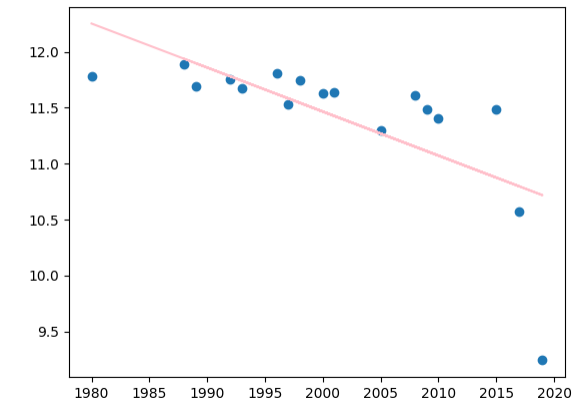

Figure 4. Shows Glacier Mean Extent with Linear Regression Line

4.3 Ocean Temperature Anomaly

In our dataset we used in this problem, there are values under the headings of "year" and "anomaly". The anomaly value takes positive and negative values, it expresses the increase and decrease in temperature. These values are categorized by years. Since the dataset we use is a csv file, we turned it into a dataframe object with pandas. Afterwards, we determined the number of clusters to be used and transferred this to the algorithm. We created the structure of the model and created the algorithm with the Kmeans function from the sklearn.cluster library. We

introduced the dataframe object to the algorithm and visualized it with the matplotlib library.

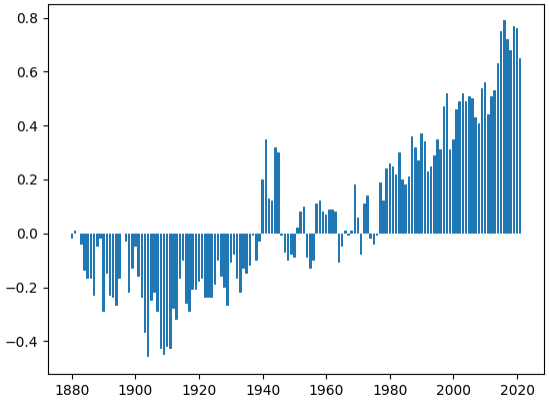

Figure 5. Bar chart showing ocean anomaly values by year

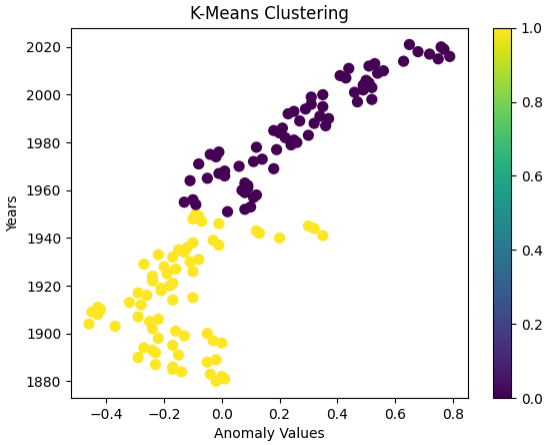

Figure 6. K-means algorithm created according to ocean anomaly values

4.4.Global Mean Sea Level

The dataset about Global Mean Sea Level shows the change in sea level with data from 1880 to 2013. Dataset contains Time,GMSL,GMSL uncertainty values. GMSL means Global Mean Sea Level and GMSL uncertainty shows sensitivity change over metrics.

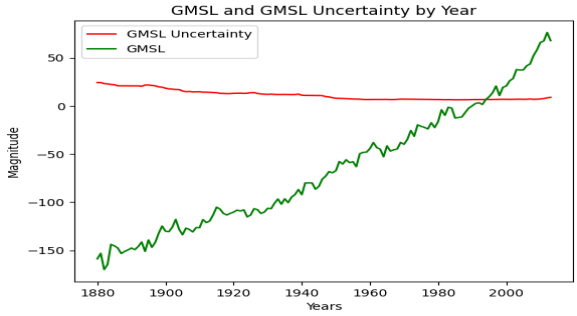

Figure 7. Shows the change in GMSL and GMSL Uncertainty value between 1880 and 2013

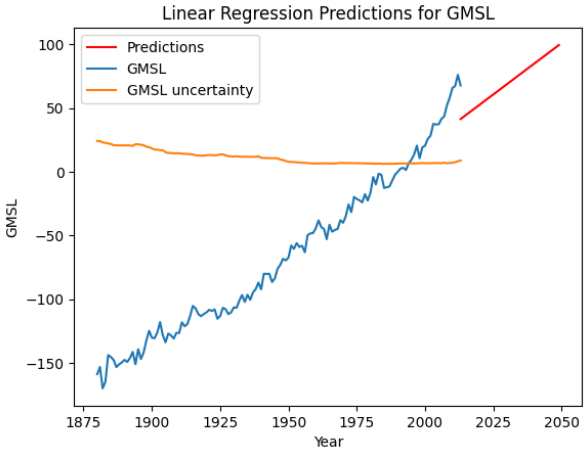

Figure 8. Based on the change in the GMSL and GMSL Uncertainty value between 1880 and 2013, shows how the GMSL value will change from 2050

Using the linear regression method, we showed how the sea level will change over the period up to 2050, with the data shown on the line chart.

4.5 Land Temperature Change

We pulled the data we used in this problem from the global-warming.org API. To pull from the API, we imported the requests library and got the data from there with the get request. Then we converted it from json format to dataframe object with the help of pandas library. Then we converted it to float form because the data is string. The data comes in the "land" and "time" columns. We used the same technique we used in the ocean temperature anomaly for the same reasons. We used the Sklearn.cluster and Kmeans function. Then we visualized it with the matplotlib library.

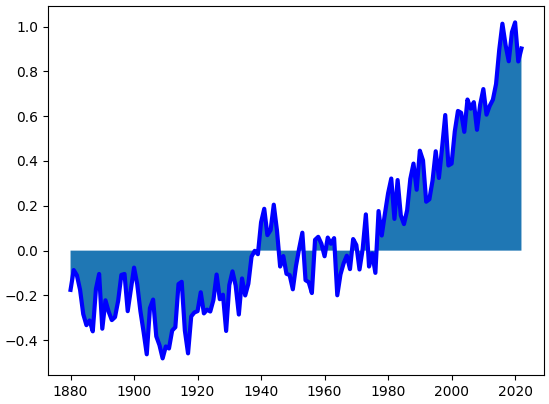

Figure 9. Display of the changing land temperature due to global warming over the years with area chart with grid

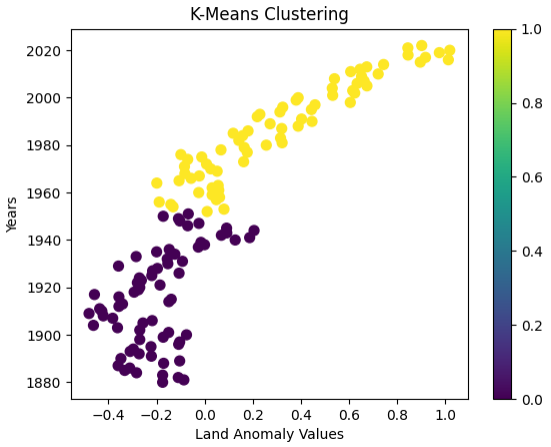

Figure 10. k-means algorithm applied to land temperature data changing due to global warming by years

5. CONCLUSION

Our analysis have shown that in dataset from 1880-2020 (generally) we have examined how the climate crisis has affected our world by addressing some specific issues. Our analyzes show that the percentage of forest area in the world is decreasing over time. When we look at the Glacier main extend graph, we see that the glaciers are getting smaller over the years. Especially the data between 2010 and 2020 reveals that there is a great loss of glaciers. When we examine our data related to sea level, we can say that the sea level has risen over the years. Our conclusion from glacial melting may be the main reason for this finding. We anticipate that the rise in sea level will continue to increase in the coming years. When we examine the ocean and land anomaly, we see that the data yield similar results. The reason for the great increase in values, especially after 1980, can be attributed to the rapid increase in global warming and the global problems created by mechanization.

6. REFERENCES

[1] *Forest area (% of land area)* (no date) *Data*. Available at: https://data.worldbank.org/indicator/AG.LND.FRST.ZS (Accessed: January 4, 2023).

[2] National Snow and Ice Data Center (2019) *Daily Sea ice extent data*, *Kaggle*. Available at: https://www.kaggle.com/datasets/nsidcorg/daily-sea-ice-extent-data (Accessed: January 4, 2023).

[3] https://www.statista.com/statistics/736147/ocean-temperature-anomalies-based-on-temperature-departure/

[4] Sharma, S. (2020) *Sea level change*, *Kaggle*. Available at: https://www.kaggle.com/datasets/somesh24/sea-level-change/versions/1?resource=download (Accessed: January 4, 2023).

[5] *Global warming data and API* (no date) *Global Warming API*. Available at: https://global-warming.org/ (Accessed: January 4, 2023).