CMSE 202 Final Project

How have carbon emissions affected climate change?

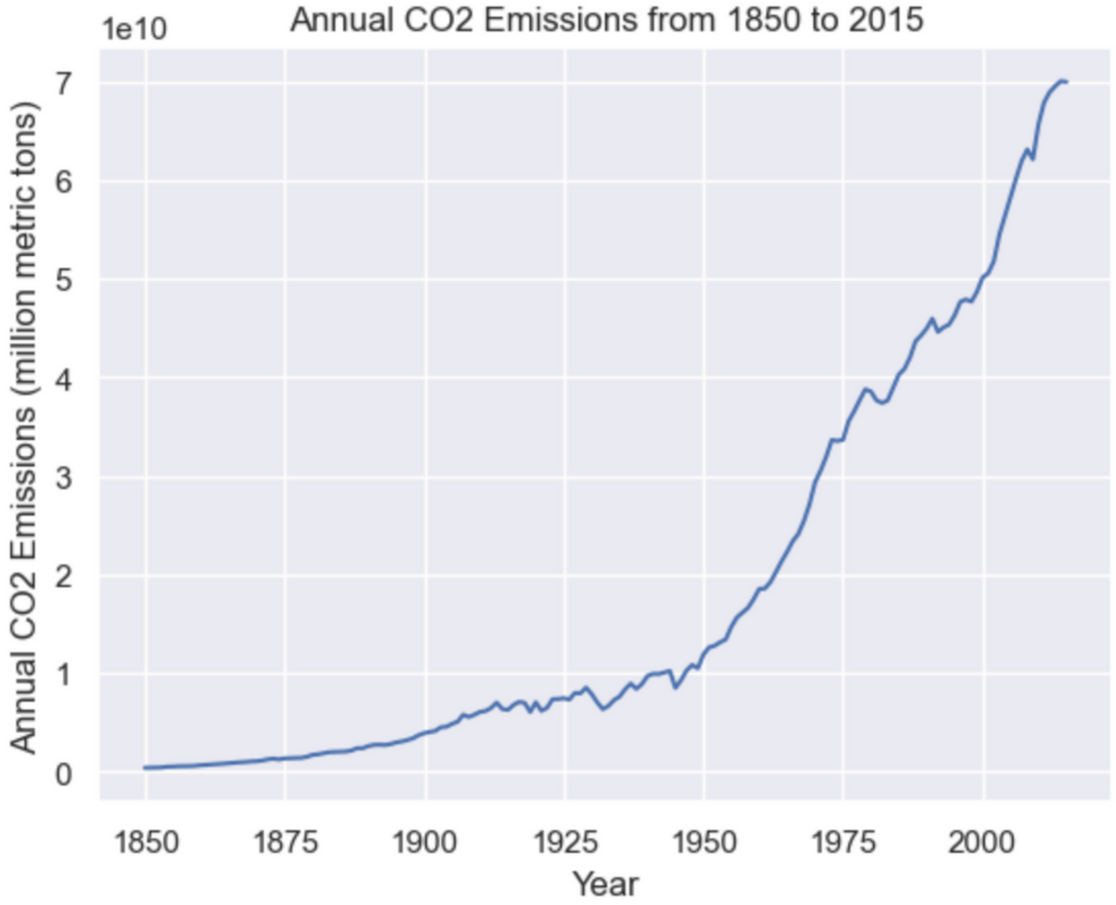
Group Members: Peter Fu Chen, Riley Millikan, Mike Liu, Alex Tran, Trevor Parcella

Background and motivation for the project

The purpose of this project is to explore the relationship between carbon emissions and climate change. Carbon emissions are known to be a leading contributor to climate change, which has significant impacts on the environment and human society. Therefore, understanding the relationship between these two variables is crucial.

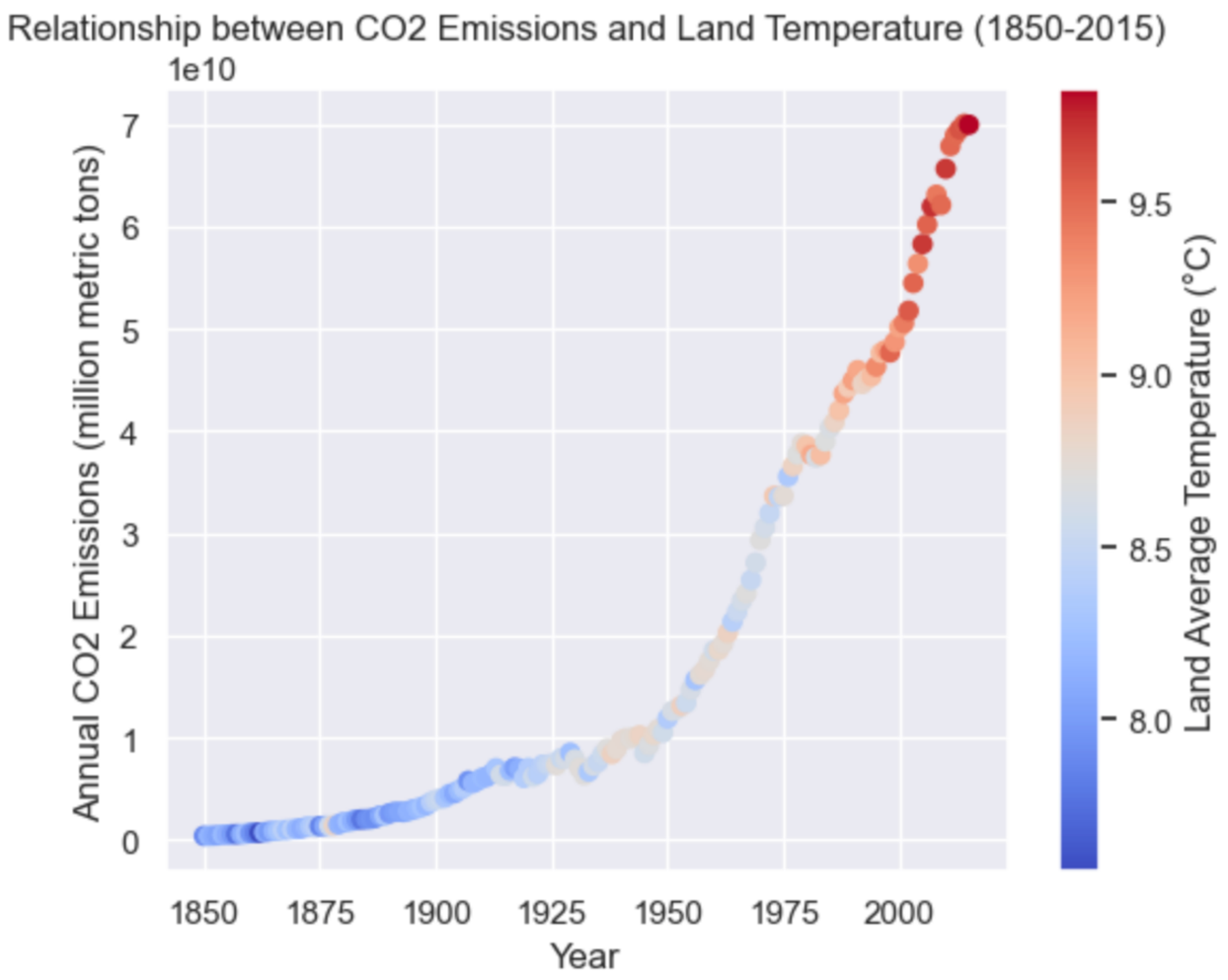
Objectives of the project

- Analyze historical trends in carbon emissions and climate change
- Examine the relationship between carbon emissions and various climate variables, such as temperature, extreme weather events, sea level, and ocean acidity
- Create visualizations to illustrate findings
- Draw conclusions about the impact of carbon emissions on climate change

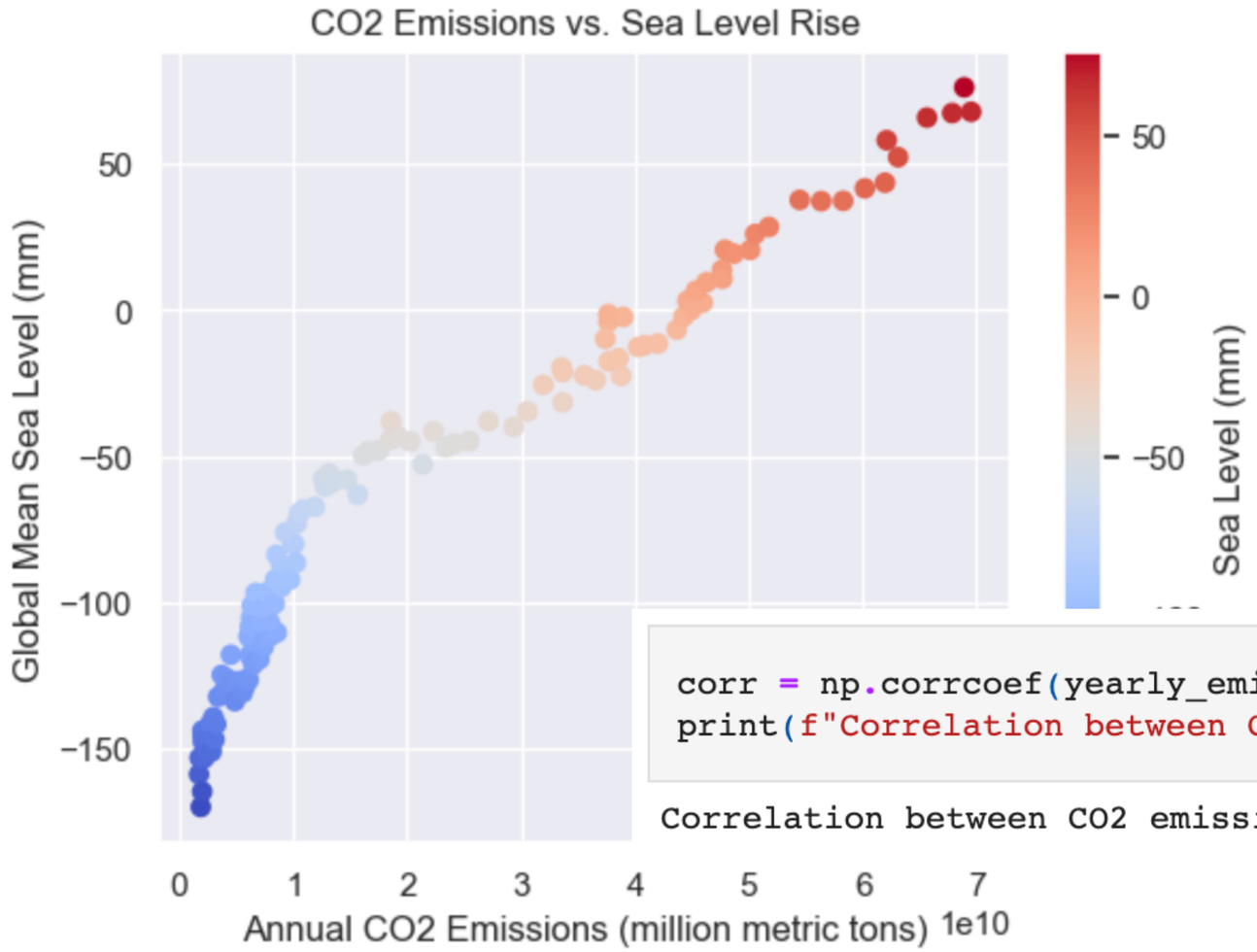


Line plot of annual CO2 emissions over time

- Historical trends of carbon emissions and climate change
- Data collection and cleaning process
- Visualization of historical trends using graphs and charts



Relationship between Carbon Emissions and Rising Temperatures



Comparative analysis

```
corr = np.corrcoef(yearly_emissions, annual_sea_level[('GMSL', 'mean')])[0, 1]
print(f"Correlation between CO2 emissions and global mean sea level: {corr:.2f}")
```

```
Correlation between CO2 emissions and global mean sea level: 0.96
```

Advanced statistical analysis and Machine Learning

```
data = pd.DataFrame({'CO2 emissions': yearly_emissions, 'Global mean sea level': annual_sea_level[('GMSL', 'mean')]})

X_train, X_test, y_train, y_test = train_test_split(data['CO2 emissions'], data['Global mean sea level'], test_size=0.3, random_state=0)

reg = LinearRegression()

reg.fit(X_train.values.reshape(-1, 1), y_train.values.reshape(-1, 1))

y_pred = reg.predict(X_test.values.reshape(-1, 1))

r_squared = reg.score(X_test.values.reshape(-1, 1), y_test.values.reshape(-1, 1))

print('R^2:', r_squared)
```

```
R^2: 0.9288360675445367
```

R-squared value of 0.93 suggests that the linear regression model provides a good fit to the data, indicating a strong relationship between CO2 emissions and sea level. We can interpret the R-squared value as the proportion of the variance in the dependent variable (sea level) that is explained by the independent variable (CO2 emissions).

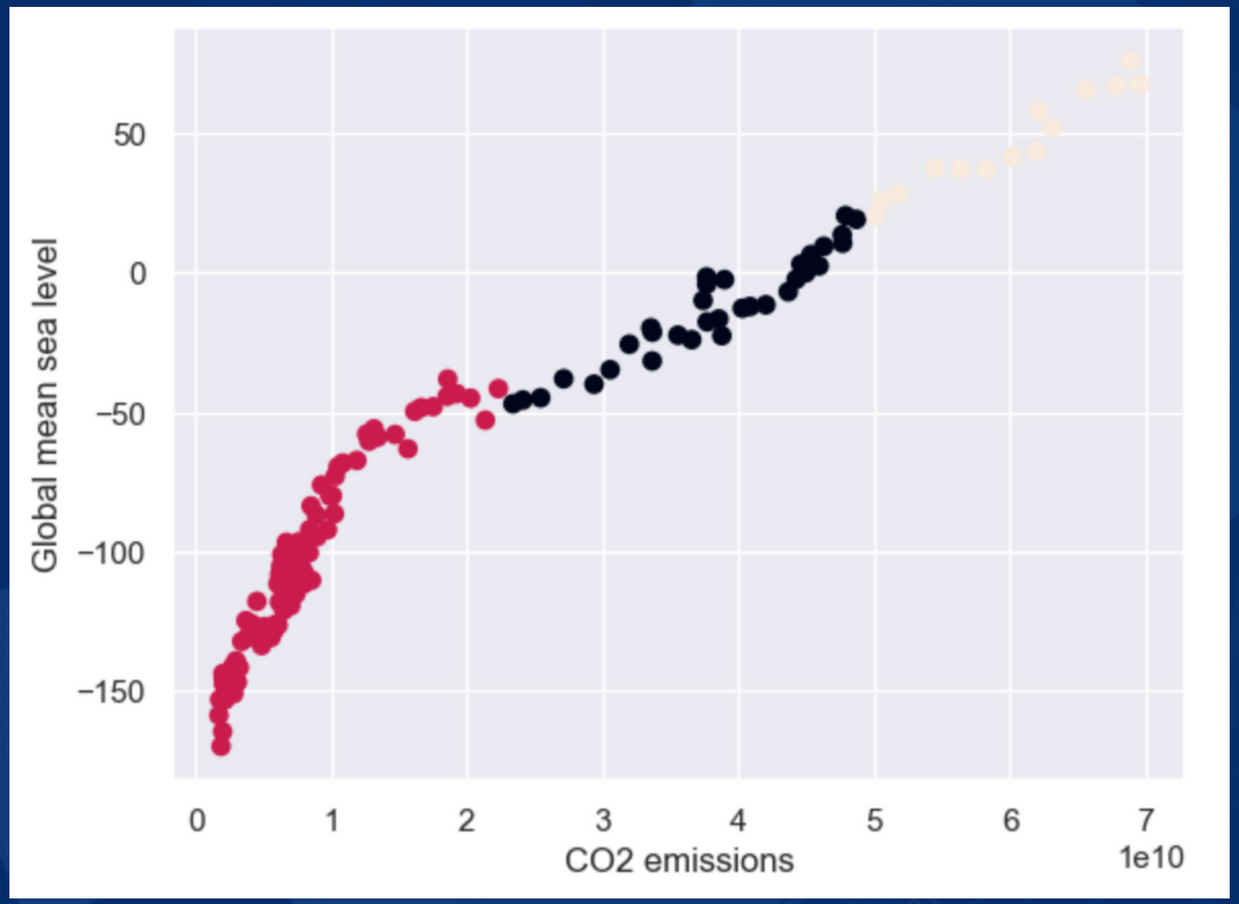
Advanced statistical analysis and Machine Learning

```
# Polynomial regression model
data = pd.DataFrame({'CO2 emissions': yearly_emissions, 'Global mean sea level': annual_sea_level[('GMSL', 'mean')]})

X_train, X_test, y_train, y_test = train_test_split(data['CO2 emissions'], data['Global mean sea level'], test_size=0.2, random_state=42)

X_train_reshape = X_train.values.reshape(-1, 1)
X_test_reshape = X_test.values.reshape(-1, 1)

poly = PolynomialFeatures(degree=2)
X_train_poly = poly.fit_transform(X_train_reshape)
X_test_poly = poly.transform(X_test_reshape)
reg = LinearRegression().fit(X_train_poly, y_train)

y_test_pred = reg.predict(X_test_poly)

r2_test = reg.score(X_test_poly, y_test)

print('R-squared (testing data):', r2_test)
```

```
R-squared (testing data): 0.9496302690577819
```

The R-squared value for the testing data is quite high, which suggests that the polynomial regression model is able to capture the nonlinear relationship between CO2 emissions and sea level more accurately than a simple linear regression model.

K-means clustering in sklearn



K-means clustering is an unsupervised machine learning algorithm that groups similar data points together in a given dataset.

Summarize

- Analyzed relationship between carbon emissions and climate change
- Used statistical and machine learning techniques for analysis and data visualization
- Found strong positive correlation between carbon emissions and sea level, as well as positive correlation between carbon emissions and land temperature
- Urgent need for policy interventions to reduce carbon emissions and mitigate impacts of climate change

Plans for future

- To increase the complexity of our model, we can add more layers or utilize advanced architectures
- Experimenting with various machine learning methods can help improve prediction accuracy
- It is important to investigate potential sources of causation in our data to ensure reliable predictions

Thank You