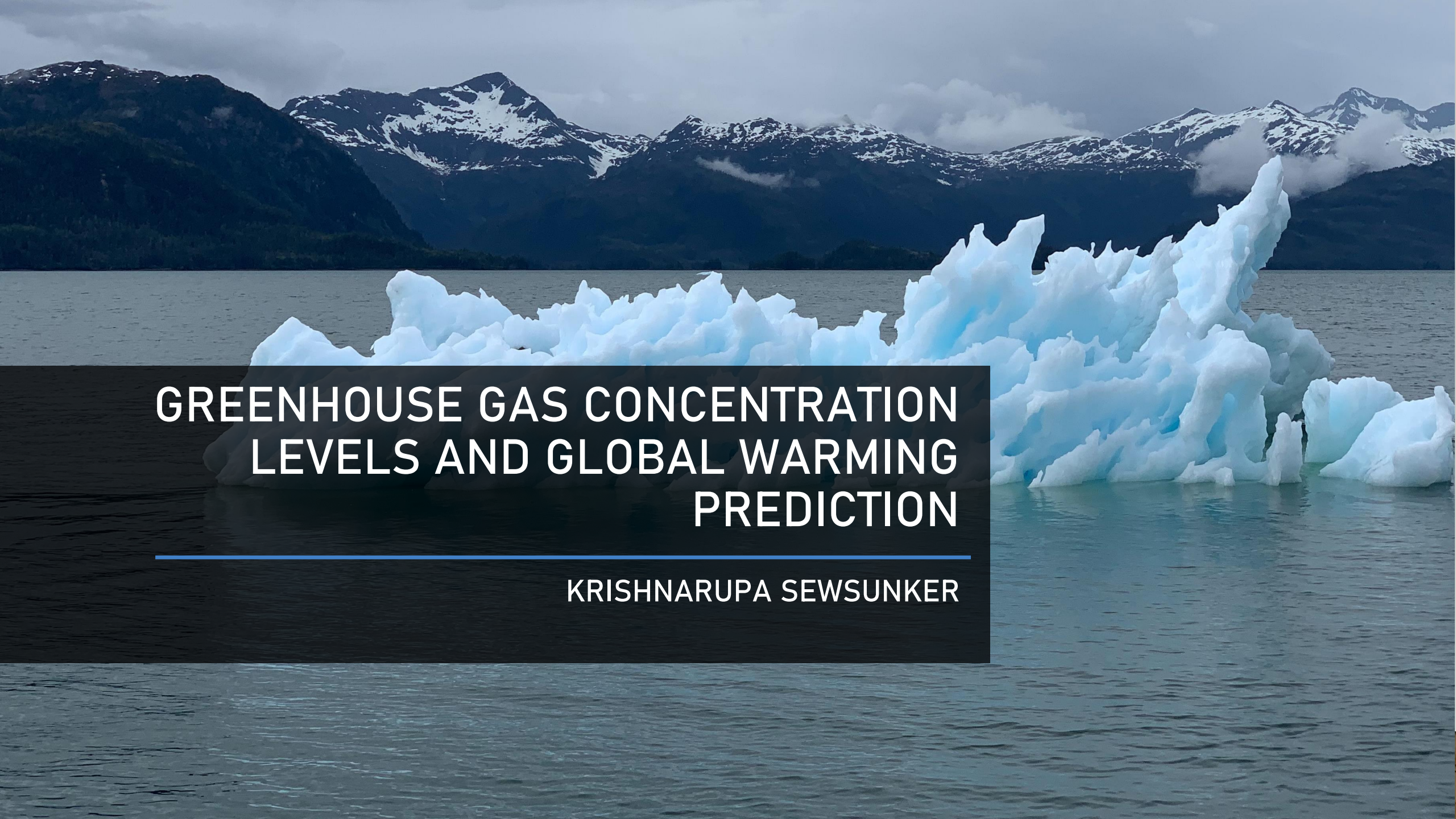
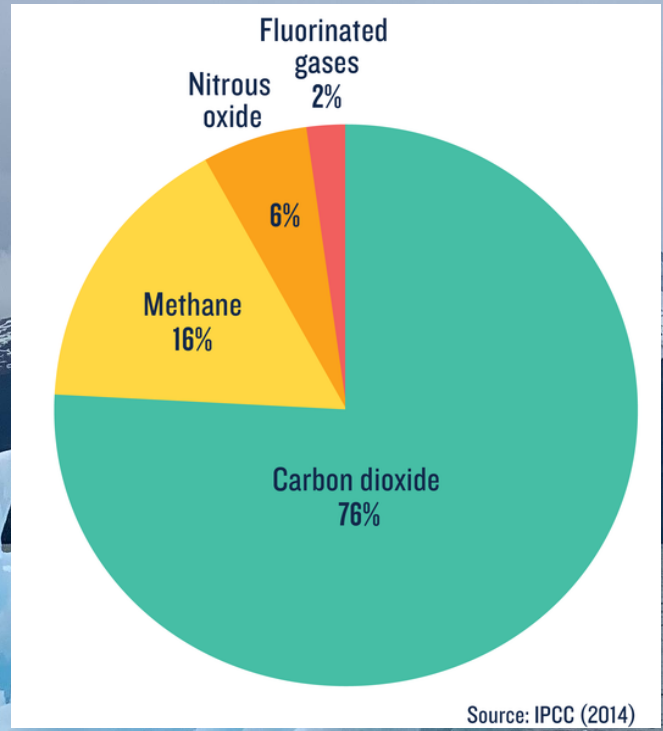


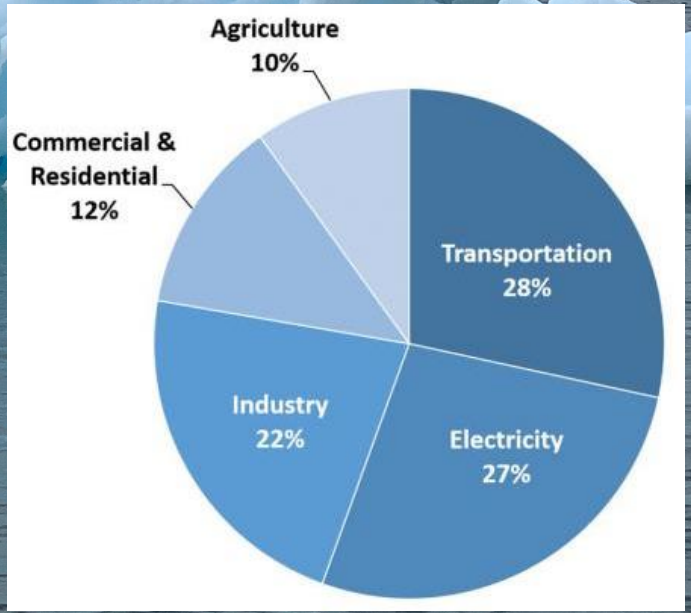
GREENHOUSE GAS CONCENTRATION LEVELS AND GLOBAL WARMING PREDICTION

KRISHNARUPA SEWSUNKER

GREENHOUSE GAS CONCENTRATION LEVELS AND GLOBAL WARMING PREDICTION: **TOPIC AND OBJECTIVE**

- As we observe more and more instances of environmental tragedies occurring globally – from wildfires to glaciers and ice melting to species extinction – we continue to become increasingly aware of how closely and directly climate change is starting to impact our lives.
- Repercussions and symptoms of climate change manifest in diverse ways. Similarly, the causes contributing to climate change are also numerous.
- However, some causes are unquestionably more serious and detrimental. The most severe and liable cause is the increasing concentration levels of greenhouse gases in the atmosphere.
- Carbon dioxide. Methane. Nitrous oxide.
- This problem is anthropogenic: generated by humans, with the largest three reasons for greenhouse gas emission being the burning of fossil fuels for transportation, electricity and industry, respectively.
- Objective: create a model that can use yearly data about the atmospheric concentration levels of the three main greenhouse gases and the average global surface temperature over a large timespan to predict how greenhouse gas concentration levels contribute to global warming. Such a model could also predict the direction we are heading in in the years to come by extrapolating the trends extracted from the large time period during which data was collected.





GREENHOUSE GAS CONCENTRATION LEVELS AND GLOBAL WARMING PREDICTION: **DESCRIPTION OF DATA**

- So far, 5 features used, all of which were numerical features.
  - Year
  - Atmospheric Concentration Level of CO_2
  - Atmospheric Concentration Level of CH_4
  - Atmospheric Concentration Level of N_2O
  - Average Global Surface Temperature
- Consideration: Include one more feature – Annual Rise in Sea Levels
- Number of samples: 1750 to present, annually
- Compiled data set manually, using separate data sets for each features
- Preprocessing steps were very involved. Several missing values – not every year was included in the three data sets of greenhouse gas concentration levels.
- Filling the missing values with the mean value for each feature did not make sense in this case, because most of the missing values were for earlier years (18^{th} century) and the concentration levels were substantially lower at that time than in the 20^{th} and 21^{st} centuries, where data was consistent.
- Other steps in preprocessing: adjusting the units across various data sets with differing units, dropping rows from data frames that had more years than others, dropping columns, and merging the sets into one data frame.

GREENHOUSE GAS CONCENTRATION LEVELS AND GLOBAL WARMING PREDICTION: **DESCRIPTION OF MODEL**

- Supervised Learning
  - For all samples in the data sets, we know all features and ground truth values
- Linear Regression Model used to make predictions
  - We would like to make predictions on the continuous dependent variable which is a numerical feature, namely, global average temperatures.
- Splitting Data set: Train:test = 85:15

```
In [98]:
# DATAFRAME
df = pd.merge(co2_ch4_n2o, temps)
df[:10]
```

Out[98]:

	Year	CO2 concentrations (NOAA, 2018)	CH4 concentration (EEA & NOAA (2019))	N2O concentrations (annual average) (EEA, 2019)	Global Average Temperature
0	1780	273.07	734.57	271.80	13.42
1	1825	281.36	774.54	274.20	13.46
2	1845	283.18	795.24	275.16	13.58
3	1850	284.00	801.42	275.40	13.63
4	1855	285.57	808.50	275.90	13.63
5	1870	287.16	831.18	277.40	13.70
6	1880	287.77	847.48	278.20	13.73
7	1890	290.92	867.22	279.10	13.75
8	1900	294.22	890.30	279.80	13.74
9	1905	299.02	912.07	280.30	13.74

GREENHOUSE GAS CONCENTRATION LEVELS AND GLOBAL WARMING PREDICTION: **DESCRIPTION OF RESULTS**

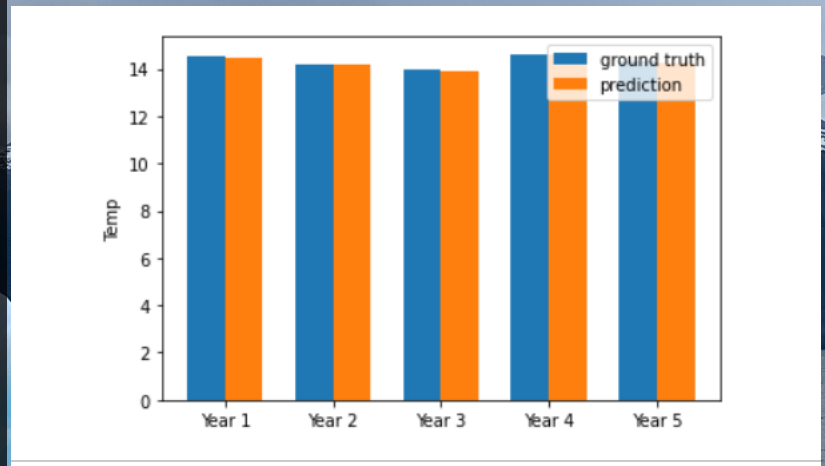
- This model can accurately make predictions. Ridge Regression will also be attempted to refine the model. Further graphing, analysis and interpretation to be carried out.
- MAE, MSE, RMSE for Training Set showed high accuracy

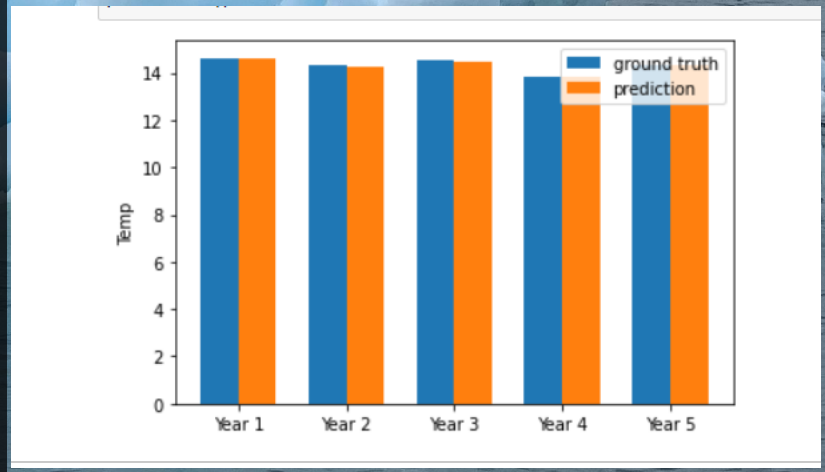
```
Bias is 14.175686274509802
Coefficients  are [ 0.13559486 -0.28334487 -0.21099596  0.67687401]
Prediction for training set:
MAE is: 0.03860427847474723
MSE is: 0.002466409549806648
RMSE is: 0.04966295953531815
```

- MAE, MSE, RMSE for Testing Set also showed high accuracy

```
Prediction for testing set:
MAE is: 0.032891897111460365
MSE is: 0.0013758896839767862
RMSE is: 0.03709298699183966
```

- The graphs on the right show the comparison between ground truth and predictions for two different 5-year samples from the testing set.





REFERENCES:

1. European Environmental Agency:

   https://www.eea.europa.eu/data-and-maps/indicators/atmospheric-greenhouse-gas-concentrations-7/assessment

2. Our World in Data:

   https://ourworldindata.org/atmospheric-concentrations

3. United States Environmental Protection Agency:

   https://www.epa.gov/climate-indicators/climate-change-indicators-atmospheric-concentrations-greenhouse-gases#ref5