GHGRP - EPA

Trends and Prediction

GHGRP

- Requires reporting of greenhouse gas (GHG) data and other relevant information
  - large GHG emission sources,
  - fuel and industrial gas suppliers,
  - and CO_2 injection sites in the United States.

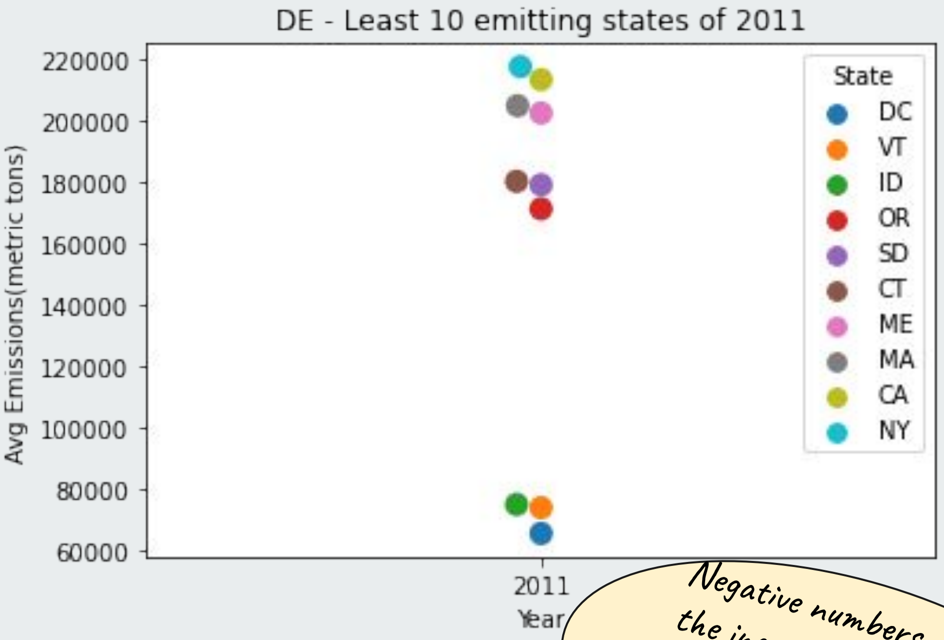
- Approximately 8,000 facilities are required to report their emissions annually

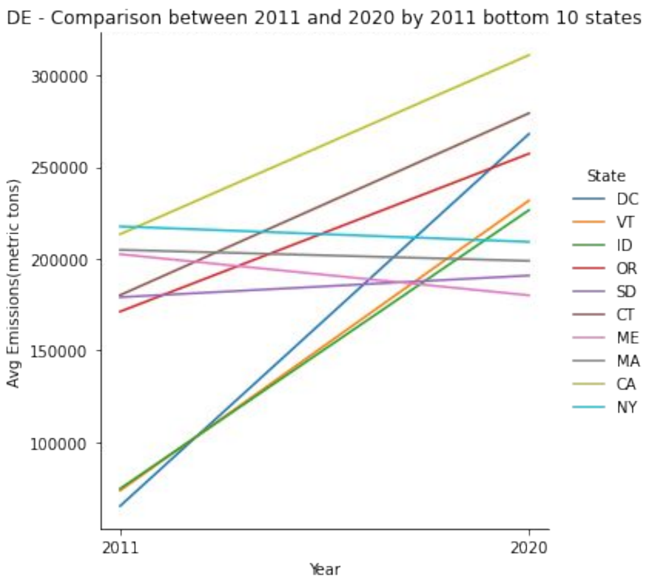
Data

- Facilities in the data we used come from the the following sectors
    - Direct emitters: the facilities that combusts fuels or otherwise put GHG into the atmosphere directly
    - SF6 - Sulphur Hexa Flouride emissions (SF6 is a highly non-decomposable gas which is thousands of times more powerful than CO_2 in holding the radiation in the atmosphere.)
    - Direct Emissions by Local distribution companies of natural gases.
    - Emission by Onshore Oil and Gas production.

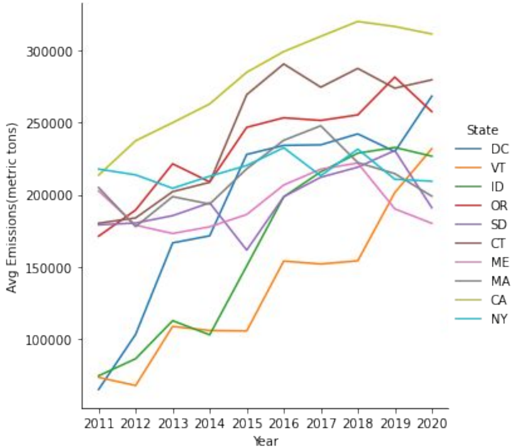
Data Wrangling

- The data in the excel sheet is imported and merged into dataframes.
- The column names are modified.
- Grouped the emissions by facilities into the State.
- Replaced the missing values by the mean emissions of that year.



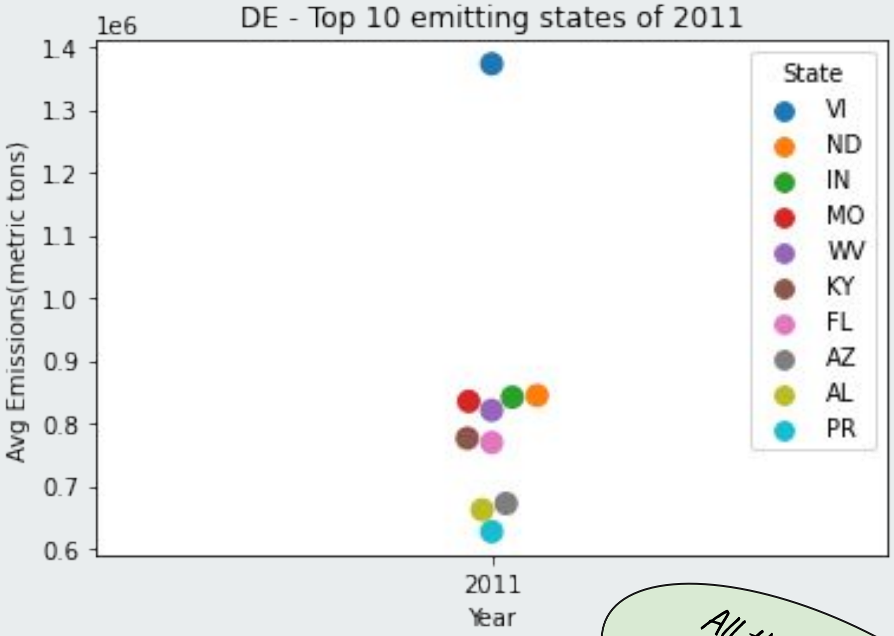


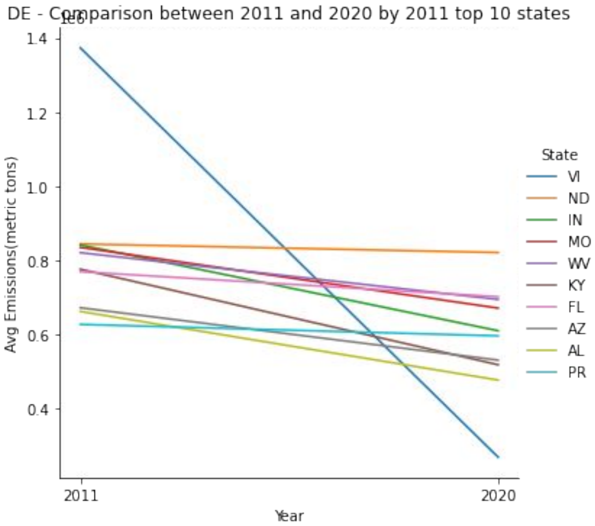
Negative numbers showing the increase in 2020 emissions. DC increased its emissions from 65K to 268K.



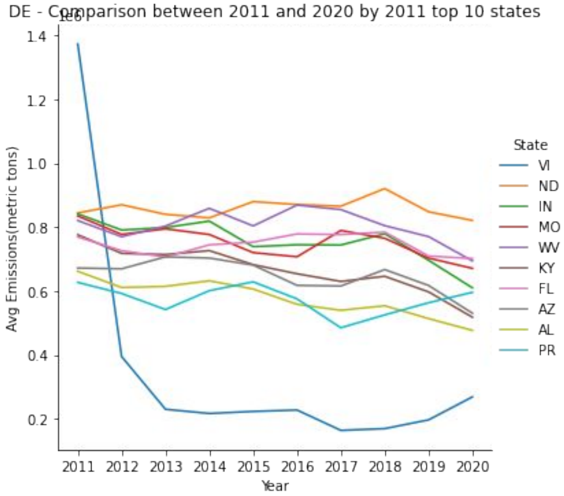
ME	11.06
NY	3.86
MA	2.93
SD	-6.6
CA	-45.76

OR	-50.26
CT	-55.02
ID	-203.27
VT	-214.72
DC	-310.95

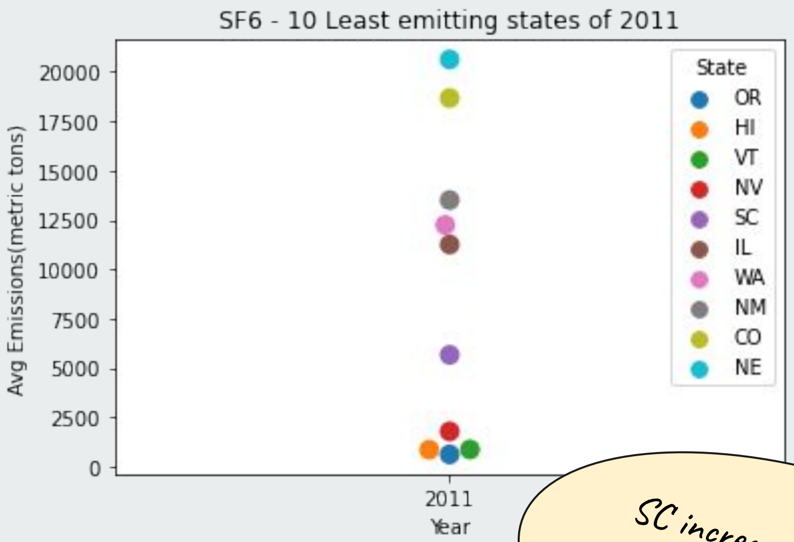


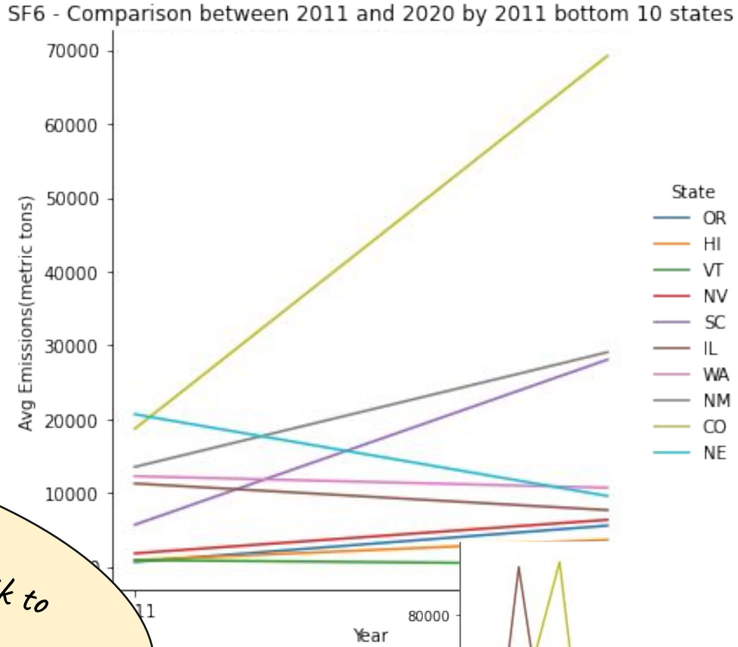


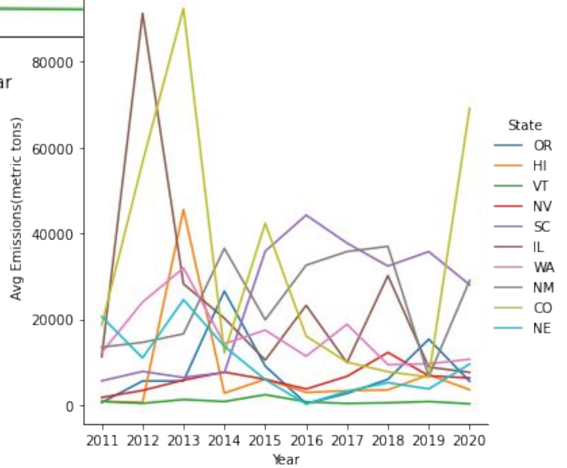
All the 10 states have reduced their emissions noticeably.



VI	80.45	MO	19.64
KY	33.25	WV	15.4
AL	27.99	FL	8.79
IN	27.53	PR	4.95
AZ	21.05	ND	2.75

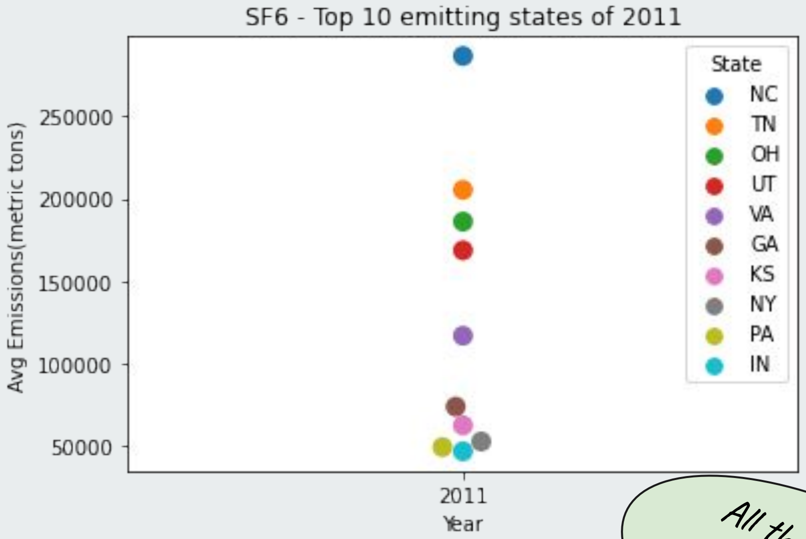


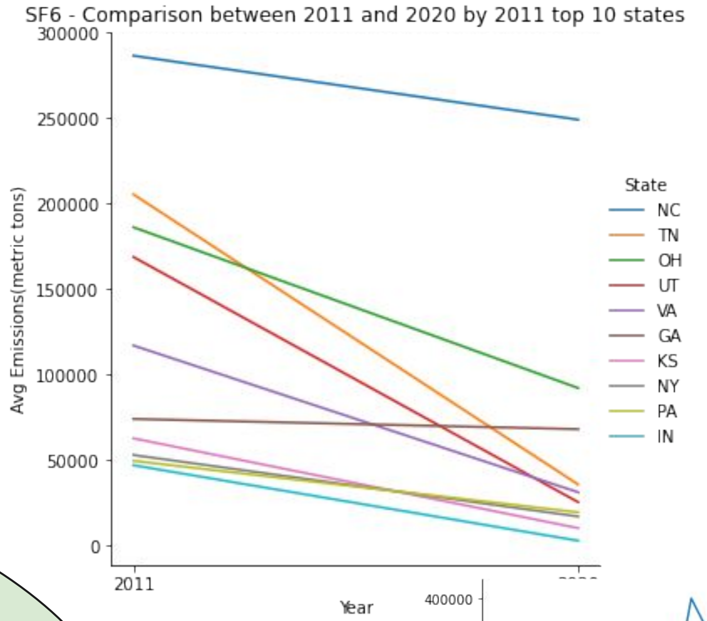


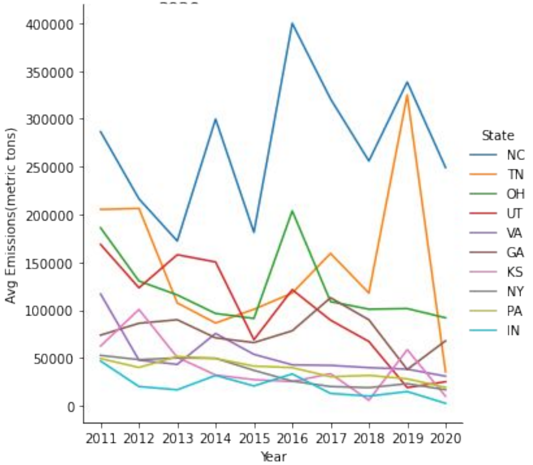


SC increased from 5.5k to 28K
and OR from 0.5K to 5.5K.

VT	63.35	NV	-254.31
NE	53.65	CO	-270.73
IL	32.06	HI	-319.95
WA	12.79	SC	-395.25
NM	-115.13	OR	-793.01



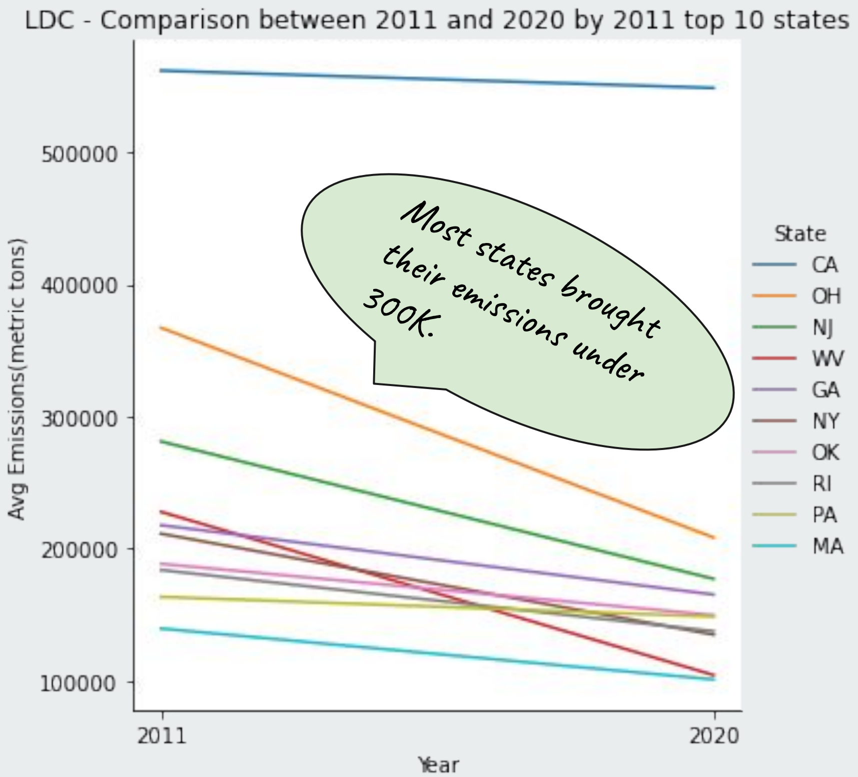




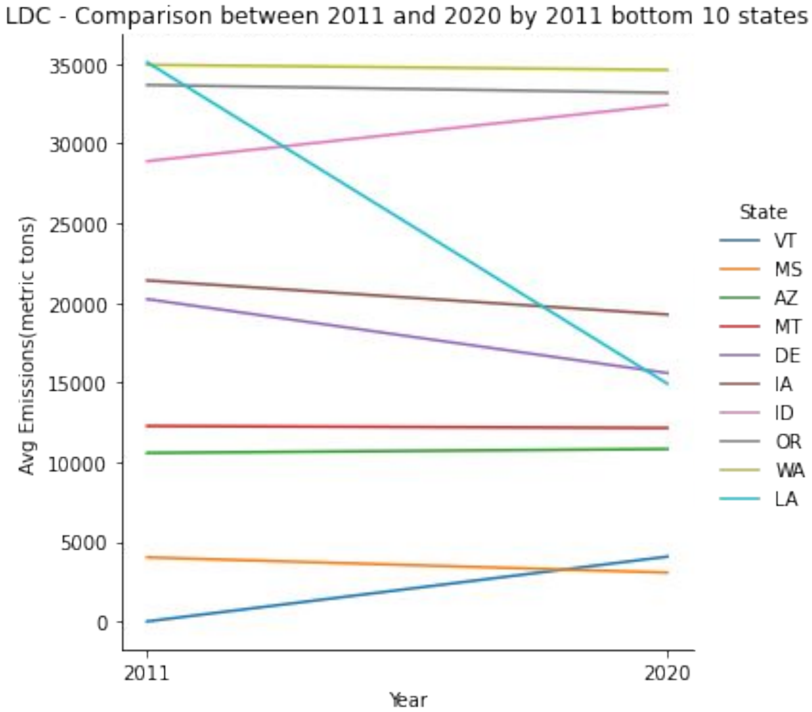
All the 10 states have reduced their emissions.

IN	94.41
UT	85.06
KS	84.01
TN	82.69
VA	73.5

NY	68.15
PA	61.01
OH	50.56
NC	13.06
GA	8.1

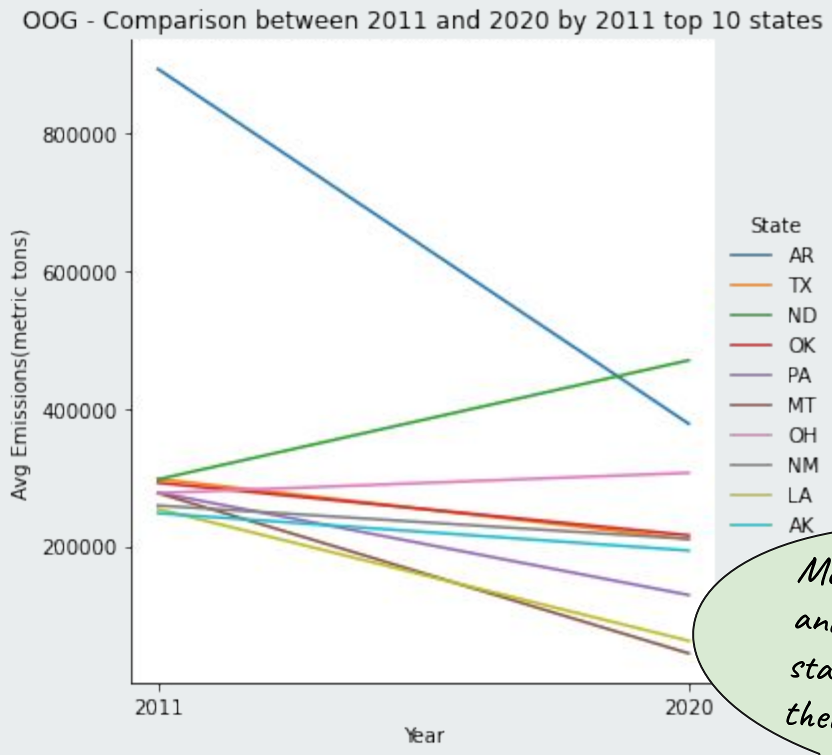


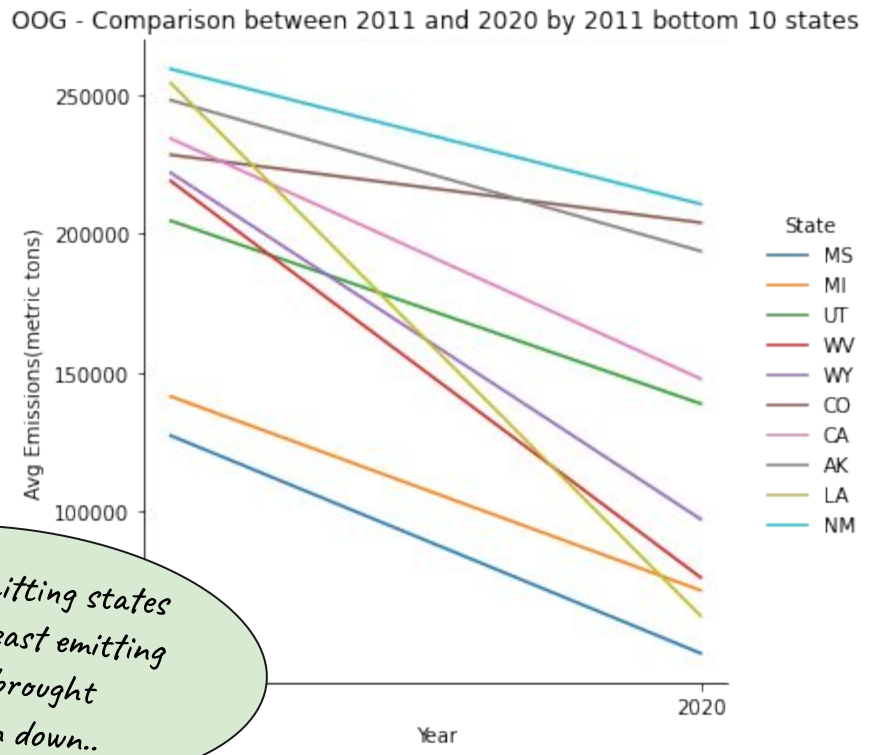
WV	54.23	RI	25.21
OH	43.25	GA	24.12
NJ	37	OK	20.61
NY	36.17	PA	9.13
MA	27.63	CA	2.34



LA	57.47
MS	23.79
DE	22.93
IA	9.99

OR	1.45
MT	1.03
WA	0.91
AZ	-2.32
ID	-12.24





Most top emitting states and all the least emitting states have brought their emission down..

MT	84.06
LA	75.41
AR	57.71
PA	53.59
TX	29.38

OK	26.14
AK	21.89
NM	18.74
OH	-10.58
ND	-58.27

LA	75.41
WV	65.14
MS	61.47
WY	56.17
MI	49.31

CA	36.91
UT	32.16
AK	21.89
NM	18.74
CO	10.68

Model Selection

- We have modelled the data with the following regression techniques:
  - OLS
  - Random Forest
  - Multi-output regressors

- Though the data is about the emissions overtime, we didn’t choose time series modelling since all we had was yearly data which wasn’t enough for a evaluating a time series model.

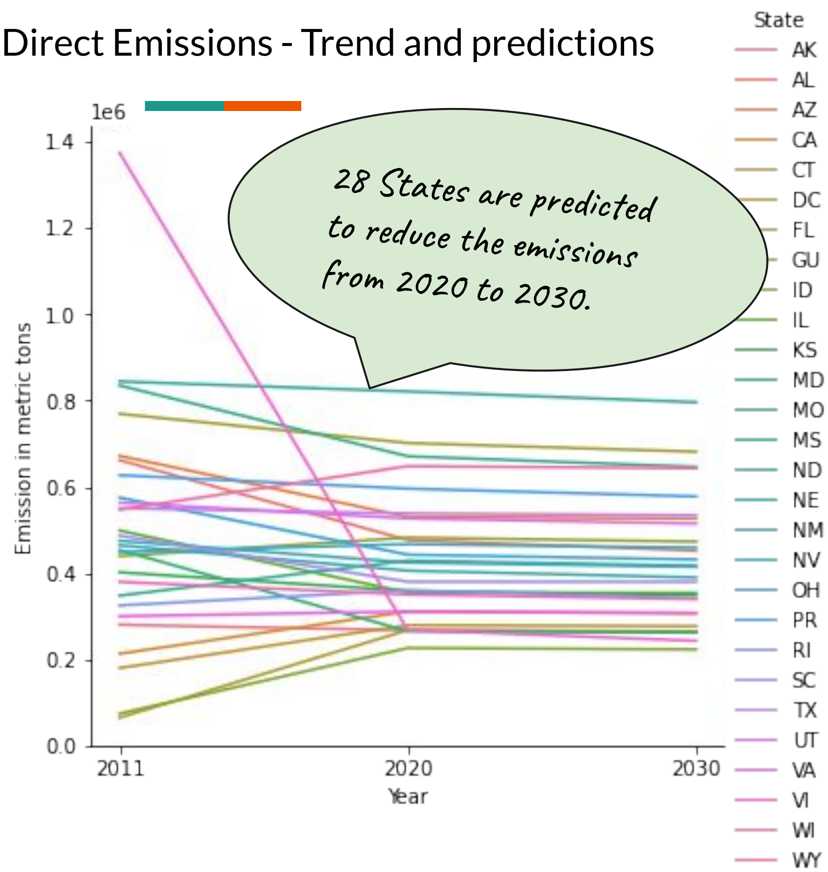
Model Metrics

Random Forest regressor and Multioutput regressors work well with the data.

We are choosing Random Forest for its good R^2 score and its comparatively reasonable fit time.

Model Name	Hyperparameters	R^2 Score	Fit Time
OLS LinMopear Regression DE		0.957	0.015
SF6		0.815	0.0009
OOG		0.775	0.0009
LDC		0.999	0.0009
RandomForest DE	n_estimators=1	0.999	0.090
SF6	1	0.999	0.005
OOG	1	0.999	0.006
LDC	1	0.999	0.005
Multioutput Ridge Regressor DE		1.0	0.018
SF6		0.999	0.008
OOG		0.999	0.007
LDC		0.999	0.008
Multioutput RF Regressor DE	n_estimators =1	1.0	0.52
SF6	1	0.999	0.05
OOG	1	0.999	0.07
LDC	1	0.999	0.06

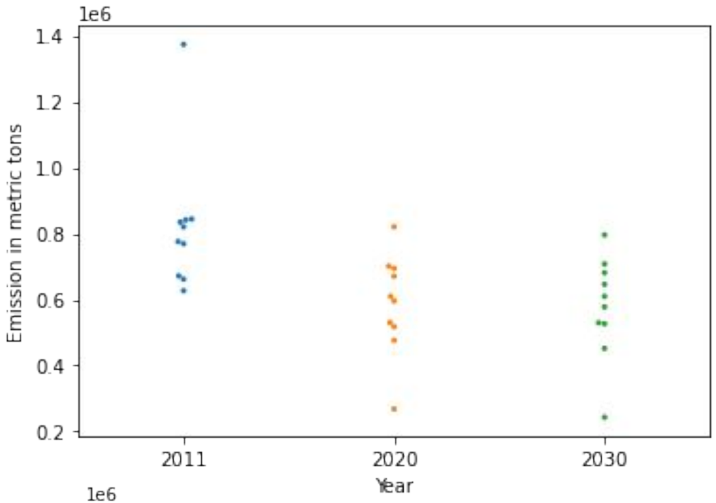
Direct Emissions - Trend and predictions

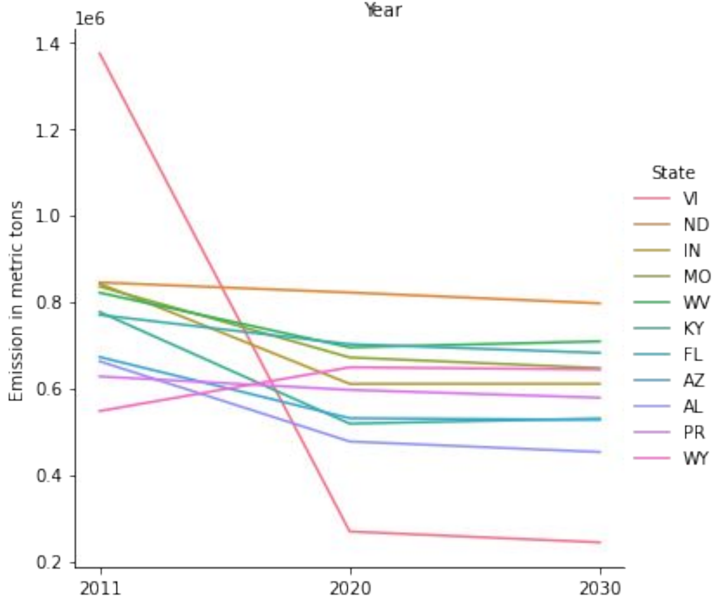


State	2011	2020	2011-20 %	2030	2020-30 %
VI	1373102.81	268431.88	80.45	243288.56	9.37
ND	844030.38	820802.66	2.75	795841.98	3.04
MO	834667.83	670748.27	19.64	646079.32	3.68
FL	769072.93	701441.68	8.79	681291.05	2.87
AZ	671738.22	530331.26	21.05	526634.34	0.7
AL	661730.26	476501.12	27.99	452171.8	5.11
PR	626840.88	595794.72	4.95	577679.48	3.04
OH	574821.24	442833.76	22.96	431610.36	2.53
UT	563220.18	526601.26	6.5	514973.74	2.21
TX	550274.19	538257.61	2.18	534233.43	0.75
WY	546886.84	647975.93	-18.48	642937.32	0.78
IL	498644.52	354906.1	28.83	354296.16	0.17
SC	486913.89	380110.29	21.93	380029.54	0.02
NM	474709.48	424592.14	10.56	414776.76	2.31
NE	464532.46	405778.85	12.65	389952.31	3.9
MD	453659.91	265426.93	41.49	262546.22	1.09
NV	449038.19	467265.3	-4.06	458919.8	1.79
GU	439623.7	482314.26	-9.71	473009.09	1.93
KS	401965.06	358883.28	10.72	349116.05	2.72
WI	379853.11	350525.39	7.72	339981.29	3.01
MS	347979.28	429031.69	-23.29	417038.02	2.8
RI	325259.05	360364.17	-10.79	341333.56	5.28
VA	300195.82	311251.98	-3.68	307567.36	1.18
AK	280444.82	267648.95	4.56	263254.74	1.64
CA	213467.89	311151.24	-45.76	307080.25	1.31
CT	180264.33	279453.64	-55.02	277167.81	0.82
ID	74725.3	226616.03	-203.27	223423.76	1.41
DC	65255.53	268166.83	-310.95	265165.12	1.12

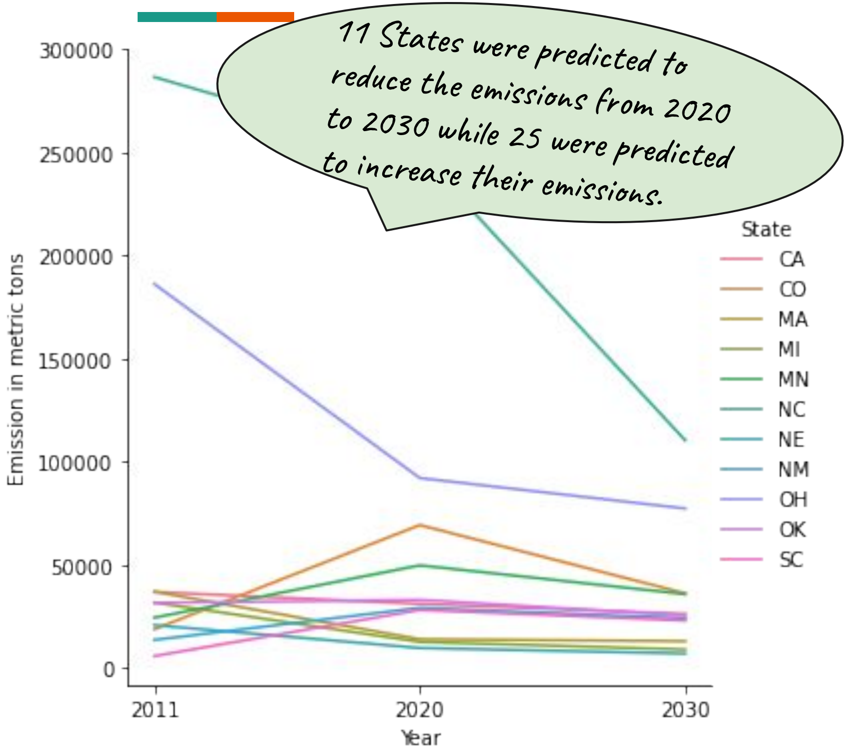
Direct Emissions - Story of top 10 emitting states of 2011

State	2011	2011-20%	2020	2030	2020-30%
VI	1373102.81	80.45	268431.88	243288.56	9.37
ND	844030.38	2.75	820802.66	795841.98	3.04
IN	841352.87	27.53	609700.12	609762.16	-0.01
MO	834667.83	19.64	670748.27	646079.32	3.68
WV	820435.24	15.4	694078.72	707939.13	-2
KY	775796.1	33.25	517822.53	529619.94	-2.28
FL	769072.93	8.79	701441.68	681291.05	2.87
AZ	671738.22	21.05	530331.26	526634.34	0.7
AL	661730.26	27.99	476501.12	452171.8	5.11
PR	626840.88	4.95	595794.72	577679.48	3.04





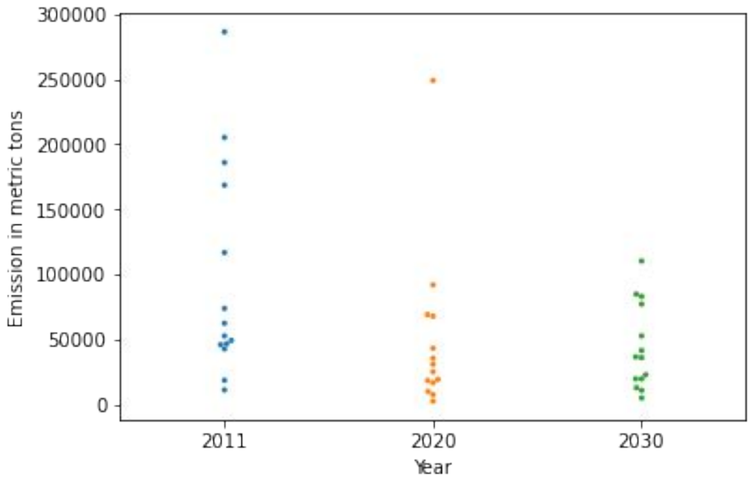
SF6 - Trend and predictions

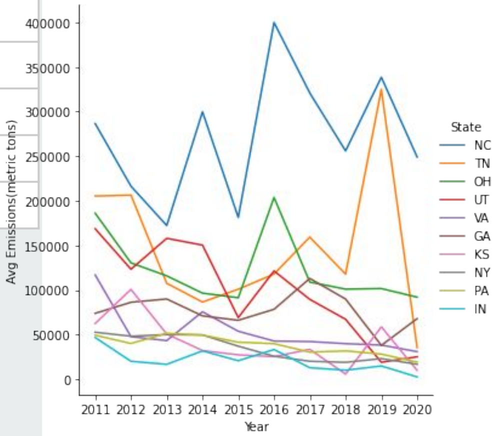


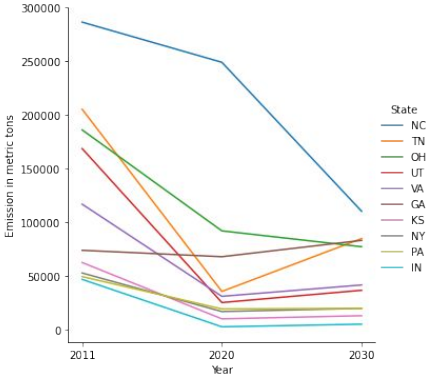
State	2011	2020	2011-20 %	2030	2020-30 %
NC	286543.56	249121.92	13.06	110320.03	55.72
OH	186095.12	92007.88	50.56	77240.16	16.05
MA	36890.4	13816.8	62.55	12875.07	6.82
CA	36746.84	30986.29	15.68	26105.54	15.75
MI	31629.4	12520.05	60.42	8850.38	29.31
OK	31493.64	32902.68	-4.47	25395.53	22.82
MN	24282.91	49673.79	-104.56	35693.69	28.14
NE	20627.16	9560.04	53.65	6895.32	27.87
CO	18670.92	69218.52	-270.73	36081.1	47.87
NM	13502.16	29047.2	-115.13	24046.73	17.21
SC	5661.24	28037.16	-395.25	22892.89	18.35

SF6 Emissions - Story of most emitting states of year 2011

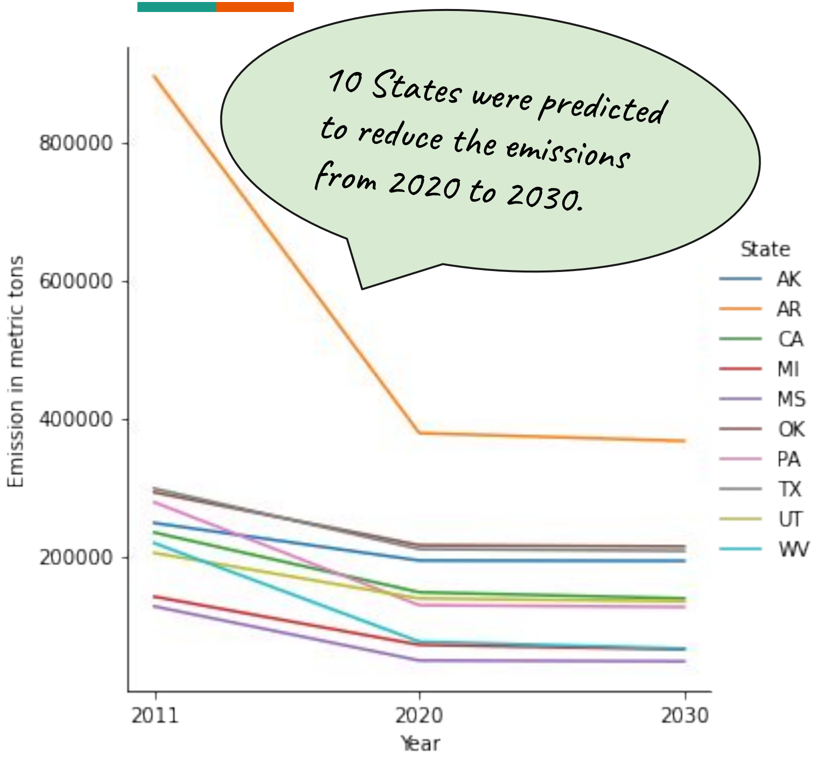
State	2011	2011-20 %	2020	2030	2020-30 %
NC	286543.56	13.06	249121.92	110320.03	55.72
TN	205327.68	82.69	35547.48	84838.91	-138.66
OH	186095.12	50.56	92007.88	77240.16	16.05
UT	168672.12	85.06	25203.12	36631.06	-45.34
VA	116872.8	73.5	30973.8	41595.3	-34.29
GA	73897.08	8.1	67909.8	83151.62	-22.44
KS	62515.32	84.01	9993.24	12934.44	-29.43
NY	52750.15	68.15	16802.08	19808.91	-17.9
PA	49367.8	61.01	19250.9	19798.52	-2.84
IN	46781.94	94.41	2616.68	5059.63	-93.36







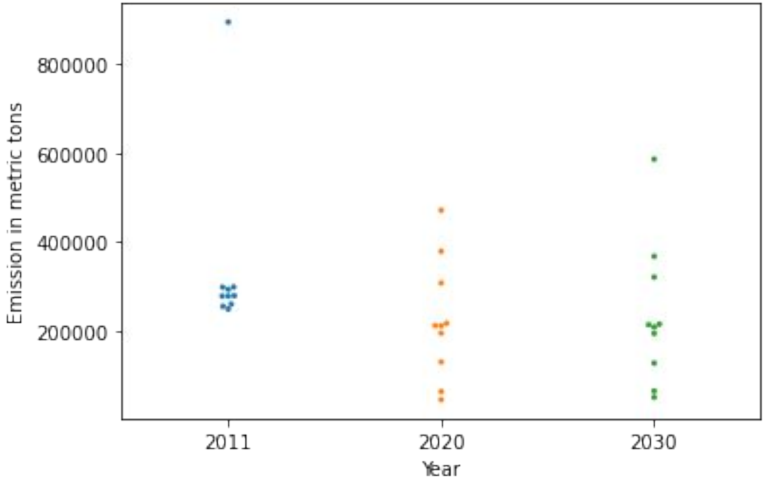
OOG - Trend and predictions

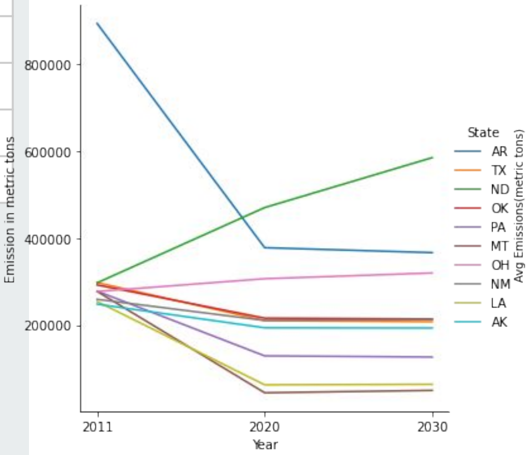


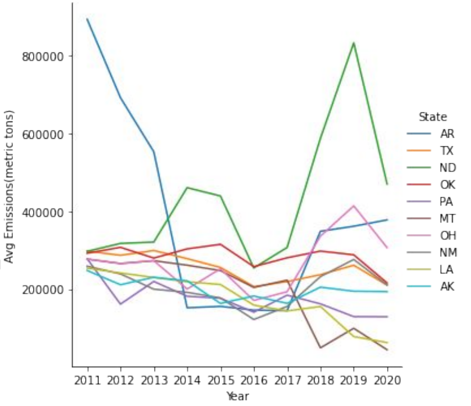
State	2011	2020	2011-20 %	2030	2020-30 %
AR	894210.02	378190.31	57.71	366835.27	3
TX	297923.84	210382.77	29.38	207723.43	1.26
OK	292430.26	215994.58	26.14	214101.09	0.88
PA	278211.23	129120.73	53.59	126544.96	1.99
AK	248107.55	193788.34	21.89	193291.3	0.26
CA	234308.41	147824.75	36.91	138949.74	6
WV	219026.37	76350.07	65.14	66412.34	13.02
UT	204738.64	138898.7	32.16	134812.19	2.94
MI	141544.66	71751.19	49.31	65173.06	9.17
MS	127445.07	49101.21	61.47	47941.63	2.36

OOG Emissions - Story of most emitting states of year 2011

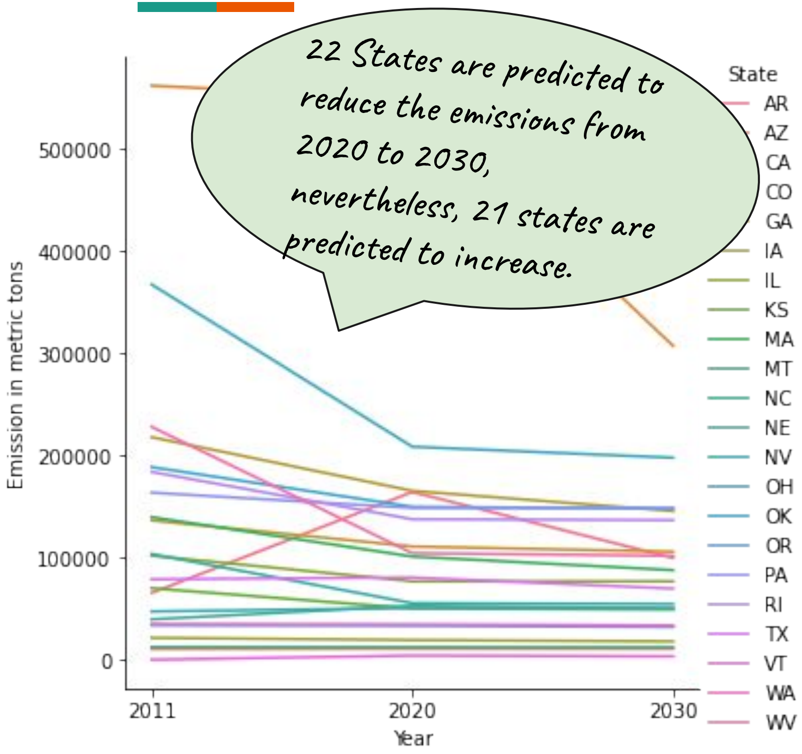
State	2011	2011-20 %	2020	2030	2020-30 %
NC	286543.56	13.06	249121.92	110320.03	55.72
TN	205327.68	82.69	35547.48	84838.91	-138.66
OH	186095.12	50.56	92007.88	77240.16	16.05
UT	168672.12	85.06	25203.12	36631.06	-45.34
VA	116872.8	73.5	30973.8	41595.3	-34.29
GA	73897.08	8.1	67909.8	83151.62	-22.44
KS	62515.32	84.01	9993.24	12934.44	-29.43
NY	52750.15	68.15	16802.08	19808.91	-17.9
PA	49367.8	61.01	19250.9	19798.52	-2.84
IN	46781.94	94.41	2616.68	5059.63	-93.36







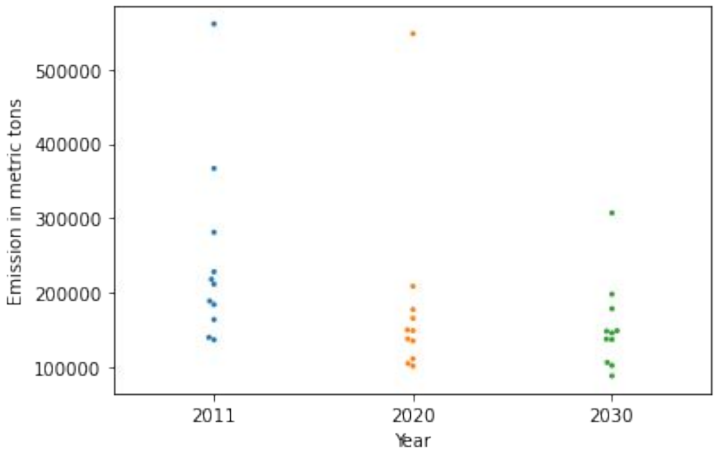
LDC- Trend and predictions

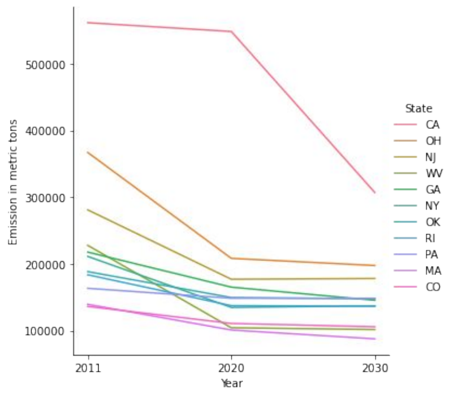


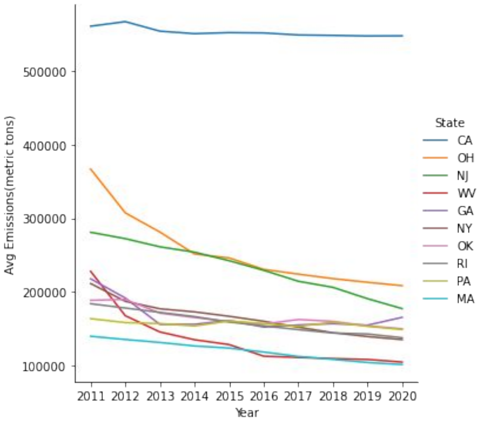
State	2011	2020	2011-20 %	2030	2020-30 %
CA	561602.05	548482.12	2.34	307037.12	44.02
OH	367033.67	208308.8	43.25	197585.61	5.15
WV	227769.1	104260.42	54.23	101620.8	2.53
GA	217658.9	165161.15	24.12	145443.03	11.94
OK	188410.3	149571.85	20.61	147715.47	1.24
RI	183717.65	137393.4	25.21	136570.7	0.6
PA	163401.57	148489.66	9.13	148404.41	0.06
MA	139428.91	100909.54	27.63	87619.09	13.17
CO	136264.5	110734.68	18.74	105701.96	4.54
NE	103434.27	55295.3	46.54	54419.01	1.58
IL	101732.19	76699.79	24.61	76680.7	0.02
TX	78728.03	80444.26	-2.18	69595.55	13.49
KS	69939.12	50142.35	28.31	49052.75	2.17
AR	65322.55	164066.75	-151.16	99574.05	39.31
NV	47225.32	50544.97	-7.03	49894.07	1.29
NC	39449.78	52776.63	-33.78	50279.16	4.73
WA	34945.1	34626	0.91	33580.68	3.02
OR	33678.45	33190.2	1.45	32072.03	3.37
IA	21407.83	19270.14	9.99	17922.97	6.99
MT	12271.47	12145.13	1.03	12013.91	1.08
AZ	10580.48	10825.92	-2.32	10790.51	0.33
VT	0	4067.15	-inf	3368.73	17.17

LDC Emissions - Trend and predictions

State	2011	2011-20 %	2020	2030	2020-30 %
CA	561602.05	2.34	548482.12	307037.12	44.02
OH	367033.67	43.25	208308.8	197585.61	5.15
NJ	280975.98	37	177008.61	178122.09	-0.63
WV	227769.1	54.23	104260.42	101620.8	2.53
GA	217658.9	24.12	165161.15	145443.03	11.94
NY	211198.61	36.17	134802.81	137162.13	-1.75
OK	188410.3	20.61	149571.85	147715.47	1.24
RI	183717.65	25.21	137393.4	136570.7	0.6
PA	163401.57	9.13	148489.66	148404.41	0.06
MA	139428.91	27.63	100909.54	87619.09	13.17
CO	136264.5	18.74	110734.68	105701.96	4.54







Conclusion and Recommendations

- The trends and predictions made are based on the available data and how the emissions are handled in the decade 2011 - 2020.

- They show that many states are heading positively towards reducing the emissions.

- Also. with the advanced techniques and equipments that will be available in the near future, the trend can be changed and so the emissions can be brought down more rapidly.

Further Study

- So far, the study is based on State level. We can further drill down the investigation to industrial sectors and the cities in each state.
- Also, we can collect data related to GHG emissions from other resources to find the reasons on studied emission levels in the facilities and find ways to bring them down.

Data Citation

U.S. Environmental Protection Agency Office of Atmospheric Programs Greenhouse Gas Reporting Program (GHGRP) [Compressed file contains a multi-year data summary spreadsheet containing the most important, high-level information for facilities, as well as yearly spreadsheets containing slightly more detailed information than the multi-year summary, including reported emissions by greenhouse gas and process.] Available at https://www.epa.gov/ghgreporting/data-sets
Date accessed: [September, 2021]