& Electric Vehicles Global Emissions



TEAM GREEN WAVE



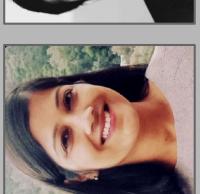
























Tarini Pal

Avneesh Sawhney

Ruchita Agarwal

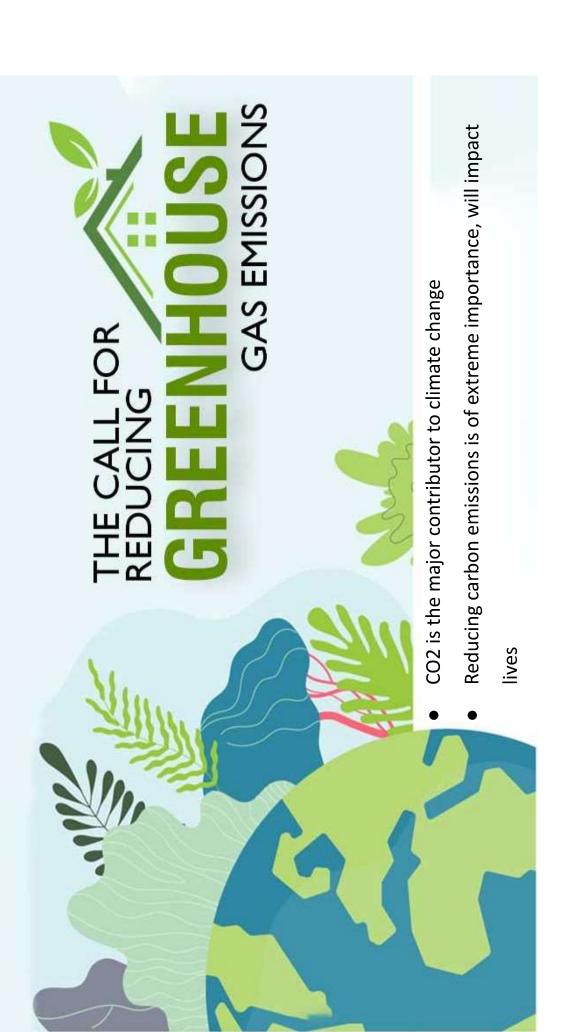
Kobe Pho

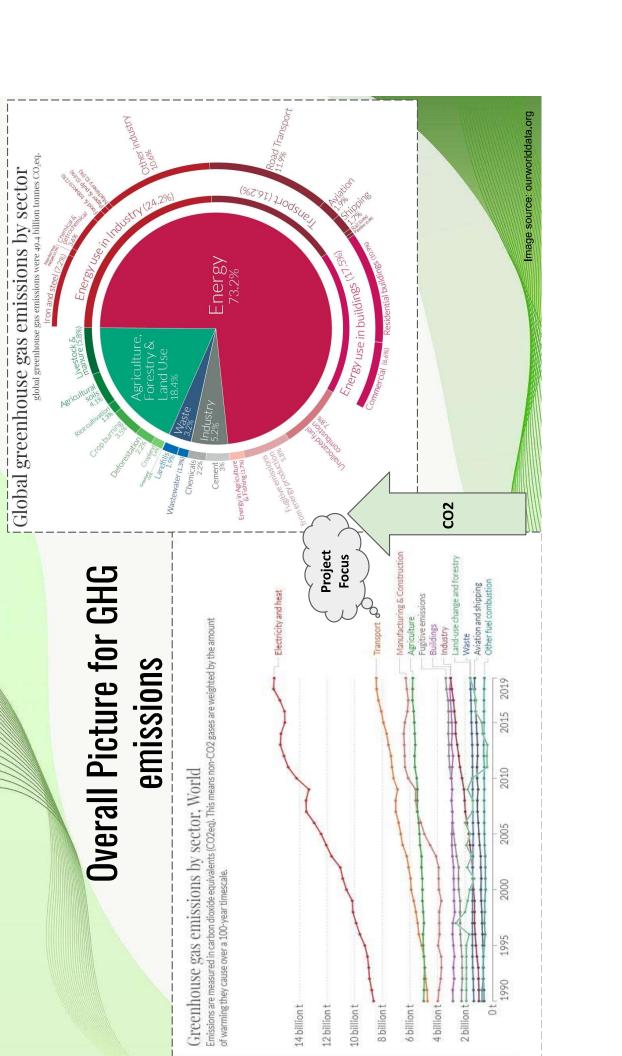
Agenda

- Project Objectives
- Understanding Carbon/GHG emissions and Leading Factors
- Data Collection and Preprocessing
- Data Analysis and Outcomes
- Machine Learning Design & Dashboard

Main Objectives

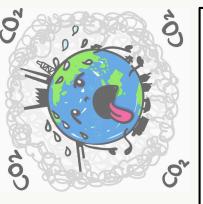
- Analyzing what factors impact carbon emissions globally
- Finding correlations between emission factors
- Determining if EV introduction results in significant positive impacts on reducing carbon transportation emissions
- Corroborating our hypothesis with a peer research review on use of EV and its impact in Norway



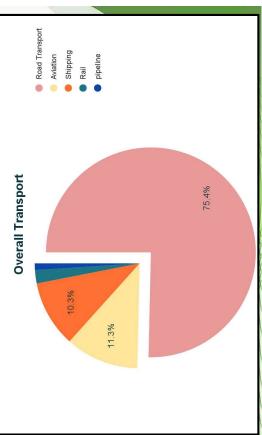


Impacts of Road Transportation

- Majority of global GHG emissions within transport industry come from road vehicles
- For each gallon of gas that a car burns, it releases about 19 pounds of carbon dioxide through tailpipe emissions







Data Collection / Sources

Data Description / Preparation

- Overall GHG emissions by sector
- Market demand of all vehicles types (Diesel, Electric, Hybrid, etc)
- Merge multiple datasets based on country and year as join keys

Data Source:

https://ourworldindata.org/

https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

https://www.iea.org/articles/global-ev-data-explorer

Database: pgAdmin



Tools:

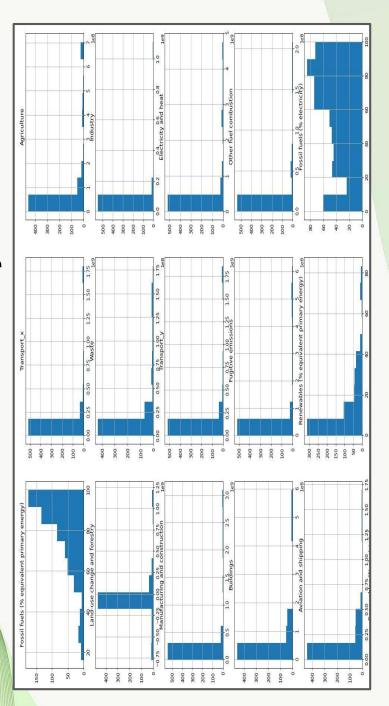
- PostgreSQL
- Scikit-learn

- pgAdmin

Data Cleaning/Wrangling

- Dropped empty values (200 Nan values)
- Data type corrections(object to int)
- Matched mapping of country names and year range (Entity, 2015-2019)
- different countries and emissions data for each As a result, our final dataset consists of 59 individual year (200+ rows)

Univariate Analysis

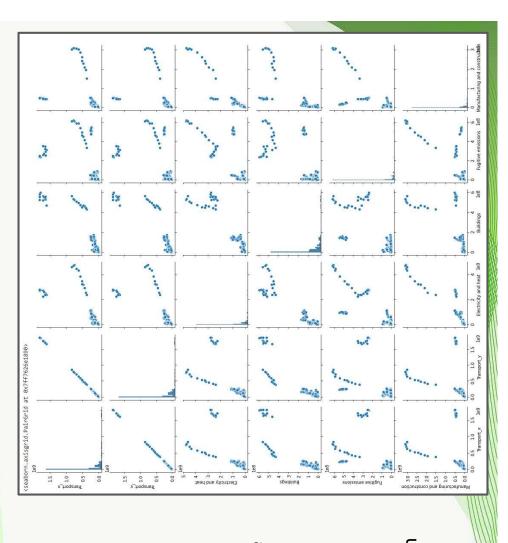


Used univariate analysis to detect and remove outliers and applied log transformation to skewed features



Multivariate Analysis

- Used multivariate analysis to detect feature significance and correlation across features
- Transport_x is only CO2 emissions while Transport_y is overall GHG emissions including CO2
 - This is evident from the scatter plots
- We can clearly see correlations between various sectors that emit greenhouse gases



Correlation Matrix

0.93 -0.79	0.13 -0.13	0.120.00043	-0.13 0.22	023 -0.14 -0.50	0.16 -0.1	0.33.0.0950.15-0.091	013 0.13	02 -0.16 -0.00	0.17 -0.18	0.19 0.13	0.16-0.093	016-021	-0.75 0.95	1 -0.77	0.75	Fossil fuels (% electricity) -
Fossil fuels (% equivalent primary energy) - 1 0.11 0.11 0.11 0.21 0.14 0.13 0.11 0.17 0.16 0.18 0.14 0.14 0.87	Tansport x - 0.11 1 0.53 -0.34 0.71 0.51 0.46 1 0.78 0.9 0.58 0.54 0.87 -0.12	Agriculture - 0,11 0.53 1 -0,031 0.76 0.68 0.7 0.53 0.72 0.67 0.54 0.81 0.38 -0,03	Land-use change and forestry - 0.11-0.34-0.031 1 0.13-0.38-0.36-0.34-0.48-0.45-0.57-0.34-0.32 0.18	Waste - 0.21 0.71 0.76 0.13 1 0.66 0.65 0.71 0.79 0.76 0.72 0.57 0.16	Industry - 014 0.51 0.68 -0.38 0.66 1. 0.99 0.51 0.92 0.77 0.74 0.96 0.38 -0.11	Manufacturing and construction - 0.13 0.46 0.7 0.36 0.65 0.99 1 0.45 0.9 0.74 0.73 0.97 0.33 0.09	Tansport y - 011 1 053 034 071 051 045 1 078 089 058 054 087 012	Electricity and heat - 017 0.78 0.72 0.48 0.79 0.92 0.9 0.78 1 0.94 0.83 0.92 0.63 0.15	Buildings - 0.16 0.9 0.67 0.45 0.79 0.77 0.74 0.89 0.94 1 0.74 0.8 0.78 0.17	Fugitive emissions - 0.18 0.58 0.54 0.57 0.76 0.74 0.73 0.58 0.83 0.74 1 0.72 0.44 0.14	Other fuel combustion - 0.14 0.54 0.81 0.34 0.72 0.96 0.97 0.54 0.92 0.8 0.72 1 0.4 -0.1	Aviation and shipping - 0.14 0.87 0.38 0.32 0.57 0.38 0.87 0.63 0.78 0.44 0.4 10.18	Renewables (% equivalent primary energy) -0.87-0.12-0.03 018 -0.16-0.11 0.095 0.12-0.15-0.17-0.14-0.1 -0.18	Fossil fuels (% electricity) - 0.93 0.13 0.12 0.13 0.23 0.16 0.15 0.13 0.2 0.17 0.19 0.16 0.15	Renewables (% electricity) -0.79-0.18 000043022 0.14 -0.1 0.091-0.13 -0.16 -0.18 -0.13-0.093-0.21 0.95	rels (% equivalent primary energy) - Tansport_x - Agriculture - Agriculture - Land-use change and forestry - Maste - Industry - Manufacturing and construction - Tansport_y - Telectricity and heat - Electricity and heat - Electricity and heat - Tombustron - Buildings - Fugitive emissions - Tombustron - Aviation and shipping - Aviation and shipping - Pother fuel combustion - Pother fuel compusion - Pother fuel combustion - Pother fuel co



Machine Learning Model Choice

Linear Regression Model

Limitations:

- Accuracy was initially low when including electric bike data
- For some of the features, standard error was really high, which required dropping columns

Benefits:

- Accuracy increased by 20%
- R-squared variance also increased by 20%

Machine Learning Design

Problem Statement:

Create a model that predicts how much CO2-eq by volume is offset by introducing EVs in a market (country+year).

Hypothesis: EV vehicles lead to reduction in overall Greenhouse gases.

X: GHG emissions by sector and energy production by fuel

types

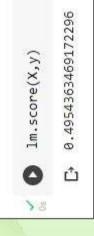
Y: CO2 offset due to EV

Target Variable(Y):

gas_vehicle_demand_percent ev_demand_percent $ev_offset = transport_co2eq_emissions*$

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 80% of the data in the training set, 10% in the validation set, **Predictions** Machine Learning Design Transport demand data Machine Learning Design, Train test split data and 10% in the test set Merged data -Year **GHG Emissions Data** Data lost while merging

Linear Regression Model Performance



Dropping Electric bike data improved performance by 10%



Train-Test-Validation Splits



Train-Test-Validation split improved performance by further 12%

0.7284700452410147

Hypothesis Testing

With a P-value threshold of 5 percent, we can see that the expected sectors are significant while predicting CO2 offset from EVs(P-value *less than* 5%)

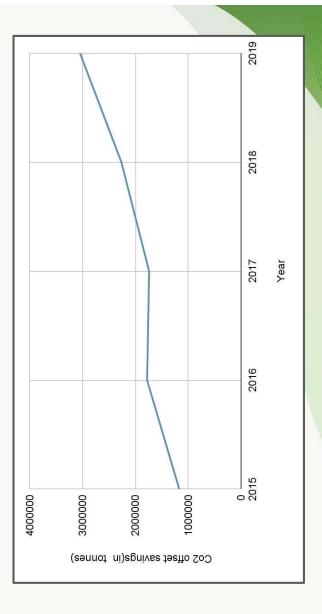
- Transport
 - Industry
- Manufacturing
- Electricity and Heat
 - Buildings
- **Fugitive Emissions**

We can also see that Primary and Secondary Energy production from Renewables have some significance (P-value *close to* 5%)

Dep. Variable: Model: Method: Date: Time: No. Observation Df Residuals: Df Model: Covariance Type Fossil fuels (% e Land-use	Dep. Variable: ev_saving Model: OLS Method: Least Squ Date: Sat, 10 Se Time: 07:01:15 No. Observations: 576 Df Residuals: 560 Df Model: 16 Covariance Type: nonrobust Transport_x Agriculture Land-use change and	ares	R-squared (uncentered): 0.606 Adj. R-squared (uncentered): 0.594 F-statistic: 53.74 Prob (F-statistic): 6.63e Log-Likelihood: -7066 AIC: 1417	entered): icentered) c:	0.606		
Mode Metho Date Time No. Observ Df Resid Df Moc Covarianc Fossil fuel Lanc	d: OLS d: Least : Sat, d: 07:01 ations: 576 uals: 560 tel: 16 a Type: nonr Transpo Agricutt -use change	ares	-squared (un F-statistic Prob (F-stati Log-Likeliho	icentered)	: 0.594		
Metho Date Time No. Observ Df Resid Df Moc Covarianc Fossil fuel Lanc	sat, sat, sat, sations: 570 uals: 560 uals: 560 tel: 16 a Type: nonro Transpo Agricutt -use change	ares p 2022	F-statistic Prob (F-stati Log-Likeliho	ö			
Date Time No. Observ Df Resid Df Moc Covarianc Fossil fuel: Lanc	sat, 1: Sat, 1: 07:01 autions: 576 uals: 560 tel: 16 a Type: nonro Transpo Agricutt -use change	p 2022	Prob (F-stati Log-Likeliho		53.74		
Time No. Observ Df Resid Df Moc Covarianc Fossil fuel: Lanc	ations: 576 uals: 560 tel: 66 e Type: nonro Transpo Agricutt	info	Log-Likeliho	stic):	6.63e-102	-102	
No. Observ Df Resid Df Moc Covariance Fossil fuel: Lanc	ations: 576 uals: 560 lel: 16 a Type: nonro Transpo Agricult I-use change	obust	AIC.	:pod:	-7066.8	∞.	
Df Mod Covariance Fossil fuel: Lanc	uals: 560 lel: 16 Pype: nonro Transpo Agricult I-use change	obust			1.417e+04	e+04	
Df Mod Covariance Fossil fuel: Lanc	e Type: nonro	obust	BIC:		1.424e+04	e+04	
Covariance Fossil fuel: Lanc	s (% equivale Transpo Agricult I-use change	phust					
Fossil fuel: Lanc	s (% equivale Transpo Agricult I-use change						
Fossil fuel: Land	s (% equivale Transpo Agricult Luse change		coef	std err	+	P> t [0.025	0.975]
Lanc	Transpo Agricult I-use change Wast	Fossil fuels (% equivalent primary energy) 171.3201		188.151 0.911		0.363 -198.248	540.888
Lanc	Agricult I-use change Wast	×_f		900'0	16.298	0.006 -16.2980.000-0.106	-0.083
Land	I-use change Wast	ure	3.493e-05 4.07e-05 0.858	4.07e-05 C		0.391 -4.51e-05 0.000	0.000
Manu	Wast	Land-use change and forestry	-1.824e-05 1.96e-05 -0.932	1.96e-05 -		0.352 -5.67e-05	2.02e-05
Manu		a	-0.0002	0.000	-0.993	0.321 -0.001	0.000
Manu	Industry	,	0.0000	0.000	1.898	4.898 (0.00) 0.001	0.001
	facturing and	Manufacturing and construction	-0.0003	7.58e-05 -	3.361	7.58e-05 -3.361 (0.00) -0.000	-0.000
	Transport_y	ת'א	0.0915	0.006	16.380	16.380 0.0000.081	0.103
	Electricity and heat	nd heat	-0.0001	4.06e-05 -	3.303	4.06e-05 -3.303 (0.001)-0.000	-5.44e-05
	Buildings	gs	900000	0.000	5.249 (0.00000000	0.001
	Fugitive emissions	issions	0.0004	5.71e-05 7.740		0.00000000	0.001
5	Other fuel combustion	nbustion	0.0004	0.001	0.765	0.444 -0.001	0.002
	Aviation and shipping	shipping	-3.848e-05 0.000		-0.179	0.858 -0.000	0.000
Renewable	s (% equivale	Renewables (% equivalent primary energy) 921.3666 473.462	1) 921.3666		1.946 (0.052-8.612	1851.346
-Fo	Fossil fuels (% electricity)	electricity)	12.0675	205.053 0.059	0.059	0.953 -390.699	414.834
Re	Renewables (% electricity)	electricity)	-675.0721	332.980 -	2.027	-675.0721 332.980 -2.027 0.043-1329.114 -21.030	-21.030
Omnibus:	s: 417.162	417.162 Durbin-Watson: 0.531	0.531				
Prob(Omni	Prob(Omnibus): 0.000	Jarque-Bera (JB): 37756.473	: 37756.473				
Skew:	2.397	Prob(JB):	0.00				
Kurtosis:	s: 42.373	Cond. No.	2.41e+08				

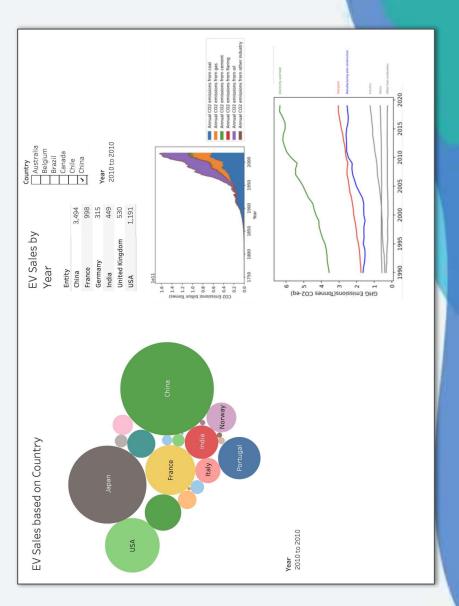
Forecasting/Predictions

- Although we do not have the y-values of the target variable(co2_offset/ev_savings) from 2015-19, we do have the input features(X values)
- We used the model to generate predictions and see that the CO2 offset due to EVs is predicted to be increasing year over year, as expected



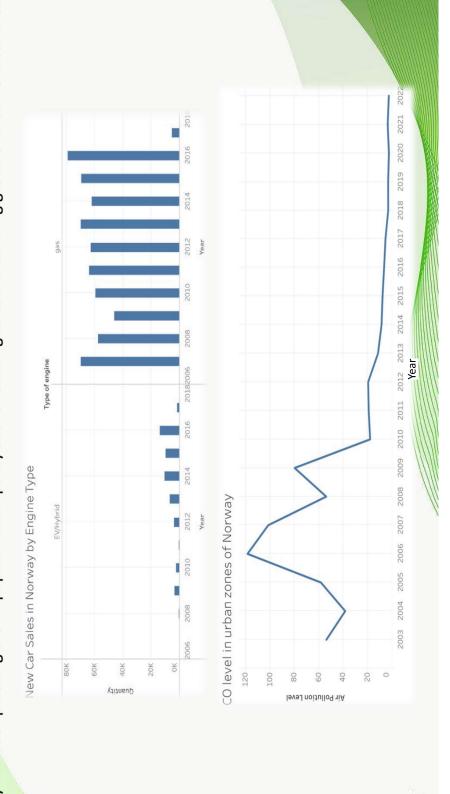


Dashboard: Tableau



Norway EV Sales: Real World Example

Norway has been pursuing the EV popularization policy such a reducing taxes and banning gas cars for some urban areas





Conclusion

- Contradicting theories around negating impact due to energy in recharging
- EV infrastructure is still at its nascent stage and evolving
- Charging your EV with renewable sources results in zero emissions for both when the car is operating and during power generation

Appendix

Database: Tables, Join, Connection String

```
ev_data
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     merged ==
                                 engine = psycopg2.connect(

    □ Tables (7)

                                                       database="postgres",
                                                                            user="postgres",
                                                                                                                                                  port='5432'
                                                           Country year" character varying COLLATE pg_catalog."default" NOT NULL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONSTRAINT temp_data_pkey PRIMARY KEY ("Country_year")
CREATE TABLE IF NOT EXISTS public.temp_data
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALTER TABLE IF EXISTS public.temp_data
                                                                                                                    '2001" double precision,
                                                                                                                                                                                '2003" double precision,
                                                                                                                                                                                                                                             '2005" double precision,
                                                                                                                                                                                                                                                                                                        '2007" double precision,
                                                                                                                                                                                                                                                                                                                                        '2008" double precision,
                                                                                                                                                                                                                                                                                                                                                                                                 '2010" double precision,
                                                                                                                                                                                                                                                                                                                                                                      '2009" double precision,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OWNER to postgres;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TABLESPACE pg_default;
```

```
host="database-ev.cirez3rkweci.us-west-1.rds.amazonaws.com",
password="qt3WxeF5SzWtQV3",
```

```
Norway_co_urban
                    ■ Norway_sales
                                                                                                                       E temperature
                                                                                                    temp_data
                                       car_sales
```

Recommendations for Future Analysis

- Time permitting we would like to integrate the CO2 emission data with the EV sales data and show that on global map
- Studies at length the assumptions around future dominating country in EV demand
- Studied Tesla's infrastructure to prove net impact of Carbon footprint due to it's energy consumed in charging EVs

Questions?