

Exploring the Determinants of Biogas Generation Potential in California

Web address for GitHub repository

Jibikeoluwa Faborode, Yinan Ding, Abhay Rao

Contents

1	Table of figures	11
1.1	Simple feature collection with 6 features and 14 fields	16
1.2	Geometry type: MULTIPOLYGON	16
1.3	Dimension: XY	16
1.4	Bounding box: xmin: -122.7851 ymin: 37.45447 xmax: -119.5423 ymax: 40.15203	16
1.5	Geodetic CRS: WGS 84	16
1.6	ObjectID NAME STATE_NAME FIPS OWCH4t AMCH4t WWTPCH4t .	16
1.7	1 184 Alameda California 06001 5700.695190 0.367359 10461.687700	16
1.8	2 185 Alpine California 06003 6.483819 0.000000 6.667176	16
1.9	3 186 Amador California 06005 156.310667 3.826180 624.136548	16
1.10	4 187 Butte California 06007 817.909133 20.460013 1210.759120	16
1.11	5 188 Calaveras California 06009 97.664105 11.663426 200.148617	16
1.12	6 189 Colusa California 06011 72.199705 2.057210 104.007942	16
1.13	LFGCH4t TotalCH4t COUNTY LOCATION E_TOTPOP E_POV	16
1.14	1 49311 65473.75020 Alameda Alameda County, California 1643700 170884 .	16
1.15	2 0 13.15099 Alpine Alpine County, California 1146 227	16
1.16	3 0 784.27340 Amador Amador County, California 37829 3323	16
1.17	4 0 2049.12827 Butte Butte County, California 227075 44410	16
1.18	5 0 309.47615 Calaveras Calaveras County, California 45235 5242	16
1.19	6 0 178.26486 Colusa Colusa County, California 21464 2929	16
1.20	E_MINRTY geometry	16
1.21	1 1120309 MULTIPOLYGON (((-122.313 37...	16
1.22	2 468 MULTIPOLYGON (((-120.0726 3...	16
1.23	3 8066 MULTIPOLYGON (((-121.0276 3...	16
1.24	4 62685 MULTIPOLYGON (((-122.0573 3...	16
1.25	5 8330 MULTIPOLYGON (((-120.9955 3...	16
1.26	6 13792 MULTIPOLYGON (((-122.7851 3...	16
1.27	18
1.28	Pearson's product-moment correlation	18
1.29	18
1.30	data: svi_sf_joinTotalCH4tandsvi_sf_joinE_TOTPOP	18
1.31	t = 10.655, df = 56, p-value = 4.365e-15	18
1.32	alternative hypothesis: true correlation is not equal to 0	18
1.33	95 percent confidence interval:	18
1.34	0.7101412 0.8887681	18
1.35	sample estimates:	18
1.36	cor	18
1.37	0.818333	18
1.38	19
1.39	Call:	19
1.40	lm(formula = svi_sf_joinTotalCH4t svi_sf_joinE_TOTPOP)	19
1.41	19
1.42	Residuals:	19

1.43	Min 1Q Median 3Q Max	19
1.44	-34706 -8182 -6090 -3565 76153	19
1.45	19
1.46	Coefficients:	19
1.47	Estimate Std. Error t value Pr(> t)	19
1.48	(Intercept) 5.836e+03 2.956e+03 1.974 0.0533	19
1.49	svi_sf_join\$E_TOTPOP 1.962e-02 1.841e-03 10.655 4.36e-15 ***	19
1.50	—	19
1.51	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	19
1.52	19
1.53	Residual standard error: 20430 on 56 degrees of freedom	19
1.54	Multiple R-squared: 0.6697, Adjusted R-squared: 0.6638	19
1.55	F-statistic: 113.5 on 1 and 56 DF, p-value: 4.365e-15	19
1.56	20
1.57	Pearson’s product-moment correlation	20
1.58	20
1.59	data: svi_sf_joinOWCH4tandsvi_sf_joinE_TOTPOP	20
1.60	t = 82.353, df = 56, p-value < 2.2e-16	20
1.61	alternative hypothesis: true correlation is not equal to 0	20
1.62	95 percent confidence interval:	20
1.63	0.9930490 0.9975794	20
1.64	sample estimates:	20
1.65	cor	20
1.66	0.9958969	20
1.67	21
1.68	Call:	21
1.69	lm(formula = svi_sf_joinOWCH4t svi_sf_joinE_TOTPOP)	21
1.70	21
1.71	Residuals:	21
1.72	Min 1Q Median 3Q Max	21
1.73	-1801.49 -77.14 28.39 100.93 1809.74	21
1.74	21
1.75	Coefficients:	21
1.76	Estimate Std. Error t value Pr(> t)	21
1.77	(Intercept) -4.289e+01 7.117e+01 -0.603 0.549	21
1.78	svi_sf_join\$E_TOTPOP 3.650e-03 4.432e-05 82.353 <2e-16 ***	21
1.79	—	21
1.80	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	21
1.81	21
1.82	Residual standard error: 491.8 on 56 degrees of freedom	21
1.83	Multiple R-squared: 0.9918, Adjusted R-squared: 0.9917	21
1.84	F-statistic: 6782 on 1 and 56 DF, p-value: < 2.2e-16	21
1.85	22
1.86	Pearson’s product-moment correlation	22
1.87	22

1.88	data: svi_sf_join <i>AMCH4tandsvisfjoinE</i> _TOTPOP	22
1.89	t = -0.02427, df = 56, p-value = 0.9807	22
1.90	alternative hypothesis: true correlation is not equal to 0	22
1.91	95 percent confidence interval:	22
1.92	-0.2613200 0.2552665	22
1.93	sample estimates:	22
1.94	cor	22
1.95	-0.003243157	22
1.96	23
1.97	Call:	23
1.98	lm(formula = svi_sf_join <i>AMCH4tandsvisfjoinE</i> _TOTPOP)	23
1.99	23
1.100	Residuals:	23
1.101	Min 1Q Median 3Q Max	23
1.102	-5569 -5558 -5523 -3877 76775	23
1.103	23
1.104	Coefficients:	23
1.105	Estimate Std. Error t value Pr(> t)	23
1.106	(Intercept) 5.569e+03 2.053e+03 2.712 0.00885 **	23
1.107	svi_sf_join\$E_TOTPOP -3.103e-05 1.279e-03 -0.024 0.98072	23
1.108	—	23
1.109	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	23
1.110	23
1.111	Residual standard error: 14190 on 56 degrees of freedom	23
1.112	Multiple R-squared: 1.052e-05, Adjusted R-squared: -0.01785	23
1.113	F-statistic: 0.000589 on 1 and 56 DF, p-value: 0.9807	23
1.114	24
1.115	Pearson’s product-moment correlation	24
1.116	24
1.117	data: svi_sf_join <i>WWTPCH4tandsvisfjoinE</i> _TOTPOP	24
1.118	t = 57.684, df = 56, p-value < 2.2e-16	24
1.119	alternative hypothesis: true correlation is not equal to 0	24
1.120	95 percent confidence interval:	24
1.121	0.9859425 0.9950932	24
1.122	sample estimates:	24
1.123	cor	24
1.124	0.9916898	24
1.125	25
1.126	Call:	25
1.127	lm(formula = svi_sf_join <i>WWTPCH4tandsvisfjoinE</i> _TOTPOP)	25
1.128	25
1.129	Residuals:	25
1.130	Min 1Q Median 3Q Max	25
1.131	-6663.5 -147.9 65.7 168.0 3785.4	25
1.132	25

1.133	Coefficients:	25
1.134	Estimate Std. Error t value Pr(> t)	25
1.135	(Intercept) -1.466e+02 1.637e+02 -0.895 0.375	25
1.136	svi_sf_join\$E_TOTPOP 5.882e-03 1.020e-04 57.684 <2e-16 ***	25
1.137	—	25
1.138	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	25
1.139	25
1.140	Residual standard error: 1131 on 56 degrees of freedom	25
1.141	Multiple R-squared: 0.9834, Adjusted R-squared: 0.9832	25
1.142	F-statistic: 3327 on 1 and 56 DF, p-value: < 2.2e-16	25
1.143	26
1.144	Pearson’s product-moment correlation	26
1.145	26
1.146	data: svi_sf_joinLFGCH4tandsvi_sf_joinE_TOTPOP	26
1.147	t = 9.6217, df = 56, p-value = 1.815e-13	26
1.148	alternative hypothesis: true correlation is not equal to 0	26
1.149	95 percent confidence interval:	26
1.150	0.6670743 0.8702275	26
1.151	sample estimates:	26
1.152	cor	26
1.153	0.7893612	26
1.154	27
1.155	Call:	27
1.156	lm(formula = svi_sf_joinLFGCH4tandsvi_sf_joinE_TOTPOP)	27
1.157	27
1.158	Residuals:	27
1.159	Min 1Q Median 3Q Max	27
1.160	-31207 -3285 -1030 -501 46121	27
1.161	27
1.162	Coefficients:	27
1.163	Estimate Std. Error t value Pr(> t)	27
1.164	(Intercept) 4.564e+02 1.688e+03 0.270 0.788	27
1.165	svi_sf_join\$E_TOTPOP 1.012e-02 1.052e-03 9.622 1.81e-13 ***	27
1.166	—	27
1.167	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	27
1.168	27
1.169	Residual standard error: 11670 on 56 degrees of freedom	27
1.170	Multiple R-squared: 0.6231, Adjusted R-squared: 0.6164	27
1.171	F-statistic: 92.58 on 1 and 56 DF, p-value: 1.815e-13	27
1.172	28
1.173	Pearson’s product-moment correlation	28
1.174	28
1.175	data: svi_sf_joinTotalCH4tandsvi_sf_joinE_POV	28
1.176	t = 11.12, df = 56, p-value = 8.476e-16	28
1.177	alternative hypothesis: true correlation is not equal to 0	28

1.17895 percent confidence interval:	28
1.1790.7271725 0.8959424	28
1.180sample estimates:	28
1.181cor	28
1.1820.8296405	28
1.183	29
1.184Call:	29
1.185lm(formula = svi_sf_join <i>TotalCH4t svi_sf_joinE_POV</i>)	29
1.186	29
1.187Residuals:	29
1.188Min 1Q Median 3Q Max	29
1.189-23578 -8007 -6799 -4377 74610	29
1.190	29
1.191Coefficients:	29
1.192Estimate Std. Error t value Pr(> t)	29
1.193(Intercept) 6.684e+03 2.834e+03 2.358 0.0219 *	29
1.194svi_sf_join\$E_POV 1.310e-01 1.178e-02 11.120 8.48e-16 ***	29
1.195—	29
1.196Signif. codes: 0 ‘ ’ 0.001 ’’ 0.01 ’’ 0.05 ‘.’ 0.1 ’’ 1	29
1.197	29
1.198Residual standard error: 19840 on 56 degrees of freedom	29
1.199Multiple R-squared: 0.6883, Adjusted R-squared: 0.6827	29
1.200F-statistic: 123.7 on 1 and 56 DF, p-value: 8.476e-16	29
1.201	30
1.202Pearson’s product-moment correlation	30
1.203	30
1.204data: svi_sf_join <i>LFGCH4tandsvi_sf_joinE_POV</i>	30
1.205t = 9.3128, df = 56, p-value = 5.653e-13	30
1.206alternative hypothesis: true correlation is not equal to 0	30
1.20795 percent confidence interval:	30
1.2080.6526255 0.8638755	30
1.209sample estimates:	30
1.210cor	30
1.2110.7795177	30
1.212	31
1.213Call:	31
1.214lm(formula = svi_sf_join <i>LFGCH4t svi_sf_joinE_POV</i>)	31
1.215	31
1.216Residuals:	31
1.217Min 1Q Median 3Q Max	31
1.218-24928 -3488 -1556 -1095 47103	31
1.219	31
1.220Coefficients:	31
1.221Estimate Std. Error t value Pr(> t)	31
1.222(Intercept) 1.059e+03 1.700e+03 0.623 0.536	31

1.223	svi_sf_join\$E_POV 6.581e-02 7.067e-03 9.313 5.65e-13 ***	31
1.224	—	31
1.225	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	31
1.226		31
1.227	Residual standard error: 11900 on 56 degrees of freedom	31
1.228	Multiple R-squared: 0.6076, Adjusted R-squared: 0.6006	31
1.229	F-statistic: 86.73 on 1 and 56 DF, p-value: 5.653e-13	31
1.230		32
1.231	Pearson’s product-moment correlation	32
1.232		32
1.233	data: svi_sf_joinWWTPCH4tandsvi_sf_joinE_POV	32
1.234	t = 30.671, df = 56, p-value < 2.2e-16	32
1.235	alternative hypothesis: true correlation is not equal to 0	32
1.236	95 percent confidence interval:	32
1.237	0.9521271 0.9831025	32
1.238	sample estimates:	32
1.239	cor	32
1.240	0.9715012	32
1.241		33
1.242	Call:	33
1.243	lm(formula = svi_sf_joinWWTPCH4t svi_sf_joinE_POV)	33
1.244		33
1.245	Residuals:	33
1.246	Min 1Q Median 3Q Max	33
1.247	-8111.2 -525.9 -266.3 200.7 8372.5	33
1.248		33
1.249	Coefficients:	33
1.250	Estimate Std. Error t value Pr(> t)	33
1.251	(Intercept) 2.329e+02 2.977e+02 0.782 0.437	33
1.252	svi_sf_join\$E_POV 3.795e-02 1.237e-03 30.671 <2e-16 ***	33
1.253	—	33
1.254	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	33
1.255		33
1.256	Residual standard error: 2084 on 56 degrees of freedom	33
1.257	Multiple R-squared: 0.9438, Adjusted R-squared: 0.9428	33
1.258	F-statistic: 940.7 on 1 and 56 DF, p-value: < 2.2e-16	33
1.259		34
1.260	Pearson’s product-moment correlation	34
1.261		34
1.262	data: svi_sf_joinOWCH4tandsvi_sf_joinE_POV	34
1.263	t = 34.256, df = 56, p-value < 2.2e-16	34
1.264	alternative hypothesis: true correlation is not equal to 0	34
1.265	95 percent confidence interval:	34
1.266	0.9612260 0.9863552	34
1.267	sample estimates:	34

1.268	cor	34
1.269	0.9769612	34
1.270	35
1.271	Call:	35
1.272	lm(formula = svi_sf_joinOWCH4t svi_sf_joinE_POV)	35
1.273	35
1.274	Residuals:	35
1.275	Min 1Q Median 3Q Max	35
1.276	-2379.5 -363.1 -195.7 96.0 4351.1	35
1.277	35
1.278	Coefficients:	35
1.279	Estimate Std. Error t value Pr(> t)	35
1.280	(Intercept) 1.895e+02 1.656e+02 1.144 0.257	35
1.281	svi_sf_join\$E_POV 2.359e-02 6.885e-04 34.256 <2e-16 ***	35
1.282	—	35
1.283	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	35
1.284	35
1.285	Residual standard error: 1160 on 56 degrees of freedom	35
1.286	Multiple R-squared: 0.9545, Adjusted R-squared: 0.9536	35
1.287	F-statistic: 1174 on 1 and 56 DF, p-value: < 2.2e-16	35
1.288	36
1.289	Pearson’s product-moment correlation	36
1.290	36
1.291	data: svi_sf_joinAMCH4tandsvi_sf_joinE_POV	36
1.292	t = 0.43506, df = 56, p-value = 0.6652	36
1.293	alternative hypothesis: true correlation is not equal to 0	36
1.294	95 percent confidence interval:	36
1.295	-0.2033043 0.3116628	36
1.296	sample estimates:	36
1.297	cor	36
1.298	0.05803922	36
1.299	37
1.300	Call:	37
1.301	lm(formula = svi_sf_joinAMCH4t svi_sf_joinE_POV)	37
1.302	37
1.303	Residuals:	37
1.304	Min 1Q Median 3Q Max	37
1.305	-10228 -5295 -5206 -3608 76704	37
1.306	37
1.307	Coefficients:	37
1.308	Estimate Std. Error t value Pr(> t)	37
1.309	(Intercept) 5.202e+03 2.023e+03 2.572 0.0128 *	37
1.310	svi_sf_join\$E_POV 3.658e-03 8.408e-03 0.435 0.6652	37
1.311	—	37
1.312	Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1	37

1.313	37
1.314Residual standard error: 14160 on 56 degrees of freedom	37
1.315Multiple R-squared: 0.003369, Adjusted R-squared: -0.01443	37
1.316F-statistic: 0.1893 on 1 and 56 DF, p-value: 0.6652	37
2 Conclusion	43
3 References	44

List of Tables

1 Table of figures

Figure 1: Population density across counties in California

Figure 2: Impoverished population density across counties in California

Figure 3: Total Methane potential across counties in California

Figure 4: Waste water Methane Potential across counties in California

Figure 5: Landfill methane potential across counties in California

Figure 6: Organic Waste methane potential across counties in California

Figure 7: Animal manure methane potential across counties in California

Figure 8: Total Methane Generation Potential

Figure 9:

Figure 10:

Figure 11:

Figure 12:

Figure 13:

Figure 14:

\newpage

Background

Rationale of the study

The transition to cleaner and renewable sources of energy requires more work to scale already existing gains, for example in the area of biogas utilization, while also seeking out new opportunities. Biogas is a form of renewable energy produced by anaerobic decomposition or thermochemical conversion of biomass such as agricultural waste, manure, municipal waste sewage, green waste and food waste. Biogas is composed mostly of methane, alongside carbon dioxide, water vapor and other gases. Similar to natural gas, biogas can be burned directly as a fuel or treated to remove the CO₂ and other gases before for being used in the form of biomethane.

Utilizing biogas as an energy source helps to transform harmful gases from decomposing waste into positive use. Methane is a powerful greenhouse gas that traps heat in the atmosphere, with a global warming potential estimated to be over 25 times as potent as that of Carbon Dioxide. Transforming waste into biogas therefore reduces greenhouse gas emissions and the risk of pollution to waterways. The [UNECE] (<https://unece.org/challenge#:~:text=Methane%20is%20a%20powerful%20greenhouses,gr>) estimates that over a 20 year period, this ratio increases to 84-86 times. However, stored biogas limits the amount of methane released into the atmosphere and reduces dependence on fossil fuels. When stored , biogas serves as a renewable and reliable baseload power source and can even be used to rapidly meet peak power demands. When biogas is used for energy generation in place of fossil fuels, it enables even more emission reductions, sometimes resulting in carbon negative systems. According to the [Environmental and Energy Study Institute, EESI] (<https://www.eesi.org/papers/view/fact-sheet-biogasconverting-waste-to-energy>) ,the reduction of methane emissions derived from tapping all the potential biogas in the United States would be equal to the annual emissions of 800,000 to 11 million passenger vehicles.

Unfortunately, despite the various benefits offered by biogas energy, the United States currently only has 2,200 operating biogas systems. The EESI estimates this current capacity to be less than 20% of the total potential. It is on this basis that this study attempts to examine the factors contributing to biogas generation potential, particularly in

California. California was selected because of its status as a leading state in biogas generation potential based on estimates by the National Renewable Energy Laboratory (NREL). According to the [American Biogas Council](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A California can power nearly 200,000 homes if this biogas is utilized appropriately. The results of our analysis for California can provide the basis for further analysis across several other US states.

Research Questions

The analysis conducted was done based on two research questions:

1. Is biogas generation potential correlated with population density across California counties?
2. What additional factors must be considered in determining biogas generation potential?

\newpage

Dataset Information

the datasets used include the following:

1. Methane Generation Potential The NREL's dataset on methane generation potential covers data for the entire United States of America. It particularly providing information for Methane generation potential across all states in metric tons from sources which include landfills, industrial organic waste, animal manure, and wastewater . All of these individual sources were aggregated to estimate total methane generation potential by state. The dataset contained estimates for 2009 to 2012, being the latest available data we could find online.
2. Population data Population data was pulled from the latest available the Social Vulnerability Index data compiled by the Center for Disease Control (CDC) and Agency for Toxic Substance and Disease Registry (ATSDR). This dataset also includes details on impoverished populations and spans from 2017-2018. Since the SDI offers data that helps to effectly plan towards meeting the needs of improverished populations, this dataset also helps to dds another dimension to our analysis as it enables the exploration of impacts on socially vulnerable people.

\newpage

Exploratory Analysis

The two datasets were wrangled and joined both datasets to have a unified dataframe that captured data on methane generation potential, population and poverty levels by county. An initial exploration of the data was done to provide a visual assessment of spatial data and determine the emergence of any trends.

- 1.1 Simple feature collection with 6 features and 14 fields
- 1.2 Geometry type: MULTIPOLYGON
- 1.3 Dimension: XY
- 1.4 Bounding box: xmin: -122.7851 ymin: 37.45447 xmax: -119.5423 ymax: 40.15203
- 1.5 Geodetic CRS: WGS 84
- 1.6 ObjectID NAME STATE_NAME FIPS OWCH4t AMCH4t WWTPCH4t
- 1.7 1 184 Alameda California 06001 5700.695190 0.367359 10461.687700
- 1.8 2 185 Alpine California 06003 6.483819 0.000000 6.667176
- 1.9 3 186 Amador California 06005 156.310667 3.826180 624.136548
- 1.10 4 187 Butte California 06007 817.909133 20.460013 1210.759120
- 1.11 5 188 Calaveras California 06009 97.664105 11.663426 200.148617
- 1.12 6 189 Colusa California 06011 72.199705 2.057210 104.007942
- 1.13 LFGCH4t TotalCH4t COUNTY LOCATION E_TOTPOP E_POV
- 1.14 1 49311 65473.75020 Alameda Alameda County, California 1643700 170884
- 1.15 2 0 13.15099 Alpine Alpine County, California 1146 227
- 1.16 3 0 784.27340 Amador Amador County, California 37829 3323
- 1.17 4 0 2049.12827 Butte Butte County, California 227075 44410
- 1.18 5 0 309.47615 Calaveras Calaveras County, California 45235 5242
- 1.19 6 0 178.26486 Colusa Colusa County, California 21464 2929
- 1.20 E_MINRTY geometry
- 1.21 1 1120309 MULTIPOLYGON (((-122.313 37...
- 1.22 2 468 MULTIPOLYGON (((-120.0726 3...
- 1.23 3 8066 MULTIPOLYGON (((-121.0276 3...

shaded yellow, Los Angeles, has the highest methane (CH₄) generation potential. The density of impoverished population appears to be similarly distributed.

<!-- -->

The total Methane generation potential is greatest at the most densely populated areas. This trend appears to continue with wastewater derived methane.

<!-- -->

Industrial Organic waste, like wastewater, is highly correlated with population. With landfill-based methane, some counties in central CA, which do not appear to have a high population, have comparatively high methane generation potential. There may be some non-population related factors in play.

<!-- -->

The trend with animal-manure derived CH₄ however substantially differs from the previous cases. Population does not appear to be a decisive factor.

<!-- -->

\newpage

Analysis

Question 1: What factors affect biogas generation potential in different counties of

1.27

1.28 Pearson's product-moment correlation

1.29

1.30 data: svi_sf_join*TotalCH4tandsvi_sf_join*E_TOTPOP

1.31 $t = 10.655$, $df = 56$, $p\text{-value} = 4.365e-15$

1.32 alternative hypothesis: true correlation is not equal to 0

1.33 95 percent confidence interval:

1.34 0.7101412 0.8887681

1.35 sample estimates:

1.36 cor

1.37 0.818333

1.38

1.39 Call:

1.40 `lm(formula = svi_sf_joinTotalCH4t svi_sf_joinE_TOTPOP)`

1.41

1.42 Residuals:

1.43 Min 1Q Median 3Q Max

1.44 -34706 -8182 -6090 -3565 76153

1.45

1.46 Coefficients:

1.47 Estimate Std. Error t value Pr(>|t|)

1.48 (Intercept) 5.836e+03 2.956e+03 1.974 0.0533 .

1.49 svi_sf_join\$E_TOTPOP 1.962e-02 1.841e-03 10.655 4.36e-15

1.50 —

1.51 Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1

1.52

1.53 Residual standard error: 20430 on 56 degrees of freedom

1.54 Multiple R-squared: 0.6697, Adjusted R-squared: 0.6638

1.55 F-statistic: 113.5 on 1 and 56 DF, p-value: 4.365e-15

![] (FDR_Final_files/figure-latex/unnamed-chunk-7-1.pdf)<!-- --> ![] (FDR_Final_files/figure-latex/unnamed-chunk-7-1.pdf)

1.56

1.57 Pearson's product-moment correlation

1.58

1.59 data: svi_sf_joinOWCH4tandsvi_sf_joinE_TOTPOP

1.60 $t = 82.353$, $df = 56$, $p\text{-value} < 2.2e-16$

1.61 alternative hypothesis: true correlation is not equal to 0

1.62 95 percent confidence interval:

1.63 0.9930490 0.9975794

1.64 sample estimates:

1.65 cor

1.66 0.9958969

```

1.67
1.68 Call:
1.69 lm(formula = svi_sf_joinOWCH4t svi_sf_joinE_TOTPOP)
1.70
1.71 Residuals:
1.72 Min 1Q Median 3Q Max
1.73 -1801.49 -77.14 28.39 100.93 1809.74
1.74
1.75 Coefficients:
1.76 Estimate Std. Error t value Pr(>|t|)
1.77 (Intercept) -4.289e+01 7.117e+01 -0.603 0.549
1.78 svi_sf_join$E_TOTPOP 3.650e-03 4.432e-05 82.353 <2e-16
***
1.79 —
1.80 Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1
1.81
1.82 Residual standard error: 491.8 on 56 degrees of freedom
1.83 Multiple R-squared: 0.9918, Adjusted R-squared: 0.9917
1.84 F-statistic: 6782 on 1 and 56 DF, p-value: < 2.2e-16

```

1.85

1.86 Pearson's product-moment correlation

1.87

1.88 data: svi_sf_joinAMCH4tandsvi_sf_joinE_TOTPOP

1.89 $t = -0.02427$, $df = 56$, $p\text{-value} = 0.9807$

1.90 alternative hypothesis: true correlation is not equal to 0

1.91 95 percent confidence interval:

1.92 -0.2613200 0.2552665

1.93 sample estimates:

1.94 cor

1.95 -0.003243157

```

1.96
1.97  Call:
1.98  lm(formula = svi_sf_joinAMCH4t svi_sf_joinE_TOTPOP)
1.99
1.100  Residuals:
1.101  Min 1Q Median 3Q Max
1.102  -5569 -5558 -5523 -3877 76775
1.103
1.104  Coefficients:
1.105  Estimate Std. Error t value Pr(>|t|)
1.106  (Intercept) 5.569e+03 2.053e+03 2.712 0.00885 **
1.107  svi_sf_join$E_TOTPOP -3.103e-05 1.279e-03 -0.024
1.108  0.98072
1.109  —
1.110  Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1
1.111  Residual standard error: 14190 on 56 degrees of freedom
1.112  Multiple R-squared: 1.052e-05, Adjusted R-squared:
1.113  -0.01785
1.114  F-statistic: 0.000589 on 1 and 56 DF, p-value: 0.9807

```

1.114

1.115 Pearson's product-moment correlation

1.116

1.117 data: svi_sf_joinWWTPCH4tandsvi_sf_joinE_TOTPOP

1.118 $t = 57.684$, $df = 56$, $p\text{-value} < 2.2e-16$

1.119 alternative hypothesis: true correlation is not equal to 0

1.120 95 percent confidence interval:

1.121 0.9859425 0.9950932

1.122 sample estimates:

1.123 cor

1.124 0.9916898


```

1.125
1.126 Call:
1.127 lm(formula = svi_sf_joinWWTPCH4t svi_sf_joinE_TOTPOP)
1.128
1.129 Residuals:
1.130 Min 1Q Median 3Q Max
1.131 -6663.5 -147.9 65.7 168.0 3785.4
1.132
1.133 Coefficients:
1.134 Estimate Std. Error t value Pr(>|t|)
1.135 (Intercept) -1.466e+02 1.637e+02 -0.895 0.375
1.136 svi_sf_join$E_TOTPOP 5.882e-03 1.020e-04 57.684 <2e-16
***
1.137 —
1.138 Signif. codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
1.139
1.140 Residual standard error: 1131 on 56 degrees of freedom
1.141 Multiple R-squared: 0.9834, Adjusted R-squared: 0.9832
1.142 F-statistic: 3327 on 1 and 56 DF, p-value: < 2.2e-16

```

1.143

1.144 Pearson's product-moment correlation

1.145

1.146 data: svi_sf_join*LFCH4tandsvi_sf_join*E_TOTPOP

1.147 $t = 9.6217$, $df = 56$, $p\text{-value} = 1.815e-13$

1.148 alternative hypothesis: true correlation is not equal to 0

1.149 95 percent confidence interval:

1.150 0.6670743 0.8702275

1.151 sample estimates:

1.152 cor

1.153 0.7893612

```

1.154
1.155 Call:
1.156 lm(formula = svi_sf_joinLFGCH4t svi_sf_joinE__TOTPOP)
1.157
1.158 Residuals:
1.159 Min 1Q Median 3Q Max
1.160 -31207 -3285 -1030 -501 46121
1.161
1.162 Coefficients:
1.163 Estimate Std. Error t value Pr(>|t|)
1.164 (Intercept) 4.564e+02 1.688e+03 0.270 0.788
1.165 svi_sf_join$E__TOTPOP 1.012e-02 1.052e-03 9.622 1.81e-13
***
1.166 —
1.167 Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1
1.168
1.169 Residual standard error: 11670 on 56 degrees of freedom
1.170 Multiple R-squared: 0.6231, Adjusted R-squared: 0.6164
1.171 F-statistic: 92.58 on 1 and 56 DF, p-value: 1.815e-13
![] (FDR_Final_files/figure-latex/unnamed-chunk-7-6.pdf)<!-- -->

```

1.172

1.173 Pearson's product-moment correlation

1.174

1.175 data: svi_sf_join*TotalCH4tandsvi_sfjoin*E_POV

1.176 $t = 11.12$, $df = 56$, $p\text{-value} = 8.476e-16$

1.177 alternative hypothesis: true correlation is not equal to 0

1.178 95 percent confidence interval:

1.179 0.7271725 0.8959424

1.180 sample estimates:

1.181 cor

1.182 0.8296405

```

1.183
1.184 Call:
1.185 lm(formula = svi_sf_joinTotalCH4t svi_sf_joinE_POV)
1.186
1.187 Residuals:
1.188 Min 1Q Median 3Q Max
1.189 -23578 -8007 -6799 -4377 74610
1.190
1.191 Coefficients:
1.192 Estimate Std. Error t value Pr(>|t|)
1.193 (Intercept) 6.684e+03 2.834e+03 2.358 0.0219 *
1.194 svi_sf_join$E_POV 1.310e-01 1.178e-02 11.120 8.48e-16
***
1.195 —
1.196 Signif. codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
1.197
1.198 Residual standard error: 19840 on 56 degrees of freedom
1.199 Multiple R-squared: 0.6883, Adjusted R-squared: 0.6827
1.200 F-statistic: 123.7 on 1 and 56 DF, p-value: 8.476e-16

```

1.201

1.202 Pearson's product-moment correlation

1.203

1.204 data: svi_sf_join*LFGCH4tandsvi_sf_join*E_POV

1.205 $t = 9.3128$, $df = 56$, $p\text{-value} = 5.653e-13$

1.206 alternative hypothesis: true correlation is not equal to 0

1.207 95 percent confidence interval:

1.208 0.6526255 0.8638755

1.209 sample estimates:

1.210 cor

1.211 0.7795177

```

1.212
1.213 Call:
1.214 lm(formula = svi_sf_joinLFGCH4t svi_sf_joinE_POV)
1.215
1.216 Residuals:
1.217 Min 1Q Median 3Q Max
1.218 -24928 -3488 -1556 -1095 47103
1.219
1.220 Coefficients:
1.221 Estimate Std. Error t value Pr(>|t|)
1.222 (Intercept) 1.059e+03 1.700e+03 0.623 0.536
1.223 svi_sf_join$E_POV 6.581e-02 7.067e-03 9.313 5.65e-13 ***
1.224 —
1.225 Signif. codes: 0 ‘ ’ 0.001 ’ ’ 0.01 ’ ’ 0.05 ‘ ’ 0.1 ’ ’ 1
1.226
1.227 Residual standard error: 11900 on 56 degrees of freedom
1.228 Multiple R-squared: 0.6076, Adjusted R-squared: 0.6006
1.229 F-statistic: 86.73 on 1 and 56 DF, p-value: 5.653e-13

```

1.230

1.231 Pearson's product-moment correlation

1.232

1.233 data: svi_sf_joinWWTPCH4tandsvi_sf_joinE_POV

1.234 $t = 30.671$, $df = 56$, $p\text{-value} < 2.2\text{e-}16$

1.235 alternative hypothesis: true correlation is not equal to 0

1.236 95 percent confidence interval:

1.237 0.9521271 0.9831025

1.238 sample estimates:

1.239 cor

1.240 0.9715012


```

1.241
1.242 Call:
1.243 lm(formula = svi_sf_joinWWTPCH4t svi_sf_joinE_POV)
1.244
1.245 Residuals:
1.246 Min 1Q Median 3Q Max
1.247 -8111.2 -525.9 -266.3 200.7 8372.5
1.248
1.249 Coefficients:
1.250 Estimate Std. Error t value Pr(>|t|)
1.251 (Intercept) 2.329e+02 2.977e+02 0.782 0.437
1.252 svi_sf_join$E_POV 3.795e-02 1.237e-03 30.671 <2e-16 ***
1.253 —
1.254 Signif. codes: 0 ‘ ’ 0.001 ’ ’ 0.01 ’ ’ 0.05 ‘ ’ 0.1 ’ ’ 1
1.255
1.256 Residual standard error: 2084 on 56 degrees of freedom
1.257 Multiple R-squared: 0.9438, Adjusted R-squared: 0.9428
1.258 F-statistic: 940.7 on 1 and 56 DF, p-value: < 2.2e-16

```

1.259

1.260 Pearson's product-moment correlation

1.261

1.262 data: svi_sf_joinOWCH4tandsvi_sf_joinE_POV

1.263 $t = 34.256$, $df = 56$, $p\text{-value} < 2.2e-16$

1.264 alternative hypothesis: true correlation is not equal to 0

1.265 95 percent confidence interval:

1.266 0.9612260 0.9863552

1.267 sample estimates:

1.268 cor

1.269 0.9769612

```

1.270
1.271 Call:
1.272 lm(formula = svi_sf_joinOWCH4t svi_sf_joinE_POV)
1.273
1.274 Residuals:
1.275 Min 1Q Median 3Q Max
1.276 -2379.5 -363.1 -195.7 96.0 4351.1
1.277
1.278 Coefficients:
1.279 Estimate Std. Error t value Pr(>|t|)
1.280 (Intercept) 1.895e+02 1.656e+02 1.144 0.257
1.281 svi_sf_join$E_POV 2.359e-02 6.885e-04 34.256 <2e-16 ***
1.282 —
1.283 Signif. codes: 0 ‘ ’ 0.001 ’ ’ 0.01 ’ ’ 0.05 ‘ ’ 0.1 ’ ’ 1
1.284
1.285 Residual standard error: 1160 on 56 degrees of freedom
1.286 Multiple R-squared: 0.9545, Adjusted R-squared: 0.9536
1.287 F-statistic: 1174 on 1 and 56 DF, p-value: < 2.2e-16

```

1.288
1.289 Pearson's product-moment correlation
1.290
1.291 data: svi_sf_join*AMCH4tandsvi_sf_join*E_POV
1.292 $t = 0.43506$, $df = 56$, $p\text{-value} = 0.6652$
1.293 alternative hypothesis: true correlation is not equal to 0
1.294 95 percent confidence interval:
1.295 -0.2033043 0.3116628
1.296 sample estimates:
1.297 cor
1.298 0.05803922

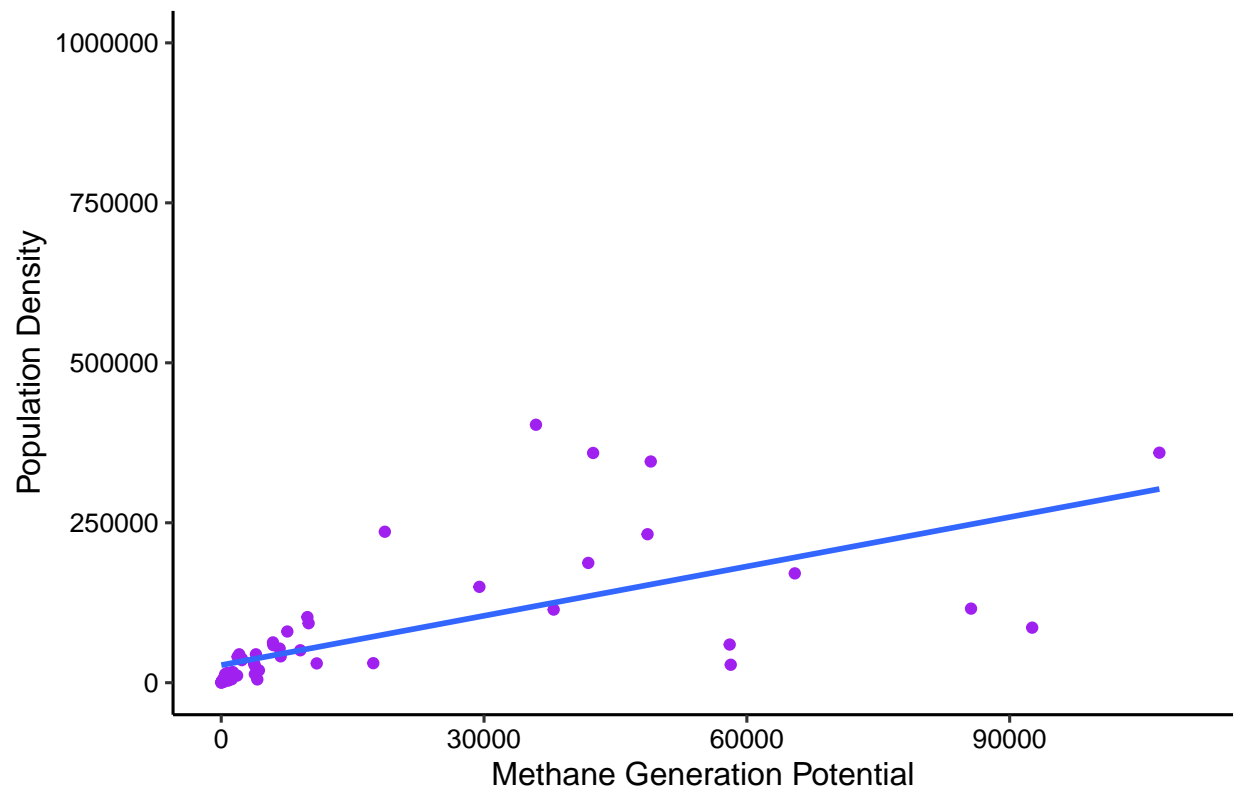
```

1.299
1.300 Call:
1.301 lm(formula = svi_sf_joinAMCH4t svi_sf_joinE_POV)
1.302
1.303 Residuals:
1.304 Min 1Q Median 3Q Max
1.305 -10228 -5295 -5206 -3608 76704
1.306
1.307 Coefficients:
1.308 Estimate Std. Error t value Pr(>|t|)
1.309 (Intercept) 5.202e+03 2.023e+03 2.572 0.0128 *
1.310 svi_sf_join$E_POV 3.658e-03 8.408e-03 0.435 0.6652
1.311 —
1.312 Signif. codes: 0 ‘’ 0.001 ’’ 0.01 ’’ 0.05 ‘’ 0.1 ’’ 1
1.313
1.314 Residual standard error: 14160 on 56 degrees of freedom
1.315 Multiple R-squared: 0.003369, Adjusted R-squared:
1.316 -0.01443
1.316 F-statistic: 0.1893 on 1 and 56 DF, p-value: 0.6652

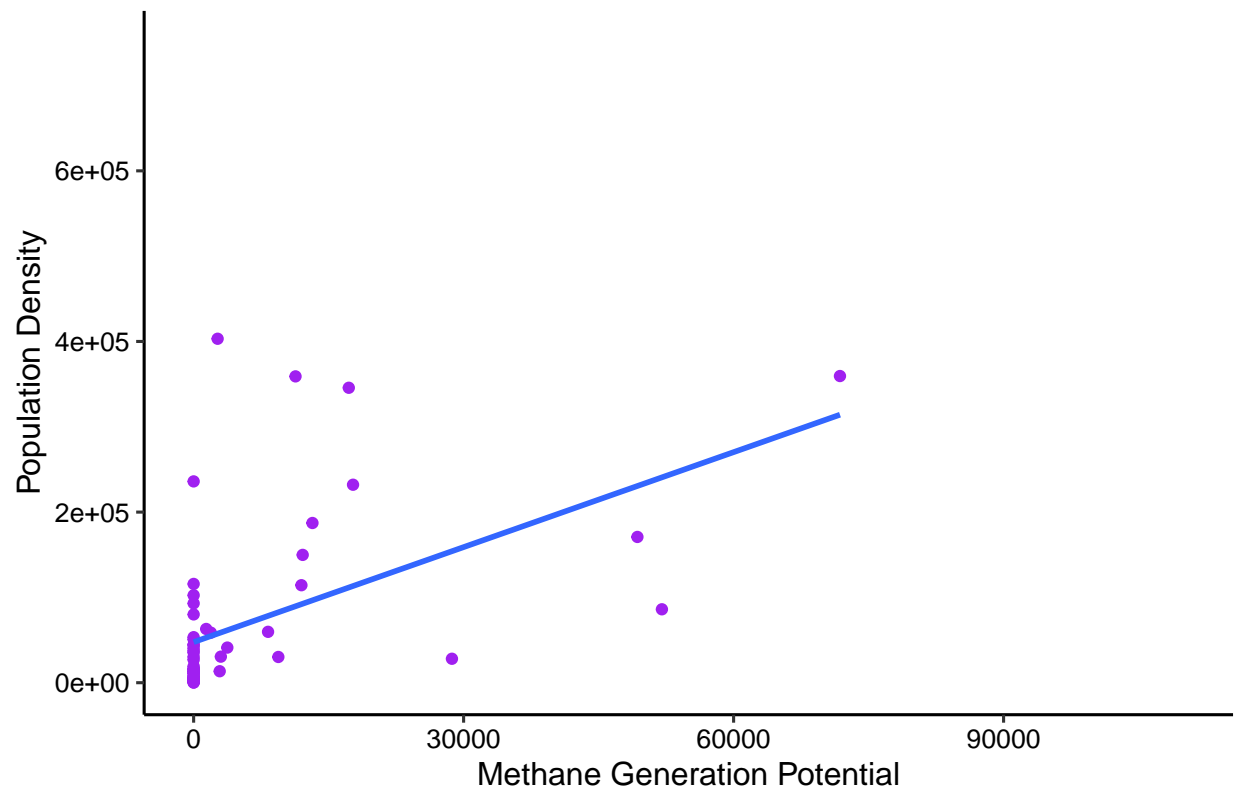
```

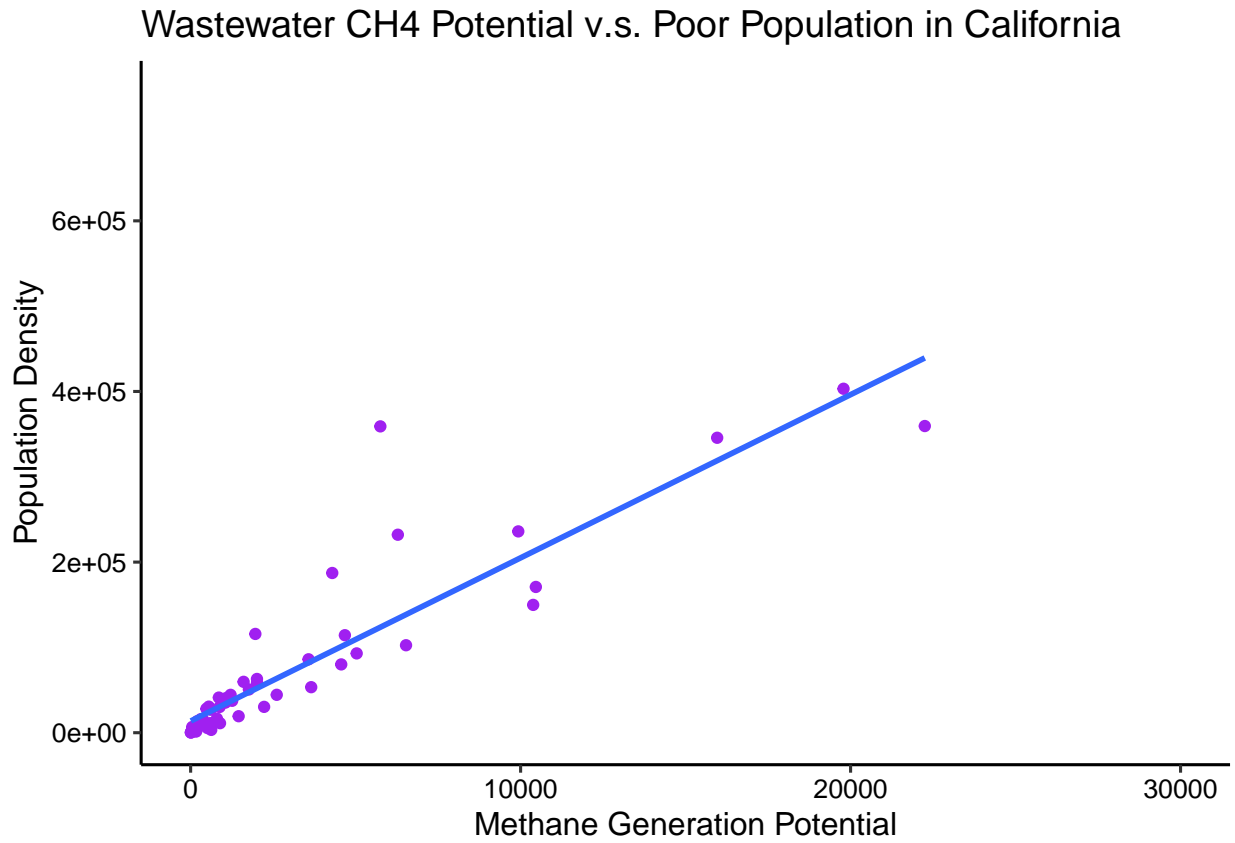
““

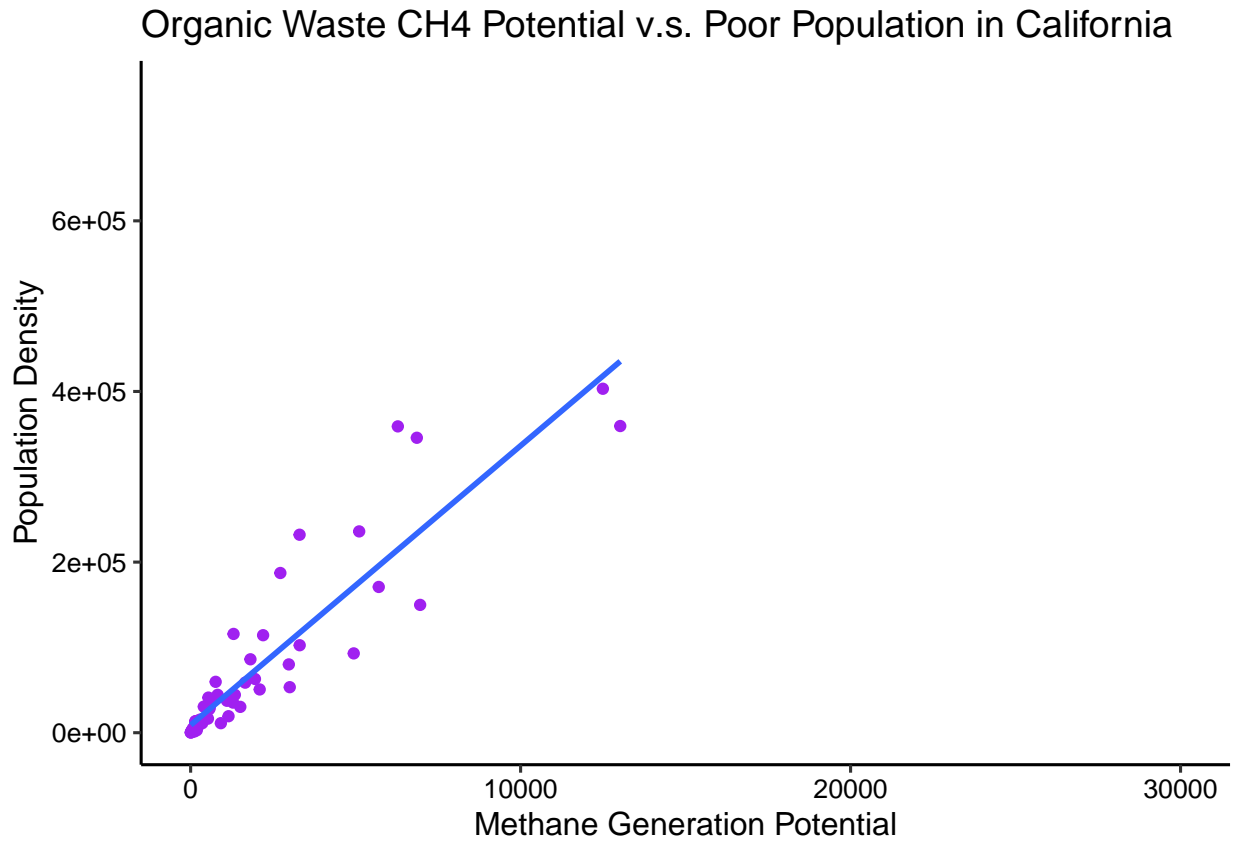
Total Methane Generation Potential v.s. Poor Population in Califor

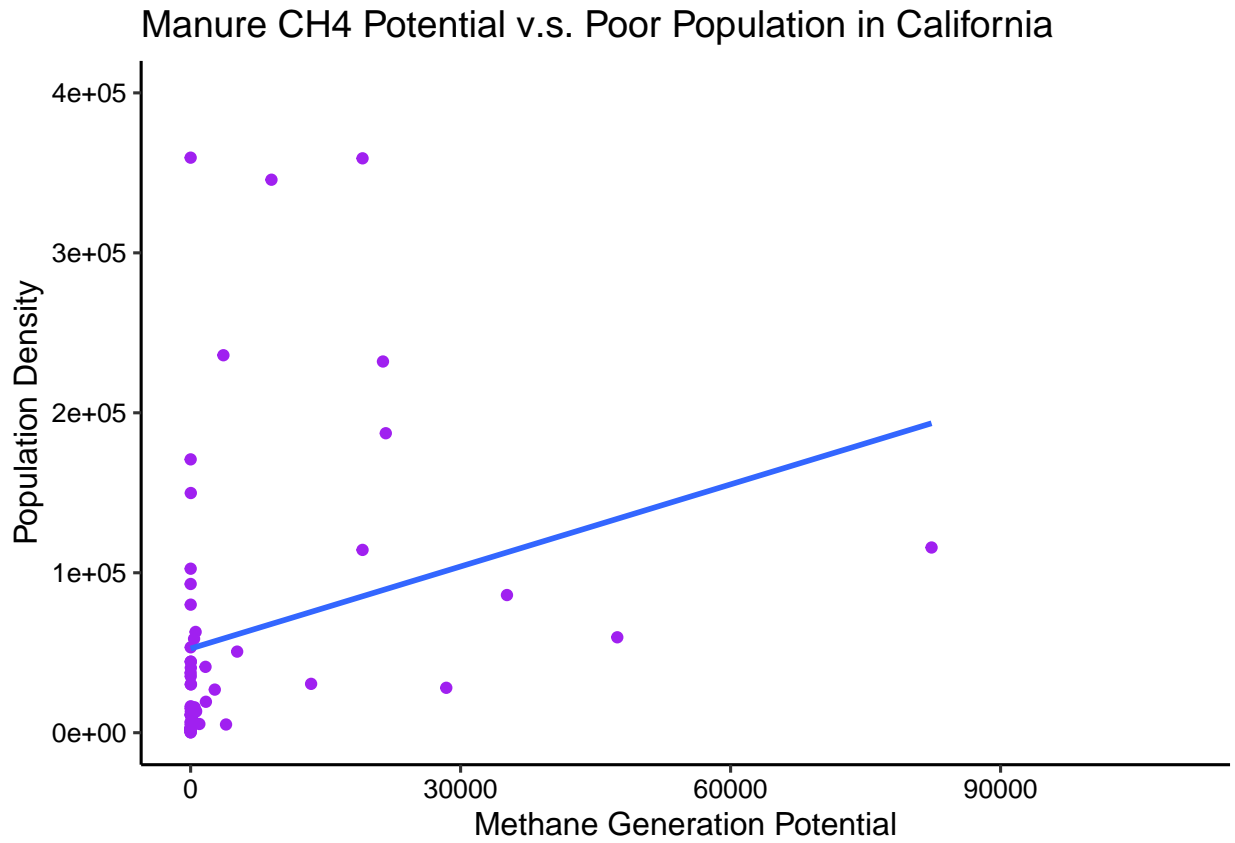


Landfill CH₄ Generation Potential v.s. Poor Population in California









2 Conclusion

The null hypothesis to our research question is that there is no correlation between population density and biogas generation potential. We ran a correlation test to test the strength and direction between California population and the total biogas generation potential. We also ran a simple linear regression model to investigate whether these two variables involves a linear relationship. By looking at the residual plots, we can tell that the regression model fits our data but has outliers that disrupt the model accuracy. The p-value is $4.22e-15$, indicating that the result is statistically significant. The multiple R-squared is 0.67, meaning that the independent variables can explain 67% of the variations in biogas generation potential. Therefore, we are confident in rejecting the null hypothesis. Now we can say the total biogas generation potential in California does have a linear relationship with the population density. However, we still want to know how and what different sources is correlated with population. Subsequently, we ran the same test for each individual biogas sources. Based on the results from running correlation test and linear regression model for each of the biogas generation sources: Organic waste, Animal manure, Wastewater, and Landfill. We found out the organic waste and wastewater contributes to the highest correlation to the population density, while Animal manure and landfill are less correlated.

Overall, the analysis conducted show that biogas generation potential across counties in California is positively correlated with the population density. This implies that the higher the population in a specific area, the higher the biogas generation potential. However, the analysis also recognises that generation potential and correlation with population density varies according to the biomass source across counties. As such, results show that animal based methane appeared to be less correlated with population. Assumptions that best explain this result are that more animals will contribute to animal-based biomass and that human habitation is often less where animal population is dense. Results of our analysis also show that the potential of landfill-based biogas was not limited to densely or sparsely populated areas, while organic waste and waste water represented the most significant sources of biogas generation potential across California.

Future analysis can focus on improving the model presented in this report by eliminating some outliers and using more recent datasets. This study made use of older data sets because those were the ones most recently made available on open and credible data portals. Other areas of future analysis include exploring the level and process of utilization of California's biogas generation potential, as well as how utilization can be scaled. For example, the potential of leveraging the biogas generation potential by enabling counties with lower generation potentials to tap into the resources for those with higher generation. This also includes scoping the potential to utilize biogas energy for base load power or peaking, using various analyses including a cost-benefit analysis.

3 References

<add references here if relevant, otherwise delete this section>