MSIN0093 Business Strategy and Analytics

GN02 - Group Coursework

California Air Resources Board Consulting Report on California's Zero-Emission Vehicle Popularization Scheme

Group Name: BrilliantConsultingGroup

Team Members (Student Number):

Jiawen (Vicky), Wang (22182987)

Jia Siang (Jay), Wu (22040941)

Sirui (Cindy), Yan (22061914)

Chun Ho Travis, Lui (22085872)

Chih Yu (Jerry), Pan (22181021)

Word Count: 3,589

1. Introduction

The state of California, US is the most aggressive leading state in the move to deliver cleaner air and massive reductions in climate-warming pollution. It has recently adopted a mission to achieve 100% new zero-emission vehicle (ZEV) sales by 2035 as one of its measures to popularise ZEV, which includes full battery-electric (EV), Plug-in hybrid and hydrogen fuel cell vehicles (California Air Resources Board, 2022). As the California Air Resources Board (CARB), we create value by setting up rules and funding programs to pave the way for the popularisation of ZEV in California. We serve the general public of California, especially focusing on disadvantaged people with low income. However, various hurdles have made our value-capturing process harder than expected. We are encountering the depletion of funding for the various incentive schemes which were set up to provide financial aid for low-income drivers to purchase ZEV. The state has already spent over $400 million on these programs, such as the CC4A program, which offers up to $9,500 towards a down payment for a ZEV if the applicant turns in a vehicle older than a 2005 model. "The rapid depletion of CVAP and DCAP money was large because incomes for many individuals were close to the maximum threshold – 400% of the federal poverty level – applied that no funds were left for people with the highest need" (Aoun, 2022). With the state running out of funding and these programs temporarily shutting down, we are faced with the urgent question: Is California on the right track to meet its mission? How should we improve the current funding scheme to ensure funds are being received by those who need them the most? What else can we do to adopt the entire ZEV-popularisation plan?

Based on the 2020 survey results published by Consumer Reports, the main reasons holding respondents back from purchasing or leasing a plug-in EV are "purchase price" (49%) and "not enough public charging stations" (39%) (Consumer Report, 2021). Therefore, the main areas of focus for this report will be the economy (indicated by income and oil price), the number of EV charging stations, as well as the environment (indicated by air quality), which is the primary focus of CARB and the state. From these three perspectives, this report will draw on various datasets to gain a better image of California's current situation. Improvements would be suggested to ameliorate the state's current funding schemes and its decision on further actions, in order to facilitate the popularisation of ZEV in California.

The first part of the report analyses whether the current funding schemes impacted ZEV sales and whether they addressed the needs of the most disadvantaged groups. Then, a comparison of the 58 counties in California is made based on two metrics: air quality and income, to identify the counties to be prioritised in our full zero-emission plan and reallocation of the grants. In the second part of the report, a causal relationship regression analysis of the effect of the number of EV charging stations on the sales of electric vehicles (EV) is conducted based on the DAG diagram to determine if California should build more electric charging stations to incentivise the purchase of EVs. Lastly, we will discuss other measures and policies to gear California towards its full ZEV mission.

2. Dataset Description

Vehicle population (2010-2021)

The data is obtained from the California Energy Commission, containing the number of vehicles in use from 2010 to 2021, with information regarding fuel type, make, models and county. To find the relation between the vehicle in use and charging stations, the fuel type is restricted to Electric vehicles and PHEVs, with the data year being 2021.

New ZEV Sales (1998-2022)

This is an official government dataset published by California Energy Commission. This county-level dataset contains information regarding different types of ZEV sales in California from 1998 to 2022, including the vehicle's fuel type, the brand and model, and the number of sales of that vehicle. The variable of interest is the number of ZEVs sold for each county in each year. Due to incomplete data for 2022, only data up to 2021 will be analysed.

Population by County (2021)

This dataset is obtained from the California Demographics website, which contains the population of each county in 2021.

Income Per Capita (2021)

This is an official government dataset from the US Bureau of Economic Analysis, which contains the per capita personal income for each county in the US from 2019 to 2021.

Domestic Crude Oil First Purchase Prices (2021)

This dataset from California U.S. Energy Information Administration (EIA) shows the crude oil price in the U.S. by area in 2021. The oil price for California state is compared to The Petroleum Administration for Defense Districts (PADDs), which EIA uses to facilitate the comprehension of petroleum data across regions.

Air Quality Index (2021):

The official government dataset from the US Environmental Protection Agency (EPA) contains information on the air quality measurement of the counties in California in 2021. Air Quality Index (AQI) measures the daily air quality from 0 to 500, with a higher number representing worse air quality.

Electric Vehicle Charging Stations (2021)

This data is obtained from the California Energy Commission, containing the number of different types of charging stations for electric vehicles in California, last updated in 2021.

3. Data and Analysis

3.1 Understanding California's Current Situation

California citizens have experienced high gas prices recently (Figure 1). As petrol cars and ZEVs are substitutes, the higher-than-average crude oil price than any other PADDs is considered a catalyst that encouraged the total number of ZEV sales in California over the years. Despite the gradual increase in raw ZEV sales over the years, the proportion of ZEV to the total number of vehicles in use is only about 3% in 2021, suggesting that there is much room for promoting the popularisation of ZEV (Figure 2).

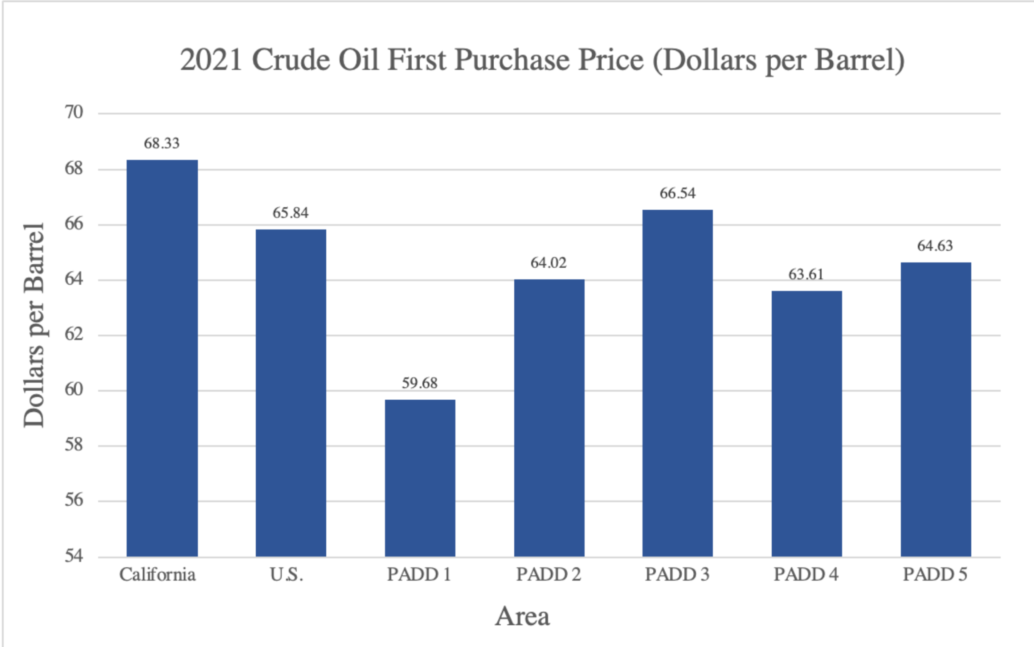

Figure 1. Comparison of Crude Oil Prices in California and Other Regions

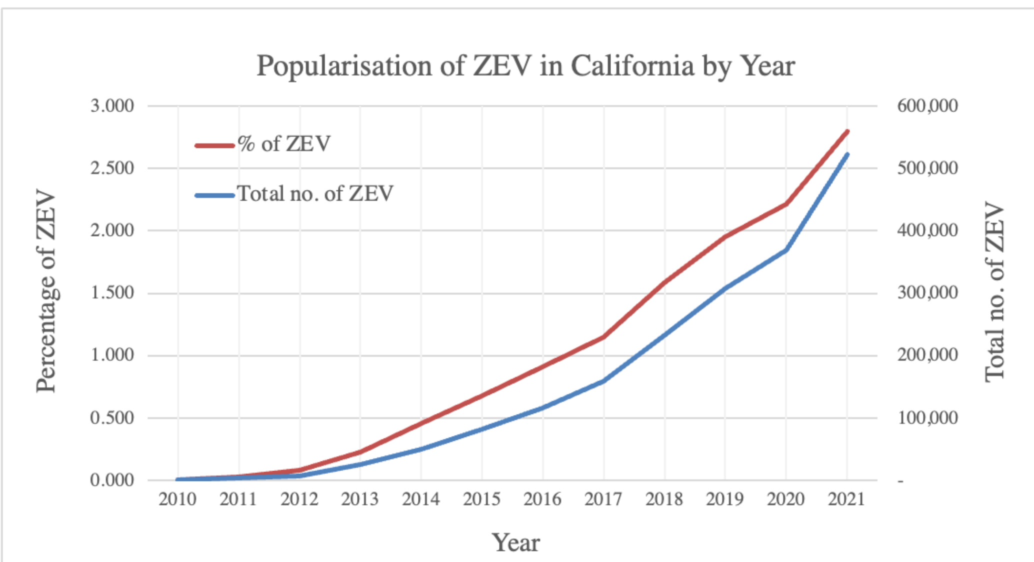

Figure 2. Total Number of ZEV and Percentage of ZEV in California by Year

3.2 Impacts of Income on Total ZEV Sales

The high price of ZEV is surveyed to be the main reason that is holding people back from purchasing. Our hypothesis assumes that personal income should be positively correlated with the sales of ZEV. If so, implementing the funding schemes for lower income groups could hypothetically increase ZEV sales, advocating the state's current plan.

To test the hypothesis, we performed a regression analysis using the 2019-2021 California personal income (county level) normalised data and the 2019-2021 California ZEV sales (county level) normalised data.

```
                                 OLS Regression Results
=========================================================================================
Dep. Variable:     Renormalized_Vehicles   R-squared (uncentered):                  0.097
Model:                               OLS   Adj. R-squared (uncentered):             0.092
Method:                    Least Squares   F-statistic:                             18.28
Date:                   Sun, 04 Dec 2022   Prob (F-statistic):                   3.17e-05
Time:                           11:47:09   Log-Likelihood:                        -233.40
No. Observations:                    171   AIC:                                     468.8
Df Residuals:                        170   BIC:                                     471.9
Df Model:                              1
Covariance Type:               nonrobust
======================================================================================
                         coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------
Renormalized_Income    0.3116      0.073      4.276      0.000       0.168       0.455
==============================================================================
Omnibus:                      212.780   Durbin-Watson:                   2.085
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             8295.812
Skew:                           5.052   Prob(JB):                         0.00
Kurtosis:                      35.592   Cond. No.                         1.00
==============================================================================
```

Figure 3. OLS Regression Results for ZEV sales with Income

The coefficient indicates that, on average, every unit increase in normalised income would increase the normalised number of ZEV sales by 0.3116, and the coefficient is statistically significant since the P-value is close to 0.

$$\textit{Sales of ZEV} = \textit{const.} + 0.31 * \textit{Income} + e$$

To better interpret the coefficients, we denormalised the data using mean and standard deviation. The regression shows that in the raw data, for every $1,000 increase in average household income, ZEV sales will increase by 59 units, which is economically significant. Thus, the presence of funding schemes is likely to alleviate Californians' financial burden and boost ZEV ownership.

3.3 Effectiveness of Existing Funds

To assist all Californians in the transition to ZEVs, the state has provided subsidies to residents since 2010. The first subsidy, Clean Vehicle Rebate Project (CVRP), still exists today along with three new funding programmes. Together they support individuals, especially the low-income ones, in ZEV purchases. Among the programmes, two are statewide, and the others are regional funds to target buyers from the most polluted and the poorest areas. A brief description of each fund is shown in Table 1.

Fund Name	Location Launch Year	Fund Limit (2022)	Income Cap (2022)	Total Fund Up Till 2022
Clean Vehicle Rebate Project (CVRP)	Statewide 2010	$7,000	$135,000	$1.6 billion
Clean Cars for All (CC4A)	South Coast Area (2015) San Joaquin Valley (2015) Bay Area (2019) Sacramento (2020)	$9,500	400% of the federal poverty level - $54,360	$190 million
Driving Clean Assistance Program (DCAP)	Bay Area Sacramento (9 counties[1]) 2016	$35,000 ($30,000 from Travis Credit Union, $5,000 from CARB	400% of the federal poverty level - $54,360	$12.5[2] million from CARB
Clean Vehicle Assistance Program (CVAP)	Statewide 2018	$7,000	400% of the federal poverty level - $54,360	$51.9 million

Table 1. Description of the Four Main EV Funds

[1] Alameda County, Contra Costa County, Marin County, Napa County, Sacramento County, San Francisco County, San Mateo County, Santa Clara County, Santa Cruz County, Solano County, Sonoma County, Yolo County, Certain California Native American Tribes.

[2] This is an estimated value for the total fund as of 29 November 2022. The total funding allocation from FY2016 - 2017 to FY2019 - 2020 (four periods) was $7.02 million. Assume the funding allocation for FY2020 - 2021 was the average of previous year's, $1.755 million. Given the total funding for DCAP and CVAP was $23.5 million (California Air Resources Board, 2021) and the allocation ratio of DCAP to CVAP was 1:5.24, funding allocation to DCAP for FY2021-2022 is $3.76 million. Hence, the estimated total fund for DCAP from FY2016-2022 is 7.02+1.755+3.76 = $12.5 million.

Except for CVRP which accepts higher-income applicants, the latter three subsidies only accept low-and-moderate-income individuals with annual income capped at $54,360. As CVRP does not target any specific disadvantaged counties and has a high-income cap, we speculate that CVRP has raised the application barrier of low-income individuals. Given that CVRP is on a first-come-first-served basis (Aoun, 2022), not only do low-income individuals struggle to complete the paperwork quickly, but they also face increased competition from the higher-income and more-educated groups. This is ineffective in achieving our mission and will further consume our financial resources. On the other hand, regional funds were intentionally designed to benefit the less privileged people. A greater amount of grants is expected to encourage more unprivileged drivers to prioritise ZEVs over gas-powered cars.

One way to evaluate the effectiveness of the current regional funds is to match the counties that are included in existing funds with the counties with disadvantaged communities (DACs) to investigate any possible resource misallocation. We follow the CalEnviroScreen 4.0 scores defined and calculated by the California Environmental Protection Agency to identify DACs[3]. The map on the left of Figure 4 illustrates counties with regional funds, where counties with darker shades receive more regional funds. The heatmap on the right shows DACs in each county, where redder colours mean the more burdened the DACs. Based on the visualisation, we discovered a resource misallocation. In Figure 4, the dark-blue squares indicate an excessive funding allocation to counties with fewer or without DACs, while the light-blue squares indicate an absence of funding for counties (San Bernardino & Imperial) with burdened DACs. Hence, two regression models[4] are run to check whether statewide funds are ineffective in counties with DACs, and whether regional funds are effective in counties with DACs given the resource misallocation. Data from 2013 to 2019 at the county level is selected. Since the total ZEV sales after 2020 were greatly affected by the pandemic regardless of the funding, data from 2020 onwards were excluded. Additionally, data were separated into two groups: counties with DACs and counties without DACs[5].

[3] The California Environmental Protection Agency (CalEPA) defines DACs based on geographic, socioeconomic, public health, and environmental hazard criteria. The higher CalEnviroScreen 4.0 score, the more burdened community.
[4] For detailed explanation of variables and the regression models, please see Appendix 1.
[5] California Office of Environmental Health Hazard Assessment has a list of DACs that corresponds to the right-hand-side heatmap.

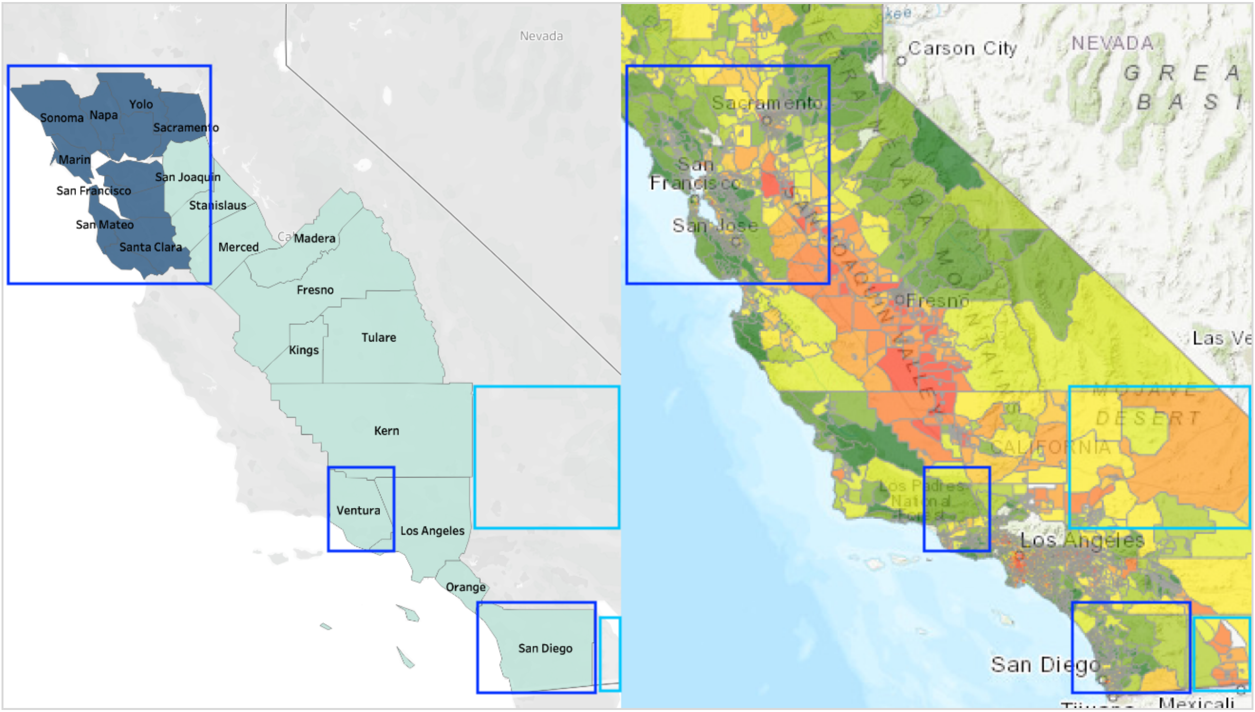

Figure 4. Counties with Regional Funds vs. Counties with Disadvantaged Communities

The regression results for counties with DACs show that impacts of the regional fund amount and the regional fund income cap on the ZEV sales are statistically significant at a 5% significance level. A $5,000 increase in regional funding is associated with a 5.8 unit increase in the total ZEV sales. For the regional fund income cap, a $500 growth is associated with a 31.4 units decrease in the total ZEV sales. Therefore, the income cap is negatively correlated with the total sales. If California plans to take an aggressive move to boost ZEV sales in the burdened counties, we have to ensure that the funding income threshold is low enough to benefit only the needed drivers. While regional funds are effective in raising ZEV sales in counties with DACs, the statewide fund is ineffective given the insignificant coefficient of the statewide fund amount. This echoes our speculation that the current statewide funding CVRP is ineffective in helping disadvantaged Californians. A possible explanation is the greater competition caused by the fund's high-income cap and broad-targeting nature, making unprivileged individuals difficult to fight for the benefits.

*Sales of ZEV = const. + 0.06*Statewide Fund Amt (in $5k) + 5.81* Regional Fund Amt (in $5k) - 0.66*Statewide Fund Income Cap (in $500) - 31.35*Regional Fund Income Cap ($500) + e*

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            total_sales   R-squared:                       0.271
Model:                            OLS   Adj. R-squared:                  0.260
Method:                 Least Squares   F-statistic:                     26.33
Date:                Sun, 04 Dec 2022   Prob (F-statistic):           1.58e-14
Time:                        11:38:38   Log-Likelihood:                -2148.3
No. Observations:                 217   AIC:                             4305.
Df Residuals:                     213   BIC:                             4318.
Df Model:                           3
Covariance Type:            nonrobust
======================================================================================
                       coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------
S_Fund_Amt(5k)       0.0599      0.037      1.638      0.103      -0.012       0.132
R_Fund_Amt(5k)       5.8139      0.800      7.268      0.000       4.237       7.391
S_Fund_ICap(500)    -0.6644      3.392     -0.196      0.845      -7.350       6.021
R_Fund_ICap(500)   -31.3501     11.171     -2.806      0.005     -53.370      -9.330
==============================================================================
Omnibus:                      146.482   Durbin-Watson:                   0.507
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1239.534
Skew:                           2.598   Prob(JB):                    6.90e-270
Kurtosis:                      13.493   Cond. No.                     1.06e+03
==============================================================================
```

Figure 5. OLS Regression Results for Counties with DACs

The regression results for counties without DACs show that only the coefficient for regional fund amount is statistically significant. This implies that for counties without DACs but receiving regional funds, regional funds are also effective in boosting their total ZEV sales. Other variables are statistically insignificant, meaning that the total ZEV sales in these counties are mostly uncorrelated with the funding. Possible reasons are the high income among residents and the low urgency to purchase ZEVs, making the residents less reliant on funding.

*Sales of ZEV = const. + 0.004*Statewide Fund Amt (in $5k) + 0.92* Regional Fund Amt(in $5k) - 0.005*Statewide Fund Income Cap (in $500) - 4.16*Regional Fund Income Cap ($500) + e*

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            total_sales   R-squared:                       0.301
Model:                            OLS   Adj. R-squared:                  0.290
Method:                 Least Squares   F-statistic:                     26.60
Date:                Sun, 04 Dec 2022   Prob (F-statistic):           2.36e-14
Time:                        11:45:28   Log-Likelihood:                -1350.1
No. Observations:                 189   AIC:                             2708.
Df Residuals:                     185   BIC:                             2721.
Df Model:                           3
Covariance Type:            nonrobust
=====================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------
S_Fund_Amt(5k)        0.0043      0.002      1.914      0.057      -0.000       0.009
R_Fund_Amt(5k)        0.9234      0.436      2.116      0.036       0.062       1.784
S_Fund_ICap(500)     -0.0046      0.232     -0.020      0.984      -0.462       0.452
R_Fund_ICap(500)      4.1613      2.683      1.551      0.123      -1.132       9.455
==============================================================================
Omnibus:                      144.058   Durbin-Watson:                   0.569
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1718.901
Skew:                           2.794   Prob(JB):                         0.00
Kurtosis:                      16.677   Cond. No.                     3.75e+03
==============================================================================
```

Figure 6. OLS Regression Results for Counties without DACs

To conclude, California is on the right track towards the mission but in an ineffective way. The statewide fund is not a strategic resource for the mission as it is not durable and appropriable in both types of counties. Counties deplete the finances but contribute little to the total ZEV sales. Consequently, the government failed to capture any significant social and environmental values as expected. While regional funds are helpful in both regions, they can be improved through resource reallocation and eligibility modification to maximise the benefits of the limited funding.

3.4 Finding the Target Counties

Besides allocating the funding to the most needed people by lowering the income cap and introducing more regional funding to replace the statewide funding, another strategy that the state can take to make effective use of the limited funding resources is to target the most burdened counties. Air quality and income are the main factors to be assessed to boost the effectiveness of our aid programmes environmentally and economically.

Transportation accounts for 80% of air pollutants in California, including nearly 80% of nitrogen oxide emissions and 90% of diesel particulate matter pollution (California Energy Commission, 2022). Targeting the most polluted counties could maximise the effect of our programmes to mitigate the detrimental effect on California’s air quality caused by petrol cars. Figure 7 illustrates that in 2021, San Bernardino, Riverside, Kern, Tulare, Los Angeles, Fresno and San Diego have recorded the highest median AQI and underperformed California's average. Compared to the “Good” level of concern for California average in the range below 50, the 7 counties have fallen into the "Moderate" level of concern, with an index value in the range of 51 to 100, meaning the air quality is risky for people who are sensitive to air pollution. The highly polluted counties indicated in red in the heat map (Figure 8) are mainly located in the South-East region of California.

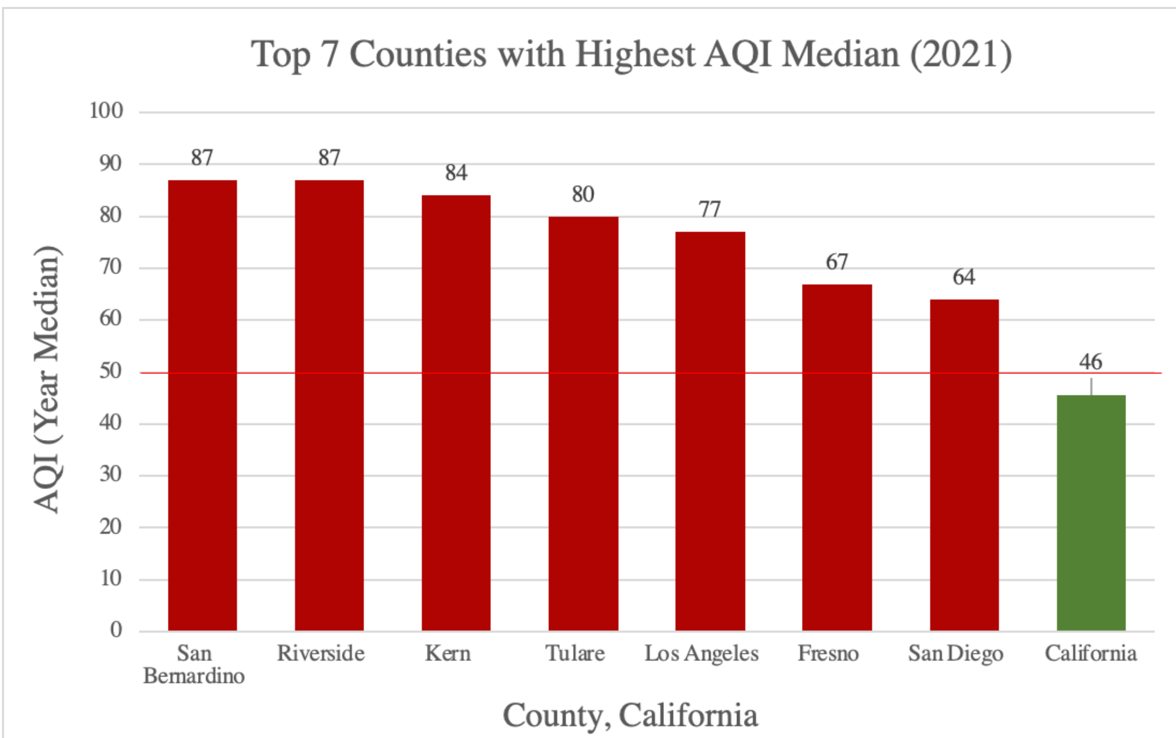

Figure 7. Top 7 Counties with the Highest Air Quality Index (2021)

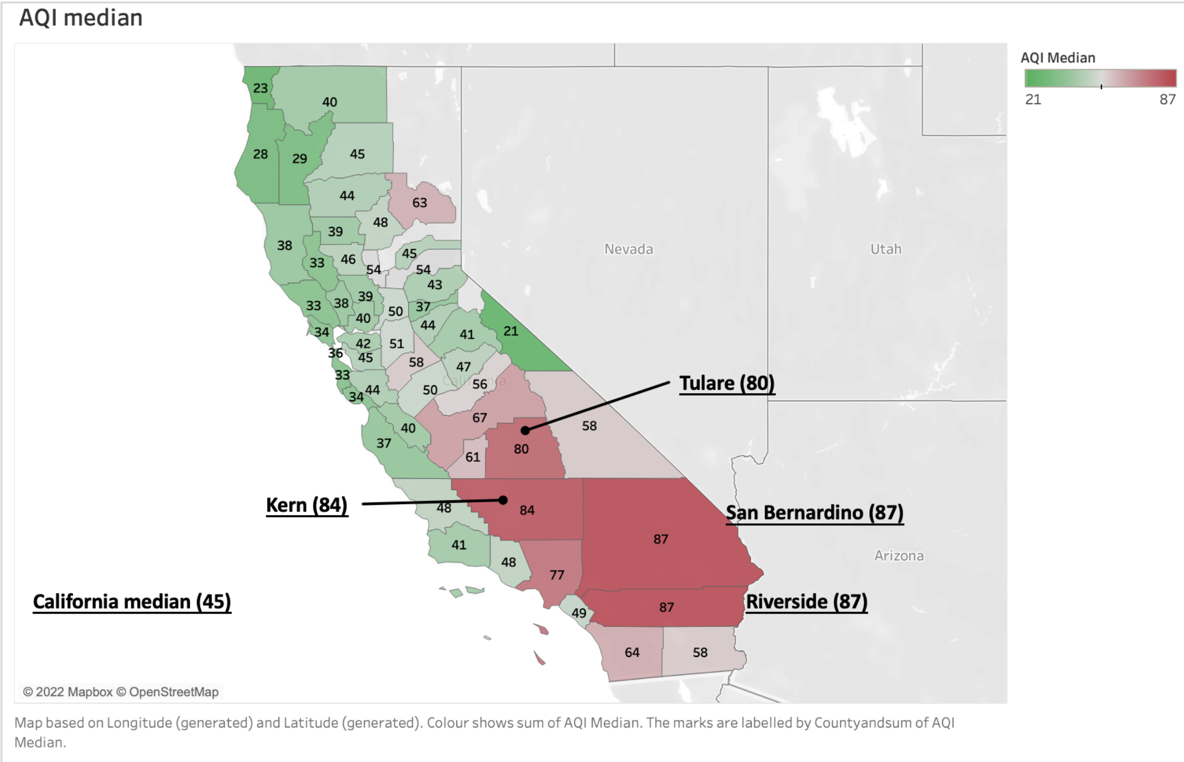

Figure 8. AQI Heat Map by County (2021)

In addition to air quality, income per capita is another factor to prioritise counties for the allocation of resources. According to recent research, the monetary cost is the most significant concern and barrier for California residents to switch to ZEV (Preston, 2020). Since our state has limited resources to subsidise every resident, priority should also be given to lower-income areas, which is indicated as red in the heat map (Figure 9). To summarise the effect of both, Figure 10 is a scatter plot that shows each county's position in terms of air quality and income level position. Counties located in the top-left corner have low-income levels and lousy air quality; thus, more urgently need to adopt the full ZEV plan and require more funding support, while the counties in the bottom-right corner could be deferred. Combined with the conclusion from part 3.3, the state should introduce more regional funding schemes in San Bernardino, Riverside, Kern and Tulare. On the other hand, San Francisco, Marin, and San Mateo can be put on the waiting list until the needs of the other counties are met and the state has extra funding.

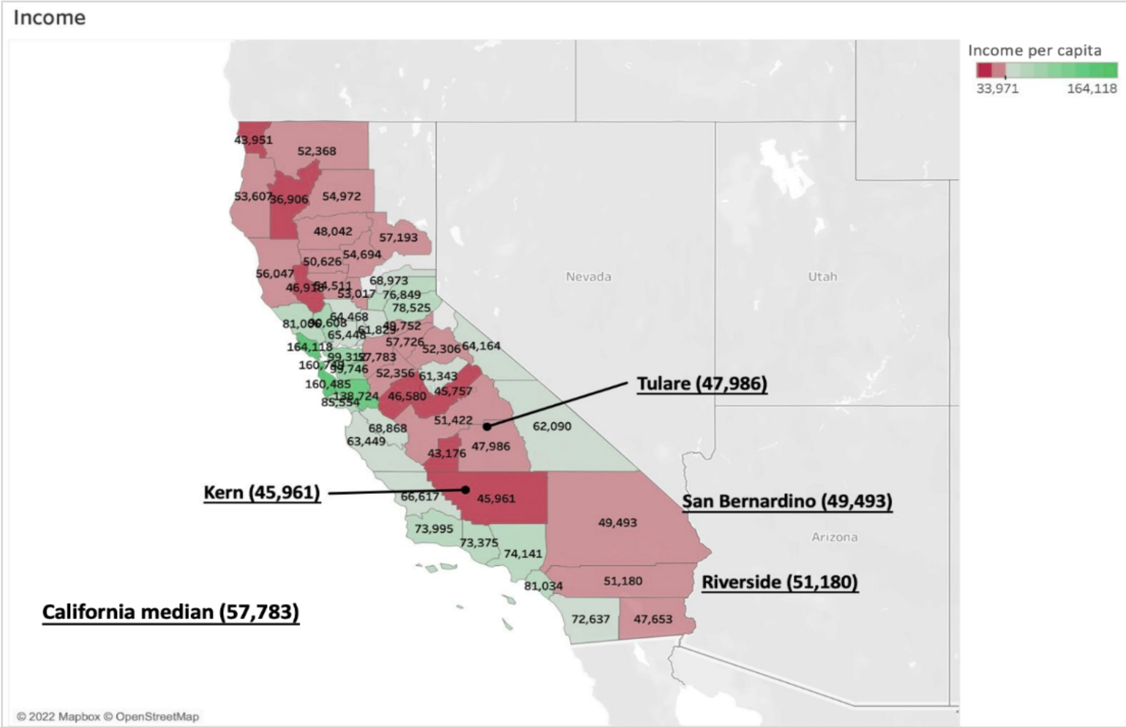

Figure 9. Income Per Capita Heat Map by County (2021)

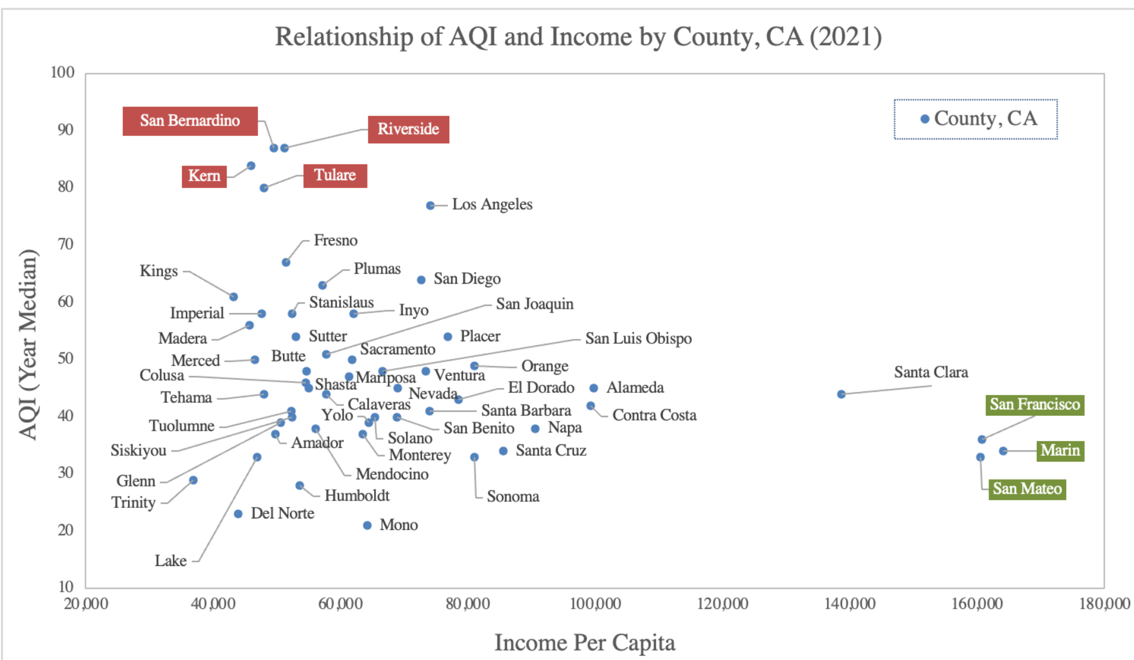

Figure 10. Scatter Plot of AQI and Income of Citizens by County (2021)

3.5 Causal Relationship between Charging Stations and Electric Vehicles in Use

According to a survey conducted by Consumer Reports, 39% of people feel that the insufficiency of charging stations would affect their willingness to buy electric vehicles (Consumer Reports, 2021). Thus intuitively, building more charging stations could increase sales. To confirm that there is a necessity to build more charging stations in the area, the causal relationship between the number of charging stations and the number of ZEV in use needs to be proven.

To examine the relationship between the number of charging stations and the sales of electric vehicles, a Directly Acyclic Graph (DAG) is applied to inspect the causal effect by controlling other potential factors such as Electric Vehicles in Use, Income, Population of each county and total area of the city that could affect the sales of the ZEV. The variables list and the DAG are shown respectively on Table 2 and Figure 11.

Firstly, the population is found to have an impact on S, U and C, since having more people in the county would increase the number of charging stations, sales and ZEV in use, with A being P's ancestor affecting both P and C. Moreover, S would also affect U and C, since the more ZEV are sold, the more charging stations there should be and more ZEV would be in use, with I being S's ancestor affecting both S and U.

Factor	Denotation	Type
Number of ZEV in Use	U	Outcome
Number of Sales of ZEV	S	Confounder
Number of Charging Station	C	Treatment
Income	I	Ancestor
Population	P	Confounder
Area	A	Ancestor

Table 2. Variable list and type

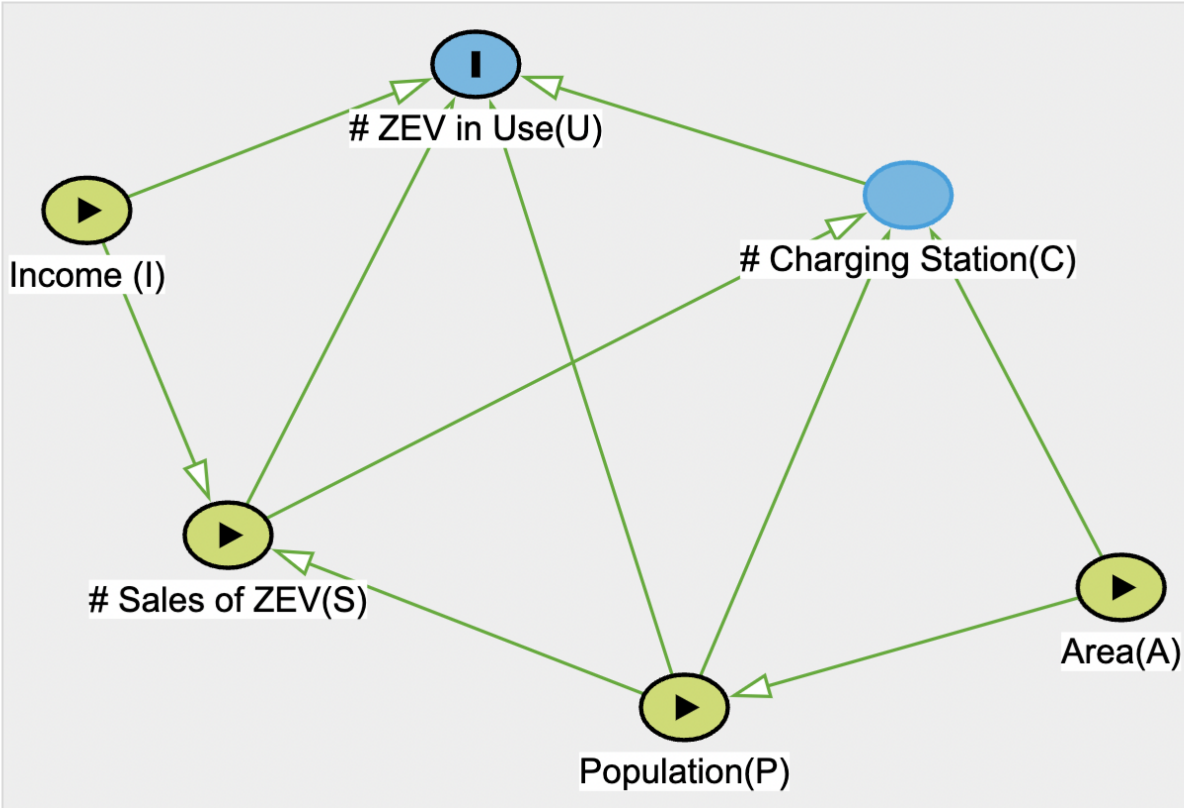

Figure 11. Causal Effect DAG Diagram of
Number of Charging Stations and Number of EVs in Use

In the DAG above, there are several causal paths from C to U. With P and S being the confounders of C and U, to block other paths that would be affecting the causal effect from C to U, P and S are included in the regression as control variables. By controlling P, the paths C <-A ->P->U and C <-P->U are blocked. As S is being controlled, C <- S->U and C <- A->P->S->U are also blocked. In this case, although Income affects both S and U, there is no causal path from C to U that includes it; Hence, there is no need to set it up as a control variable in the regression. As for Area, though it is on the causal path, it is always linked to population, which means that by controlling the population, the path would already be blocked. Therefore, it is unnecessary to block Area as well. Considering the situation above, the regression is shown below.

$$U = \beta 0 + \beta 1C + \beta 2P + \beta 3S$$

The estimated regression model is shown below using the data from 2021, grouping by counties.

$$U = 1737.76 + 11.66C + 0.1P + 21.08S$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:          Vehicle_in_use   R-squared:                       0.989
Model:                             OLS   Adj. R-squared:                  0.989
Method:                  Least Squares   F-statistic:                     1639.
Date:                 Wed, 30 Nov 2022   Prob (F-statistic):           5.50e-53
Time:                         11:08:02   Log-Likelihood:                -699.78
No. Observations:                   58   AIC:                             1408.
Df Residuals:                       54   BIC:                             1416.
Df Model:                            3
Covariance Type:            nonrobust
==========================================================================================
                             coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------
const                   1737.7573   6427.001      0.270      0.788   -1.11e+04    1.46e+04
Total_charging_station    11.6574      4.168      2.797      0.007       3.301      20.013
Population                 0.1048      0.017      6.301      0.000       0.071       0.138
Total_sales               21.0756      3.009      7.005      0.000      15.043      27.108
==============================================================================
Omnibus:                       18.144   Durbin-Watson:                   2.176
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               33.616
Skew:                          -0.961   Prob(JB):                     5.02e-08
Kurtosis:                       6.196   Cond. No.                     1.78e+06
==============================================================================
```

Figure 12. OLS Regression Results for DAG

As shown in Figure 12, the coefficients of those three variables could all significantly affect the total number of electric vehicles in use. By holding other variables constant, the coefficient of the number of charging stations symbolises the causal effect of the number of charging stations on electric vehicles in use. The coefficient of 11.66 further represents that if the number of charging stations increases by 1 unit, the number of electric vehicles in use will increase by 11.66 units.

Knowing that there is a causal effect from the number of charging stations to the number of ZEV in use, the state should consider it as an imperative factor in improving the usage rate of electric vehicles. The establishment of charging stations would increase the convenience to ZEV owners, while the completeness of the infrastructure would also attract more people to use ZEVs, building up a network effect to drag even more people into the cycle to drive the success of the government's mission.

4. Conclusion

The series of analyses conducted in this report have proven that the idea of implementing funding schemes is setting California on the right track to popularise ZEV. However, the results of the analysis also shed light on various problems in the design of the funding scheme programs and revealed the reasons why the funds are running out faster than expected. Based on the data analysis in part 3, the mismatch between local needs and the current funding programmes is the prominent problem. Concurrently, regional funds are proven to be more effective than statewide funds in incentivising the purchase of ZEV because they are designed to target the most needed groups. Another problem is that the income cap set by the funds is too high. While the low-income people have higher elasticity of demand on ZEV, the funds are not properly allocated to the hands of the disadvantaged group. As a result, most of the funding is not being effectively used and is barely contributing to the increase in the sale of ZEV. When a resource is not appropriable nor durable, it is not value-creating and should be reallocated elsewhere. Hence, a reallocation of resources to the prioritised counties with poor air quality and low average income would boost the cost-effectiveness massively. Our analysis also indicated that more charging stations would increase the sales of ZEV. We strongly suggest the state invest in building more EV charging stations to accelerate its mission.

With the strategic advice proposed in this report, we hope to see the current problems regarding the funding schemes improve and the use of ZEV increase significantly over the next period. California will continue to lead the move to deliver cleaner air and massive reductions in climate-warming pollution.

5. Limitations and Improvements

Owing to the assumptions behind our analytical models, below are the limitations associated. In our analysis, oil prices and income are assumed to be good representations of the economy. In our follow-up report with sufficient time and resources, we could include other economic indicators such as income inequality, spending on transportation and employment rate. The report also assumed that average household income represents the entire population. It was assumed that the county with the lowest household average income is the county with the biggest low-income group. Yet, the income distribution factor is neglected. Counties with huge income inequality and counties with the majority of the population being mid/low-income will have different needs and should be addressed differently. Instead of finding the lowest average income, one way to improve the accuracy of the analysis is to acquire inside data from the Census or conduct surveys to gather data on each household in each county, such as household income, spending on transportation, ways of commute and car type. We can then analyse the income distribution by county and find the county with the biggest low-income group. We can also obtain datasets about each person that successfully applied for the fund schemes, including income, county, the price of the ZEV that they purchased, their old car model type and year purchased to perform a cluster analysis. We could then find trends of which type of low-income people are more likely to benefit from the scheme, as well as estimating the suitable amount of funding for the underprivileged. As a result, we could advise the state with better budget planning to avoid wasting funds.

Regarding the incompleteness of the AQI dataset, Alpine, Lassen, Modoc, Sierra and Yuba counties are excluded due to the lack of monitoring stations. They are assumed to be indifferent and hence are excluded from the analysis.

Another limitation is that the DAG Diagram did not include all factors that could potentially be confounders, which should be controlled in the regression model to predict a more accurate causal analysis. Due to the unavailability of data, it also could not take into account the fixed effect variable that contributes to the existing difference between each county's EV sales regardless of the number of charging stations. Eventually, the coefficient of correlation is not purely the relationship. To improve this, panel data should be collected over years so that each county has multiple data points for the model to capture the fixed effect variable.

6. Attribution Statement

Everyone in the team was involved in brainstorming ideas on problems, preparing datasets and identifying the limitations of the report. All members also described the data and made references for resources that are used respectively.

Cindy was responsible for setting out the structure of the report and assisting each group member with their parts. She also wrote the introduction, conclusion and limitations.

Vicky was responsible for part 3.3: Effectiveness of Existing Funds and Appendix 1. She dug out numbers from funding scheme reports and added them to the *New ZEV Sales (1998-2022)* dataset for the use of regression analysis. She also formatted all references.

Jay was responsible for part 3.2: How does income affect the sales of ZEV, as well as the data visualisation of part 3.4: Finding the Target Counties. He cleaned and normalised the data, analysed the data through an OLS regression model and visualised the data in heatmaps.

Travis was responsible for part 3.1: Understanding California's Current Situation, and part 3.4: Finding the Target Counties. He cleaned and organised all the underlying datasets that were relevant, including vehicle population, oil price, income and AQI, and performed the data visualisation and analysis. He also did proofreading and paraphrasing for all written analysis.

Jerry was responsible for part 3.5: Causal Relationship between Charging Stations and Electric Vehicles in Use. He came up with the DAG explaining the causal relation between number of charging stations and the number of electric vehicles in use, and also combined data from different sources to form the regression for DAG.

7. References

Aoun, G. (2022, October 13). California wants everyone to drive EVs. How will low-income people afford them? *The Guardian*. Retrieved December 4, 2022, from https://www.theguardian.com/us-news/2022/oct/13/electric-vehicles-evs-california-low-income

Beneficial State Impact. (2022). Clean Vehicle Assistance Program. Retrieved December 4, 2022, from https://impact.beneficialstate.org/clean-vehicle-assistance-program/

California Air Resources Board. (2016, October 3). 2016-17 Grant Solicitation: Low Carbon Transportation Greenhouse Gas Reduction Fund (GGRF) Investments and Air Quality Improvement Program (AQIP). Retrieved December 4, from https://ww2.arb.ca.gov/sites/default/files/2020-05/fy1617_cvrp_solicitation.pdf

California Air Resources Board. (2018, September 21). Proposed Fiscal Year 2018-19 Funding Plan for Clean Transportation Incentives For Low Carbon Transportation Investments and the Air Quality Improvement Program. Retrieved December 4, from https://ww2.arb.ca.gov/sites/default/files/classic/msprog/aqip/fundplan/proposed_1819_funding_plan.pdf

California Air Resources Board. (2021, February 25). Scrap-and-replace incentive program Clean Cars 4 All reaches 10,000th participant milestone. Retrieved December 4, from https://ww2.arb.ca.gov/news/scrap-and-replace-incentive-program-clean-cars-4-all-reaches-10000th-participant-milestone

California Air Resources Board. (2021, June 17). Vehicle Purchase Incentive Projects For Low-Income Consumers. Retrieved December 4, 2022, from https://ww2.arb.ca.gov/sites/default/files/2021-06/fa_wkgrp_pres_06172021.pdf

California Air Resources Board. (2021, October 8). Proposed Fiscal Year 2021-22 Funding Plan for Clean Transportation Incentives. Retrieved December 4, 2022, from https://ww2.arb.ca.gov/sites/default/files/2021-10/fy21-22_fundingplan.pdf

California Air Resources Board. (2022). Clean Cars 4 All. Retrieved December 4, 2022, from https://ww2.arb.ca.gov/sites/default/files/movingca/vehiclescrap.html

California Air Resources Board. (2022). Financing Assistance for Low-Income Consumers. Retrieved December 4, 2022, from https://ww2.arb.ca.gov/sites/default/files/movingca/vehiclefinancing.html

California Air Resources Board. (2022, August 25). California moves to accelerate to 100% new zero-emission vehicle sales by 2035. Retrieved December 4, 2022, from https://ww2.arb.ca.gov/news/california-moves-accelerate-100-new-zero-emission-vehicle-sales-2035

California Clean Vehicle Rebate Project. (2022). Eligibility & Requirements. Retrieved December 4, 2022, from https://cleanvehiclerebate.org/en/eligibility-guidelines

California Energy Commission. (2022). Transforming transportation. Retrieved December 4, 2022, from https://www.energy.ca.gov/about/core-responsibility-fact-sheets/transforming-transportation

California Energy Commission. (2022). ZEV and Infrastructure Stats Data. Retrieved December 4, 2022, from https://www.energy.ca.gov/files/zev-and-infrastructure-stats-data

California Office of Environmental Health Hazard Assessment. (2022). CalEnviroScreen 4.0. Retrieved December 4, 2022, from https://advocacy.consumerreports.org/wp-content/uploads/2021/03/California-EV-FE-Survey-Report-3.8.21.pdf

California U.S. Energy Information Administration. (2022). Domestic Crude Oil First Purchase Prices by Area. Retrieved December 4, 2022, from https://www.eia.gov/dnav/pet/pet_pri_dfp1_k_a.htm

Clean Vehicle Assistance Program. (n.d.). Final Report | FY2015-2016. Retrieved December 4, 2022, from https://cleanvehiclerebate.org/sites/default/files/attachments/CVRP_FY2015-16_FinalReport.pdf

Clean Vehicle Assistance Program. (2022). Eligibility. Retrieved December 4, 2022, from https://cleanvehiclegrants.org/eligibility/

Community Housing Development Corporation. (2022). How the Program Works. Retrieved December 4, 2022, from https://communityhdc.org/dcap/#Submit

Consumer Attitudes Towards Electric Vehicles and Fuel Efficiency in California: 2020 Survey Results. (2021, March). *Consumer Reports*. Retrieved December 4, 2022, from https://advocacy.consumerreports.org/wp-content/uploads/2021/03/California-EV-FE-Survey-Report-3.8.21.pdf

Environmental Protection Agency. (2022). Air Quality Index Report. Retrieved December 4, 2022, from https://www.epa.gov/outdoor-air-quality-data/air-quality-index-report

Personal Income by County, Metro, and Other Areas | U.S. Bureau of Economic Analysis (BEA). (2022). Retrieved December 4, 2022, from https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas

Preston, B. (2020, December 18). Consumer reports survey shows strong interest in Electric Cars. *Consumer Reports*. Retrieved December 4, 2022, from https://www.consumerreports.org/hybrids-evs/cr-survey-shows-strong-interest-in-evs-a1481807376/

U.S. Department of Energy. (2022). Alternative Fuels Data Center: Maps and Data - Electric Vehicle Registrations by State. Retrieved December 4, https://afdc.energy.gov/data/10962

8. Appendix 1

Two datasets were built on top of the *New ZEV Sales (1998-2022)* dataset. We calculated the sum of ZEV sales for each county in each year from 2013 to 2019. Samples of the built datasets are shown below, they are for the use of regression analysis in section 3.3.

Year	County	total_sales	S_Fund_Amt(5k)	R_Fund_Amt(5k)	S_Fund_ICap(500)	R_Fund_ICap(500)	S_Fund_Limit(500)	R_Fund_Limit(500)	Post	Treat
2013	Alameda	2943	15816.5	0	300	0	10	0	0	1
2014	Alameda	4966	22240.5	0	300	0	10	0	0	1
2015	Alameda	5536	22240.5	0	300	0	14	0	0	1
2016	Alameda	5327	26600	180	300	95.04	14	10	1	1
2017	Alameda	7056	28000	400	300	96.48	14	10	1	1
2018	Alameda	12578	41000	800	300	97.12	14	10	1	1
2019	Alameda	11781	47600	781.116	300	99.92	14	19	1	1

Figure 13. A Sample of Dataset of Counties with DACs

Year	County	total_sales	S_Fund_Amt(5k)	R_Fund_Amt(5k)	S_Fund_ICap(500)	R_Fund_ICap(500)	S_Fund_Limit(500)	R_Fund_Limit(500)	Post	Treat
2013	Alpine	4	15816.5112	0	300	0	10	0	0	0
2014	Alpine	1	22240.48	0	300	0	10	0	0	0
2015	Alpine	0	24724.4	0	300	0	14	0	0	0
2016	Alpine	0	26600	0	300	0	14	0	0	0
2017	Alpine	1	28000	0	300	0	14	0	0	0
2018	Alpine	1	41000	0	300	0	14	0	0	0
2019	Alpine	3	47600	0	300	0	14	0	0	0

Figure 14. A Sample of Dataset of Counties without DACs

Variable	Description
Year	ZEV sales year
County	Name of a California county
total_sales	Total ZEV sales in the year, including PHEV, Hydrogen, Electric
S_Fund_Amt(5k)	Amount of statewide fund of the year in $5,000 - proposed funding allocation or the actual funding amount for that year was used - numbers were dug out from CARB/CVRP annual reports - sources are cited in the references
R_Fund_Amt(5k)	Amount of regional fund of the year in $5,000

	<u>For DCAP</u> - funding allocation was used, and it can be found from CARB "Vehicle Purchase Incentive Projects For Low-Income Consumers" source <u>For CC4A</u> - actual funding figures were used, and they can be found from CARB "Scrap-and-replace incentive program Clean Cars 4 All reaches 10,000th participant milestone" source - amount was calculated by taking the ratio and averages - i.e. as of September 30, 2020, there were 10,379 participants in all subsidised counties, and the total spending was $82.5 million. Number of participants in each region is - South Coast Air Quality Management District: 6,510 (5 years), San Joaquin Valley Air Pollution Control District: 2,894 (5 years), Bay Area Air Quality Management District: 953(2 years), Sacramento Metropolitan Air Quality Management District: 22 (few months). - then, the regional funding for CC4A in South Coast Air Quality Management District in 2015 is (82.5 million/10,379)*(6,510/5) = $10,349,262 (2,569 in 5k)
S_Fund_ICap(500)	Income cap for statewide funding of the year in $500
R_Fund_ICap(500)	Income cap for regional funding of the year in $500
S_Fund_Limit(500)	Maximum statewide subsidy a participant can get in that year is $500.
R_Fund_Limit(500)	Maximum regional subsidy a participant can get in that year is $500.
Post	Whether the county has received any regional funding in that year. 1 if yes, 0 if no.
Treat	If the county has received any regional funding from 2013 to 2019. 1 if yes, 0 if no.

Table 3. Description of Variables