

Understanding the alternative fuel vehicles market in the US

Introduction

Business Context. You are an analyst at the ICCT (International Council on Clean Transportation (<https://theicct.org/mission-history>)), an NGO that provides research to environmental authorities around the world. You are part of a team that is preparing a report on the alternative fuel vehicles (AFV) market in the United States, with a particular focus on the state of Washington.

Business Problem. Your job is to **provide a characterization of the AFV market in the US and in Washington, including data about energy efficiency of the vehicles over time, geographical distribution, etc.**

Analytical Context. You will be working with the `vehicles.xlsx` (/extended.eda_basic_data_wrangling_excel_fellow/files/data/vehicles.xlsx) Excel file that contains four tabs of raw data. The first one is the `sales` tab, which contains aggregated sales figures for the US; the second one is the `models` tab, which contains data on the EV type. The third tab is the `range_data` tab, which has data about ranges of electrical vehicles. These data files come from the US Department of Energy. The last one is `ev_data` tab, which is a census of all the electric vehicles in circulation in Washington state. This census data comes from the Washington State Department of Licensing.

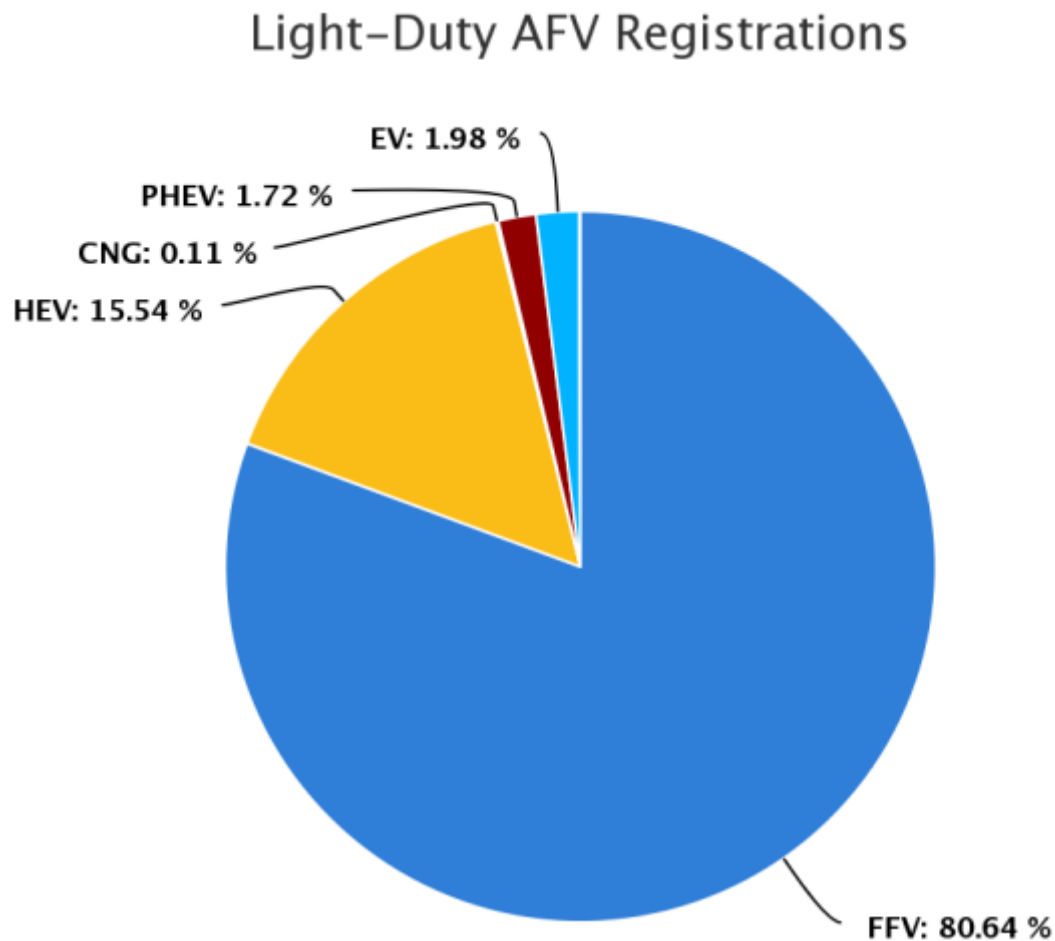
Instructions: For this extended case, you will only submit one Excel file with your answers written in the `ANSWER` tab. Each question will have its specific designated field that you will be placing your answers into. Continue for further instructions.

Alternative fuel vehicles in the US

Alternative fuel vehicles (AFV) (https://en.wikipedia.org/wiki/Alternative_fuel_vehicle) are vehicles that run either partially or totally on an energy source that is not gasoline or diesel, such as ethanol, solar power, electricity, or hydrogen. Some AFVs are highly efficient and, perhaps more

importantly, have reduced greenhouse gas emissions relative to traditional vehicles, which makes them an important part of environmental recovery and protection plans around the globe.

This chart shows the distribution of light-duty AFVs in the US as of Dec. 31, 2018:



Legend: *FFV*: Flexible fuel vehicle (cars that run on biofuels or mixtures of gas and biofuels - most of the FFVs in the US operate primarily on gasoline, see here (https://afdc.energy.gov/vehicles/flexible_fuel.html) for more information). *HEV*: Hybrid electric vehicle. *CNG*: Compressed natural gas. *EV*: Electric vehicle. *PHEV*: Plug-in hybrid electric vehicle. PHEVs and EVs are often counted as a single category, Plug-in electric vehicles (*PEVs*). **Source:** Alternative Fuels Data Center, US Department of Energy.

Note: Chart totals add to 99.99 due to rounding. This is not uncommon, but a good reason to avoid pie charts.

Exercise 1

Write a paragraph with an interpretation of this chart in your own words.

Instructions: Write your analysis in cell X6 in the ANSWER tab.

Use at least two evidence-based examples to support arguments. Your response should be clear and concise without grammatical and spelling errors.

Note: For response-based questions, answers should be clear and your reasoning should be supported with evidence. For instance, you can compare and contrast various types of EVs based on proportions of the market, and state reasons why these types of vehicles comprised the market in 2018.

Answer.

Exercise 2

Read this excerpt from this page (https://afdc.energy.gov/vehicles/flexible_fuel.html) about flexible fuel vehicles:

Flexible fuel vehicles (FFVs) have an internal combustion engine and are capable of operating on gasoline and any blend of gasoline and ethanol up to 83%. E85 (or flex fuel) is a gasoline-ethanol blend containing 51% to 83% ethanol, depending on geography and season. According to IHS Markit, as of 2017, there were more than 21 million FFVs in the United States. Because FFVs are factory made and are capable of operating on gasoline and gasoline-ethanol blends, many vehicle owners don't realize their car is an FFV and that they have a choice of fuels to use.

Given the additional information you gained from the excerpt, does the meaning you took from the graph change? Explain your answer.

Instructions: Write your analysis in cell X36 in the ANSWER tab.

Answer.

Exercise 3

Your team would like to know how the sales of fully electric vehicles compare to those of plug-in hybrid electric vehicles over time. The models worksheet has a list of all the models in the dataset with their type (either EV or PHEV) and the sales worksheet shows how many units were sold of each model in each year between 2011 and 2019.

Instruction:

1. Make a pivot table that has these rows and columns described above (insert it in tab Pivot Table 1 at cell A1).
2. Complete "Column D" in the sales tab.

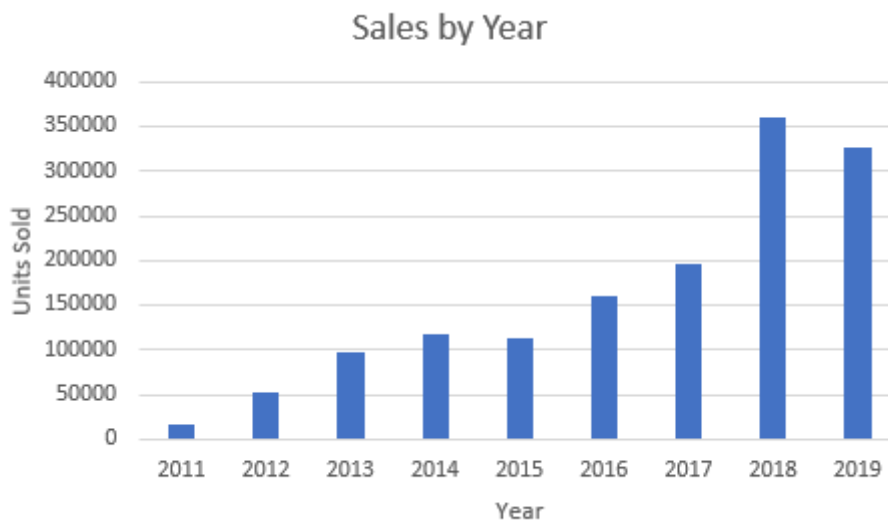
	A	B	C	D
1				
2	Sum of Units sold	Column Labels		
3	Row Labels	EV	PHEV	Grand Total
4	2011			
5	2012			
6	2013			
7	2014			
8	2015			
9	2016			
10	2017			
11	2018			
12	2019			
13	Grand Total			
14				

Hint: You will need to use the INDEX() and MATCH() functions.

Answer.

Exercise 4

The figures you have just calculated can be represented using the below plots. Provide a paragraph in which you explain in your own words what can be inferred from them:



US sales bar plot by type

Instructions: Write your analysis in cell X118 in the ANSWER tab.

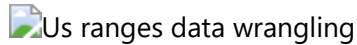
Respond to the question with at least two logical arguments. Mention derivations of two of the key points. Use at least two evidence-based examples to support arguments. Paragraph should be logical, easy to understand, and generally free of spelling errors.

Answer.

Exercise 5

Let's now investigate the ranges of electric vehicles in the US. The range of an electric vehicle is the theoretical number of miles it can travel without needing to be recharged.

The data can be found in the `range_data` tab. It contains range data for a number of EV models in the form of text strings. For example, `2013 Toyota - Scion Iq 038` means that the range of a Toyota Scion Iq model 2013 is 38 miles. Your job is to use Excel functions to create a table out of this data with the four columns `YEAR` , `MAKE` , `MODEL` , and `RANGE` , to end up with something like this:



Instructions:

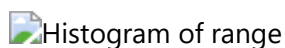
1. Split Column A using Text-to-Column Feature
2. Complete Column C
3. Complete Column D
4. Complete Column E
5. Complete Column F

Hint: Use the `FIND()` , `LEN()` , `RIGHT()` , `LEFT()` , `TRIM()` , and `VALUE()` functions to wrangle the data. It is also useful to start with `Text to Columns` , available under the `Data` tab, and split by `-` . Notice that all years have four characters and all ranges have three. You must recreate all four columns for credit, but as long as you can extract the year and mileage you will be able to complete exercises 6 and 7.

Answer.

Exercise 6

This is a histogram made with the data from Exercise 5:



6.1

Recreate this histogram on the same sheet where you have `range_data`.

Instructions: Be sure to consider each feature of the histogram:

1. Number of Bins (10)
2. Axis Labels
3. Title

Once the chart is complete. Copy the chart and paste it in the `ANSWER` tab, cell `X198` .

6.2

There is a gap in the interval from ~147 to ~228 miles. What does that mean? What could be causing it?

Instructions: Write your analysis in cell X228 in the ANSWER tab.

Respond to the question with at least two logical arguments. Mention the derivations of two of the key points. Use at least two evidence-based examples to support arguments. Paragraph should be clear, easy to understand, and generally free of spelling errors.

Answer.

Exercise 7

7.1

Create a pivot table showing the average range for each year.

Instructions: Make a pivot table that has these rows and columns (insert it in Pivot Table 2 at cell A1).

7.2

Use your pivot table to answer this question: How has the average range changed over time? (Do newer cars have longer ranges?)

Instructions: Write your analysis in cell X295 in ANSWER tab.

Respond to the question with one or two logical arguments. The response should be clear, easy to understand, and generally free of spelling errors.

Answer.

Electric vehicle population in Washington state

Let's now move on to studying the population of electric vehicles in Washington State using pivot tables. The relevant tab is `ev_data` tab.

Note: There are vehicles with states other than Washington in the "State" field. There is no error in the dataset. It contains all EVs that were *registered* in Washington in the given period. The "State" field is the state where the owner *resides*. There are vehicles that were registered in Washington but whose owners live somewhere else. See the data dictionary here (<https://web.archive.org/web/20220610203945/https://data.wa.gov/Transportation/Electric-Vehicle-Population-Data/f6w7-q2d2>).

Exercise 8

Complete the following two questions by creating a pivot table for each question. Both pivot tables are placed in their own separate sheet titled `Pivot Table 3` and `Pivot Table 4`. Use the output to answer each question in a complete sentence.

8.1

What are the 5 most popular makes in Washington?

Instructions:

1. Make a pivot table in the `Pivot Table 3` sheet at cell `A1`.
2. Provide a quick summary of the top 5 makes in Washington in cell `X325` of the `ANSWER` tab.

Answer.

8.2

What are the most common model years in Washington? (The answer to this question would be the 5 years that had the most models.) Does this result and that of the previous exercise suggest that cars in Washington tend to have long ranges?

Instructions:

1. Make a pivot table in the `Pivot Table 4` sheet at cell `A1`.
2. Provide a quick summary of the top 5 makes in Washington in cell `X362` of the `ANSWER` tab.

Answer.

Exercise 9

This is a histogram of vehicles in Washington state according to their ranges.



Histogram of ranges (Washington)

9.1

Recreate this histogram and place it to the right of the data on the `ev_data` sheet.

Instructions: Be sure to consider each feature of the histogram:

1. Number of Bins (10)
2. Axis Labels
3. Title

Once the chart is complete, copy the chart and paste it in the `ANSWER` tab, cell `X399` .

9.2

Is this histogram consistent with the hypothesis that cars there tend to have long ranges?

Note: While the histogram you analyzed in Exercise 6 showed the distribution of ranges from a list of *models*, this histogram shows you the distribution from a list of *individual vehicles*, and therefore the two histograms are not comparable. The dataset from Exercise 6 had 65 models, while the Washington state dataset has 76,767 vehicles.

Instructions: Write your analysis in cell `X429` in the `ANSWER` tab.

Answer.

Submission Instructions

Be sure to turn in a single Excel file. Your excel files should only have the following tabs:

`ev_data` , `range_data` , `models` , `sales` , `Pivot Table 1` , `Pivot Table 2` , `Pivot Table 3` , `Pivot Table 4` , `ANSWER` , `TA Feedback` and `TA ONLY` .

Attribution

"Light-Duty AFV Registrations", Aug 2020, Alternative Fuels Data Center, US Department of Energy, Public Domain (<https://www.energy.gov/about-us/web-policies>),
<https://afdc.energy.gov/data/10861> (<https://afdc.energy.gov/data/10861>)

"Electric Vehicle Population Data", Aug 2021, Washington State Department of Licensing (DOL), DOL Conditions of Use (<https://www.dol.wa.gov/conditions.html>)
<https://catalog.data.gov/dataset/electric-vehicle-population-data>
(<https://catalog.data.gov/dataset/electric-vehicle-population-data>)

"U.S. Plug-in Electric Vehicle Sales by Model", Jan 2020, Alternative Fuels Data Center, US Department of Energy, Public Domain (<https://www.energy.gov/about-us/web-policies>),
<https://afdc.energy.gov/data/10567> (<https://afdc.energy.gov/data/10567>)

"Average Range and Efficiency of U.S. Electric Vehicles", Jan 2020, Alternative Fuels Data Center, US Department of Energy, Public Domain (<https://www.energy.gov/about-us/web-policies>),
<https://afdc.energy.gov/data/10963> (<https://afdc.energy.gov/data/10963>)