Term Project:

Washington State’s Electric Vehicle Population Data Repository

Marissa Fischer

LIS 545: Data Curation I: Fundamentals

February 23, 2024

### Data and Metadata Profile

The [Electric Vehicle Population Data](https://data.wa.gov/d/f6w7-q2d2), published by the Washington State Department of Licensing, is a public data set consisting of data on registered vehicles in Washington that are Battery Electric Vehicles (BEVs) or Plug-in Hybrid Electric Vehicles (PHEVs). This data set was created in April 2019 and was updated as recently as last week. Despite being consistently updated by Washington DOL, there is no indication of how far back registration data is being compiled in this data set.

Key stakeholders for this data set include Washington Department of Licensing and Washington Department of Transportation, as well as the departments on the federal level, for tracking vehicle registration trends. This data is also relevant for the state and federal Departments of Ecology or the EPA to help monitor the rate of transition to “clean fuel” that residents in certain areas are making.

There are four data files listed on Data.gov for this particular data set, in CSV, RDF, JSON, and XML formats, and all have the option to download them directly from the page. There is also a link to the Data.WA site that initially published this data set and gives more information, including an easily read table version of the data set that does not require any downloads (see Fig. 1 below). The .csv file will open in Notepad if Excel is not available on your computer and both the JSON and XML files will open in an internet browser window. The RDF file was able to be downloaded, but it requires either the Firefox internet browser or additional programs to open it. There are no listed restrictions for this data set and both Data.gov and Data.WA indicate that it is publicly available and does not contain any federal information that would be subject to more extensive restrictions.

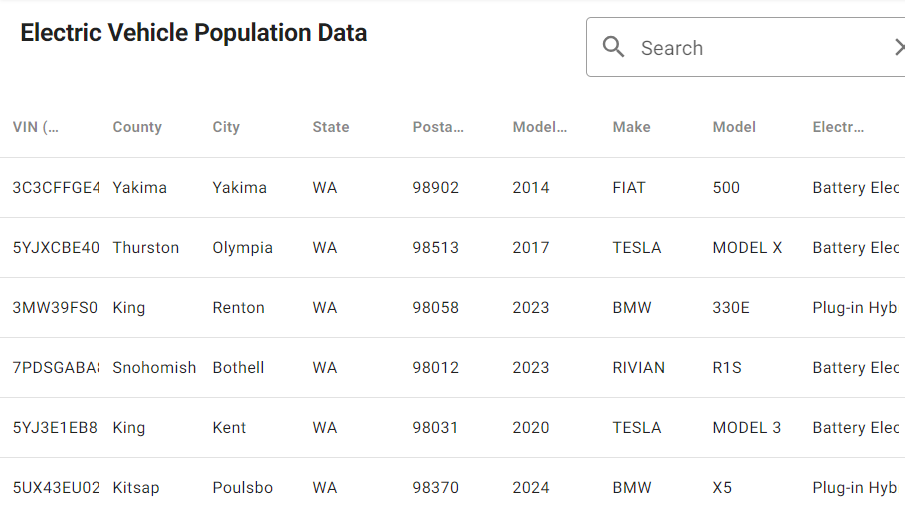


Fig. 1: Snapshot of data set table provided on Data.WA

Available metadata on the data set includes created date, updated date, publisher and maintainer information, public access level, schema version, licensing information, links for cataloging and context, etc. All of this metadata is mapped out in a table on Data.gov and Data.WA, but does not exist within the downloadable files. Further context on the data is given in the data set files through attributes, such as VIN, County, Make, Model, etc., for the records.

While the data set seems very straightforward, no context is given for what the data is used for or by whom, other than Washington DOL. Some metadata components, e.g., Harvest Source, are detailed enough for data scientists and curators, but would not be as easily understood by general users who were browsing the data set for various referencing reasons. The metadata is structured according to the Project Open Data Metadata Schema v1.1. The schema is linked in the metadata of the data set so that it can be easily referenced or replicated.

The current data could be improved by including the most recent date of registration, helping to show the scope of how electric cars are increasing in prevalence in the state and could help show movements patterns of BEV and PHEV owners if the registration was not renewed. Including the registration date would also allow other states to compare the data set with their own data based on the latest year of registration rather than having to rely on the model year of the car. This would remove any uncertainty of if the car was registered in Washington the year it was made or if it was moved to Washington or was bought as a used car years later. The data or metadata for this data set also does not indicate what updates are being made, beyond adding new records of registrations. Indicating in the description of the data set if inactive registrations are being left in or removed would help users better understand what sort of historical data is being retained and increase the potential for drawing any conclusions from the data.

On the Data.WA site page for this data set, there are other datasets listed that were made by referencing the Electric Vehicle Population Data. These datasets include ones, such as Electric Vehicle Population Maps, Most Common Registered Electric Vehicle Models, etc., that help expand on and try to make sense of the patterns in the original data set. Additional publications that reference Data.WA’s data set include briefs and grant documentation from Washington Department of Commerce, Washington State Department of Transportation, and the Washington Utilities and Transportation Commission. These publications use the data set to forecast clean fuel trends (WDOC 2022), establish a zero-emission vehicle infrastructure for the State of Washington (WSDOT 2023), and identify trends that could affect Washington’s energy supply (UTC 2023). There was also a published article that uses the data set as a repository for an MIT study on energy resources for the United States (Haider, Xu, Yang 2023). All of these publications were found through searching the URL of the data set or the data set name in Google. No results were found through Google Scholar or any UW databases, such as EBSCO or Environment Complete.

As a whole, Washington’s Electric Vehicle Population Data is a good data set to use for creating visualizations and connecting with other data to see greater trends. With some improvements, such as including more context on historical data or the date of the record, this data set could be applied to a broader selection of studies conducted by other states on the existence of BEV and PHEV registrations in their regions. It is surprising that this data set has not been leveraged more by the State of Washington, beyond the publications that were referenced previously, as there is extensive data in this data set that could bolster an increase in support for electric vehicles and their owners at the state level.

**References**

Data.gov (2024, January 19). *Electric Vehicle Population Data: Data.gov: United States General*

*Services Administration.* <https://catalog.data.gov/dataset/electric-vehicle-population-data>

Haider, R., Xu, Y., & Yang, W. (2023, June). Towards closing the data gap: A project-driven

distributed energy resource data set for the US Grid. In *Companion Proceedings of the*

*14th ACM International Conference on Future Energy Systems* (pp. 52-58).

Washington Department of Commerce (2022, September 23). *Clean Fuel Forecast:*

*Ecology.WA.gov: State of Washington.* [https://ecology.wa.gov/getattachment/4ada812c-](https://ecology.wa.gov/getattachment/4ada812c-a8b5-41c5-bcb8-3a724f8e5b27/20220923CleanFuelForecastRpt.pdf)

[a8b5-41c5-bcb8-3a724f8e5b27/20220923CleanFuelForecastRpt.pdf](https://ecology.wa.gov/getattachment/4ada812c-a8b5-41c5-bcb8-3a724f8e5b27/20220923CleanFuelForecastRpt.pdf)

Washington State Department of Licensing (2024, January 16). *Electric Vehicle Population*

*Data: Data.WA: State of Washington*. <https://data.wa.gov/d/f6w7-q2d2>

Washington State Department of Transportation (2023, March 9). *Zero-Emission Vehicle*

*Infrastructure Partnership Program, 23-25 Grant Guidelines: WSDOT.WA: State of Washington*. [https://wsdot.wa.gov/sites/default/files/2023-03/WSDOT-ZEVIP-Grant-](https://wsdot.wa.gov/sites/default/files/2023-03/WSDOT-ZEVIP-Grant-Webinar-Slides.pdf)

[Webinar-Slides.pdf](https://wsdot.wa.gov/sites/default/files/2023-03/WSDOT-ZEVIP-Grant-Webinar-Slides.pdf)

Utilities and Transportation Commission (2023, February 7). *Issue Brief #6, The Mobility*

*Transition: UTC.WA: State of Washington.* [https://www.utc.wa.gov/sites/default/files/](https://www.utc.wa.gov/sites/default/files/2023-02/Issue%20Brief%20%236%20-%20The%20Mobility%20Transition.pdf)

[2023-02/Issue%20Brief%20%236%20-%20The%20Mobility%20Transition.pdf](https://www.utc.wa.gov/sites/default/files/2023-02/Issue%20Brief%20%236%20-%20The%20Mobility%20Transition.pdf)

### Repository Profile

The [Alternative Fuels Data Center (AFDC)](https://afdc.energy.gov/) is a data repository hosted by the United States Department of Energy’s Office of Energy Efficiency and Renewable Energy. It was released in 1991 in response to legislation surrounding alternative fuel and the Clean Air Act Amendments of 1990 and hosts a variety of data centered around advanced transportation technologies. The repository stores and presents data, maps, publications, case studies, etc., to track alternative energy measures and support the case for reduced petroleum use in the United States.

The AFDC is best suited for hosting Washington State’s *Electric Vehicle Population Data* because of the existing content and focus of the repository. Existing data on the AFDC site includes profiles by state, including one for Washington, with fueling station data, laws impacting the use of alternative fuels, fuel compensation over time, 2022 vehicle registration totals, etc. Figure 1 below is a snapshot of Washington’s profile page and the visualizations that the National Renewable Energy Laboratory (NREL) creates based on the data compiled in the repository.

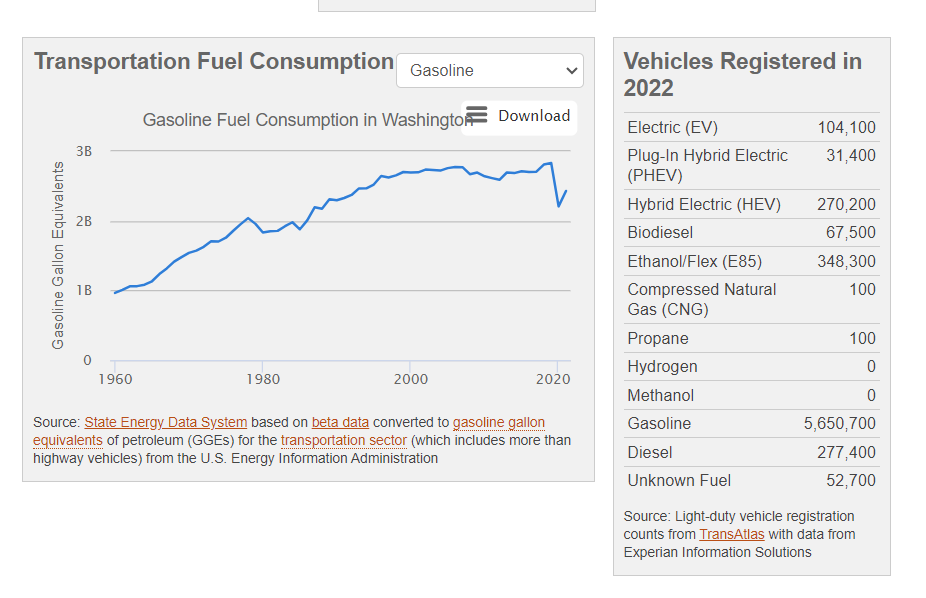


Fig. 1: Example of information provided on Washington’s profile in the AFDC Repository.

Other general visualizations are included in the repository, covering information for all of the United States, such as [*Electric Vehicle Registrations by State*](https://afdc.energy.gov/data/10962), but they do not go into detail about the models registered or the trends outside of the previous year’s data. The incorporation of *Electric Vehicle Population Data* into the AFDC repository would help supplement the more generalized information on the site to help improve the ability to draw connections between renewable energy data within Washington and with other states across the country.

The AFDC repository does not accept submissions from anyone, as their data is gathered from projects completed in the National Renewable Energy Laboratory (NREL) and by partners of the Clean Cities coalition. Projects from outside these groups can apply to receive funding and data publication through their project assistance program that is offered with Clean Cities. Each piece of data featured in the repository comes from a different source that has been analyzed by the NREL. The visualization mentioned in the above paragraph, *Electric Vehicle Registrations by State*, is an example of this, as it is cited as coming from the National Renewable Energy Laboratory with data collection by Experian Information Solutions.

No account is needed to download or view the data, but there are certain steps to go through to get access to the raw metadata that is used and stored in the repository. On the AFDC site, visualizations and other data have the option to be viewed as a graph, image, etc., or be downloaded as an Excel file directly from the page they are being viewed from. These Excel downloads are not datasets, but lists of the values, totals, etc., used to create the corresponding visualization, including the visualization that was generated to represent the data. For the full data set of all records contributing to the analysis presented in AFDC, a separate webpage can be navigated to and the form (shown in Figure 2 below) needs to be filled out before being able to download the data. Datasets for fuel stations, laws, and vehicle registrations can be downloaded through this form submission as a CSV file, but the download does take a long time depending on the different options that were selected on the form. There seems to be no approval or waiting period and AFDC indicates that the collection of contact and use information is for tracking the use agreement conditions listed in the terms that are linked on the form.

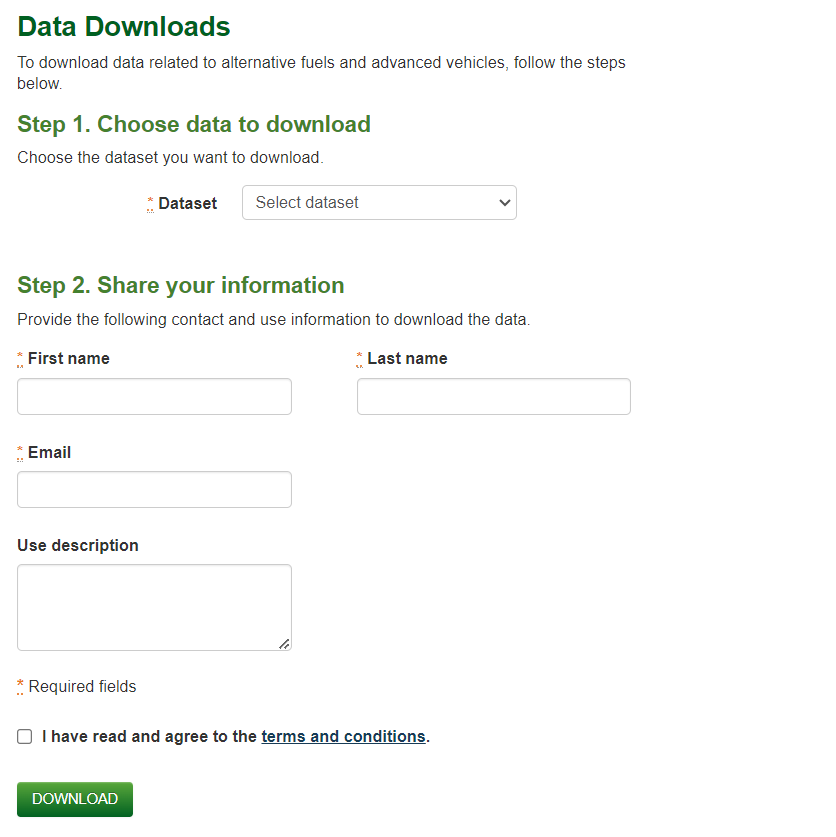


Fig. 2: Basic form for downloading datasets as .csv files from AFDC.

A sample of the data in the repository was downloaded using the form provided on the AFDC site, focusing on light-duty vehicle registrations for Battery Electric Vehicles and Plug-in Hybrid Electric Vehicles (the types of vehicles featured in *Electric Vehicle Population Data*) for all makes and models for 2021. The form does not allow the query to be drilled down to a specific state or region and upon looking at the data set that was downloaded, no indication of location is given for any of the records. It is not clear if this data does not exist in the AFDC or if it is not available for download because it is considered “identifying information” that the data use terms prohibit. This method of direct file download is the only option provided on the repository and no metadata standard is given. The page for AFDC on re3data.org does have a link to a “data citation guideline” for the repository on Energy.gov, but when clicking the link, it results in an access denied page.

As a whole, the Alternative Fuels Data Center is a suitable home for the specific focus of the *Electric Vehicle Population Data* data set, but it is by no means a perfect repository. The scope and potential for use for government departments is aligned for both the repository and data set, but the non-existent submission access does make it difficult for datasets to be added to AFDC. While this may pose a problem for expanding data in AFDC in the future, *Electric Vehicle Population Data* is a data set published by a state government, with connections to the U.S. Department of Energy, so it should be fairly easy for Washington to have this data uploaded to AFDC.

**References**

Alternative Fuels Data Center (2023, July). *Electric Vehicle Registrations by State*:

AFDC.Energy.gov: U.S. Department of Energy. <https://afdc.energy.gov/data/10962>.

Registry of Research Data Repositories. *Alternative Fuels Data Center:* re3data.org.

<https://www.re3data.org/repository/r3d100010226>.

Registry of Research Data Repositories. *Washington Transportation Data for Alternative Fuels*

*and Vehicles:* AFDC.Energy.gov: U.S. Department of Energy.

<https://afdc.energy.gov/states/wa>.

Washington State Department of Licensing (2024, January 16). *Electric Vehicle Population*

*Data: Data.WA: State of Washington*. <https://data.wa.gov/d/f6w7-q2d2>.

### GitHub Repository

Recommended properties for data citations from *Electric Vehicle Population Data* (the data set and its corresponding GitHub repository) include, publisher, year published, title, type, date accessed, and identifier. The publisher information, title, and type, are necessary to help orient the data as belonging to a state government published data set on registered electric vehicles. Since this data set is continuously updated, including the accessed date helps indicate what version of the data set was referenced and including a permanent link to the data will help locate and update the data set in the future. The formatted example below would be a complete citation for this data set based on the recommendations discussed in this paragraph.

Publisher (Publication Year). *Title* [Type]. Accessed Date. Identifier.

E.g., Washington State Department of Licensing (2023). *Electric Vehicle Population Data* [data set ]. Accessed February 22, 2024. <https://data.wa.gov/api/views/f6w7-q2d2>.

Preservation recommendations for this repository include regular updating of data set files and increased documentation on the full scope of the data based on additional research/outreach to the data set publisher(s) (e.g., if records are removed, what the full coverage of dates include, etc.). The uploaded data set is continuously updated on Data.gov and Data.WA, making the current data in the GitHub repository out of date as soon as additional data is added on those sites. Since the data is included in CSV, RDF, JSON, and XML file types, the data has more versatility, as far as what software is needed to view the data, and this will increase data longevity. Something of note for this specific data set are the repository file size limitations. All file formats, except for the RDF file, needed to be uploaded in a ZIP file so that they were under the maximum file size that GitHub allows. This could pose an issue if users do not extract the files once they have been downloaded from the repository.

The copyright license that is listed for this data set on Data.gov and Data.WA is the Open Data Commons Open Database License (ODbl). This license allows for the free use, modification, and sharing of the licensed database. The licensing notes that this covers only rights and usage of the database, not its contents, which means the data in this data set has no current copyright licensing attributed to it. Since the data set is public and meant to be openly accessed and used for various different purposes, the same ODbL copyright license should be applied to the GitHub repository so that it remains accessible to anyone that wants to use it. The Open Data Commons Open Database License also requires acknowledgement of where the data comes from, which would help link any data usage.

No personally identifiable data about the owners of the vehicles exist in this data set. VIN and the DOL Vehicle ID for each record are included in the data set, but require further inquiry at the Washington Department of Transportation, including the submission of a request form and payment of a fee, are needed to retrieve any associated ownership data that is tied to the vehicle. There are additional sites that the Washington Department of Licensing links to that provide VIN lookup services, but they are either available for purchase with limited information (e.g., the number of owners, accident history, etc.) with full reports available through licensed car dealerships only. Based on this factor and that Washington State publicly published the data set with no PII restrictions, there was no need to adjust the data before uploading it into the repository.

**References**

Data.gov (2024, February 17). *Electric Vehicle Population Data: Data.gov: United States*

*General Services Administration.* [https://catalog.data.gov/dataset/electric](https://catalog.data.gov/dataset/electric-vehicle-population-data)

[-vehicle-population-data](https://catalog.data.gov/dataset/electric-vehicle-population-data).

National Motor Vehicle Title Information System. *Research Vehicle History: NMVTIS*:

<https://vehiclehistory.bja.ojp.gov/nmvtis_vehiclehistory>.

Open Data Commons. *Open Data Commons Open Database License (ODbL) v1.0:* Open Data

Commons: The Open Knowledge Foundation. [https://opendatacommons.org/licenses](https://opendatacommons.org/licenses/odbl/1-0/)

[/odbl/1-0/](https://opendatacommons.org/licenses/odbl/1-0/).