

Dataset 1: vehicle_images

This dataset is too large to upload to GitHub, but it can be accessed and downloaded at Mendeley Data [1]. Upon downloading the dataset, it is in a zip folder, with 5 subfolders. These subfolders contain image data for different vehicle classes. These classes are detailed below, with example images for each class. The unit of observation of this dataset is simply an image classified by the subfolder.

Hatchback:

■



Other:

■



Pickup:

■



Sedan:

■



SUV:

■



Dataset 2: vehicle_images_dataset.csv

Since the previous dataset is too large to import into GitHub, it was mounted to Google Drive and then connected to Google Colab's "content" folder to save resources. A variable was then created which is a file path directing to the content folder holding the image, which is the "image_path" variable, and the corresponding image's subfolder was then created as a new column titled "label". The unit of observation is a vehicle image, where each row pertains to one vehicle image.

Below, there is an example subset of the dataset, vehicle_images_dataset.csv. The specific variables will then be discussed.

1	image_path	label
2	/content/vehicleimages/hatchback/PHOTO_89.jpg	hatchback
3	/content/vehicleimages/hatchback/PHOTO_279.jpg	hatchback
4	/content/vehicleimages/hatchback/PHOTO_315.jpg	hatchback

image_path:

This variable takes on string values and represents the path to access the image content when the previous dataset (Dataset 1) is mounted to Google Colab. Summary statistics for this variable can be found below.

	image_path
count	4356
unique	4356
top	/content/vehicleimages/sedan/PHOTO_744.jpg
freq	1
dtype: object	

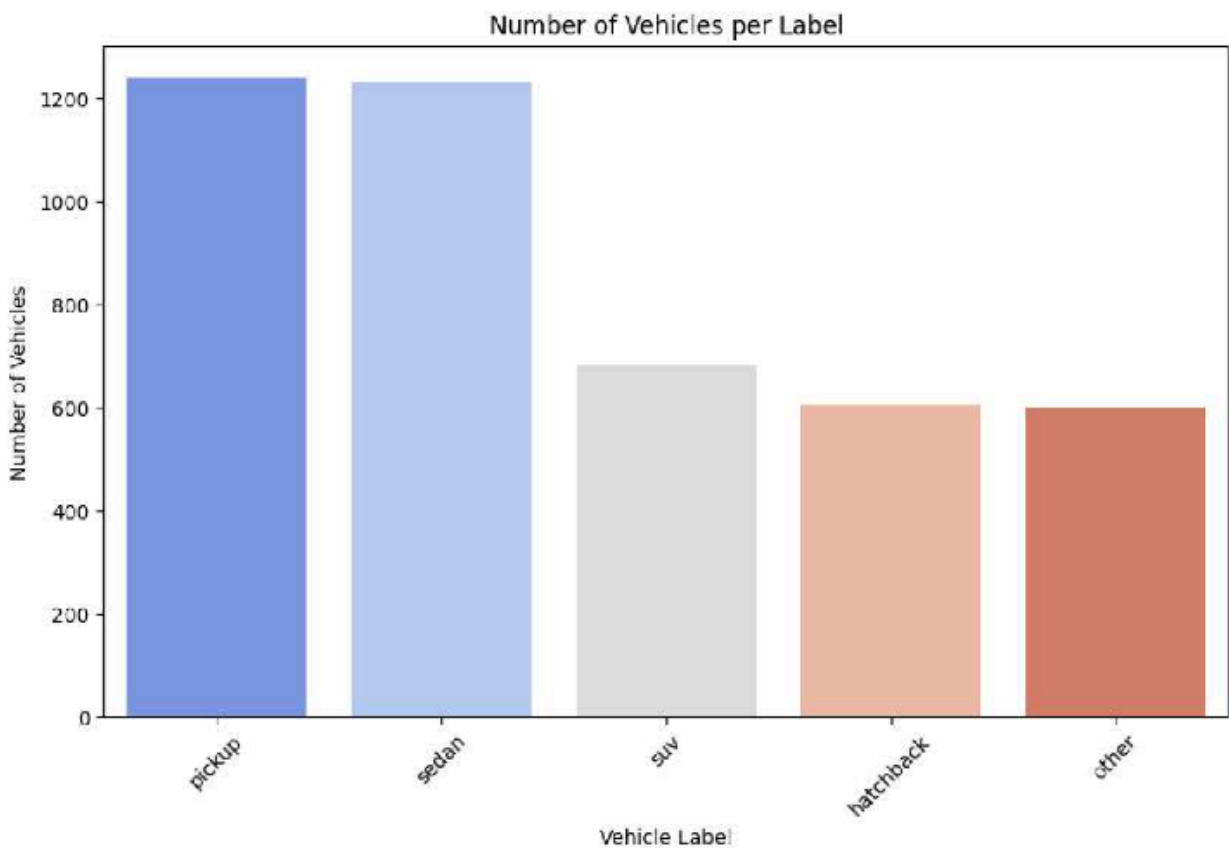
label:

This variable takes on string values representing the vehicle type of the image being represented. Summary statistics for this variable can be found below.



	label
count	4356
unique	5
top	pickup
freq	1240

dtype: object



References:

- [1] Narong Boonsirisumpun, "Vehicle Type Image Dataset (Version 2): VTID2," *Mendeley Data*, vol. 2, no. 10.17632/htsngg9tpc.2, Nov. 2021, Accessed: Nov. 06, 2024. [Online]. Available: <https://data.mendeley.com/datasets/htsngg9tpc/2>