Question/Need:

#vanlife is all the rage these days; I tried to buy into it in 2019, and now two-and-a-half years later, I am able to accept that it really isn't for me. I own a Sprinter that I converted myself over 2019 and 2020, and now I am ready to sell it and move on. The trouble is that I have no idea what it's really worth, between all the features it has and doesn't have, it's vehicular characteristics, and so on. Therefore, I am proposing to build a regression model that will allow me to predict the value of my van as it is, comparing it's load out to others currently for sale.

Data Description:

•What dataset(s) do you plan to use, and how will you obtain the data?

- I plan to create my own data set from vanlifetrader.com.
- I will scrape the entirety of the postings available on the site, currently sitting at ~510 rows representing more than 50 features.
- I will obtain this data by creating a web-scraping pipeline for vanlifetrader to collect and store the data as a .csv.
- If I have the time, I will create an additional scraping pipeline for thevancamper.com, following the same process as for vanlifetrader. This addition will (probably) get me close to 1000 data points combined.

•What is an individual sample/unit of analysis in this project? What characteristics/features do you expect to work with?

- Location
- Make
- Model
- Year
- Mileage
- Battery capacity
- Drive-train
- Heater present?
- Running water?
- Sleeping capacity
- Solar capacity
- Fuel type/MPG
- Asking price
- Freshwater capacity
- Potentially many more

•If modeling, what will you predict as your target?

- Asking price

Tools:

•How do you intend to meet the tools requirement of the project?

- I intend to use Selenium + BeautifulSoup to scrape vanlifetrader, as well as clean data for Pandas
- Pandas to explore and further clean/prepare data
- statsmodels to create regression metrics
- Seaborn to visualize

•Are you planning in advance to need or use additional tools beyond those required?

- I sure hope not, this sounds like a lot already.

MVP Goal:

•What would a MVP Example look like for this project?

- I'll use two or three features – simple ones like mileage, location, and make/model – compared against asking price to model the relationship between these two features.
- I hope to include a prediction interval.