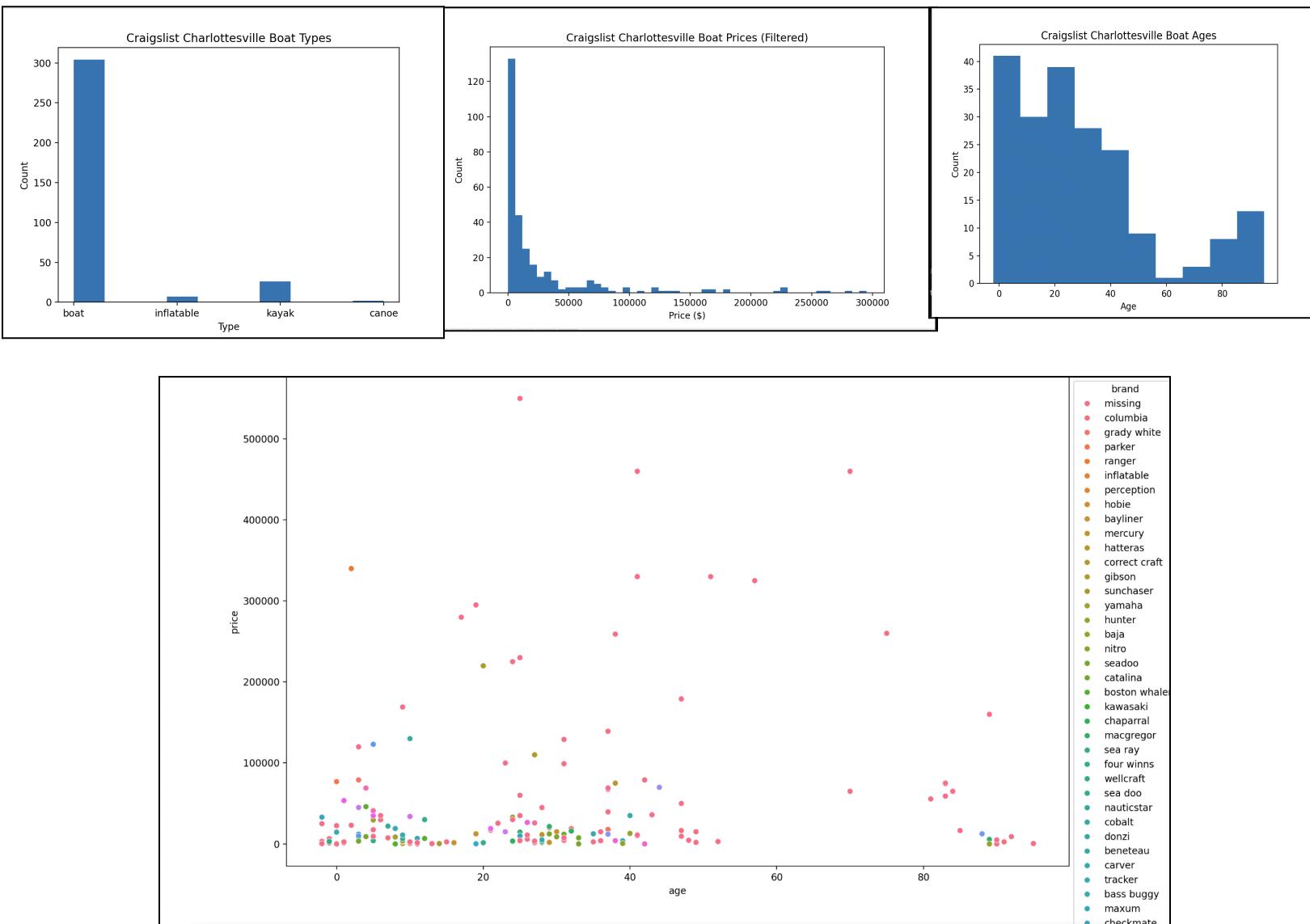


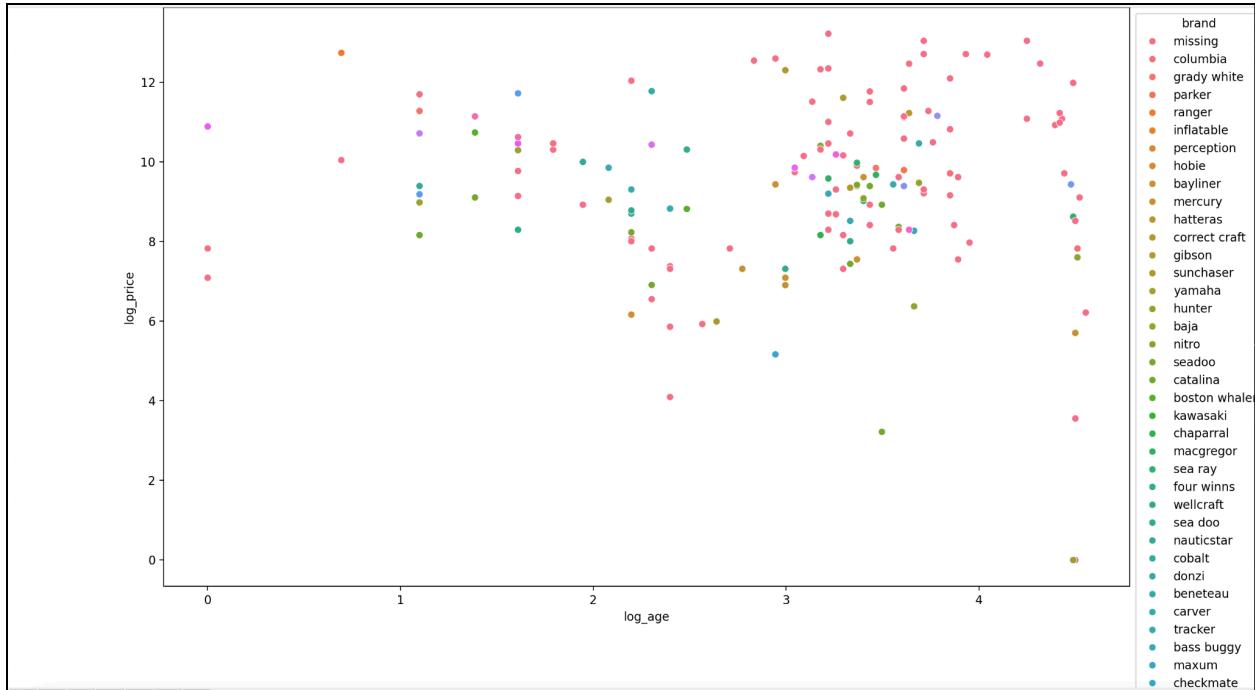
Graphs + Charts Produced for Lab 1: Scraping of Craigslist's Charlottesville Boats

Ananya Kumar



Scatterplot of Age vs Price for Craigslist Charlottesville Boats

Continue down



Scatterplot of Log_Age vs Log_Price for Craigslist Charlottesville Boats

```
log_price      log_price      log_age
log_price    6.332891 -0.082554
log_age     -0.082554  1.009104
```

```
log_price      log_price      log_age
log_price    1.000000 -0.038481
log_age     -0.038481  1.000000
```

brand	count	price	mean	std	min	25%	50%	75%	max
alberg	2.0	8250.000000	6010.407640	4000.0	6125.0	8250.0	10375.0	12500.0	
aqua patio	1.0	33900.000000	NaN	33900.0	33900.0	33900.0	33900.0	33900.0	33900.0
baja	1.0	33000.000000	NaN	33000.0	33000.0	33000.0	33000.0	33000.0	33000.0
bass buggy	1.0	19000.000000	NaN	19000.0	19000.0	19000.0	19000.0	19000.0	19000.0
bayliner	4.0	7650.000000	7122.967546	1200.0	1725.0	7200.0	13125.0	15000.0	
... tidewater	2.0	44250.000000	13081.475452	35000.0	39625.0	44250.0	48875.0	53500.0	
tracker	5.0	8820.000000	3865.488326	5000.0	6800.0	6800.0	11000.0	14500.0	
trophy	4.0	19000.000000	0.000000	19000.0	19000.0	19000.0	19000.0	19000.0	
wellcraft	3.0	4250.000000	2165.063509	3000.0	3000.0	3000.0	4875.0	6750.0	
yamaha	7.0	16657.142857	16468.010115	700.0	7950.0	8500.0	21750.0	48000.0	

[61 rows x 8 columns]

```
count      297.000000
mean     27410.552189
std      50449.083929
min     125.000000
25%    1500.000000
50%    7950.000000
75%    25500.000000
max    29500.000000
Name: price, dtype: float64
```

Analysis below:

Analysis/Story:

Most listings are for boats, rather than more specific categories like kayaks, canoes, or inflatables, which means that Craigslist is likely less for smaller vehicles and more for mainstream/motorized boats. Prices are greatly right-skewed, with most boats on the lower-priced end, with the most expensive boat at the \$295,000 price, as indicated by the chart on the page above. Most boats are around 10-40 years old and then this peaks again around the 80 year old end, showing that Craigslist has both modern/new boats on sale and also more vintage boats. The correlation between age and price is -0.03 to -0.08 , which shows almost no relationship between age and price, as supported by the two scatterplots, as well (no visible trend). So older boats can be expensive or cheaper; newer boats can also be expensive or cheaper. This probably means that brand is likely a stronger indicator of price, with Aqua Patio, Baja, and Tidewater being noticeably on the more expensive side. However, there is a pretty wide variety of boats offered on Craigslist!