

Jacob Clement

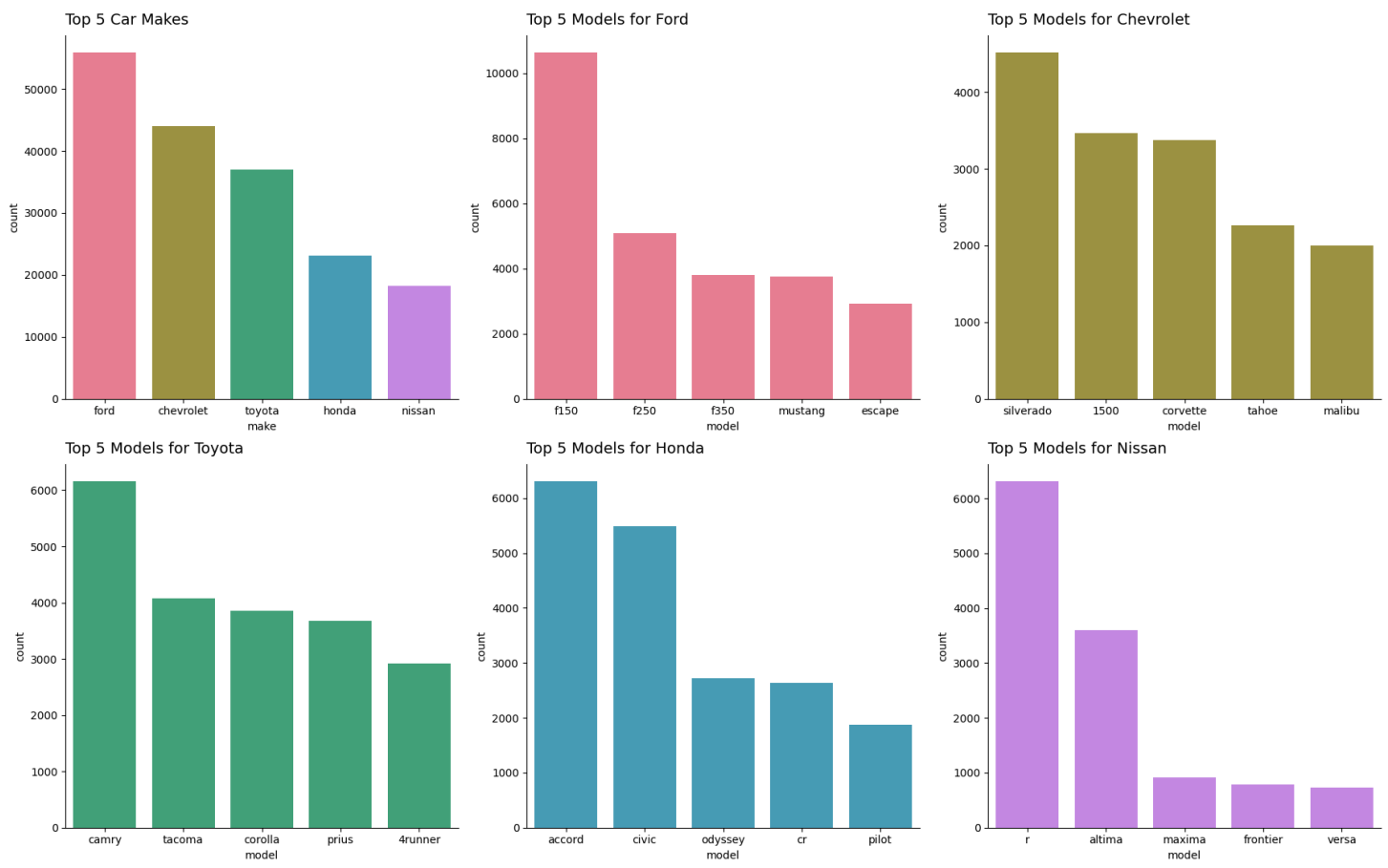
Data Visualization

9/23/2024

The submissions for: (Easy) Build a visualization that illustrates what makes and models of cars are most popular. is graphs\top_makes_and_models_combined.png.

<https://chatgpt.com/share/66f118dd-41e0-800a-b8a6-3ca307669811>

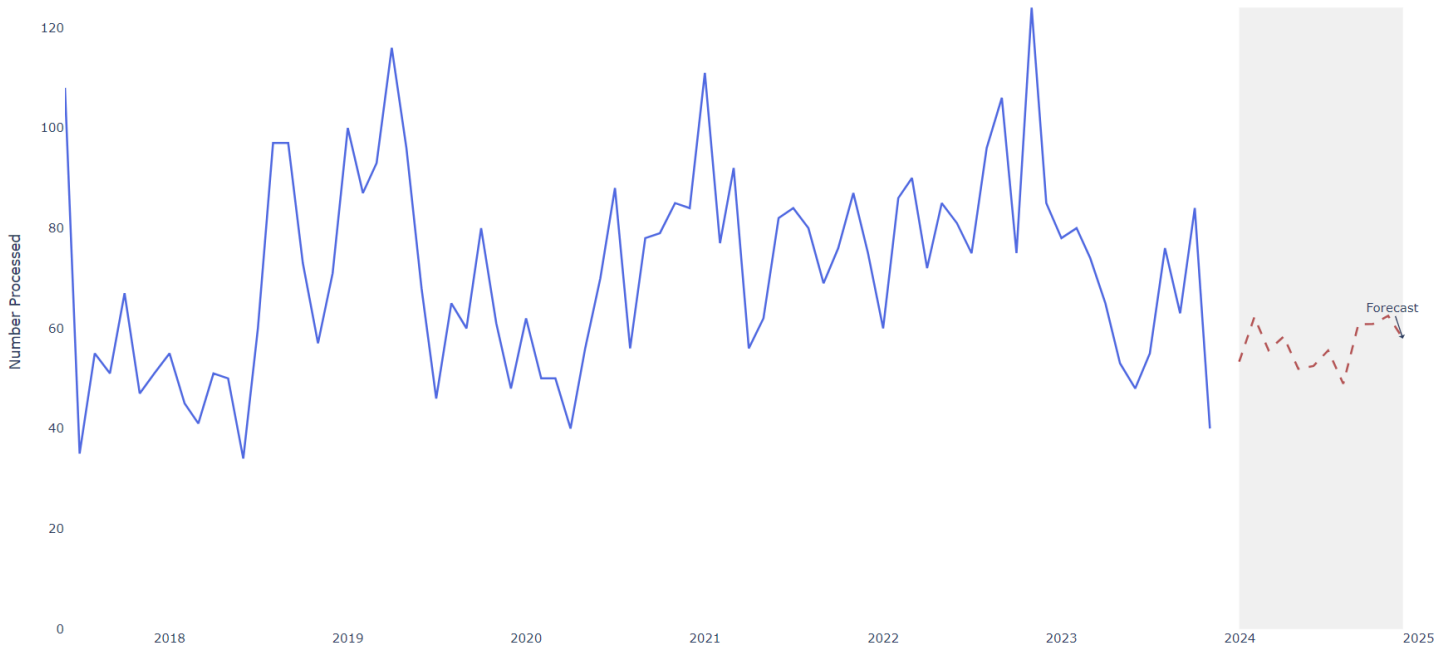
Top 5 Makes and Their Top 5 Models



I used plotly to create the submissions for: (Easy) Build a visualization that visualizes the annual pattern of entries into the HMIS. The product is html\better_line_chart.html. I took a snip as well and saved it in the graphs folder

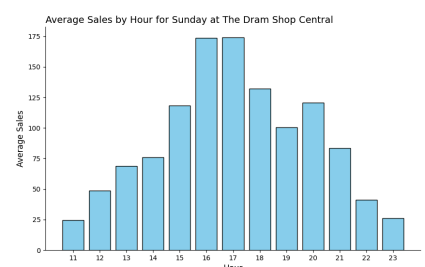
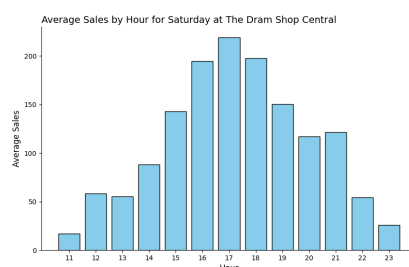
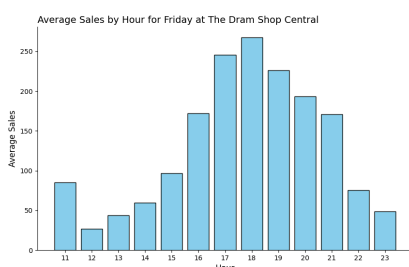
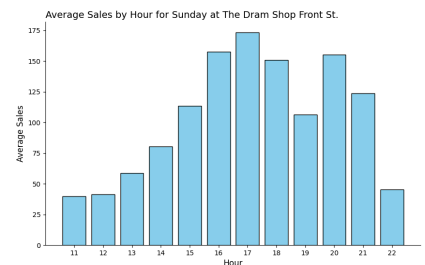
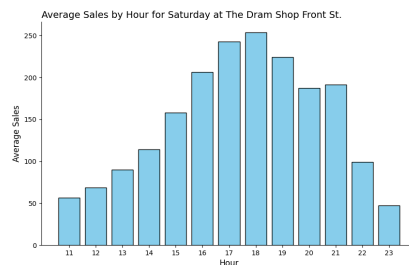
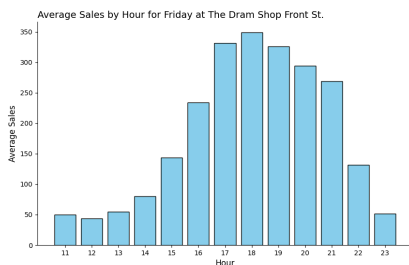
<https://chatgpt.com/share/66f19463-ebd0-800a-bec1-e114d331529f>

Trends in Homeless Management Information System (HMIS) Entry



For (Easy) Build a visualization that illustrates the patterns in purchases by time of day for both the Front Street and Central locations. I made way too many graphs. I just used matplotlib and loops. I did discover several interesting characteristics of the data. There are some hours with negative sales. Also, there are some transactions outside of normal business hours. To make the visualization consistent and simple, I did some filtering and included only two of the locations for a limited set of hours. I want to add the overall title "Similar Weekend Business Patterns Across Locations" but neither ChatGPT or I could figure it out at the time of the writing.

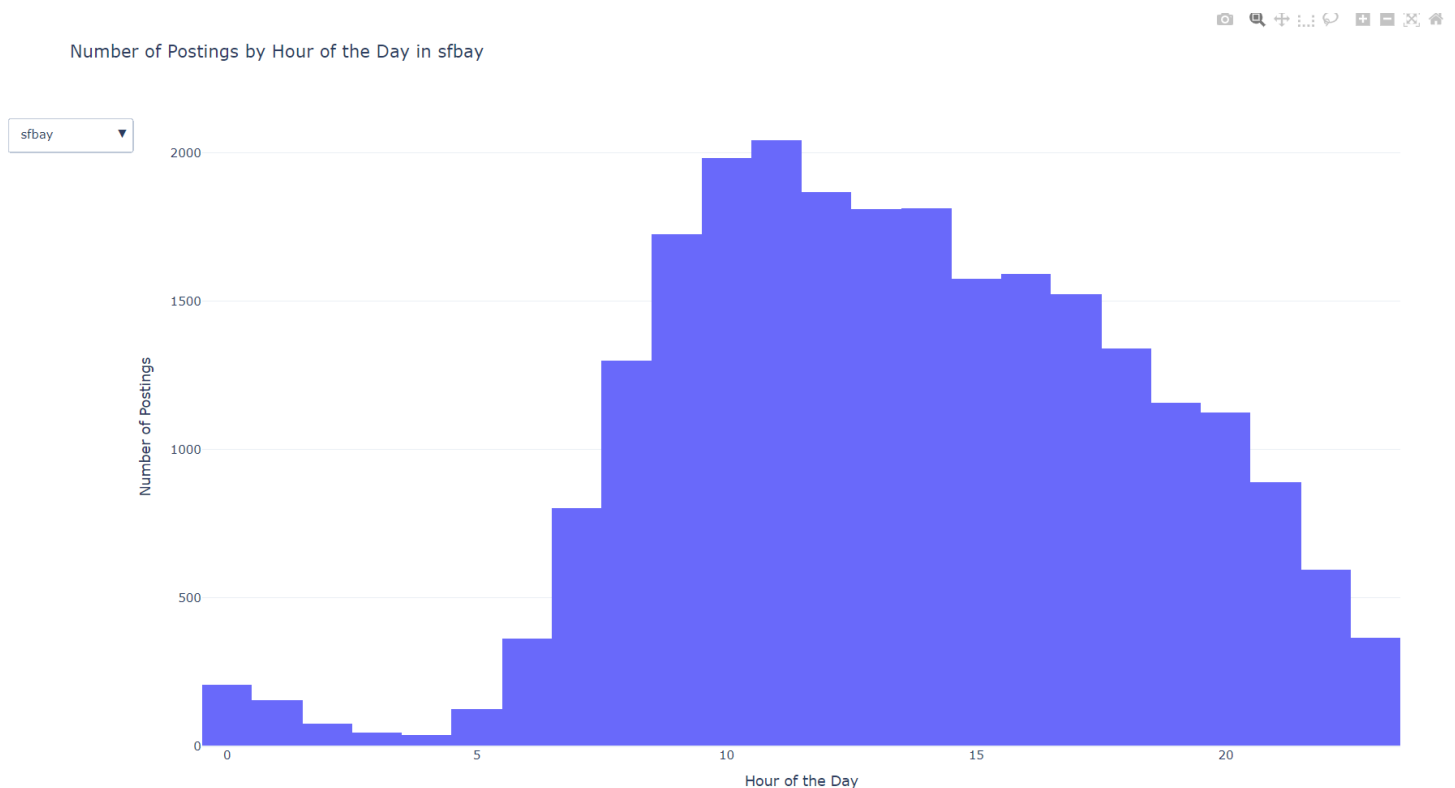
<https://chatgpt.com/share/66f1aed3-6c2c-800a-9dbf-364ba58009de>



On (Medium) Build a visualization that attempts to answer this question: When do cars tend to be posted by location? I once again made way too many graphs but this time they put together in an effective way. A look at the overall data shows that the distributions of postings by day of the month and day of the week are relatively uniform. There might be an increase in postings on Saturday and there is possibly a spike in posting near the end of the month. The histogram of postings by hour shows a near normal distribution of postings centered on the daylight hours. Therefore, I created visualizations by hour for the final product as it seems to be the most interesting.

The screen capture below is just a static view of the interactive plot. You can select from the top 25 locations and the view that locations histogram of postings. The html is here:
html\interactive_histogram_by_location.html.

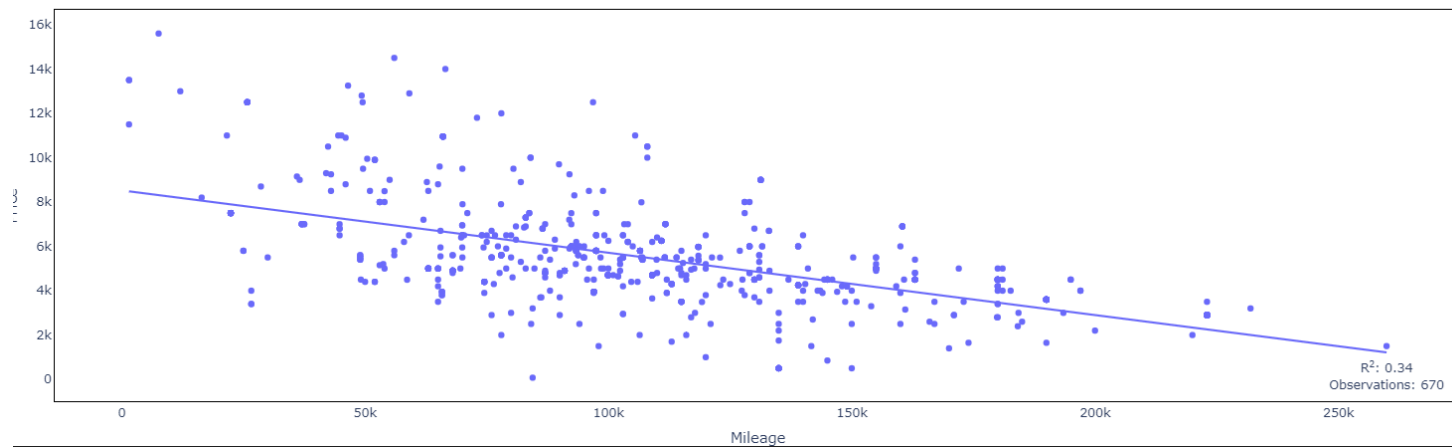
<https://chatgpt.com/share/66f1c967-e524-800a-ab04-ecb887bea409>



(Difficult) Create a visualization that captures the interplay between age and mileage for a given make and model. For the final submission, I present another dynamic visualization using plotly out of jupyter notebooks. I had this one running and then broke it by having ChatGPT adjust too much. I reverted to an earlier script and the program appears to be working. There are certainly issues with the graphs and program functionality but I am submitting the working version and will continue to work on it for revision. The screen cap below is one static visual from the program at data_vis.ipynb.

<https://chatgpt.com/share/66f22288-88d4-800a-9ab0-066ad3ca222a>

ford fiesta Price vs Mileage



I created a Github repo here <https://github.com/jacobchrono/data-visualization>. I just made the repo public so you should be able to access and see the python code if you would like.