**Car Ownership Trends, and their Impact on Accidents**

Introduction:

I was watching a documentary about self driving cars the other day, called ‘autonomous’. Orchestrated, or directed, by self-styled intellectual and futurist, Malcolm Gladwell. The same guy that really popularized the case approach to non-fiction. In other words, the compendium of quirky essays.

The man is clearly not a statistician, and it is always hard to please an actual nobel laureate like Daniel Kahneman, who summarily rejects Gladwell as ‘a great disservice to the study of’ psych, stats, you name it. But I will say, that while he is not as high-minded, his dialogue is infused with a sort of consciousness. And now that he is in the public eye, he is likely to produce more methodologically sound narratives. Nonetheless, it is a bunch of opinion, and ‘Autonomous’, featuring executives from car companies, to DARPA challenge winners, raises the question of where we are headed with self-driving cars, what is their current perception, and how will they affect motor safety, and jobs.

Data:

On to the post. The data for this post comes primarily from the Bureau of Transportation Statistics. But is supplemented by data from the FED. I was heartbroken to know we couldn’t find any data on car price statistics. The very data every intro stats class uses “the car dataset”, is impossible to find in real life in the public domain. You can probably purchase such data though.

The post is divided into 3 parts.

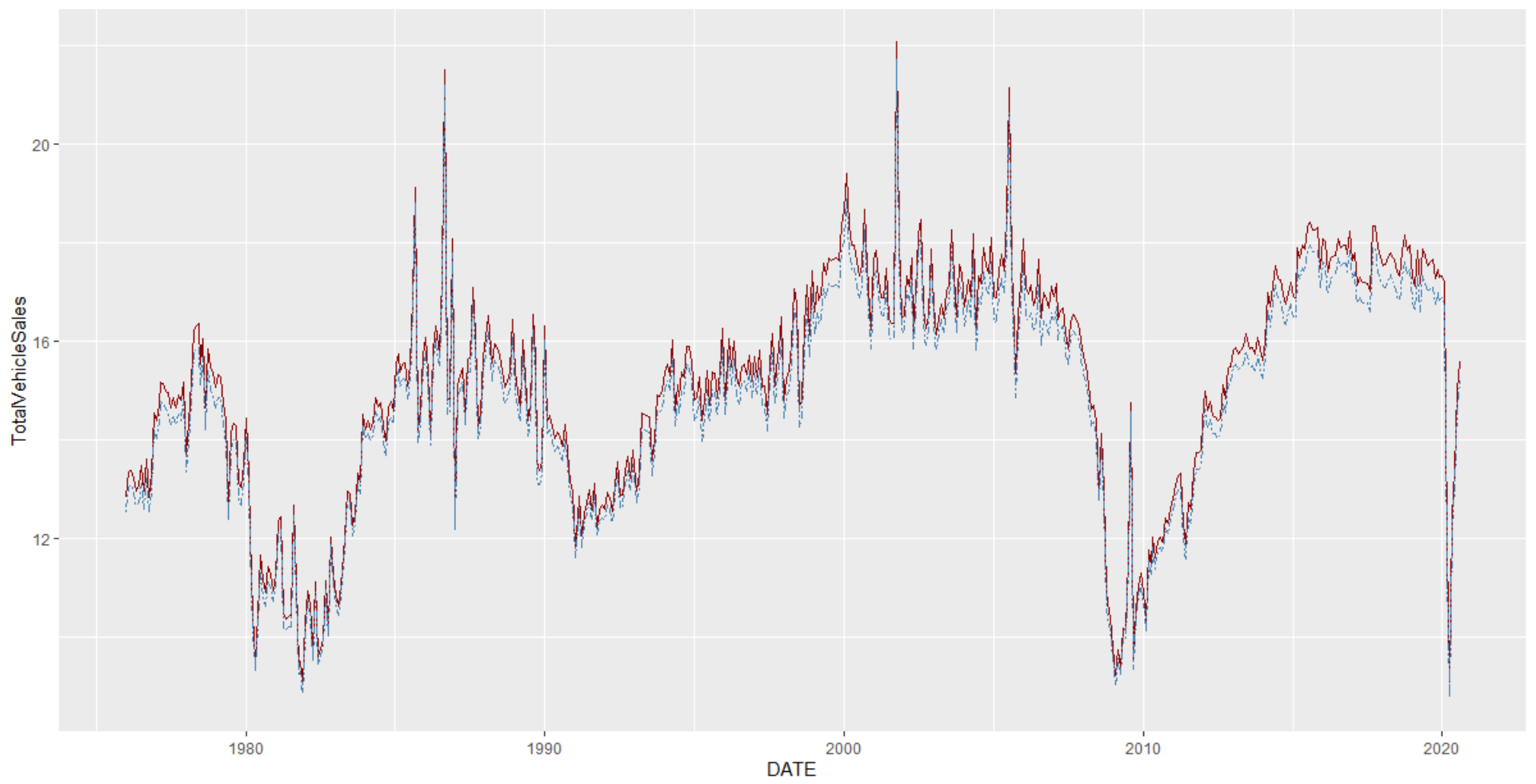
1. Car sales
2. Fatalities and accidents
3. Means, methods, and infrastructure in place

The infrastructure portion, collected at the county level, is meant to be a brief tabulation of regional resources of commuting. This in turn also affects the job opportunities we can avail to, and speaks to the work clusters in certain regions of the country. For instance, DC is not just DC, but is in fact DMV. Dallas is no longer just Dallas, but is Dallas-Fort Worth. These regions also doubtless have high connectivity in terms of transportation and are thus better able to facilitate such continued growth.

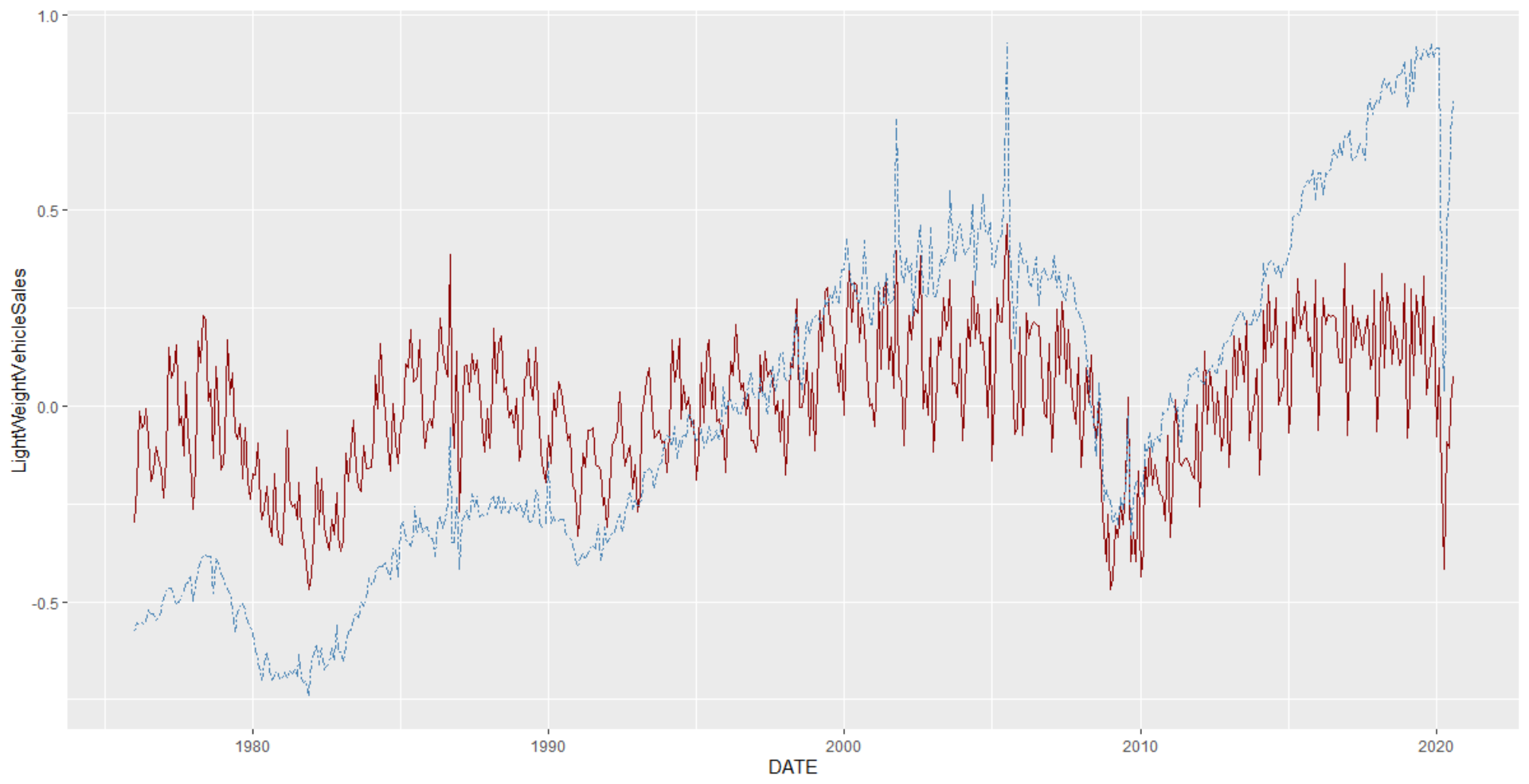
Car Sales, Compiled By the Fed from the BEA

The insights gleaned from the following data are very surface-level. As a matter of fact, the Federal reserve’s portal has dashboards that render the very same graphs online for you. However, I’ve chosen to normalize and reproduce the graphs here. The numbers and metrics for each measure are in the millions. However, since we are only interested in linear trends in this blog post, we stick to normalized data, and the direction of the movement.

In the below graph, we see a near congruous relationship between the two linear trends. The trends of course are total vehicular sales, in red, and total lightweight-automotive-vehicle-and-truck sales. This excludes heavyweight trucks. And shows us that lightweight auto are driving the bulk of total sales.



In the next graph, however, we see a breakdown of the lightweight auto sales. What we see is a more telling story. While the red line – lightweight auto sales – does not exhibit much volatility, or expansion beyond it’s current sales patterns, lightweight autos seem to have surged in popularity. Moreover, what is also interesting to note is the marked seasonality present in the light weight auto (red line) sales. It’s tough to explain what is driving the incline in lightweight truck (dotted blue line) sales for the past 30 years, maybe simply the fact that they have become so much more compact and efficient than in the past.



FARS Data

The FARS data comes from a comprehensive survey run by the BTS, and is based on data from various sources, including death certificates, and coronary reports. Although a bit grim in its subject matter, it tells a better story than we would imagine. Traffic safety has improved overall, the average number of deaths per accident has reduced marginally over time. Because this data is so comprehensive, and somewhat normally distributed, a lot of the collapsed data and means centered around a few numbers – 1 death on average, afternoon time, 2-3 people in the vehicle, and so on. However, cutting the graph multiple ways is bound to present a more comprehensive picture.

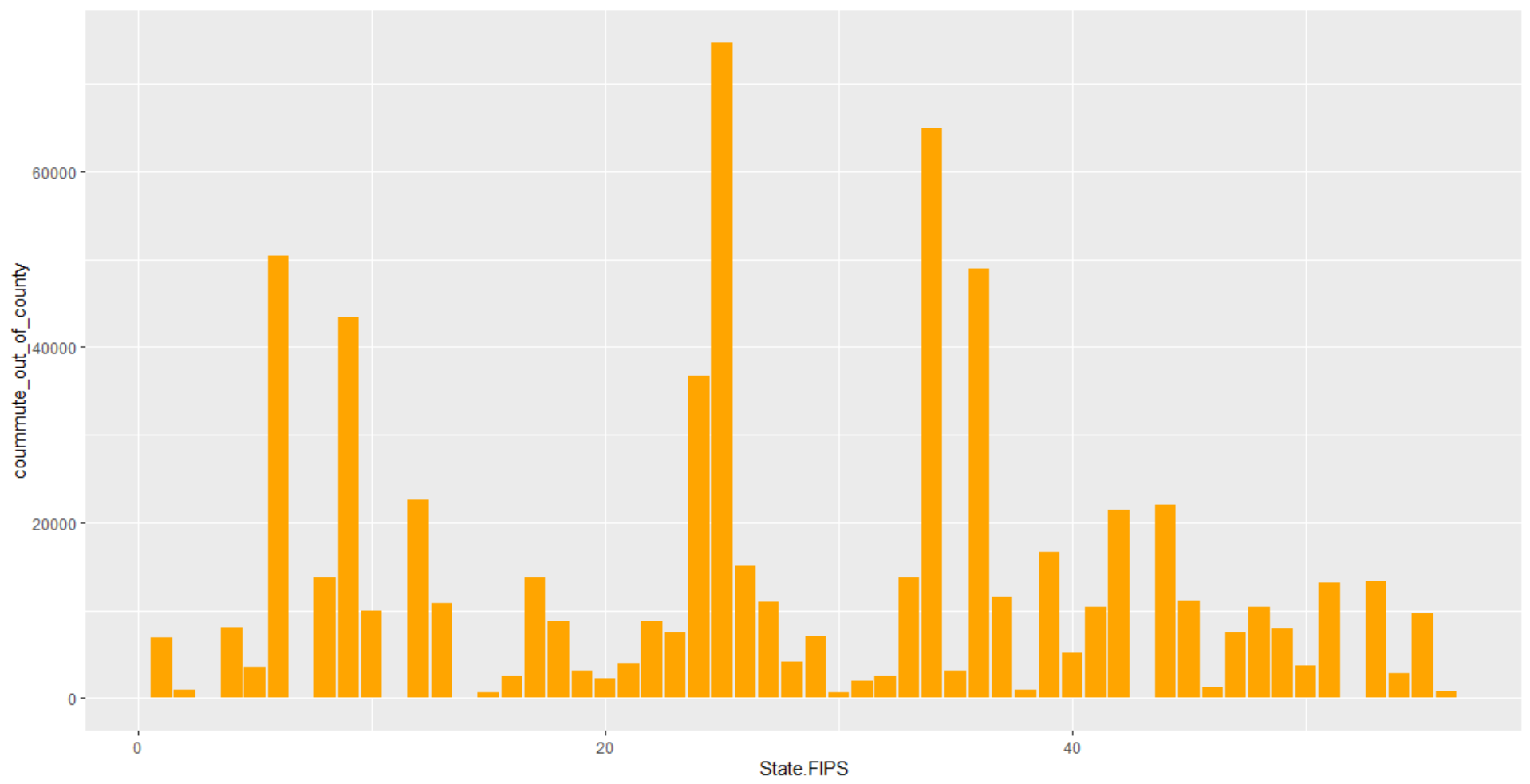


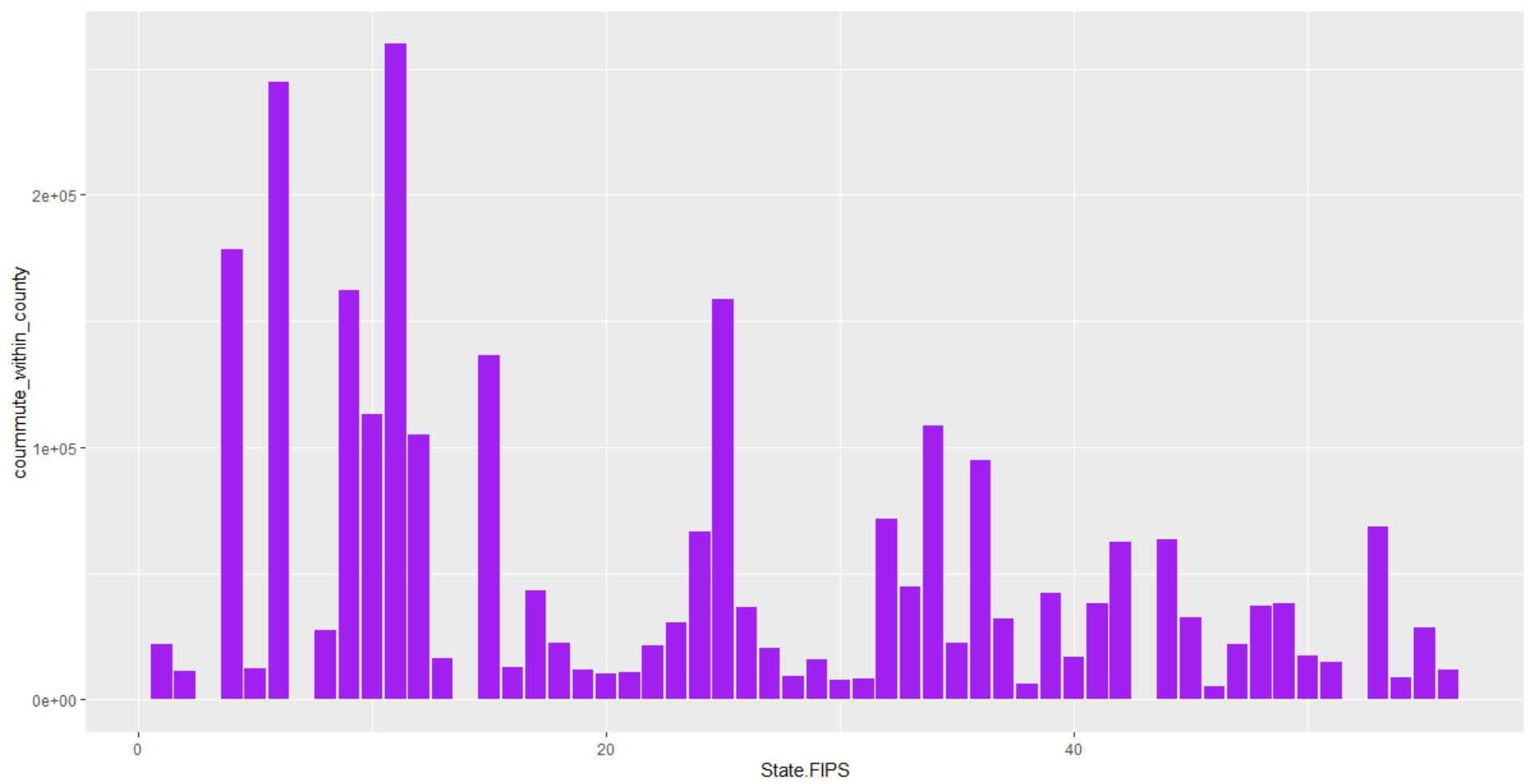
All this is bound to change with the advent of more automatable features in vehicles.

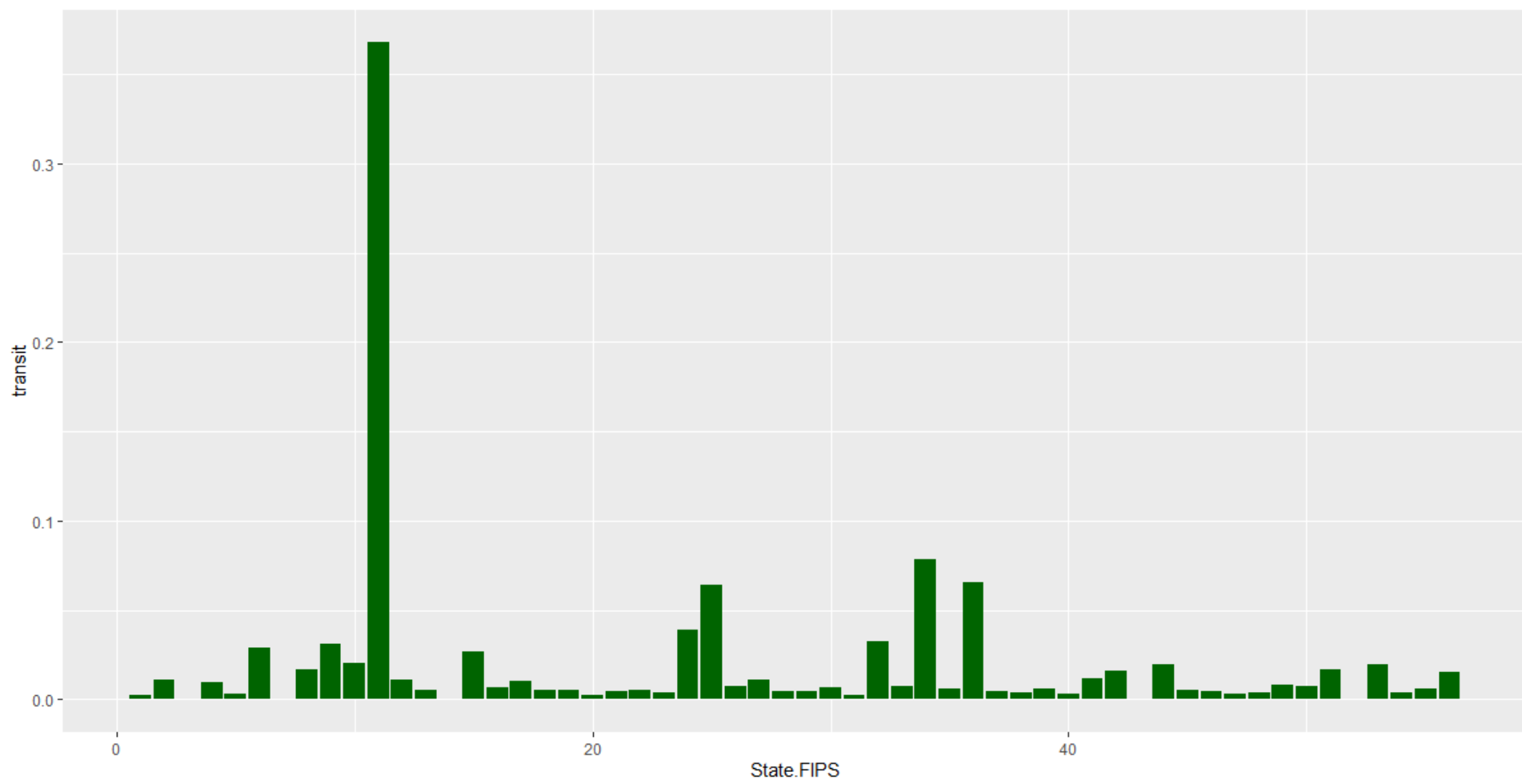
BTS County level data on Commuters:

And lastly, we map the county level data according to states. I aggregated or collapsed the data in this dataset to a row for each state, using the means for each column. Importantly, this is cross-sectional data, from 2018. What the graphs below represent are: (i) average numbers of people in each state commuting outside of their counties for work. (ii) average numbers of people in each state commuting within of their counties for work. (iii) the percentage of people per county using their transit system for commuting.

Admittedly, the district of Columbia is an outlier here, as it probably often is. Where state, and map, city, and county all collide. DC probably reflects higher percentages for transit usage.







All in all however, what I want to point out from a comparison of graphs i, ii, and iii, is what they can tell us when we overlay them. Without bothering to label the State FIPS in question, we can discern that the states with higher across county commuters (in graph i) are somewhat complementary to the states with higher within county commuters (in graph ii). However, when we compare the transit graph (graph iii) to the two above graphs, discounting the outlier of dc, we find that it aligns more closely with graph i.

Conclusion

This rudimentary cross graph analysis implies that the states with a better transit system are more conducive to commuting across counties. And that conversely, states with counties that are high in numbers of businesses, tightly clustered together, tend to be more amenable to building thriving transit systems.

In conclusion, our 3 key takeaways were that lightweight truck sales have been increasing over the past 30 years, average fatalities per accident have been declining, and cross county commuter states have better transit systems. It would be all too tempting to link these 3 findings together, but I’ll leave that for another time.