Sum Cars Purchased, Per State and Year, By Vin Year

This section displays each the new registration of each VinYear vehicle, over time. Here, we see not only the initial spike of new car sales, but also the secondary spike which shows the count of new cars that return to the pool and are again purchased. Because registration implies purchase, we are using this data to inform our understanding of the pattern of used-vehicle purchasing for each VinYear. The pattern appears consistent in frequency.



This graph depicts the totals of all new vehicle registration, over time, by VInYear.



Yearly sums for each VinYear model registration count, displayed over time.

This graph appears to indicate that the number of new cars purchased is increasing since 2008, and is even surpassing (what we can assume is) the pre-recession threshold. This graph also follows each year's new car, over time. Here, we see an indication, consistent in all cars that age long enough, a secondary, and tertiary purchasing peak.



The primary and secondary peak for new-in-2007 vehicles – the secondary apex emerges 3 years after the initial wave of new-car purchases.

(These happen to be sorted by state, but this doesn't matter – California is at the top)



New-in-2008 vehicles – again three years between the first and second waves.



New-in-2009 vehicles – again, three years between first and second purchasing waves – counts at the first peak have dropped dramatically, and at the second peak, more steadily.

Here is the effect of the 2008 economic crisis.



2010 Vin Year vehicles -- same three year frequency between crests – counts of both waves are increasing.



VinYear 2011 – same three year frequency between crests – counts of both waves are steadily increasing.



VinYear 2012 – from closer examination of the data, the value in 2015 appears to include the max values for this crest. Here, we see the same three-year frequency between crests – counts of both waves continue to increase. This is the last VinYear from our dataset that containing the second purchasing wave.

Trend Line Analysis

In this section, data discovery is applied to create a forecasting capability that, when annually update with the most recent year's data, supports the future-year predictions of the number of available cars to used-car dealers. Because these trend lines follow car purchasing patterns over time, what we are predicting is the total number of vehicles that are *likely* to return to market – not the total number of previously produced cars. We anticipate that this distinction is important, and the difference more useful to dealers.



Each row in this graph shows a VinYear car, starting with 2007 at the bottom, and continuing to 2015 at the top. Each month is depicted across the horizontal axis, starting at January 2007 and continuing to July 2016.

The blue represent an increase, the red represents a decrease. The red and the blue together are the wave.

The orange lines delineate the peaks of the crests of both the fist and second waves.

This graph informs the model on the next slide...



Voila, recreation of each VinYear, in context, over time



The second and third purchasing waves, when the VinYears are shifted to reflect the same annual timeline, depicted in quarters.

This shift is important because it gives us the best model for predicting future count estimates.



The forecasting model and its equations.



Application of the model yields the following estimated 2017 total for our given set of states



Total expanded across all states – regions can be defined freely since we have state specific totals. The method for reaching this total is the important thing because this same method is how you would define state totals. If you want region groupings, send me the states for each group, and I'll put them in a table and someone can run this through Tableau for a beautiful graphic... To me, this seem unnecessary for our purposes this Friday.

It would be more useful to look at the distribution of Makes and Models within a region to determine consumer trends over time, but that is not going to answer their question, so I didn't focus on it. Remember, we were "hired" to answer their question. This model accomplishes both. We have a way to forecast if we keep current with the last year's data.

To forecast car production, we need to look at the correlation between the past 20 years of US economic data (like the S&P index) and new car purchases. Even the, we would only be able to accurately predict as far as we can accurately predict future US economic trends.

Currently, we see several indication that the auto industry is approaching a flooded

market. Currently, new car supply supersedes demand, but this is because the market is still stabilizing after the 2008 economic effects. What this means is that CarMax will have to get VERY precise about the cars they purchase in order to withstand the upcoming dramatic swells.