FA49QRS 18JUN21

MEMORANDUM FOR U.S. Army Analytic Agency

ATTN: Ms. Janet Smith

SUBJECT: United States Border Crossing Information Paper

1. Purpose: To provide insights gathered from the border crossing/entry dataset.

2. Background: The Border Crossing/Entry dataset contains 364,510 records, pertaining to 8 distinct measures (variables). The dataset specifically measures border crossing at the U.S.-Canada and the U.S.-Mexico border, at the port level[[1]](#footnote-1). Data are available for the following crossing methods: Personal Vehicle Passengers, Personal Vehicles, Truck Containers Empty, Truck Containers Loaded, Trucks, Pedestrians, Rail Containers Empty, Rail Containers Loaded, Train Passengers, Trains, Buses, Bus Passengers[[2]](#footnote-2). Data are available from 1996, for 103 of 117 port cities, with multiple additional dates for the remaining ports.

3. Data Preparation: Using the SODA API provided by the Bureau of Transportation, the data was integested into R using the following API call:

data\_pull <- read.socrata("https://data.bts.gov/resource/keg4-3bc2.json")  
data\_pull <- data\_pull %>% mutate(., value = as.numeric(value))  
data\_pull <- data\_pull %>% mutate(., measure = as.factor(measure))  
data\_pull <- data\_pull %>% mutate(year = format(date, format="%Y"))

Once ingested, the variable ‘value’ was transformed to numeric, ‘measure’ transformed to a factor, and a new variable, ‘year’, created for aggregation of years for statistical comparison.

4. Discussion: The border crossing dataset presents various interesting facets, to include periodicity of the port cities, differences between northern and southern borders, anomolies, and future forecasts.

1. Every border crossing city experiences periodicity of some sort, especially those where Personal Vehicle Passengers and pedestrian traffic is predominant. This phenomenon is evident in Piegan, MT (Figure A), where both personal vehicles and Personal Vehicle Passengers are the predominant method of transportation, and significant spikes occur each year. Similarly, we see that a similar trend is present for Skagway, AK, with buss passengers as the predominant mode of transportation (Figure B).
2. There exists a significantly larger traffic footprint at the U.S.-Mexico border, than the U.S.-Canada border. Although inference is not made with regards to reasons, in the largest metric, Personal Vehicle Passengers, the largest yearly value for the US-Mexico border is nearly doube that of the US-Canada border. Within the US-Mexico subset, even greater variation exists between port cities, where Personal Vehicle Passengers range from zero(not allowed) to in excess of 23 million per year.
3. Multiple anomolies exist within the dataset that warrant further examination, outside of the current research. In the port city of Van Buren, Maine (Figure C), the metric for Personal Vehicle Passengers decreases to zero, causing large downward spikes in the time series graph. Similarly, in Point Roberts, Washington (Figure D) there is a signficant spike in Personal Vehicle Passengers, well beyond the normal seasonl variation. Finally, for all port cities, significant declines in all metrics occur around 2020, which can be attributed to COVID restrictions, but is still anomalous given the entire time series presented.
4. Forecasts for any given port may produce inaccurate results, given the significant decrease in port traffic at the beginin of 2020, as shown in Figures A through D. Using Personal Vehicle Passenger data for the port city of Del Rio, TX, the expected forecast would continue trends seen in previous months (Figure E). This however, is not representative of the data, and in Figure E, there is a clear difference between the forecast and true data. The forecast in Figure E was created using an ARIMA(0,1,1) model, but could not account for the rapid decline in traffic due to COVID.

5. Findings: From the data, it is evident there are differences between port cities and both borders. If the necessity arose for US Army support to Homeland Security (DHS) at a given port, greater impact would likely come from deployment to the southern border. Similarly, multiple ports present greater overall volume of traffic than others, and would therefore need additional support. From Figure F, we see the vast difference between southern border ports in places like San Ysidro, and potential support planning could easily account for ports with a greater volume of traffic.

Along with site selection, the time of year must be taken into account, given the periodicity of port traffic. Although variation exists, most ports experience spikes in traffic during certain months of the year, which would allow further granular support planning, if necessary.

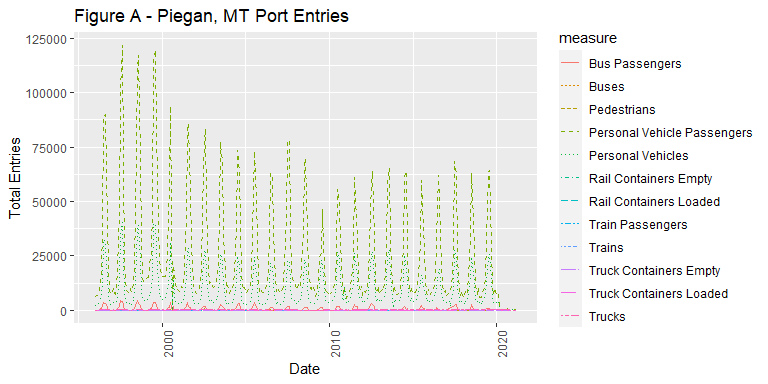
Finally, though numbers are rebounding, caution should be taken to forecast future port traffic based on data collected begining in 2020, due to the COVID epidemic. Forecasts for future traffic could discount or mitigate the decrease in 2020, given the eventual return to normalcy post COVID.

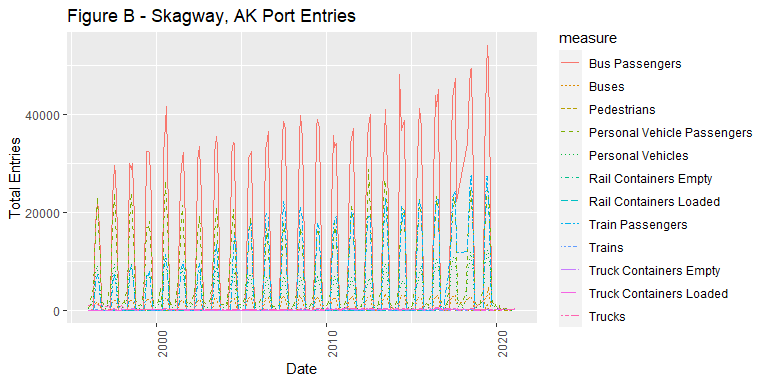
All the factors presented above give various analytical insights to leaders, should the need arise for the Army to provide support to DHS and border security.

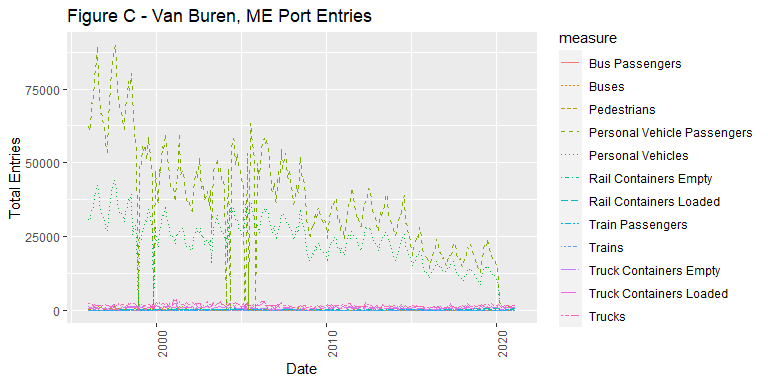
MAJ Nicholas Uhorchak / 254-317-9339

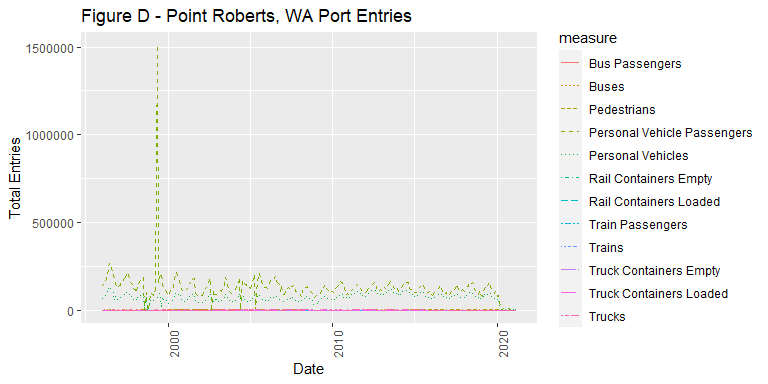
Official: LTC Spivey

Appendix 1 - Figures









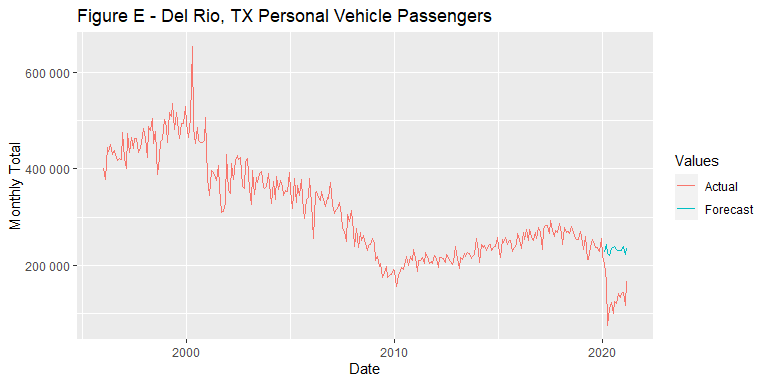
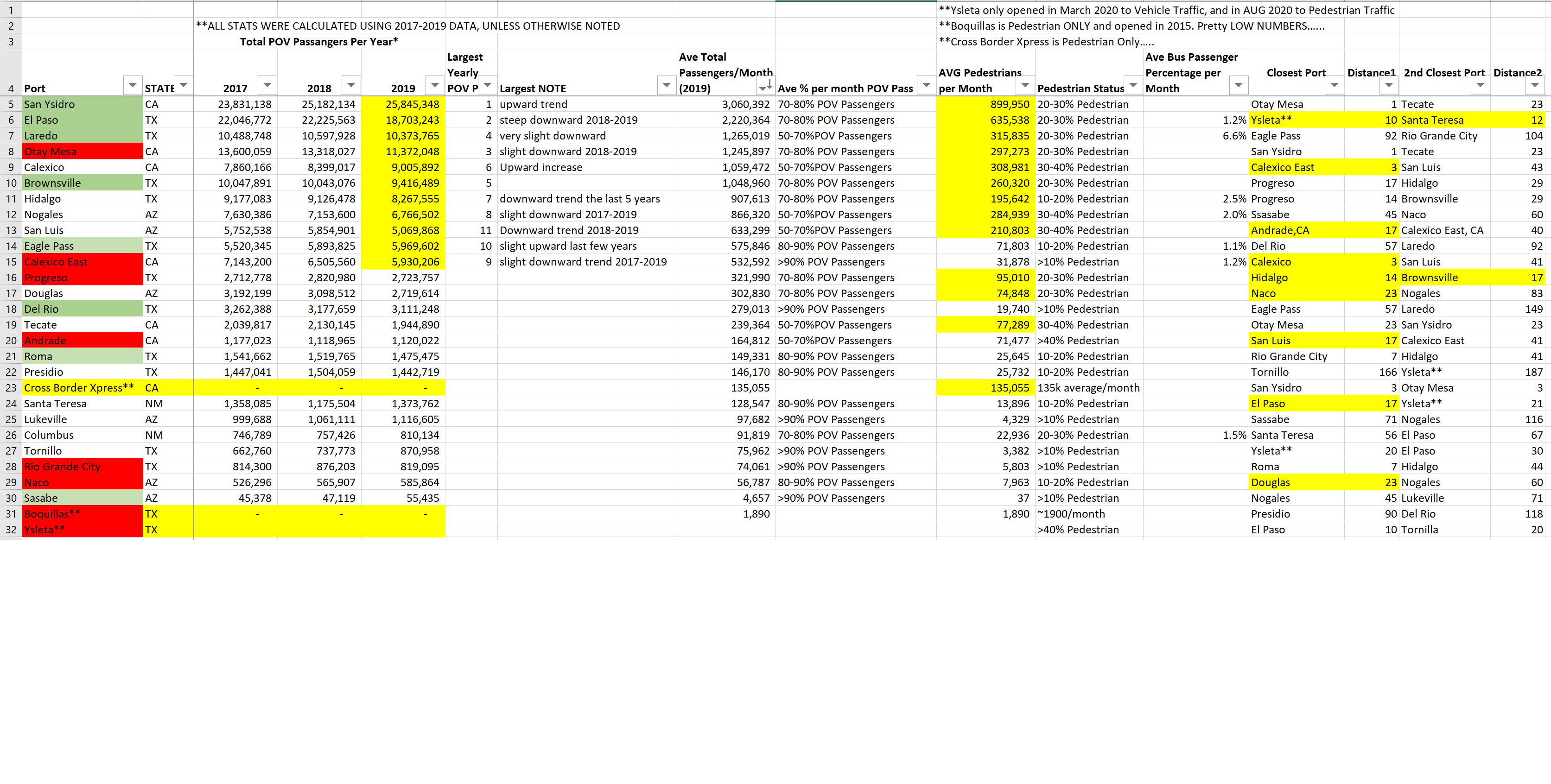


Figure F - US-Mexico Border Stats



1. Border Crossing/Entry Data. Accessed on 11Jun21. <https://www.bts.gov/browse-statistical-products-and-data/border-crossing-data/border-crossingentry-data> [↑](#footnote-ref-1)
2. Ibid. [↑](#footnote-ref-2)