Mustafa Abdullayev

USA Cars Assignment

AY6010 – Probability Theory and Introductory Statistics

Term: Fall 2019

Instructor: Tom Breur

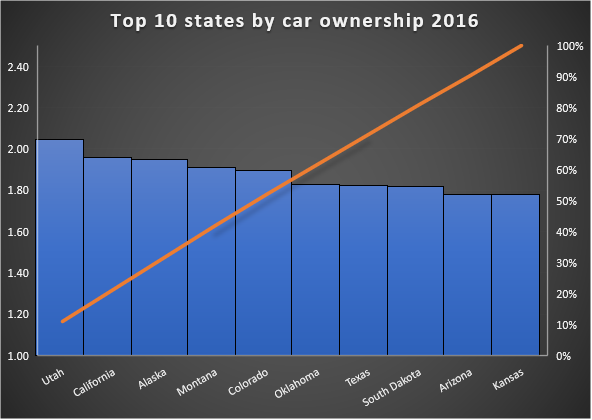
09/21/2019

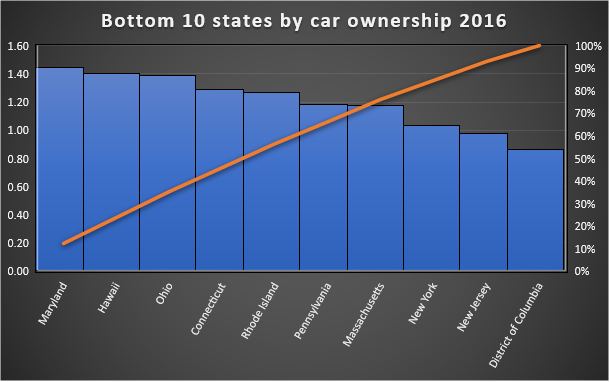
Introduction

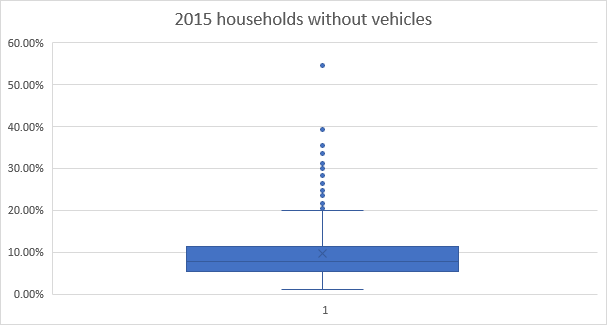
This is Microsoft Word Report accompanying Microsoft Excel Workbook. In my workbook, I analyzed USA cars dataset. Data was provided by class instructor, Tom Breur, and derived from USA government database. It consisted of information about percentage of household without vehicle and average number of vehicles in 317 USA cities for years 2015 and 2016. Data was cleaned, organized and ready for analysis. I used descriptive statistics to gain some insights. Also, I utilized powerful Excel built-in functions and graphs to dive deeper to observe hidden patterns and visually communicate my findings to audience.

Analysis

According to 2016 number of vehicles per household data (it is the latest we have), top 10 states per number of vehicles are usually located in the west side of USA. On the other all 9 of 10 states in the bottom part of that table is from north-east region of United States. Moreover 8 of these 9 states is the part of outlier group for the data of 2016 percentage of household without vehicles. There is no clear way to define outliers. But in my analysis, I defined an outlier as data point which is 1.5IR (Interquartile Range) above third percentile or 1.5 IR below first percentile. Excel boxplot uses this criterion to define an outlier. Reasons which I mentioned above can signal us that north-east regions have very good public transportation and people prefer them to owning car.







The interesting point is that in Utah average household has 2 cars in both years and it was only state like that. The other interesting fact is that, New York, NY was extreme outlier in both years with more than half of households without a car.

Examining data for 2015 yield almost same results. This also demonstrate that there is something different in the peoples’ view about car ownership and good public transport is a good candidate to explain that. But without having any information about wellness of public transport, it is impossible to conclude that. Moreover 8 of these states with lowest car ownership is among top 10 densely populated states of USA (2019, World Population Review). This is also important because it signal that car ownership and population density is in negative correlation. Maybe high population density is a cause for high number of traffic jams and eventually people prefer to not own a car.

Moreover, I used descriptive statistics to analyze both of 2015 and 2015 average car ownership data. Results were almost same. Both data had almost -0.5 skewness and 0.65 kurtosis. This statistic demonstrates our data is skewed to the right. That is also backed by our frequency distribution histogram for the same data. It is clearly visible that relative frequency distribution for percentage of households without a car for both years is right skewed. Also, standard error was 0.016 for both. Mean and median values were approximately same.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *2015 Vehicles per Household* | |  | *2016 Vehicles per Household* | |
|  |  |  |  |  |
| Mean | 1.691041009 |  | Mean | 1.715362776 |
| Standard Error | 0.01639595 |  | Standard Error | 0.016791423 |
| Median | 1.68 |  | Median | 1.72 |
| Mode | 1.6 |  | Mode | 1.69 |
| Standard Deviation | 0.291921593 |  | Standard Deviation | 0.298962781 |
| Sample Variance | 0.085218217 |  | Sample Variance | 0.089378745 |
| Kurtosis | 0.682849762 |  | Kurtosis | 0.627163335 |
| Skewness | 0.504133096 |  | Skewness | -0.512692645 |
| Range | 1.76 |  | Range | 1.73 |
| Minimum | 0.63 |  | Minimum | 0.63 |
| Maximum | 2.39 |  | Maximum | 2.36 |
| Sum | 536.06 |  | Sum | 543.77 |
| Count | 317 |  | Count | 317 |

Finally, with the help of Excel’s scatter plot I analyzed the similarities of 2015 and 2016 average vehicle per household numbers. This analysis present that, there is an obvious correlation between these two series which quite reasonable because people usually keep their cars for at least 5 years. When they become car owner, it gradually increases the average car ownership. But since correlation does not imply causation, we cannot conclude that there is causation between these two.

Conclusion

To conclude, I used Excel and its libraries to analyze a USA car ownership data. Data had information about 2015 and 2016 car ownership statistics for USA states and cities. Key finding was that, north-east regions of USA was massively different than other states. People in these states usually preferred to not own a car. Also, these states were among the most densely populated USA states. New York, NY was the most extreme one with almost 54% households without a car. In order to conclude that this was due to good public transportation system or population density we need to analyze data about transportation and population in these states. Moreover, average car ownership for United States slightly increased from 1.69 to 1.72. Finally, in 2015 and 2016, percentage of cities in which 90% or more households owned a car were 67.51 and 70.98, in respectively.

Reference

(2019, August 28). World Population Review. Retrieved from [http://worldpopulationreview.com](http://worldpopulationreview.com/states/state-densities/)