**CLOUDWALK CASE – MAY 2022**

From the csv file, it’s mostly about the time to deliver a product. This time is divided in two parts: from the central storage or production place to the local distribution; and from the local distribution to the delivery address. Starting from this, the first step is to look at each of this parts by group, in this case by state, and look out for outliers, as shown in the two graphs below:

From this first look, it’s possible to infer that the Local Distribution is not the main issue and it has 3 outliers that could be looked at more in-depth. The product’s transit to the Local Distribution is the part with the most variation from state to state. For better insights, we need to compare if this variations derives from the transport company or actual logistics problems, like lack of a road or accessibility in general. By comparing the Estimated delivery time and the Real delivery time (total) it becomes easier to make reasonable assumptions.

With the new data, it’s possible to divide the focus on two main types of problems: the “3 days delivery time” stated as a priority on the case and the high time variation from estimated to actual delivery cases. For this analysis, we’ll assume the delivery time mentioned refers to the time an item arrives at local distribution.

Starting at the last image, we have 7 states to worry about: AC, DF, ES, MA, PE, PI and RJ with 3.50, 0.98, 2.09, 1.20, 0.89, 1.47 and 0.75 days late, respectively, in comparison to the estimated time. Crossing this information with the one from the second image, it becomes clear that only Espírito Santo has a big problem with the delivery time, as always considering the mean values. This would be the best bet for a fast gain.

The estimated time for a product to arrive at the local distribution should be standardized by state and more or less known by the company, with little deviations. With this in mind, there shouldn’t be many states with high differences between real and estimated delivery time (considering the mean value) that also happen to have acceptable delivery time. This could mean two main things, both complicated to solve: first one being poor understanding of the “sending the product to the local distribution” process; the other one being excessive variation on both sides, above and below, masking itself with an acceptable mean value, also referred to as not having a standard.

The next steps would be to check if any of the two cases cited is happening. As the analysis of means shows, a more detailed look is needed with boxplot being an excellent option. Also, more data is probably needed for most states to define a pattern. Changes/problems within the region, like a road being closed for hours/days could impact the end result and must be included in the analysis.