Business case: CSI currently has to throw out any product that does not pass inspection. Product can fail by being over or under weight or by visual inspection of its color. When a product has failed weight it could still be usable if the ratio of the fluids dispensed is still within spec. In cases like these sending the product to a rework station could reduce waste by allowing someone to add or remove some of the product to bring the product to the correct weight.

Unfortunately, for our dataset there were 88 inspection failures that represent product that would have to be thrown out. After inspecting the data, we found 21 failures where the vial failed weigh inspection indicating it was either over filled or underfilled. From that small sample a further 6 did not complete visual inspection. That left a small sample of 15 records to train our logistic regression on. For this project we decided to focus on a binary classification of reworkable or not reworkable. Had there been a larger set of data we would have been able to add more categories to determine over/underfilled sub categories for the reworks as well as classifying the vials that fail visual inspection into categories that could allow them to be reworked.

An example of how this might apply to a real business case would be a shampoo factory that has a recipe with specific ratios of soap, moisturizer, and fragrance. If after filling a bottle is found to be underweight, but had the correct ratios you could just add more of the finished shampoo product to the bottle to make it the correct weight. If however, it only dispensed half the soap but the normal amount of everything else, then you would need to have a separate category “needs soap” to flag that particular product for the rework station.