The first step I did in this project is to group the dataset by pitcher, pitch type, and whether there is a base runner or not. Then I calculate the average and standard deviation of spin rate, horizontal movement, and vertical movement under these conditions. The next step I used the left_join function to match every pitch's average and standard deviation. Based on the average and standard deviation, I calculated the z-score of every pitch. I used the z-score to calculate the percentile of every pitch. The condition I set that has no negative influence on the pitcher is when the spin rate is above 50% percentile and one of the vertical and horizontal movement is above the 50% percentile. The data that meets the condition I mentioned early will have a 0% probability of getting affected by dew point. For the rest of the data, I added up the percentile and divided it by three (since there are three factors to add up) to see how far away it is to the average. I used the average as the probability of a pitcher getting affected by the dew point.