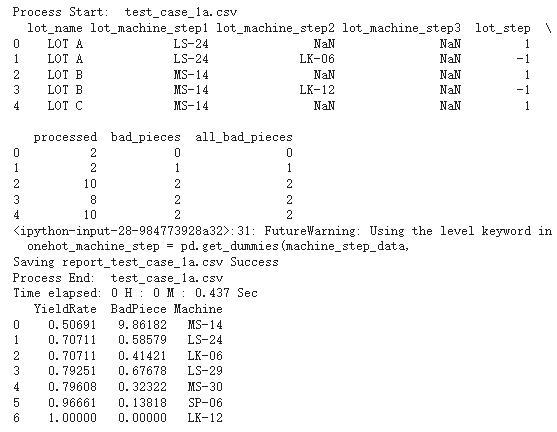
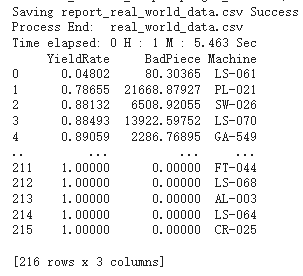
**EM Real Work Lab Report**

* Screenshot of the test report



* Screenshot of the real-world report



* Explain the reason why we use the “EM algorithm” to estimate machine yield rate?

Since our goal is to obtain results for different machines. However, there are some missing values in two files. Therefore, following the lesson, we need to employ the EM algorithm to predict the missing data. So that we can use the updated dataset to determine which machine has the lowest yield rate or produces the highest number of defective pieces, similar to the random coins scenario.

* After the manufacturer has the results of machine yield rate estimation, what can they do using that result?

These results show us the approximate yield rate and the number of defective pieces in each machine. Since they are arranged in order, we can easily identify the machine with the lowest yield rate in the first row.

Because we now have the performance rates of each machine, therefore we can replace the machines with low yield rates and try to improve the overall yield rate based on these results, which was our original goal for the experiment.

* What do you learn from this lab?

Actually, I felt it somewhat challenging to grasp each concept in the teaching materials overall for this section. Hence, I searched for something related to EM algorithm on the Internet to understand it much clearer to help finish the lab.

To summarize the EM methods briefly, we obtain the likelihood exception in the E step and find the maximum, the best likelihood value in the M step.

* Any questions and/or difficulties in following the EM Algorithm topic?

In the process of finishing lab, I thought it was hard to finish before finding the solution about the version of pandas from the senior.