Questions and notes for the JHU data correction

* There are non-unique JHU.num values. Some are flagged as “rerun of the same sample of gas” but others have no comment. How do you want these treated?
  + For example, I could automatically assume you only want the original analysis or keep both by auto-assigning a new JHU.num. If the JHU.num was 225, I would assign one 225 and the rerun 225.1.
  + However, if an analysis actually is a rerun, we should delete it so it isn’t double counted in data analysis.
  + This also comes into play with the assigned “group number.” Because this is ultimately for troubleshooting and not actually part of the reduction, it doesn’t really matter if you choose to do nothing – the data will still be normalized following our standard scheme.
* To force spelling standardization for standards, the UM code copies the names from Type.2 over to sample.ID (this is not done for samples, which do not have individual Type.2 values). Double check with Naomi that this did not drop any information (i.e., if there were two CO2 tanks that both used the same Type.2 value, they will now incorrectly have the same name).
* Date times need to have hms information – this is how the drift is done.
  + The SMOW-based trends will still likely be better than the JHU-number, but if this information is around let’s use it.
  + There are also some “DD\_MMM” and 5-digits day codes that we need to fix
* We need to know how many primes were done before each sample so that the memory affect can be compensated for. Right now, I’m assuming no primes were ever done.
  + And need to get rid of any samples that are placeholders (i.e., if a sample was pumped away and never analyzed, serving only as a prime, etc.)
* Are these all of the JHU data? If so, it might be convenient to assign them static “reactor numbers” just for programming stuff. But my notation tends to get overly complicated…
* Reactors 1, 2, 12, and 13 fail to run.
  + 1 – does not have analysis time in correct format
  + 2 – all data is flagged to have analytical problems
  + 12 – does not have SLAPS (but few analyses, all are standards)
  + 13 – Only has two SMOWs on the same day. As they both have the same analysis time (which is just the date), the program can’t fit a line to the data. 5 total unknowns.