



But with batch names, it can be tricky. After browsing through TXT, you will have an idea of how the sample was processed and where it came from.

A simple way to determine whether each sample is in the same batch is to determine whether they are in the same experiment site, whether the time period between experiments is short (less than 3 weeks, things like that), and whether the machines and methods used are consistent. If the answer is no, you can assume that they are in different batches. After you have defined your batches, name them. (e.g. if 3 batches are found, identify them as 1,2,3. These names are not particularly meaningful, they are just used to distinguish batches. You could also say a, b, c)

For features, you can choose a lot of features to put in the metadata, not just the 'Normal VS. Cancer' feature like the example above. All the features you need should be included in the metadata so that you can easily call them in the analysis. Therefore, you can increase the number of columns as needed.

### STEP3

Import your Excel spreadsheet into R.

Tip: put samples in rows will make your life easier, just like example above.