**Use MATLAB to have a quick basic look into the training dataset.**

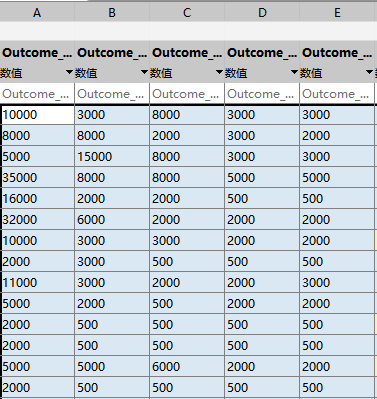


Fig. 1

Get:

1. The data was preprocessed already.
2. Too many features.
3. Different feature type.
4. PCA might be necessary.
5. More than 1 label.
6. Dataset not very large.

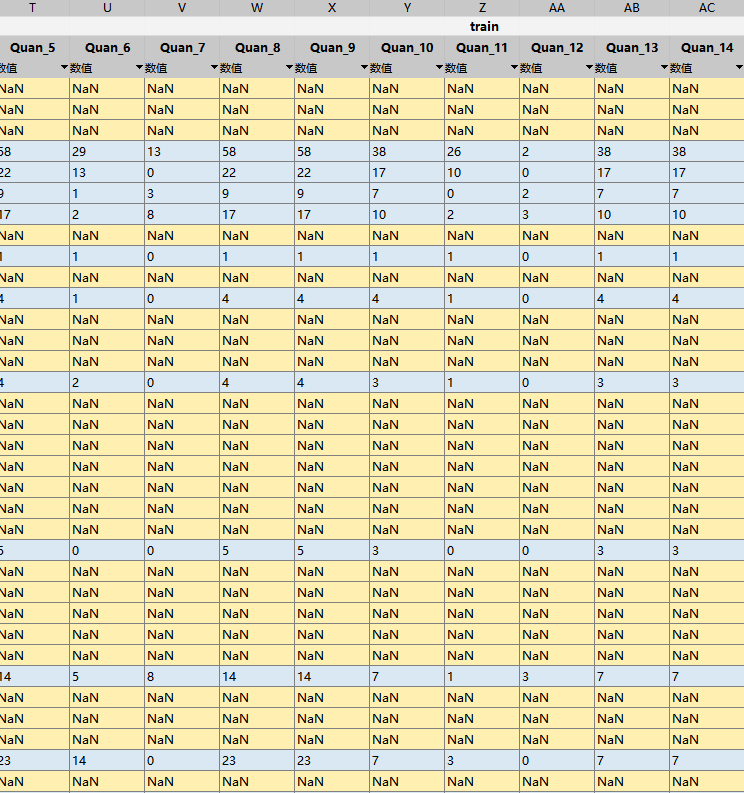


Fig. 2

1. Lots of missing values formatted as: ‘NaN’

**Use MATLAB to have a quick basic look into the sample submission dataset.**

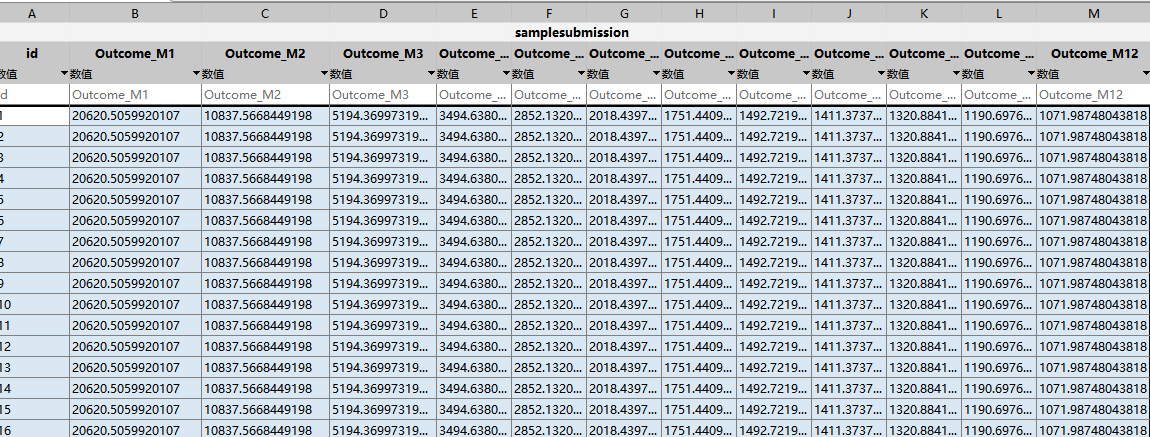


Fig. 3

Get:

1. More than 1 response/label needs to be predicted.

Ok, now, some pandas/seaborn based EDA.