Write a Python script named process\_16plex\_tmt\_multiple\_groups.py to handle regular 16plex TMT data. I have attached proList.csv and Sample\_info.csv. Follow these exact steps:

1. Use proList.csv as the main input file.
2. In this file:

* Remove all columns whose header contains "avg".
* Remove all rows where the column "accession" contains "contaminant\_".
* Remove all rows where the column "accession" contains "Reverse\_".

1. Add a new column called "GeneName" at the leftmost position. The values should be gene names extracted from the column "description". For example, from:  
   Metallothionein-1 OS=Mus musculus OX=10090 GN=Mt1 PE=1 SV=1  
   extract "Mt1" (the value after "GN=" and before the next space).

Here is an example function to use:  
def GN\_select(names):  
new = []  
for i in names:  
if i.find('GN=') > 0:  
s = i.split('GN=')[1]  
e = s.find(' ')  
new.append(s[:e])  
else:  
new.append(i)  
return new

1. Use Sample\_info.csv to rename the TMT intensity columns (e.g.: header like ‘norm int m/z\_127.131081’). In Sample\_info.csv, there are columns: TMT\_channels, Sample\_ids, Group, Sex. For each matching column in proList.csv, rename it in this format:  
   Sample\_Sample\_ids\_Group\_Sex.
2. Add a new column called "One-Way ANOVA". Calculate p-values across all the groups using only columns with the renamed format. Groups are defined by the "Group" column in Sample\_info.csv (in this case four groups: A, B, C, D. The group number may be different but will always names in a similar way).
3. Add new columns to show post hoc p-values between each pair of groups (e.g., A vs B, A vs C, etc.).
4. Save the final processed data to a new CSV file.

Make sure to handle all edge cases and not drop any additional columns or rows beyond what is described.