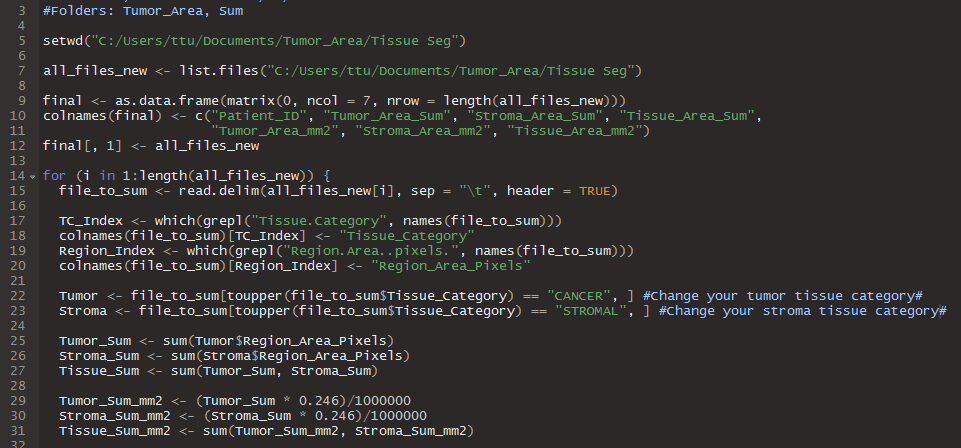
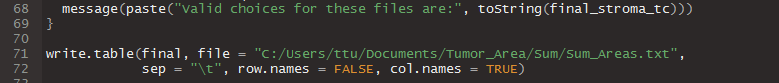
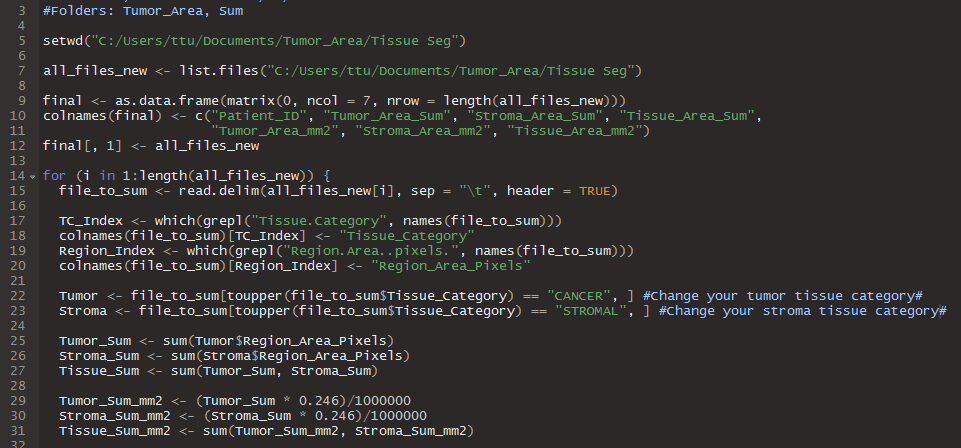
The purpose of this R script is to calculate the total tumor, stroma, and tissue pixel area of a sample by summing up the pixel areas of each tissue category in every high powered field in a **tissue\_seg\_data\_summary.txt** file. The tissue pixel area is the sum of the tumor and stroma area. The pixel areas are also converted into the area for each tissue category in mm2. The output is a text file that has tumor, stroma, and tissue pixel area and mm2 area for each tissue\_seg\_data\_summary.txt file run. If the pixel area of either the tumor or stroma is 0, then the script will print a message listing the files the script failed on and list valid tissue category choices for those files. The user should then remove the files the script succeeded on, change the name of final output file, change the failed tissue category names, and rerun.

1. The folder **Tumor\_Area** should have two subfolders: **Tissue Seg** and **Sum**.
   1. Check to make sure the Tumor\_Area folder with the two subfolders in it is in the working directory (the stuff inside the **setwd()** and **write.table()** functions). There may be other subfolders but these are the only necessary ones.
      1. If it is not in the working directory, change your working directory in the areas boxed in red below to wherever the Tumor\_Area folder is.
   2. Make sure to keep the slashes **“/”** at the end of the string (the stuff inside the quotes in green) when there is a **paste** function or the directory will be incorrect
   3. You can change the file name at the end of write.table() directory. It is set by default to be “Sum\_Areas.txt”. You can change it to anything provided it ends with a “.txt” and is written after the last “/” in that line within the string (the stuff within the quotes which should be green at least in my view of RStudio).





1. The **Tissue Seg** subfolder should contain **tissue\_seg\_data\_summary.txt** files Before running the **Sum** subfolder should be empty. This is where your results will go.
   1. Make sure the **tissue\_seg\_data\_summary.txt** files are “Tab Delimitated”. If they are not (if they were exported from Spotfire they might be Unicode), ask Anthony or Ting-Fang for a script to change them or you can open them in Excel and “Save As”, “Save as type: **Text (Tab delimited) (\*.txt)**”
2. To change the tissue category the script is searching for:
   1. The line with the variable Tumor in it is for the tumor tissue category
      1. To change what it is searching for, change the string (the green stuff in the quotes). **Make sure everything you type in here is UPPERCASE**. The script converts everything to uppercase to avoid potential misses due to case sensitivity. The example here is “TUMOR”, boxed in yellow.
   2. The line with the variable Stroma in it is for the stroma tissue category
      1. To change what it is searching for, change the string (the green stuff in the quotes). **Make sure everything you type in here is UPPERCASE**. The script converts everything to uppercase to avoid potential misses due to case sensitivity. The example here is “STROMAL”, boxed in yellow.
   3. **DO NOT CHANGE THE COMMA THE COMES AFTER THE END OF THE STRING.**



1. Click “Source” in the upper right corner of the RStudio scripting window to run
2. If a message pops out in the console:
   1. **Tumor Area was 0 for**:
      1. A list of file names will appear for which the strings boxed in yellow above have failed. Keep those files in the Tissue Seg folder.
      2. **Valid choice for these files are**:
         1. A list of potential strings to replace the ones in the yellow box for the line corresponding to the Tumor variable will appear. Choose the one that signifies some sort of cancer (i.e. “Cancer Island”).
   2. **Stroma Area was 0 for**:
      1. A list of file names will appear for which the strings boxed in yellow above have failed. Keep those files in the Tissue Seg folder.
      2. **Valid choice for these files are**:
         1. A list of potential strings to replace the ones in the yellow box for the line corresponding to the Stroma variable will appear. Choose the one that signifies some sort of cancer (i.e. “Stroma”).
   3. Remove the files that succeeded in calculating both tumor and stroma area. **CHANGE THE NAME OF THE FILE BEING CREATED IN THE WRITE.TABLE FUNCTION (see 1c)**. Failure to do so will result in previous run being overwritten. After changing the strings in the yellow box, click “Source again to rerun”
3. Repeat Step 5 until the area for all the remaining files has been successfully calculated.
   1. **NOTE: You will save yourself a lot of pain by just opening all the tissue\_seg\_data\_summary.txt files before running them and checking to make sure that they all have the same Tissue Category elements** (i.e. all of them use “Cancer”, “Stromal”, and “Other”) **and removing the files that do not match to run after the first batch.** The messages in Step 5 are really just meant to help in case you miss something.