```bash
#!/bin/bash -l


#  Written by EBR, 2022
#
#
#  This script takes the output of processed pupillometry data and matches it,
   trial by trial, to fMRI beta series
#  Then, we compute the correlation between pupil size and beta series
#
#


data_dir=/home/shared/aclab-fmri/Studies/33_MOTIP2018/MRI_data
output_dir=/home/shared/aclab-fmri/Studies/33_MOTIP2018/MRI_data/ABE_analysis/beta_series
behav_dir=/home/shared/aclab-fmri/Studies/33_MOTIP2018/behavioral_data/ABE/mid_task
output_dir2=/home/shared/aclab-fmri/Studies/33_MOTIP2018/MRI_data/ABE_analysis/pupil_vectors
pupil_dir=/home/shared/aclab-fmri/Studies/33_MOTIP2018/eye_data/ABE_firstlevel_analysis

read -p 'What metric? (AUC,max,deriv): ' extract


cd $output_dir
#get subject list from whatever folder is holding beta series bucket files
subject_list=`ls beta_series_???+tlrc.BRIK | cut -d'_' -f3 | cut -d'+' -f1`
###########this call depends on the name of the beta series datasets
for sub in $subject_list;
do
    outfile=${output_dir}/beta_series_4D_${sub}+tlrc.BRIK
    if [[ -f $outfile ]]
        then
        echo "Skipping 3dTcat for ${sub}..."
        else
      echo "Convert beta series format for ${sub}..."
        #convert from bucket format to 4d format
        3dTcat -overwrite -prefix $output_dir/beta_series_4D_${sub}
$output_dir/beta_series_${sub}+tlrc
    fi
done

rm $output_dir2/pupil_vector_${extract}*.txt #delete every file in the output
directory since we will be rewriting with this script

for sub in $subject_list;
do


    #now we find how many trials are included in the final betaseries - everybody
```

```bash
saw 576 but since the timing is pseudorandom and we dropped 4 TRs, people will have lost 0-1 trials per run
    ntrials=`3dinfo -nv ${output_dir}/beta_series_4D_${sub}+tlrc`
    echo '***'
    echo $sub
    #echo $ntrials

    #now we need to get the special ABE subject number to be able to open the pupil files properly
    if [ "$sub" -gt 199 ]
    then
        subtract_amount=150
        ABE_special_subject_number=$(($sub - $subtract_amount))
    elif [ "$sub" -gt 149 ] && [ "$sub" -lt 199 ]
    then
        ABE_special_subject_number=$sub
    else
        subtract_amount=100
        ABE_special_subject_number=$(($sub - $subtract_amount))
    fi
    #echo $ABE_special_subject_number

    #in order to figure out which runs had a dropped first trial, so we can effectively drop the first trial from our summary pupil data, we need to look at the behavioral output files

    for run in 1 2 3 4;
    do

        all_trials_file=$behav_dir/tempAttnAudTME-${ABE_special_subject_number}.*-1.${run}.*ef.txt
        onset_scrambled=`cut -f10 ${all_trials_file} | sed -n '2 p'` #onset time for first scrambled trial

        #now we need to get the onset times for the first 2 "real" trials (nonscrambled)
        onset_nonscrambled=`awk '($4!=0) {print $10}' $all_trials_file | head -n 3 | tail -n +2 | sed -n '1 p'`
        onset_nonscrambled2=`awk '($4!=0) {print $10}' $all_trials_file | head -n 3 | tail -n +2 | sed -n '2 p'`

        criteria=$(echo "($onset_nonscrambled - $onset_scrambled)" | bc -l)
        criteria2=$(echo "($onset_nonscrambled2 - $onset_scrambled)" | bc -l) #this is the only way (?) to subtract floats in a shell script, apparently
        #echo $criteria
        #echo $criteria2

        #if the non-scrambled trial onset (or the 2nd one) is less than 10 seconds after the scrambled onset, it's less than 4 2.5s TRs out from the beginning of acquisition, and we will need to drop that trial from our pupil metrics for that run
        #it's columns 6 (AUC), 7 (max) and 11 (pos deriv) of the pupil summary
```

```bash
output files that we might be interested in using as a regressor for future analyses
        #they are comma-separated files with the first line being labels
        cutoff=10
        compare=`echo "$criteria<$cutoff" | bc`
        compare2=`echo "$criteria2<$cutoff" | bc`


        if [[ $extract == 'deriv' ]] #depending on which variable we want to pull
out, find the correct field to extract from pupil summary data
        then
            field=f11
        elif [[ $extract == 'AUC' ]]
        then
            field=f6
        elif [[ $extract == 'max' ]]
        then
            field=f7
        fi

        if [[ $compare -gt 0 && $compare2 == 0 ]] #if the start of the first real
trial is less than 10 seconds after the first trigger
        then #we want to drop that trial
            towrite=`cut -d ',' -${field}
${pupil_dir}/s${ABE_special_subject_number}_run${run}_summary_statistics.csv | tail
-n +3` #we're starting on line 3, skipping both the header and the first entry
            echo "The first trial was removed from run ${run}"
        elif [[ $compare -gt 0 && $compare2 -gt 0 ]]
        then
            towrite=`cut -d ',' -${field}
${pupil_dir}/s${ABE_special_subject_number}_run${run}_summary_statistics.csv | tail
-n +4` #we're starting on line 3, skipping both the header and the first two entries
            echo "The first 2 trials were removed from run ${run}"

        else #if the first trial happened to be rather late and doesn't need to be
dropped
            towrite=`cut -d ',' -${field}
${pupil_dir}/s${ABE_special_subject_number}_run${run}_summary_statistics.csv | tail
-n +2` #we're starting on line 2, skipping only the header
            echo "No trials removed from run ${run}"
        fi

        length_of_vector=`echo "$towrite" | wc -l`
        if [ $length_of_vector -gt 4 ] #if we have a missing data, don't make a
file with nothing but newlines in it, you'll just confuse yourself
        then
        echo "$towrite" >> ${output_dir2}/pupil_vector_${extract}_${sub}.txt
        fi

    done
```

```bash
    ntrials_pupil=`cat ${output_dir2}/pupil_vector_${extract}_${sub}.txt | wc -l`
    if [ $ntrials != $ntrials_pupil ]
    then
    echo 'UNEQUAL FILE LENGHTHS'
    fi
    echo "There are $ntrials_pupil trials in the pupil vector"
    echo "There are $ntrials trials in the beta series"
    echo "... for subject $sub"

#correlate the 1d pupil timeseries we made with the beta series
3dTcorr1D -overwrite -Fisher -pearson -prefix
${data_dir}/ABE_analysis/pupil_beta_correlations/pupil_${extract}_beta_corr_${sub}
${output_dir}/beta_series_4D_${sub}+tlrc
${output_dir2}/pupil_vector_${extract}_${sub}.txt

#uncomment this if you want to write the dataset without Fisher transform
#3dTcorr1D -overwrite -pearson -prefix
${data_dir}/ABE_analysis/pupil_beta_correlations/pupil_${extract}_beta_corr_noFisher
_${sub} ${output_dir}/beta_series_4D_${sub}+tlrc
${output_dir2}/pupil_vector_${extract}_${sub}.txt

done


3dmerge -overwrite -gmean -prefix
${data_dir}/ABE_analysis/pupil_beta_correlations/average_pupil_${extract}_beta_corr.
nii
${data_dir}/ABE_analysis/pupil_beta_correlations/pupil_${extract}_beta_corr_???+tlrc
.BRIK

#uncomment this if you want to write the dataset without Fisher transform
#3dmerge -overwrite -gmean -prefix
${data_dir}/ABE_analysis/pupil_beta_correlations/average_pupil_${extract}_beta_corr_
noFisher.nii
${data_dir}/ABE_analysis/pupil_beta_correlations/pupil_${extract}_beta_corr_noFisher
_???+tlrc.BRIK
```