MBatch 04-04 Using MBatch Assessments: Boxplot_Group_Structures Tod Casasent 2017-11-02-0905

1 Introduction

These instructions are aimed at people familiar with R and familiar with TCGA/GDC platforms and data types. They are intended to introduce the reader to producing the given assessment. These instructions will only rarely, if ever, touch on the appropriateness of the assessment algorithm or interpretation of output. See MBatch_01_InstallLinux.docx for instructions on downloading test data.

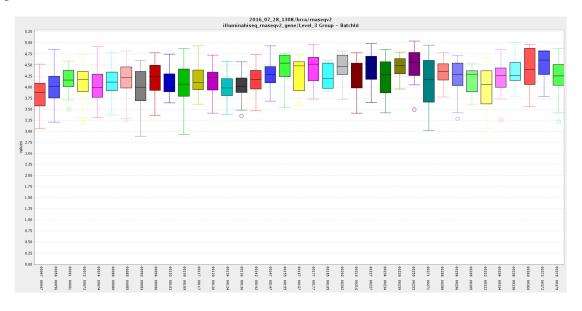
2 Algorithm

Boxplot_Group_Structures is a function used to perform batch effects assessments using the boxplots and violin plots algorithms. Assessment is performed on each batch type available by default.

3 Output

The primary output method for MBatch is to view results in the Batch Effects Website, described elsewhere. The PNG files are rough versions of the website output.

Graphical output is a set of boxplots, one for each batch, based in this case on the mean for each sample.



4 Usage

Boxplot_Group_Structures(theData, theTitle, theOutputPath, theBatchTypeAndValuePairsToRemove, theBatchTypeAndValuePairsToKeep, theListOfGroupBoxFunction, theListOfGroupBoxLabels, theMaxGeneCount=20000, theJavaParameters = "-Xms8000m")

5 Arguments

theData An instance of BEA_DATA BEA_DATA-class. This is the MBatch Data Object (of class BEA_DATA) described in MBatch_03_UserData.docx, and returned from mbatchLoadFiles or mbatchLoadStructures.

theTitle Object of class "character". Title to use in PNG files.

theOutputPath Object of class "character". Directory in which to place output PNG files and related data files used by the Batch Effects Website.

theBatchTypeAndValuePairsToRemove Object of class "list". A list of vectors containing the batch type (or * for all types) and the value to remove. list() or NULL indicate do nothing. This type of list is described in MBatch_03_ParametersBatchTypesValues.docx.

theBatchTypeAndValuePairsToKeep Object of class "list". A list of vectors containing the batch type (or * for all types) and a vector of the value(s) to keep. list() or NULL indicate do nothing. This type of list is described in MBatch_03_ParametersBatchTypesValues.docx.

theListOfGroupBoxFunction Object of class "vector" A list of functions to use for group assessments. This is a legacy argument. Always send this an argument of "c(mean)". The length must match the length for theListOfGroupBoxLabels. A future version will replace this argument with an alternative method of selecting group functions other than mean.

theListOfGroupBoxLabels Object of class "vector" A list of strings giving strings for labels to use for group assessments. This is a legacy argument. Always send this an argument of "c("Mean")". The length must match the length for theListOfGroupBoxFunction. A future version will replace this argument with an alternative method of selecting group functions other than mean.

theMaxGeneCount Integer giving maximum number of features (genes) to keep. Default is 20000. 0 means keep all.

theJavaParameters Object of class "character" String for initializing JVM. Defaults to - Xms8000m.

6 Example Call

The following code performs Boxplots and is taken from the tests/Boxplot_Group_Structures.R file. Data used is from the testing data as per the MBatch_01_InstallLinux.docx document.

```
library(MBatch)
# set the paths
theGeneFile <- "/bea_testing/MATRIX_DATA/matrix_data-Tumor.tsv"
theBatchFile <- "/bea testing/MATRIX DATA/batches-Tumor.tsv"
theOutputDir <- "/bea_testing/output/Boxplot_Group_Structures"
# make sure the output dir exists and is empty
 unlink(theOutputDir, recursive=TRUE)
 dir.create(theOutputDir, showWarnings=FALSE, recursive=TRUE)
# load the data and reduce the amount of data to reduce run time
 myData <- mbatchLoadFiles(theGeneFile, theBatchFile)
 myData@mData <- mbatchTrimData(myData@mData, 100000)
 # here, we take most defaults
 Boxplot_Group_Structures(theData=myData,
               theTitle="Test",
               theOutputPath=theOutputDir,
               theBatchTypeAndValuePairsToRemove=NULL,
               theBatchTypeAndValuePairsToKeep=NULL,
               theListOfGroupBoxFunction=c(mean),
               theListOfGroupBoxLabels=c("Mean"),
               theJavaParameters="-Xms8000m",
               theMaxGeneCount=10000)
```

6.1 Command Line Output

In the future, we plan to make the output from MBatch more user friendly, but currently, this produces the following output at the command line.

```
> library(MBatch)
> # set the paths
> theGeneFile <- "/bea_testing/MATRIX_DATA/matrix_data-Tumor.tsv"
> theBatchFile <- "/bea_testing/MATRIX_DATA/batches-Tumor.tsv"
> theOutputDir <- "/bea_testing/output/Boxplot_Group_Structures"
> # make sure the output dir exists and is empty
> unlink(theOutputDir, recursive=TRUE)
> dir.create(theOutputDir, showWarnings=FALSE, recursive=TRUE)
```

```
> # load the data and reduce the amount of data to reduce run time
> myData <- mbatchLoadFiles(theGeneFile, theBatchFile)
2017 10 17 11:45:51.392 DEBUG MachineName Changing LC COLLATE to C for duration of run
2017 10 17 11:45:51.413 INFO MachineName V V V V V V V V V V V V V V V
2017 10 17 11:45:51.413 INFO MachineName Starting mbatchLoadFiles
2017 10 17 11:45:51.414 INFO MachineName MBatch Version: 2017-09-19-1530
2017 10 17 11:45:51.415 INFO MachineName read batch file= /bea_testing/MATRIX_DATA/batches-
2017 10 17 11:45:51.432 INFO MachineName read gene file=
/bea testing/MATRIX DATA/matrix data-Tumor.tsv
Read 100000 records
2017 10 17 11:46:03.431 INFO MachineName filter samples in batches using gene samples
2017 10 17 11:46:03.432 INFO MachineName sort batches by gene file samples
2017 10 17 11:46:03.526 INFO MachineName Finishing mbatchLoadFiles
2017 10 17 11:46:03.527 INFO MachineName ^^^^^^^^^^^^^^^^^^^^^^^^^^^
> myData@mData <- mbatchTrimData(myData@mData, 100000)
2017 10 17 11:46:03.528 DEBUG MachineName Changing LC_COLLATE to C for duration of run
2017 10 17 11:46:03.529 INFO MachineName mbatchTrimData Starting
2017 10 17 11:46:03.529 INFO MachineName MBatch Version: 2017-09-19-1530
2017 10 17 11:46:11.527 INFO MachineName mbatchTrimData Finishing
2017 10 17 11:46:11.528 INFO MachineName ^^^^^^^^^^^^^^^^^^^^^^^^^^
> # here, we take most defaults
> Boxplot Group Structures(theData=myData,
              theTitle="Test",
              theOutputPath=theOutputDir,
              theBatchTypeAndValuePairsToRemove=NULL,
              theBatchTypeAndValuePairsToKeep=NULL,
              theListOfGroupBoxFunction=c(mean),
              theListOfGroupBoxLabels=c("Mean"),
              theJavaParameters="-Xms8000m",
              theMaxGeneCount=10000)
2017 10 17 11:46:11.534 DEBUG MachineName Changing LC_COLLATE to C for duration of run
2017 10 17 11:46:11.537 INFO MachineName V V V V V V V V V V V V V V V V
2017 10 17 11:46:11.538 INFO MachineName mbatchFilterData Starting
2017 10 17 11:46:11.538 INFO MachineName MBatch Version: 2017-09-19-1530
2017 10 17 11:46:11.539 DEBUG MachineName rows pre filter 1250
2017 10 17 11:46:11.820 DEBUG MachineName rows post filter 1250
2017 10 17 11:46:11.820 DEBUG MachineName mbatchFilterData Prefilter, gene data had 1250 while
post filter 1250
2017 10 17 11:46:11.822 DEBUG MachineName mbatchFilterData Prefilter, batch data had 80 while
post filter 80
2017 10 17 11:46:11.822 INFO MachineName mbatchFilterData Finishing
2017 10 17 11:46:11.823 INFO MachineName ^^^^^^^^^^^^^^^^^^^^^^^
2017 10 17 11:46:11.838 DEBUG MachineName Changing LC_COLLATE to C for duration of run
2017 10 17 11:46:11.839 INFO MachineName mbatchTrimData Starting
2017 10 17 11:46:11.840 INFO MachineName MBatch Version: 2017-09-19-1530
2017 10 17 11:46:11.840 INFO MachineName mbatchTrimData Finishing
2017 10 17 11:46:11.841 INFO MachineName ^^^^^^^^^^^^^^^^^^^^^^^^
```

```
2017 10 17 11:46:11.842 DEBUG MachineName checkCreateDir:
/bea testing/output/Boxplot Group Structures
2017 10 17 11:46:11.845 DEBUG MachineName boxplotJinit - Calling .jinit /home/linux/R/x86_64-pc-
linux-gnu-library/3.4/MBatch/BoxplotJava/jcommon-1.0.17.jar:/home/linux/R/x86 64-pc-linux-gnu-
library/3.4/MBatch/BoxplotJava/ifreechart-1.0.14.jar:/home/linux/R/x86 64-pc-linux-gnu-
library/3.4/MBatch/BoxplotJava/commons-lang3-3.3.2.jar:/home/linux/R/x86_64-pc-linux-gnu-
library/3.4/MBatch/BoxplotJava/commons-math3-3.3.jar:/home/linux/R/x86_64-pc-linux-gnu-
library/3.4/MBatch/BoxplotJava/BoxplotJava.jar:/home/linux/R/x86 64-pc-linux-gnu-
library/3.4/MBatch/BoxplotJava/LegendJava.jar
2017 10 17 11:46:12.164 DEBUG MachineName dim(theMatrixGeneData) 1250,
dim(theMatrixGeneData) 80
2017 10 17 11:46:12.165 DEBUG MachineName length(colnames(theMatrixGeneData)) 80
2017 10 17 11:46:12.165 DEBUG MachineName length(rownames(theMatrixGeneData)) 1250
2017 10 17 11:46:12.166 DEBUG MachineName dim(theDataframeBatchData) 80,
dim(theDataframeBatchData) 5
2017 10 17 11:46:12.166 DEBUG MachineName length(names(theDataframeBatchData)) 5
2017_10_17_1146.12.201
                                                   BoxplotJava 2017-07-24-1300
                             DEBUG
                                                   Prepare data for Boxplot GroupFunction
2017 10 17 1146.12.202
                             DEBUG
                                           1
2017 10 17 1146.12.202
                             DEBUG
       theOutputDir=/bea_testing/output/Boxplot_Group_Structures
2017_10_17_1146.12.202
                             DEBUG
                                                   theTitle=Test Group
2017_10_17_1146.12.204
                             DEBUG
                                           1
                                                   ReadMatrixFile::convertToMatrix
theBarcodeLength=80
2017 10 17 1146.12.204
                                           1
                                                   ReadMatrixFile::convertToMatrix
                             DEBUG
theGeneLength=1250
2017 10 17 1146.12.208
                             DEBUG
                                           1
                                                   Before Boxplot_GroupFunction
2017_10_17_1146.12.223
                                            1
                                                   BoxplotJava 2017-07-24-1300
                             DEBUG
2017 10 17 1146.12.225
                             DEBUG
                                           1
                                                   Boxplot GroupFunction::processInternal - start
2017_10_17_1146.12.225
                                                   Boxplot_GroupFunction::processInternal -
                                           1
                             DEBUG
batchTypeName=BatchId
2017 10 17 1146.12.225
                             DEBUG
                                           1
                                                   Boxplot GroupFunction::processInternal - list of
batches
2017 10 17 1146.12.226
                             DEBUG
                                           1
                                                   Boxplot GroupFunction::processInternal -
group function values for samples
2017_10_17_1146.12.280
                             DEBUG
                                           1
                                                   Boxplot_GroupFunction::processInternal -
iterate through batches
2017_10_17_1146.12.280
                             DEBUG
                                           1
                                                   Boxplot_GroupFunction::processInternal -
iterate through batches - before inside loop
2017 10 17 1146.12.280
                             DEBUG
                                           1
                                                   Boxplot GroupFunction::processInternal -
iterate through batches - after inside loop
2017 10 17 1146.12.280
                             DEBUG
                                            1
2017 10 17 1146.12.280
                                                   Boxplot GroupFunction
                             DEBUG
                                            1
2017_10_17_1146.12.281
                                           1
                             DEBUG
2017_10_17_1146.12.281
                             DEBUG
                                           1
                                                   batchTypeIndex=0
2017_10_17_1146.12.281
                                                   values.length=80
                             DEBUG
                                            1
2017_10_17_1146.12.281
                             DEBUG
                                            1
                                                   valLabels.length=80
2017_10_17_1146.12.281
                             DEBUG
                                            1
                                                   elementLabel=00304
                                           1
2017 10 17 1146.12.281
                                                   groupLabel=00304
                             DEBUG
2017 10 17 1146.12.281
                                            1
                             DEBUG
                                                   groupId=0
2017_10_17_1146.12.281
                             DEBUG
```

2017_10_17_1146.12.294	DEBUG	1	Boxplot_GroupFunction::processInternal - after
iterate through batches			
2017_10_17_1146.12.294	DEBUG	1	writeBoxplotFiles - boxplotPng
2017_10_17_1146.12.294	DEBUG	1	BoxplotImpl::boxplotPng theElements.size()=1
2017_10_17_1146.12.295	DEBUG	1	BoxplotImpl::boxplotPng
thePNGFile=/bea_testing/output	ut/Boxplot_Grou	p_Struct	ures/Group-MEAN/BoxPlot_Group-
MEAN_Diagram-BatchId.png			
2017_10_17_1146.12.295	DEBUG	1	BoxplotImpl::boxplotPng theDataLabel=values
2017_10_17_1146.12.295	DEBUG	1	BoxplotImpl::boxplotPng
theSampleLabel=BatchId			
2017_10_17_1146.12.295	DEBUG	1	BoxplotImpl::boxplotPng theTitle=Test Group
2017_10_17_1146.12.295	DEBUG	1	BoxplotImpl::boxplotPng theGroup=BatchId
LegendJava 2013_05_03_0823			
writeLegendWithSymbols theT	Title = Test Grou	p - Batcl	ıld
writeLegendWithSymbols the\		•	
		oea testi	ng/output/Boxplot_Group_Structures/Group-
MEAN/BoxPlot_Group-MEAN			
Colors is non-null	- 0	1 0	
writeLegendWithSymbols writ	e		
writeLegendWithSymbols done			
2017_10_17_1146.16.121	DEBUG	1	writeBoxplotFiles - boxplotAnnotations
2017_10_17_1146.16.170	DEBUG	1	writeBoxplotFiles - boxplotBox
2017_10_17_1146.16.171	DEBUG	1	writeBoxplotFiles - boxplotCat
2017_10_17_1146.16.172	DEBUG	1	writeBoxplotFiles - done
2017_10_17_1146.16.172	DEBUG	1	Boxplot_GroupFunction::processInternal - after
files	DEBCG	•	Boxpiot_Groupi unetionprocessinternar urter
2017_10_17_1146.16.172	DEBUG	1	Boxplot_GroupFunction::processInternal -
batchTypeName=PlateId	DLBCG	1	Boxplot_Groupf unctionprocessinternal
2017_10_17_1146.16.172	DEBUG	1	Boxplot_GroupFunction::processInternal - list of
batches	DLBCG	1	Boxplot_Groupf unctionprocessinternal list of
2017_10_17_1146.16.173	DEBUG	1	Boxplot_GroupFunction::processInternal -
group function values for samp		1	Boxplot_Groupf unctionprocessinternal
2017_10_17_1146.16.187	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches	DEBUG	1	Boxpiot_GroupFunctionprocessinternar-
2017_10_17_1146.16.187	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - before		1	Boxplot_OroupFunctionprocessinternal -
2017_10_17_1146.16.188	DEBUG	1	Boxplot_GroupFunction::processInternal -
		1	Boxplot_GroupFunctionprocessinternal -
iterate through batches - after in 2017_10_17_1146.16.188	DEBUG	1	
2017_10_17_1146.16.188		1	Downlot CrownEuration
	DEBUG	1	Boxplot_GroupFunction
2017_10_17_1146.16.188	DEBUG	1	1 . 1 Tp . T . 1
2017_10_17_1146.16.188	DEBUG	1	batchTypeIndex=1
2017_10_17_1146.16.188	DEBUG	1	values.length=80
2017_10_17_1146.16.188	DEBUG	1	valLabels.length=80
2017_10_17_1146.16.188	DEBUG	1	elementLabel=A29J
2017_10_17_1146.16.188	DEBUG	1	groupLabel=A29J
2017_10_17_1146.16.189	DEBUG	1	groupId=0
2017_10_17_1146.16.189	DEBUG	1	
2017_10_17_1146.16.189	DEBUG	1	Boxplot_GroupFunction::processInternal - after
iterate through batches			
2017_10_17_1146.16.189	DEBUG	1	writeBoxplotFiles - boxplotPng

2017_10_17_1146.16.189	DEBUG	1	BoxplotImpl::boxplotPng theElements.size()=1
2017_10_17_1146.16.189	DEBUG	1	BoxplotImpl::boxplotPng
	ut/Boxplot_Grou	p_Struct	tures/Group-MEAN/BoxPlot_Group-
MEAN_Diagram-PlateId.png	DEDUC	1	
2017_10_17_1146.16.190	DEBUG	1	BoxplotImpl::boxplotPng theDataLabel=values
2017_10_17_1146.16.190	DEBUG	1	BoxplotImpl::boxplotPng
theSampleLabel=PlateId	DEDUG		
2017_10_17_1146.16.190	DEBUG	1	BoxplotImpl::boxplotPng theTitle=Test Group
2017_10_17_1146.16.191	DEBUG	1	BoxplotImpl::boxplotPng theGroup=PlateId
LegendJava 2013_05_03_0823		DI.	Y 1
writeLegendWithSymbols theT		p - Plate	ald
writeLegendWithSymbols the			
			ing/output/Boxplot_Group_Structures/Group-
MEAN/BoxPlot_Group-MEAN	N_Legend-PlateI	d.png	
Colors is non-null			
writeLegendWithSymbols writ			
writeLegendWithSymbols done		1	. D 1 (D)1 1 1 1 . A
2017_10_17_1146.16.346	DEBUG	1	writeBoxplotFiles - boxplotAnnotations
2017_10_17_1146.16.347	DEBUG	1	writeBoxplotFiles - boxplotBox
2017_10_17_1146.16.348	DEBUG	1	writeBoxplotFiles - boxplotCat
2017_10_17_1146.16.349	DEBUG	1	writeBoxplotFiles - done
2017_10_17_1146.16.349	DEBUG	1	Boxplot_GroupFunction::processInternal - after
files	DEBLIC	1	
2017_10_17_1146.16.350	DEBUG	1	Boxplot_GroupFunction::processInternal -
batchTypeName=ShipDate	DEBLIC	1	
2017_10_17_1146.16.350	DEBUG	1	Boxplot_GroupFunction::processInternal - list of
batches	DEDLIC	1	Danielat Casarenationama assolutamal
2017_10_17_1146.16.350	DEBUG	1	Boxplot_GroupFunction::processInternal -
group function values for samp		1	Daymlat ChaupEunationuma aggsIntamal
2017_10_17_1146.16.360 iterate through batches	DEBUG	1	Boxplot_GroupFunction::processInternal -
	DEBUG	1	Payalat GroupEunationuprocessInternal
2017_10_17_1146.16.360		1	Boxplot_GroupFunction::processInternal -
iterate through batches - before	DEBUG	1	Daymlat ChaupEunationuma aggsIntamal
2017_10_17_1146.16.360		1	Boxplot_GroupFunction::processInternal -
iterate through batches - after in 2017 10 17 1146.16.360	DEBUG	1	
2017_10_17_1146.16.361		1	Poyelst Countypotion
	DEBUG DEBUG	1	Boxplot_GroupFunction
2017_10_17_1146.16.361	DEBUG	1	batchTypeIndex=2
2017_10_17_1146.16.361	DEBUG	_	▼ ±
2017_10_17_1146.16.361 2017_10_17_1146.16.361		1	values.length=80
	DEBUG	1	valLabels.length=80
2017_10_17_1146.16.361	DEBUG	1	elementLabel=2013-05-08
2017_10_17_1146.16.361	DEBUG	1	groupId=0
2017_10_17_1146.16.361	DEBUG	1	groupId=0
2017_10_17_1146.16.362	DEBUG	1	Payriot CrownEurotionumacoccInternal often
2017_10_17_1146.16.362	DEBUG	1	Boxplot_GroupFunction::processInternal - after
iterate through batches	DEDUC	1	wwitaDaymlatEilaa hawalatDa
2017_10_17_1146.16.362	DEBUG	1	writeBoxplotFiles - boxplotPng PoynlotPng theFlorents size()-1
2017_10_17_1146.16.362	DEBUG	1	BoxplotImpl::boxplotPng theElements.size()=1

2017_10_17_1146.16.362	DEBUG	1	BoxplotImpl::boxplotPng	
	•	ip_Struc	tures/Group-MEAN/BoxPlot_Group-	
MEAN_Diagram-ShipDate.png				
2017_10_17_1146.16.362	DEBUG	1	BoxplotImpl::boxplotPng theDataLabel=values	
2017_10_17_1146.16.362	DEBUG	1	BoxplotImpl::boxplotPng	
theSampleLabel=ShipDate				
2017_10_17_1146.16.363	DEBUG	1	BoxplotImpl::boxplotPng theTitle=Test Group	
2017_10_17_1146.16.363	DEBUG	1	BoxplotImpl::boxplotPng theGroup=ShipDate	
LegendJava 2013_05_03_0823	3			
writeLegendWithSymbols theTitle = Test Group - ShipDate				
writeLegendWithSymbols theV	Version =			
writeLegendWithSymbols theF	FilenamePath = /	bea_test	ing/output/Boxplot_Group_Structures/Group-	
MEAN/BoxPlot_Group-MEAN	N_Legend-ShipD	Date.png		
Colors is non-null				
writeLegendWithSymbols writ	te			
writeLegendWithSymbols don				
2017_10_17_1146.16.482	DEBUG	1	writeBoxplotFiles - boxplotAnnotations	
2017_10_17_1146.16.483	DEBUG	1	writeBoxplotFiles - boxplotBox	
2017_10_17_1146.16.484	DEBUG	1	writeBoxplotFiles - boxplotCat	
2017_10_17_1146.16.484	DEBUG	1	writeBoxplotFiles - done	
2017_10_17_1146.16.484	DEBUG	1	Boxplot_GroupFunction::processInternal - after	
files			1 - 1	
2017_10_17_1146.16.484	DEBUG	1	Boxplot_GroupFunction::processInternal -	
batchTypeName=TSS	22200	-	Zonprov_oroup: uncoronnprovossmoonum	
2017_10_17_1146.16.484	DEBUG	1	Boxplot_GroupFunction::processInternal - list of	
batches	DEBCC	•	Bonprot_Groups unerrom.processmeethur institut	
2017_10_17_1146.16.484	DEBUG	1	Boxplot_GroupFunction::processInternal -	
group function values for samp		1	Boxplot_Groupf unctionprocessmermar	
2017_10_17_1146.16.492	DEBUG	1	Boxplot_GroupFunction::processInternal -	
iterate through batches	DEDUC	1	Boxplot_Groupt unetionprocessinternal	
2017_10_17_1146.16.492	DEBUG	1	Boxplot_GroupFunction::processInternal -	
iterate through batches - before		1	Boxplot_Groupt unetionprocessinternal	
2017_10_17_1146.16.492	DEBUG	1	Boxplot_GroupFunction::processInternal -	
iterate through batches - after i		1	Boxplot_Groupt unctionprocessmermar	
2017_10_17_1146.16.492	DEBUG	1		
2017_10_17_1140.10.492	DEBUG	1		
2017_10_17_1146.16.493	DEBUG	1	Boxplot_GroupFunction	
2017_10_17_1146.16.493		Ī.	hotah Tura Inday 2	
	DEBUG	1	batchTypeIndex=3	
2017_10_17_1146.16.493	DEBUG	1	values.length=72	
2017_10_17_1146.16.493	DEBUG	1	valLabels.length=72	
2017_10_17_1146.16.493	DEBUG	1	elementLabel=OR - University of Michigan	
2017_10_17_1146.16.493	DEBUG	1	groupLabel=OR - University of Michigan	
2017_10_17_1146.16.493	DEBUG	1	groupId=0	
2017_10_17_1146.16.493	DEBUG	1		
2017_10_17_1146.16.493	DEBUG	1	Boxplot_GroupFunction::processInternal -	
iterate through batches - before				
2017_10_17_1146.16.493	DEBUG	1	Boxplot_GroupFunction::processInternal -	
iterate through batches - after i	•			
2017_10_17_1146.16.493	DEBUG	1	=======================================	
2017_10_17_1146.16.493	DEBUG	1	Boxplot_GroupFunction	
2017_10_17_1146.16.493	DEBUG	1	=======================================	

2017_10_17_1146.16.493	DEBUG	1	batchTypeIndex=3
2017_10_17_1146.16.494	DEBUG	1	values.length=1
2017_10_17_1146.16.494	DEBUG	1	valLabels.length=1
2017_10_17_1146.16.494	DEBUG	1	elementLabel=OU - Roswell Park
2017_10_17_1146.16.494	DEBUG	1	groupLabel=OU - Roswell Park
2017_10_17_1146.16.494	DEBUG	1	groupId=1
2017_10_17_1146.16.494	DEBUG	1	
2017_10_17_1146.16.494	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - before			
2017_10_17_1146.16.494	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - after i	•		
2017_10_17_1146.16.494	DEBUG	1	
2017_10_17_1146.16.494	DEBUG	1	Boxplot_GroupFunction
2017_10_17_1146.16.494	DEBUG	1	=======================================
2017_10_17_1146.16.494	DEBUG	1	batchTypeIndex=3
2017_10_17_1146.16.494	DEBUG	1	values.length=2
2017_10_17_1146.16.494	DEBUG	1	valLabels.length=2
2017_10_17_1146.16.494	DEBUG	1	elementLabel=P6 - Translational Genomics
Research Institute			
2017_10_17_1146.16.494	DEBUG	1	groupLabel=P6 - Translational Genomics
Research Institute			•
2017_10_17_1146.16.495	DEBUG	1	groupId=2
2017_10_17_1146.16.495	DEBUG	1	=======================================
2017_10_17_1146.16.495	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - before	inside loop		
2017_10_17_1146.16.495	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - after i	nside loop		• •
2017_10_17_1146.16.495	DEBUĞ	1	=======================================
2017_10_17_1146.16.495	DEBUG	1	Boxplot_GroupFunction
2017_10_17_1146.16.495	DEBUG	1	=======================================
2017_10_17_1146.16.495	DEBUG	1	batchTypeIndex=3
2017_10_17_1146.16.495	DEBUG	1	values.length=1
2017_10_17_1146.16.495	DEBUG	1	valLabels.length=1
2017_10_17_1146.16.495	DEBUG	1	elementLabel=PA - University of Minnesota
2017_10_17_1146.16.495	DEBUG	1	groupLabel=PA - University of Minnesota
2017_10_17_1146.16.495	DEBUG	1	groupId=3
2017_10_17_1146.16.495	DEBUG	1	=======================================
2017_10_17_1146.16.495	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - before	inside loop		1 – 1
2017_10_17_1146.16.496	DEBUG	1	Boxplot_GroupFunction::processInternal -
iterate through batches - after i	nside loop		1 - 1
2017_10_17_1146.16.496	DEBUG	1	=======================================
2017_10_17_1146.16.496	DEBUG	1	Boxplot_GroupFunction
2017_10_17_1146.16.496	DEBUG	1	=======================================
2017_10_17_1146.16.496	DEBUG	1	batchTypeIndex=3
2017_10_17_1146.16.496	DEBUG	1	values.length=4
2017_10_17_1146.16.496	DEBUG	1	valLabels.length=4
2017_10_17_1146.16.499	DEBUG	1	elementLabel=PK - University Health Network
2017_10_17_1146.16.499	DEBUG	1	groupLabel=PK - University Health Network
2017_10_17_1146.16.499	DEBUG	1	groupId=4
2017_10_17_1146.16.499	DEBUG	1	=======================================

EBUG	1	Boxplot_GroupFunction::processInternal - after		
		writeBoxplotFiles - boxplotPng		
EBUG	1	BoxplotImpl::boxplotPng theElements.size()=5		
EBUG	1	BoxplotImpl::boxplotPng		
oxplot_Group	_Structi	ures/Group-MEAN/BoxPlot_Group-		
EBUG	1	BoxplotImpl::boxplotPng theDataLabel=values		
EBUG	1	BoxplotImpl::boxplotPng theSampleLabel=TSS		
EBUG	1	BoxplotImpl::boxplotPng theTitle=Test Group		
EBUG	1	BoxplotImpl::boxplotPng theGroup=TSS		
= Test Group	- TSS			
ion =				
writeLegendWithSymbols theFilenamePath = /bea_testing/output/Boxplot_Group_Structures/Group-				
MEAN/BoxPlot_Group-MEAN_Legend-TSS.png				
Colors is non-null				
EBUG	1	writeBoxplotFiles - boxplotAnnotations		
EBUG	1	writeBoxplotFiles - boxplotBox		
EBUG	1	writeBoxplotFiles - boxplotCat		
EBUG	1	writeBoxplotFiles - done		
EBUG	1	Boxplot_GroupFunction::processInternal - after		
EBUG	1	Boxplot_GroupFunction::processInternal -		
EBUG	1	After Boxplot_GroupFunction		
2017 10 17 11:46:16.698 DEBUG MachineName after groupFunction call				
	EBUG EBUG EBUG EBUG EBUG EBUG EBUG EBUG	EBUG 1		

6.2 Example File Output

The above code creates the following output files. Files are named using the following naming convention:

BoxPlot_Group-MEAN_Diagram-<BatchType>.png

BoxPlot_Group-MEAN_Legend-<BatchType>.png

The diagram file contains a Boxplot for all batch types (the columns from the batches.tsv file). The legends give the list of batches within the given batch type, and the colors used for each batch.

The annotation TSV files are used internally and contain the total number of data points and the number of non-NA points..

```
linux@MachineName:/bea_testing/output/Boxplot_Group_Structures/Group-MEAN$ ls -l
total 188
-rw-r--r-- 1 linux linux
                               54 Oct 17 11:46 BoxPlot_Group-MEAN_Annotations-BatchId.tsv
-rw-r--r- 1 linux linux 53 Oct 17 11:46 BoxPlot_Group-MEAN_Annotations-PlateId.tsv
-rw-r--r-- 1 linux linux 59 Oct 17 11:46 BoxPlot_Group-MEAN_Annotations-ShipDat
-rw-r--r-- 1 linux linux 265 Oct 17 11:46 BoxPlot_Group-MEAN_Annotations-TSS.tsv
                               59 Oct 17 11:46 BoxPlot_Group-MEAN_Annotations-ShipDate.tsv
-rw-r--r- 1 linux linux 283 Oct 17 11:46 BoxPlot_Group-MEAN_BoxData-BatchId.tsv
-rw-r--r- 1 linux linux 282 Oct 17 11:46 BoxPlot_Group-MEAN_BoxData-PlateId.tsv
-rw-r--r-- 1 linux linux
                              288 Oct 17 11:46 BoxPlot_Group-MEAN_BoxData-ShipDate.tsv
-rw-r--r-- 1 linux linux 1013 Oct 17 11:46 BoxPlot_Group-MEAN_BoxData-TSS.tsv
-rw-r--r-- 1 linux linux 9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-BatchId-00304.tsv
                            9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-PlateId-A29J.tsv
9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-ShipDate-2013-05-08.tsv
9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-TSS-OR - University of
-rw-r--r-- 1 linux linux
-rw-r--r-- 1 linux linux
-rw-r--r-- 1 linux linux
Michigan.tsv
-rw-r--r-- 1 linux linux
                                9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-TSS-OU - Roswell Park.tsv
-rw-r--r-- 1 linux linux 9 Oct 1/ 11:46 BoxPlot_Group-MEAN_CatData-155-00 - Noswell Fulk.
-rw-r--r-- 1 linux linux 9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-TSS-P6 - Translational
Genomics Research Institute.tsv
-rw-r--r-- 1 linux linux
                                9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-TSS-PA - University of
Minnesota.tsv
-rw-r--r-- 1 linux linux
                             9 Oct 17 11:46 BoxPlot_Group-MEAN_CatData-TSS-PK - University Health
Network.tsv
-rw-r--r- 1 linux linux 23062 Oct 17 11:46 BoxPlot_Group-MEAN_Diagram-BatchId.png
-rw-r--r-- 1 linux linux 22846 Oct 17 11:46 BoxPlot_Group-MEAN_Diagram-PlateId.png
-rw-r--r- 1 linux linux 23394 Oct 17 11:46 BoxPlot_Group-MEAN_Diagram-ShipDate.png
-rw-r--r- 1 linux linux 24699 Oct 17 11:46 BoxPlot_Group-MEAN_Diagram-TSS.png
-rw-r--r-- 1 linux linux 2108 Oct 17 11:46 BoxPlot_Group-MEAN_Legend-BatchId.png
-rw-r--r-- 1 linux linux 2018 Oct 17 11:46 BoxPlot_Group-MEAN_Legend-PlateId.png
-rw-r--r 1 linux linux 2636 Oct 17 11:46 BoxPlot_Group-MEAN_Legend-ShipDate.png
-rw-r--r-- 1 linux linux 11139 Oct 17 11:46 BoxPlot_Group-MEAN_Legend-TSS.png
linux@MachineName:/bea_testing/output/Boxplot_Group_Structures/Group-MEAN$
```