Proof of Concept for RTF Render

#' Format Model Estimator  
#'  
#' The function assume 1 or 2 column.   
#' If there is only 1 column, only represent mean  
#' If there are 2 column, represent mean (sd) or mean(se)  
#' Decimals will understand the number will be formated as x.x(x.xx)   
fmt\_est <- function(data, columns = c("mean", "sd"), decimals = c(1,2)){  
 .mean = formatC(data[[columns[[1]]]], digits = decimals[1], format = "f", flag = "0")  
 if(length(columns) > 1){  
 .sd = formatC(data[[columns[[2]]]], digits = decimals[2], format = "f", flag = "0")  
 paste0(.mean, " (", .sd, ")")  
 }else{  
 .mean  
 }  
  
}   
  
#' Format Confidence Interval  
fmt\_ci <- function(data, columns = c("lower.CL", "upper.CL"), decimals = 2){  
 .lower = formatC(data[[columns[[1]]]], digits = decimals, format = "f", flag = "0")  
 .upper = formatC(data[[columns[[2]]]], digits = decimals, format = "f", flag = "0")  
 paste0("(", .lower, ", " , .upper, ")")  
}  
  
#' Format P-Value   
fmt\_pval <- function(data, columns = "p.value", decimals = 3){  
 scale <- 10 ^ (-1 \* decimals)  
 p\_scale <- paste0("<",scale)  
 if\_else(data[[columns[[1]]]] < scale, p\_scale,   
 formatC(data[[columns[[1]]]], digits = decimals, format = "f", flag = "0") )  
}

# RTF Table Standard

## General Guidance

* FDA Specification Document: <https://www.fda.gov/downloads/drugs/ucm120957.pdf>
* RTF Pocket Guide <http://shop.oreilly.com/product/9780596004750.do>
* RTF 1.5 specification <http://www.biblioscape.com/rtf15_spec.htm>
* RTF 1.9 specification <https://www.microsoft.com/en-us/download/details.aspx?id=10725>

## Merck SAS Macro

* Merck SAS macro %RTFtable Introduction <https://www.lexjansen.com/pharmasug/2003/TechnicalTechniques/tt038.pdf>
* Merck SAS macro %RTFsymbol Introduction <https://www.lexjansen.com/nesug/nesug02/cc/cc010.pdf>

## Summary of Table Standard

### Title

* Outside of table
* Times New Roman, 12 point (Required by FDA) {\\f0\\froman Times New Roman;}
* If a table is multi-paged, the table title/subtitles must be repeated at the top of each page of the table. (Required by FDA)
* Soft returns must be used after each title/subtitle line (Required for automated generation of List of Tables) \line
* Titles and subtitles should be centered over the top border of table \qc tools::toTitleCase

start\_rtf <- '{\\rtf1\\ansi'  
  
style <- list(  
 start\_style = '\\deff0\\deflang1033{\\fonttbl',  
 font0 = '{\\f0\\froman Times New Roman;}',  
 font1 = '{\\f1\\froman\\fcharset2\\fprq2 Symbol;}',  
 font166 = '{\\f166\\froman\\fcharset161\\fprq2 Times New Roman Greek;}',  
 font266 = '{\\f266\\fswiss\\fcharset161\\fprq2 Arial Greek;}',  
 end\_style = "}"  
)  
  
title <- list(  
 title\_format <- '\\qc \\f0\\fs24 \\lang1033\\langfe1033\\cgrid\\langnp1033\\langfenp103',  
 title\_value = 'This is Title \\line This is Subtitle \\line',  
 title\_end = '\\par'  
)

### Column Headings

* Times New Roman 8 to 12 point (Required by FDA)
* If a table is multi-paged, the column headings should be repeated at the top of each page of the table. (Required by FDA)
* Column headings should be centered
* Column headings should use Title Case

### Subheadings on Tables

* Times New Roman, 8 to 12 point
* Contained in single cell merged to the width of the table to avoid unnecessary wrapping.
* Left justified
* Format characters in Bold font style, with 3-point vertical spacing

### Cell Contents

* Times New Roman, 8 to 12 point (8 to 10 point recommended)
* Left justified with hanging indent (approximately two character spaces).

row\_start <- '\\trowd\\trgaph108\\trleft0\\trqc'  
  
row\_boarder <- list(  
 '\\trbrdrt\\brdrdb\\brdrw15\\trbrdrl\\brdrs\\brdrw15\\trbrdrb\\brdrdb\\brdrw15\\trbrdrr\\brdrs\\brdrw15',  
 '\\clbrdrt\\brdrdb\\brdrw15\\clbrdrl\\brdrs\\brdrw15\\clbrdrr\\brdrhair\\cellx1767',  
 '\\clbrdrt\\brdrdb\\brdrw15\\clbrdrb\\brdrhair\\clbrdrr\brdrs\\brdrw15\\cellx3403',  
 end = ''  
)

### Table Borders

* 3/4 point lines on all four sides of the table
* On each page of the table double lines should appear at the very top of the table. The lines at the bottom of each page should be single, except on the last page of the table. On the last page of the table, the bottom line should be a double line.
* 3/4 point vertical lines between columns
* 3/4 point horizontal lines between column header row and table text, and between footnote row and table text.

row\_cells <- list(  
 list(  
 start = '\\pard\\intbl',  
 cell\_format = '\\qc\\sb15\\sa15 \\f0\\fs16',  
 cell\_value = 'x.x',  
 cell\_split = '\\cell'  
 ),  
   
 list(  
 start = '\\pard\\intbl',  
 cell\_format = '\\qc\\sb15\\sa15 \\f0\\fs16',  
 cell\_value = 'y.y',  
 cell\_split = '\\cell'  
 )  
)  
row\_cells <- lapply(row\_cells, paste0, collapse = "")  
row\_end <- '\\intbl\\row\\pard'  
row <- c(row\_start, row\_boarder, row\_cells, row\_end)   
end\_rtf <- "}"

### Bolding

* Category heading lines on counts tables are bolded.
* If a category heading is bolded, then the entire line on the table is bolded, including the n and (%) values.

### Indenting

* On counts tables, the category sub-terms should be indented approximately two character spaces in from the left. \\sb15\\sa15

### Data Source LIne:

### Page Orientation

* The default orientation for all tables is portrait. Landscape should be used in tables and figures only if absolutely necessary. For listings, landscape is usually necessary.

### Margin Format

* Page size 8.5 by 11 inches
* Potrait Margin (inch): top 1, bottom 1, left 1.25, right 1
* Landscape Margin (inch): top 1.25, bottom 1, left 0.5, right 0.5
* Based upon these margins (and the 8.5 by 11 inch paper size), the “content area”, i.e., the area available for use, is 6.25" by 9" for portrait layout and 10" by 6.25" for landscape layout.

margin <- list(  
 landscape = '\\margl720\\margr720\\margt1843\\margb1800\\headery1800\\footery1440',  
 portrait = "\\margl1800\\margr1440\\margt2160\\margb1440\\headery720\\footery720"  
 )

### Assemble RTF

res <- unlist(c(  
 start\_rtf,   
 style,  
 margin[["portrait"]],  
 title,  
 row,  
 end\_rtf  
)  
)  
  
fileConn <- file("tmp.rtf")  
writeLines(res,fileConn)  
close(fileConn)

# Examples

## Example 1: ANCOVA analysis for HOMA data

* The data is avaliable at <https://missingdata.lshtm.ac.uk/dia-working-group/example-data-sets/>
* For example, we want to summary the analysis results using ANCOVA model

load("HAMD17.Rdata")  
ana\_week = 8 # Analysis Week   
  
high\_lmfit <- high %>% filter(week == ana\_week) %>%   
 lm(change ~ basval + TRT, data = .)

### Raw summary

t11 <- high %>% filter(week == ana\_week) %>%   
 group\_by(TRT) %>%   
 summarise( N = n(),  
 mean\_bl = mean(basval),  
 sd\_bl = sd(basval),  
 mean = mean(change),   
 sd = sd(change))

### LS mean

t12 <- emmeans(high\_lmfit, "TRT")  
t1 <- merge(t11, t12) %>%   
 mutate(emmean\_sd = SE \* sqrt(df)) %>%  
 mutate( Trt = c("Study Drug", "Placebo"),  
 N1 = N,  
 Mean1 = fmt\_est(., c("mean\_bl", "sd\_bl") ),  
 N2 = N,   
 Mean2 = fmt\_est(., c("mean", "sd")),  
 N3 = N,   
 Mean3 = fmt\_est(., c("emmean", "emmean\_sd")),   
 CI = paste(fmt\_est(., "emmean"), fmt\_ci(., c("lower.CL", "upper.CL")))  
 ) %>%   
 select(Trt:CI)

### Treatment Comparision

t2 <- data.frame(pairs(t12))  
  
t2 <- t2 %>% mutate(  
 lower = estimate - 1.96 \* SE,   
 upper = estimate + 1.96 \* SE) %>%  
 mutate(  
 comp = "Study Drug vs. Placebo",  
 mean = paste(fmt\_est(., "estimate"), fmt\_ci(., c("lower", "upper"))),   
 p = fmt\_pval(., "p.value")  
 ) %>%   
 select(comp:p)

### RMSE

t3 <- data.frame(rmse = paste0("Root Mean Squared Error of Change = ",   
 formatC(sd(high\_lmfit$residuals), digits = 2, format = "f", flag = "0")))

The prupose of this exercise is to create a table as in rtf\_example1.rtf by using the three datasets t1, t2 and t3 ## Create gt Class

t1

## Trt N1 Mean1 N2 Mean2 N3 Mean3  
## 1 Study Drug 61 16.6 (4.41) 61 -6.6 (5.95) 61 -7.0 (9.16)  
## 2 Placebo 70 18.4 (6.34) 70 -9.0 (7.04) 70 -8.7 (8.54)  
## CI  
## 1 -7.0 (-8.58, -5.38)  
## 2 -8.7 (-10.17, -7.18)

t2

## comp mean p  
## 1 Study Drug vs. Placebo 1.7 (-0.49, 3.88) 0.130

t3

## rmse  
## 1 Root Mean Squared Error of Change = 6.23

t1\_gt <- gt(t1) %>%  
 tab\_header(title = "ANCOVA of Change from Baseline at Week 8",   
 subtitle = "Missing Data Approach \n Analysis Population") %>%  
 tab\_footnote("$^\\dagger$Based on an ANCOVA model.\nANCOVA = Analysis of Covariance, CI = Confidence Interval, LS = Least Squares, SD = Standard Deviation",   
 locations = cells\_column\_labels(8)) %>%  
 cols\_label(  
 Trt = "Treatment",   
 N1 = "N",   
 Mean1 = "Mean (SD)",  
 N2 = "N",   
 Mean2 = "Mean (SD)",  
 N3 = "N",   
 Mean3 = "Mean (SD)",  
 CI = "$LS Mean (95% CI)\\dagger$" # with latex words  
 )   
  
t2\_gt <- gt(t2) %>%  
 cols\_label(  
 comp = "Pairwise Comparison",   
 mean = "Difference in LS Mean (95% CI)\\dagger$",  
 p = "p-Value"  
 )  
  
t3\_gt <- gt(t3)

### Suggested Features:

* tab\_footnote should have an option to not show footnote symbol or customize the symbol of the footnote. (e.g. , )
* one footnote could be used in multiple places in a table.
* some footnote does not require a footnote symbol.
* I hope some latex symbol could be translate to proper ANIS endoing in RTF (e.g. )
* That would be nice to assign label by a vector.
* e.g: the first value in the list attr(t1\_gt, "col\_labels") is assigned to c(“Treatment", rep( c("N", "Mean (SD)", "$LS Mean (95% CI)\\dagger$")))
* I hope to have multple rows in boxh\_df:
* “column\_width”: control the relative width of a column in the table. And can be understand by RTF render. (Translate to )
* “column\_boarder”: control the border for the right side of cells

attr(t1\_gt, "boxh\_df") <-   
 rbind(attr(t1\_gt, "boxh\_df"),   
 column\_width = rep(10, 8),   
 column\_boarder = rep("solid", 8)  
 )

* Location of Title

attr(t1\_gt, "title\_loc") <- "outside"

* RTF Pagesize

attr(t1\_gt, "pagesize") <- 42 # Number of lines shown in one page

* Repeat the title/footnote

attr(t1\_gt, "repeat\_title") <- TRUE  
attr(t1\_gt, "repeat\_footer") <- FALSE

* Multi-layer title.
* In RTF, it is just two rows, but not sure how to translate to a proper data structure in gt.
* SAS uses macro variables to define the header as in [https://www.lexjansen.com/pharmasug/2003/TechnicalTechniques/tt038.pdf](Merck%20SAS%20macro%20%25RTFtable%20Introduction)

In each header, a vertical line symbol ‘|’ is used to separate columns. If the user does not want to display a vertical line between the columns, ‘!’ should be used instead of ‘|’. If the user wants to merge cells, ‘^’ should be used to replace ‘|’. If all the columns in every header definition were lined up, the user will have very good idea about what the table header will look like. In the following example, there are two headers, defined as:

%let header1=| Laboratory AE | Total No |AE Event ^|;  
%let header2=| Term |Patients | n ! %|;

* as\_rtf(t1, t2, t3) should be able to append three tables into the table in rtf\_example1.rtf
* there should be an option to not display column label. (e.g. we would not need column label for gt\_t3).

## Example 2: Summary of LS mean change from baseline for HOMA data

high\_lm\_week <- high %>% split(.$week) %>%   
 map(~ lm(change ~ basval + TRT, data = .))  
  
  
t41 <- high\_lm\_week %>% map(emmeans, specs = "TRT" ) %>%   
 map(as.data.frame) %>% bind\_rows(.id = "week")  
t42 <- high\_lm\_week %>% map(emmeans, specs = "TRT" ) %>% map(pairs) %>%   
 map(as.data.frame) %>% bind\_rows(.id = "week") %>%   
 mutate( lower = estimate - 1.96 \* SE, upper = estimate + 1.96 \* SE)  
  
t4 <- merge(t41, t42, by = "week") %>%   
 mutate(  
 Week = paste0("Week ", week),   
 Trt = if\_else(TRT == 1, "Study Drug", "Placebo"),   
 N = 100,   
 mean = fmt\_est(., c("emmean", "SE.x")),  
 ci = fmt\_ci(c(., c("lower.CL", "upper.CL")) ),   
 ci\_diff = paste(fmt\_est(., "estimate"), fmt\_ci(., c("lower", "upper")))  
 ) %>%   
 select(Week:ci\_diff)

t4\_gt <- gt(t4) %>% tab\_header(title = "LDA of Change from Baseline over Time",   
 subtitle = "Missing Data Approach \n Analysis Population") %>%  
 tab\_footnote("$^\\dagger$Based on an LDA model.\nN = Number of subjects in the population.\nANCOVA = Analysis of Covariance, CI = Confidence Interval, LS = Least Squares, SD = Standard Deviation",   
 locations = cells\_column\_labels(3)) %>%  
 cols\_label(  
 Week = "Week",  
 Trt = "Treatment",   
 N = "N",   
 mean = "LS Mean (SE)",  
 ci = "95% CI",   
 ci\_diff = "Difference in LS Means vs. Placebo\n(95% CI)"  
 )

The exercise is to create a table in the format of rtf\_example2.rtf using t4

t4

## Week Trt N mean ci ci\_diff  
## 1 Week 1 Study Drug 100 -1.6 (0.46) (-2.55, -0.74) 0.0 (-1.23, 1.32)  
## 2 Week 1 Placebo 100 -1.7 (0.46) (-2.59, -0.78) 0.0 (-1.23, 1.32)  
## 3 Week 2 Study Drug 100 -3.4 (0.62) (-4.63, -2.17) 0.7 (-1.09, 2.40)  
## 4 Week 2 Placebo 100 -4.1 (0.63) (-5.30, -2.81) 0.7 (-1.09, 2.40)  
## 5 Week 4 Study Drug 100 -4.8 (0.68) (-6.15, -3.48) 1.3 (-0.54, 3.23)  
## 6 Week 4 Placebo 100 -6.2 (0.68) (-7.50, -4.83) 1.3 (-0.54, 3.23)  
## 7 Week 6 Study Drug 100 -5.7 (0.75) (-7.20, -4.24) 2.4 (0.29, 4.44)  
## 8 Week 6 Placebo 100 -8.1 (0.74) (-9.55, -6.62) 2.4 (0.29, 4.44)  
## 9 Week 8 Study Drug 100 -7.0 (0.81) (-8.58, -5.38) 1.7 (-0.49, 3.88)  
## 10 Week 8 Placebo 100 -8.7 (0.75) (-10.17, -7.18) 1.7 (-0.49, 3.88)

* A subheading should be identified and displayed in the subheading format.

attr(t4\_gt, "subheading\_var") <- "Week"

* Note sure how to arrange the first header in gt class.
* Need a flag to identify ci\_diff only display in the first row of each table subsection?

attr(t4\_gt, "subheading\_display\_var") <- "ci\_diff"

# Appendix:

## rtf\_example1.rtf:

ANCOVA of Change from Baseline at Week 8   
Missing Data Approach   
Analysis Population

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Baseline | | Week 8 | | Change from Baseline | | |
| Treatment | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) | LS Mean (95% CI)† |
| Study Durg | 61 | 16.6 (4.41) | 61 | -6.6 (5.95) | 61 | -7.0 (9.16) | -7.0 (-8.58, -5.38) |
| Placebo | 70 | 18.4 (6.34) | 61 | -9.0 (7.04) | 70 | -8.7 (8.54) | -7.0 (-8.58, -5.38) |
| Pairwise Comparison | | | | Difference in LS Means (95% CI)† | | | p-Value‡ |
| Study Durg vs. Placebo | | | | 1.7 (-0.49, 3.88) | | | 0.130 |
| Root Mean Squared Error of Change = 6.23 | | | |  | | |  |
| † Based on an ANCOVA model.  ANCOVA = Analysis of Covariance, CI = Confidence Interval, LS = Least Squares, SD = Standard Deviation | | | | | | | |
| Source: [study999: adam-adeff] | | | | | | | |

### rtf\_example2.rtf

LDA of Change from Baseline over Time   
Missing Data Approach   
Analysis Population

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Change from Baseline | | Difference in LS Means vs. Placebo |
| Treatment | N | LS Mean (SE)† | 95% CI† | (95% CI)† |
| **Week 6** | | | | |
| Study Durg | 100 | -3.02 (0.60) | (-3.85, -3.67) | -0.30 (-0.64, 0.04) |
| Placebo | 100 | -3.65 (0.60) | (-3.34, -2.97) |  |
| **Week 12** | | | | |
| Study Durg | 100 | -4.30 (0.60) | (-4.32, -4.01) | -0.06 (-0.37, 0.66) |
| Placebo | 100 | -4.65 (0.60) | (-4.34, -3.96) |  |
| **Week 18** | | | | |
| Study Durg | 100 | -4.31 (0.60) | (-4.80, -4.61) | -0.05 (-0.37, 0.67) |
| Placebo | 100 | -4.45 (0.60) | (-4.75, -4.16) |  |
| **Week 24** | | | | |
| Study Durg | 100 | -4.36 (0.60) | (-4.84, -4.67) | 0.09 (-0.62, 0.39) |
| Placebo | 100 | -4.64 (0.60) | (-4.93, -4.36) |  |
| **Week 30** | | | | |
| Study Durg | 100 | -4.64 (0.60) | (-4.93, -4.35) | -0.63 (-0.33, 0.08) |
| Placebo | 100 | -4.41 (0.60) | (-4.70, -4.13) |  |
| †Based on a LDA model  N = Number of subjects in the population.  CI = Confidence Interval, LS = Least Square, SE = Standard Error | | | | |
| Source: [study999: adam-adeff] | | | | |