CHER’s 2022 Health Equity Research Intensive

Josephine McKelvy

2022-09-01

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

###################  
# READ THIS FIRST #  
###################  
  
# Right-click the R icon and "Run as administrator" > Tools > Check for Updates. Session > Restart R and Clear Output.  
# Create new GitHub repository. Open New Project in RStudio w/repo URL.  
# Qualtrics: Data & Analysis/Export & Import/Export Data/CSV/Use Numeric values. Extract files.  
# Identify column numbers in Excel. Rename csv being read in.

# Overview (Steps)

* import all responses (as \*.csv & numeric values)
* subset only the data needed per chart
* sort CATA by session, rename columns, bind
* factor the categorical variables
* generate average item scores for dumbbell plots
* plot the dumbbell/diverging bar charts & save

###################  
# Import the data #  
###################  
  
# Import data, subset and name the variables in the dataframes, remove blanks (by language)  
raw.df <- read.csv("UNC Center for Health Equity Research - Health Equity Summer Intensive Evaluation\_September 6, 2022\_07.10.csv", stringsAsFactors = TRUE)  
raw.df <- raw.df[3:nrow(raw.df),] #cut rows 2-3 with question wording & "ImportID"  
raw.df$Q3 <- as.character(raw.df$Q3) #convert CATA to character and not factor  
  
#############  
# Libraries #  
#############  
library(data.table) #for %like%

## Warning: package 'data.table' was built under R version 4.1.3

library(dplyr) #for %>% ?

## Warning: package 'dplyr' was built under R version 4.1.3

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.1.3

#install.packages("ggalt")  
library(ggalt) #for geom\_dumbbell

## Warning: package 'ggalt' was built under R version 4.1.3

## Registered S3 methods overwritten by 'ggalt':  
## method from   
## grid.draw.absoluteGrob ggplot2  
## grobHeight.absoluteGrob ggplot2  
## grobWidth.absoluteGrob ggplot2  
## grobX.absoluteGrob ggplot2  
## grobY.absoluteGrob ggplot2

library(likert)

## Loading required package: xtable

##   
## Attaching package: 'likert'

## The following object is masked from 'package:dplyr':  
##   
## recode

#########################  
# Wrangle Dumbbell Data #  
#########################  
  
#slice Loop & Merge dataset by sessions for dumbbell  
ksa01.df <- raw.df[raw.df$Q3 %like% "1",c(21,22:25)] #will also catch 10,11,12,14  
ksa02.df <- raw.df[raw.df$Q3 %like% "2",c(21,90:93)] #will also catch 12  
ksa03.df <- raw.df[raw.df$Q3 %like% "3",c(21,158:161)]  
ksa04.df <- raw.df[raw.df$Q3 %like% "4",c(21,226:229)] #will also catch 14  
ksa05.df <- raw.df[raw.df$Q3 %like% "5",c(21,294:297)]  
ksa06.df <- raw.df[raw.df$Q3 %like% "6",c(21,362:365)]  
ksa07.df <- raw.df[raw.df$Q3 %like% "7",c(21,430:433)]  
ksa08.df <- raw.df[raw.df$Q3 %like% "8",c(21,498:501)]  
ksa09.df <- raw.df[raw.df$Q3 %like% "9",c(21,566:569)]  
ksa10.df <- raw.df[raw.df$Q3 %like% "10",c(21,634:637)]  
ksa11.df <- raw.df[raw.df$Q3 %like% "11",c(21,702:705)]  
ksaws.df <- raw.df[raw.df$Q3 %like% "14",c(21,770:773)]  
  
#rename the session no. for CATA obs & listwise delete missingness  
ksa01.df$Q3[ksa01.df$Q3 != "1"] <- "1" #this is key for the CATA  
ksa01.df[ksa01.df==""] <- NA  
ksa01.df <- na.omit(ksa01.df)  
  
ksa02.df$Q3[ksa02.df$Q3 != "2"] <- "2"  
ksa02.df[ksa02.df==""] <- NA  
ksa02.df <- na.omit(ksa02.df)  
  
ksa03.df$Q3[ksa03.df$Q3 != "3"] <- "3"  
ksa03.df[ksa03.df==""] <- NA  
ksa03.df <- na.omit(ksa03.df)  
  
ksa04.df$Q3[ksa04.df$Q3 != "4"] <- "4"  
ksa04.df[ksa04.df==""] <- NA  
ksa04.df <- na.omit(ksa04.df)  
  
ksa05.df$Q3[ksa05.df$Q3 != "5"] <- "5"  
ksa05.df[ksa05.df==""] <- NA  
ksa05.df <- na.omit(ksa05.df)  
  
ksa06.df$Q3[ksa06.df$Q3 != "6"] <- "6"  
ksa06.df[ksa06.df==""] <- NA  
ksa06.df <- na.omit(ksa06.df)  
  
ksa07.df$Q3[ksa07.df$Q3 != "7"] <- "7"  
ksa07.df[ksa07.df==""] <- NA  
ksa07.df <- na.omit(ksa07.df)  
  
ksa08.df$Q3[ksa08.df$Q3 != "8"] <- "8"  
ksa08.df[ksa08.df==""] <- NA  
ksa08.df <- na.omit(ksa08.df)  
  
ksa09.df$Q3[ksa09.df$Q3 != "9"] <- "9"  
ksa09.df[ksa09.df==""] <- NA  
ksa09.df <- na.omit(ksa09.df)  
  
ksa10.df$Q3[ksa10.df$Q3 != "10"] <- "10"  
ksa10.df[ksa10.df==""] <- NA  
ksa10.df <- na.omit(ksa10.df)  
  
ksa11.df$Q3[ksa11.df$Q3 != "11"] <- "11"  
ksa11.df[ksa11.df==""] <- NA  
ksa11.df <- na.omit(ksa11.df)  
  
ksaws.df$Q3[ksaws.df$Q3 != "12"] <- "12"  
ksaws.df[ksaws.df==""] <- NA  
ksaws.df <- na.omit(ksaws.df)  
  
# rename all the columns in these dumbbell dataframes  
rename\_ksa <- lapply(list(ksa01.df,  
 ksa02.df,  
 ksa03.df,  
 ksa04.df,  
 ksa05.df,  
 ksa06.df,  
 ksa07.df,  
 ksa08.df,  
 ksa09.df,  
 ksa10.df,  
 ksa11.df,  
 ksaws.df),function(x) {  
 names(x)<- c("session",  
 "prek","prea","postk","posta")  
 x})  
   
names(rename\_ksa) <- c("ksa01.df",  
 "ksa02.df",  
 "ksa03.df",  
 "ksa04.df",  
 "ksa05.df",  
 "ksa06.df",  
 "ksa07.df",  
 "ksa08.df",  
 "ksa09.df",  
 "ksa10.df",  
 "ksa11.df",  
 "ksaws.df")  
list2env(rename\_ksa, envir = .GlobalEnv)

## <environment: R\_GlobalEnv>

#bind the dumbbell dataframes into one  
wide.df <- rbind(ksa01.df,  
 ksa02.df,  
 ksa03.df,  
 ksa04.df,  
 ksa05.df,  
 ksa06.df,  
 ksa07.df,  
 ksa08.df,  
 ksa09.df,  
 ksa10.df,  
 ksa11.df,  
 ksaws.df)  
  
#convert dumbbell variables to factor or numeric type  
wide.df$session <- factor(wide.df$session,  
 levels = c(1,2,3,4,5,6,7,8,9,10,11,12),  
 labels = c("Concept of Trust",  
 "Structural Racism",  
 "Participatory Budgeting",  
 "Social Determinants of Health",  
 "Data is More than Numbers",  
 "Equitable Partnerships",  
 "Applying an Equity Lens",  
 "Race & Racism in Healthcare",  
 "Intervention Mapping",  
 "Culturally Responsive Evaluation",  
 "Co-creating Data Visualization",  
 "Co-Lab(orative) Learning Workshop"))  
wide.df$prek <- as.numeric(as.character(wide.df$prek))  
wide.df$prea <- as.numeric(as.character(wide.df$prea))  
wide.df$postk <- as.numeric(as.character(wide.df$postk))  
wide.df$posta <- as.numeric(as.character(wide.df$posta))  
  
#generate average scores for dumbbells: knowledge for all sessions  
wide.df %>%  
 group\_by(session) %>%  
 summarise\_at(vars("prek","postk"), mean)-> prepostk.df  
prepostk.df$id <- "knowledge" #this is key for aggregated dataset  
names(prepostk.df)<- c("session", "pre","post","id")  
  
#generate average scores for dumbbells: ability for all sessions  
wide.df %>%  
 group\_by(session) %>%  
 summarise\_at(vars("prea","posta"), mean)-> preposta.df  
preposta.df$id <- "ability"  
names(preposta.df)<- c("session", "pre","post","id")  
  
#knowledge & ability by session: https://r-graph-gallery.com/web-extended-dumbbell-plot-ggplot2.html  
prepost.df <- rbind(prepostk.df,preposta.df)  
  
#########################  
# Wrangle Lollipop Data #  
#########################  
  
overall01.df <- raw.df[raw.df$Q3 %like% "1",c(21,70)] #will also catch 10,11,12,14  
overall02.df <- raw.df[raw.df$Q3 %like% "2",c(21,134)] #will also catch 12  
overall03.df <- raw.df[raw.df$Q3 %like% "3",c(21,198)]  
overall04.df <- raw.df[raw.df$Q3 %like% "4",c(21,262)] #will also catch 14  
overall05.df <- raw.df[raw.df$Q3 %like% "5",c(21,326)]  
overall06.df <- raw.df[raw.df$Q3 %like% "6",c(21,390)]  
overall07.df <- raw.df[raw.df$Q3 %like% "7",c(21,464)]  
overall08.df <- raw.df[raw.df$Q3 %like% "8",c(21,518)]  
overall09.df <- raw.df[raw.df$Q3 %like% "9",c(21,582)]  
overall10.df <- raw.df[raw.df$Q3 %like% "10",c(21,646)]  
overall11.df <- raw.df[raw.df$Q3 %like% "11",c(21,710)]  
overallws.df <- raw.df[raw.df$Q3 %like% "14",c(21,775)]  
  
#rename the session no. for CATA obs & listwise delete missingness  
overall01.df$Q3[overall01.df$Q3 != "1"] <- "1" #this is key for the CATA  
overall01.df[overall01.df==""] <- NA  
overall01.df <- na.omit(overall01.df)  
  
overall02.df$Q3[overall02.df$Q3 != "2"] <- "2"  
overall02.df[overall02.df==""] <- NA  
overall02.df <- na.omit(overall02.df)  
  
overall03.df$Q3[overall03.df$Q3 != "3"] <- "3"  
overall03.df[overall03.df==""] <- NA  
overall03.df <- na.omit(overall03.df)  
  
overall04.df$Q3[overall04.df$Q3 != "4"] <- "4"  
overall04.df[overall04.df==""] <- NA  
overall04.df <- na.omit(overall04.df)  
  
overall05.df$Q3[overall05.df$Q3 != "5"] <- "5"  
overall05.df[overall05.df==""] <- NA  
overall05.df <- na.omit(overall05.df)  
  
overall06.df$Q3[overall06.df$Q3 != "6"] <- "6"  
overall06.df[overall06.df==""] <- NA  
overall06.df <- na.omit(overall06.df)  
  
overall07.df$Q3[overall07.df$Q3 != "7"] <- "7"  
overall07.df[overall07.df==""] <- NA  
overall07.df <- na.omit(overall07.df)  
  
overall08.df$Q3[overall08.df$Q3 != "8"] <- "8"  
overall08.df[overall08.df==""] <- NA  
overall08.df <- na.omit(overall08.df)  
  
overall09.df$Q3[overall09.df$Q3 != "9"] <- "9"  
overall09.df[overall09.df==""] <- NA  
overall09.df <- na.omit(overall09.df)  
  
overall10.df$Q3[overall10.df$Q3 != "10"] <- "10"  
overall10.df[overall10.df==""] <- NA  
overall10.df <- na.omit(overall10.df)  
  
overall11.df$Q3[overall11.df$Q3 != "11"] <- "11"  
overall11.df[overall11.df==""] <- NA  
overall11.df <- na.omit(overall11.df)  
  
overallws.df$Q3[overallws.df$Q3 != "12"] <- "12"  
overallws.df[overallws.df==""] <- NA  
overallws.df <- na.omit(overallws.df)  
  
# rename all the columns in these dumbbell dataframes  
rename\_overall <- lapply(list(overall01.df,  
 overall02.df,  
 overall03.df,  
 overall04.df,  
 overall05.df,  
 overall06.df,  
 overall07.df,  
 overall08.df,  
 overall09.df,  
 overall10.df,  
 overall11.df,  
 overallws.df),function(x) {  
 names(x)<- c("session","rating")  
 x})  
   
names(rename\_overall) <- c("overall01.df",  
 "overall02.df",  
 "overall03.df",  
 "overall04.df",  
 "overall05.df",  
 "overall06.df",  
 "overall07.df",  
 "overall08.df",  
 "overall09.df",  
 "overall10.df",  
 "overall11.df",  
 "overallws.df")  
list2env(rename\_overall, envir = .GlobalEnv)

## <environment: R\_GlobalEnv>

#bind the overall dataframes into one  
overall.df <- rbind(overall01.df,  
 overall02.df,  
 overall03.df,  
 overall04.df,  
 overall05.df,  
 overall06.df,  
 overall07.df,  
 overall08.df,  
 overall09.df,  
 overall10.df,  
 overall11.df,  
 overallws.df)  
  
#convert overall variables to factor or numeric type  
overall.df$session <- factor(overall.df$session,  
 levels = c(1,2,3,4,5,6,7,8,9,10,11,12),  
 labels = c("Concept of Trust",  
 "Structural Racism",  
 "Participatory Budgeting",  
 "Social Determinants of Health",  
 "Data is More than Numbers",  
 "Equitable Partnerships",  
 "Applying an Equity Lens",  
 "Race & Racism in Healthcare",  
 "Intervention Mapping",  
 "Culturally Responsive Evaluation",  
 "Co-creating Data Visualization",  
 "Co-Lab(orative) Learning Workshop"))  
overall.df$rating <- as.numeric(as.character(overall.df$rating))  
  
#generate average scores for overall Q  
overall.df %>%  
 group\_by(session) %>%  
 summarise\_at(vars("rating"), mean)-> overall.df  
  
#######################  
# Wrangle Likert Data #  
#######################  
  
#subset Likert scale items by session  
likert01.df <- raw.df[raw.df$Q3 %like% "1",c(21,71:74)] #will also catch 10,11,12,14  
likert02.df <- raw.df[raw.df$Q3 %like% "2",c(21,135:138)] #will also catch 12  
likert03.df <- raw.df[raw.df$Q3 %like% "3",c(21,199:202)]  
likert04.df <- raw.df[raw.df$Q3 %like% "4",c(21,263:266)] #will also catch 14  
likert05.df <- raw.df[raw.df$Q3 %like% "5",c(21,327:330)]  
likert06.df <- raw.df[raw.df$Q3 %like% "6",c(21,391:394)]  
likert07.df <- raw.df[raw.df$Q3 %like% "7",c(21,455:458)]  
likert08.df <- raw.df[raw.df$Q3 %like% "8",c(21,519:522)]  
likert09.df <- raw.df[raw.df$Q3 %like% "9",c(21,583:586)]  
likert10.df <- raw.df[raw.df$Q3 %like% "10",c(21,647:650)]  
likert11.df <- raw.df[raw.df$Q3 %like% "11",c(21,711:714)]  
likertws.df <- raw.df[raw.df$Q3 %like% "12",c(21,774:783)]  
  
#rename the Likert session no. for CATA obs & listwise delete missingness  
likert01.df$Q3[likert01.df$Q3 != "1"] <- "1" #this is key for CATA  
likert01.df[likert01.df==""] <- NA  
likert01.df <- na.omit(likert01.df)  
  
likert02.df$Q3[likert02.df$Q3 != "2"] <- "2"  
likert02.df[likert02.df==""] <- NA  
likert02.df <- na.omit(likert02.df)  
  
likert03.df$Q3[likert03.df$Q3 != "3"] <- "3"  
likert03.df[likert03.df==""] <- NA  
likert03.df <- na.omit(likert03.df)  
  
likert04.df$Q3[likert04.df$Q3 != "4"] <- "4"  
likert04.df[likert04.df==""] <- NA  
likert04.df <- na.omit(likert04.df)  
  
likert05.df$Q3[likert05.df$Q3 != "5"] <- "5"  
likert05.df[likert05.df==""] <- NA  
likert05.df <- na.omit(likert05.df)  
  
likert06.df$Q3[likert06.df$Q3 != "6"] <- "6"  
likert06.df[likert06.df==""] <- NA  
likert06.df <- na.omit(likert06.df)  
  
likert07.df$Q3[likert07.df$Q3 != "7"] <- "7"  
likert07.df[likert07.df==""] <- NA  
likert07.df <- na.omit(likert07.df)  
  
likert08.df$Q3[likert08.df$Q3 != "8"] <- "8"  
likert08.df[likert08.df==""] <- NA  
likert08.df <- na.omit(likert08.df)  
  
likert09.df$Q3[likert09.df$Q3 != "9"] <- "9"  
likert09.df[likert09.df==""] <- NA  
likert09.df <- na.omit(likert09.df)  
  
likert10.df$Q3[likert10.df$Q3 != "10"] <- "10"  
likert10.df[likert10.df==""] <- NA  
likert10.df <- na.omit(likert10.df)  
  
likert11.df$Q3[likert11.df$Q3 != "11"] <- "11"  
likert11.df[likert11.df==""] <- NA  
likert11.df <- na.omit(likert11.df)  
  
likertws.df$Q3[likertws.df$Q3 != "12"] <- "12"  
likertws.df[likertws.df==""] <- NA  
likertws.df <- na.omit(likertws.df)  
  
# rename all the columns in these Likert dataframes  
rename\_likert <- lapply(list(likert01.df,  
 likert02.df,  
 likert03.df,  
 likert04.df,  
 likert05.df,  
 likert06.df,  
 likert07.df,  
 likert08.df,  
 likert09.df,  
 likert10.df,  
 likert11.df),function(x) {  
 names(x)<- c("session",  
 "effective presenter",  
 "knowledgeable presenter",  
 "insightful information",  
 "relevant techniques")  
 x})  
   
names(rename\_likert) <- c("likert01.df",  
 "likert02.df",  
 "likert03.df",  
 "likert04.df",  
 "likert05.df",  
 "likert06.df",  
 "likert07.df",  
 "likert08.df",  
 "likert09.df",  
 "likert10.df",  
 "likert11.df")  
list2env(rename\_likert, envir = .GlobalEnv)

## <environment: R\_GlobalEnv>

names(likertws.df) <- c("session",  
 "effective guidance",  
 "welcoming space",  
 "structured activities",  
 "insightful formation",  
 "relevant techniques")  
   
#bind the dataframes into one  
likert.df <- rbind(likert01.df,  
 likert02.df,  
 likert03.df,  
 likert04.df,  
 likert05.df,  
 likert06.df,  
 likert07.df,  
 likert08.df,  
 likert09.df,  
 likert10.df,  
 likert11.df)  
  
#convert variables to factor or numeric type  
likert.df$session <- factor(likert.df$session,  
 levels = c(1,2,3,4,5,6,7,8,9,10,11),  
 labels = c("Concept of Trust",  
 "Structural Racism",  
 "Participatory Budgeting",  
 "Social Determinants of Health",  
 "Data is More than Numbers",  
 "Equitable Partnerships",  
 "Applying an Equity Lens",  
 "Race & Racism in Healthcare",  
 "Intervention Mapping",  
 "Culturally Responsive Evaluation",  
 "Co-creating Data Visualization"))  
likertws.df$session <- factor(likertws.df$session,  
 levels = c(12),  
 labels = c("Co(Laborative) Learning Workshop"))  
  
likertlabels <- c("Strongly Disagree",  
 "Somewhat Disagree",  
 "Neither",  
 "Somewhat Agree",  
 "Strongly Agree")  
  
likert.df$`effective presenter` <- factor(likert.df$`effective presenter`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likert.df$`knowledgeable presenter` <- factor(likert.df$`knowledgeable presenter`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likert.df$`insightful information` <- factor(likert.df$`insightful information`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likert.df$`relevant techniques` <- factor(likert.df$`relevant techniques`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
  
likertws.df$`effective guidance` <- factor(likertws.df$`effective guidance`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likertws.df$`welcoming space` <- factor(likertws.df$`welcoming space`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likertws.df$`structured activities` <- factor(likertws.df$`structured activities`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likertws.df$`insightful information` <- factor(likertws.df$`insightful information`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
likertws.df$`relevant techniques` <- factor(likertws.df$`relevant techniques`,  
 levels = c(1,2,3,4,5),  
 labels = likertlabels)  
#######################  
# Wrangle Likely Data #  
#######################  
  
#subset likelihood items by session  
likely01.df <- raw.df[raw.df$Q3 %like% "1",c(21,80:81)] #will also catch 10,11,12,14  
likely02.df <- raw.df[raw.df$Q3 %like% "2",c(21,144:145)] #will also catch 12  
likely03.df <- raw.df[raw.df$Q3 %like% "3",c(21,208:209)]  
likely04.df <- raw.df[raw.df$Q3 %like% "4",c(21,272:273)] #will also catch 14  
likely05.df <- raw.df[raw.df$Q3 %like% "5",c(21,336:337)]  
likely06.df <- raw.df[raw.df$Q3 %like% "6",c(21,400:401)]  
likely07.df <- raw.df[raw.df$Q3 %like% "7",c(21,464:465)]  
likely08.df <- raw.df[raw.df$Q3 %like% "8",c(21,528:529)]  
likely09.df <- raw.df[raw.df$Q3 %like% "9",c(21,592:593)]  
likely10.df <- raw.df[raw.df$Q3 %like% "10",c(21,656:657)]  
likely11.df <- raw.df[raw.df$Q3 %like% "11",c(21,720:721)]  
likelyws.df <- raw.df[raw.df$Q3 %like% "12",c(21,784:785)]  
  
#rename the likely session no. for CATA obs & listwise delete missingness  
likely01.df$Q3[likely01.df$Q3 != "1"] <- "1" #this is key for CATA  
likely01.df[likely01.df==""] <- NA  
likely01.df <- na.omit(likely01.df)  
  
likely02.df$Q3[likely02.df$Q3 != "2"] <- "2"  
likely02.df[likely02.df==""] <- NA  
likely02.df <- na.omit(likely02.df)  
  
likely03.df$Q3[likely03.df$Q3 != "3"] <- "3"  
likely03.df[likely03.df==""] <- NA  
likely03.df <- na.omit(likely03.df)  
  
likely04.df$Q3[likely04.df$Q3 != "4"] <- "4"  
likely04.df[likely04.df==""] <- NA  
likely04.df <- na.omit(likely04.df)  
  
likely05.df$Q3[likely05.df$Q3 != "5"] <- "5"  
likely05.df[likely05.df==""] <- NA  
likely05.df <- na.omit(likely05.df)  
  
likely06.df$Q3[likely06.df$Q3 != "6"] <- "6"  
likely06.df[likely06.df==""] <- NA  
likely06.df <- na.omit(likely06.df)  
  
likely07.df$Q3[likely07.df$Q3 != "7"] <- "7"  
likely07.df[likely07.df==""] <- NA  
likely07.df <- na.omit(likely07.df)  
  
likely08.df$Q3[likely08.df$Q3 != "8"] <- "8"  
likely08.df[likely08.df==""] <- NA  
likely08.df <- na.omit(likely08.df)  
  
likely09.df$Q3[likely09.df$Q3 != "9"] <- "9"  
likely09.df[likely09.df==""] <- NA  
likely09.df <- na.omit(likely09.df)  
  
likely10.df$Q3[likely10.df$Q3 != "10"] <- "10"  
likely10.df[likely10.df==""] <- NA  
likely10.df <- na.omit(likely10.df)  
  
likely11.df$Q3[likely11.df$Q3 != "11"] <- "11"  
likely11.df[likely11.df==""] <- NA  
likely11.df <- na.omit(likely11.df)  
  
likelyws.df$Q3[likelyws.df$Q3 != "12"] <- "12"  
likelyws.df[likelyws.df==""] <- NA  
likelyws.df <- na.omit(likelyws.df)  
  
# rename all the columns in these Likert dataframes  
rename\_likely <- lapply(list(likely01.df,  
 likely02.df,  
 likely03.df,  
 likely04.df,  
 likely05.df,  
 likely06.df,  
 likely07.df,  
 likely08.df,  
 likely09.df,  
 likely10.df,  
 likely11.df,  
 likelyws.df),function(x) {  
 names(x)<- c("session",  
 "use what you learned",  
 "recommend this session to others")  
 x})  
   
names(rename\_likely) <- c("likely01.df",  
 "likely02.df",  
 "likely03.df",  
 "likely04.df",  
 "likely05.df",  
 "likely06.df",  
 "likely07.df",  
 "likely08.df",  
 "likely09.df",  
 "likely10.df",  
 "likely11.df",  
 "likelyws.df")  
list2env(rename\_likely, envir = .GlobalEnv)

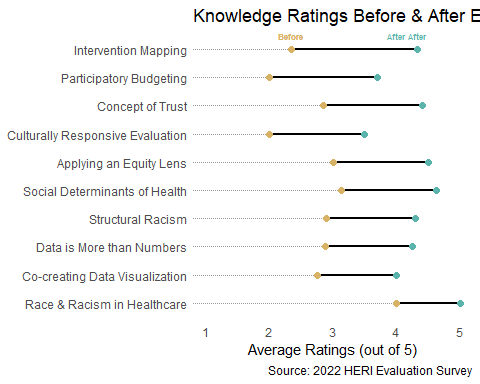
## <environment: R\_GlobalEnv>

#bind the dataframes into one  
likely.df <- rbind(likely01.df,  
 likely02.df,  
 likely03.df,  
 likely04.df,  
 likely05.df,  
 likely06.df,  
 likely07.df,  
 likely08.df,  
 likely09.df,  
 likely10.df,  
 likely11.df,  
 likelyws.df)  
  
#convert variables to factor or numeric type  
likely.df$session <- factor(likely.df$session,  
 levels = c(1,2,3,4,5,6,7,8,9,10,11,12),  
 labels = c("Concept of Trust",  
 "Structural Racism",  
 "Participatory Budgeting",  
 "Social Determinants of Health",  
 "Data is More than Numbers",  
 "Equitable Partnerships",  
 "Applying an Equity Lens",  
 "Race & Racism in Healthcare",  
 "Intervention Mapping",  
 "Culturally Responsive Evaluation",  
 "Co-creating Data Visualization",  
 "Co(Laborative) Learning Workshop"))  
likelylabels <- c("Not at all Likely",  
 "Somewhat Unlikely",  
 "Neutral",  
 "Somewhat Likely",  
 "Extremely Likely")  
likely.df$`use what you learned` <- factor(likely.df$`use what you learned`,  
 levels = c(1,2,3,4,5),  
 labels = likelylabels)  
likely.df$`recommend this session to others` <- factor(likely.df$`recommend this session to others`,  
 levels = c(1,2,3,4,5),  
 labels = likelylabels)  
  
###########  
# Comments#  
###########  
feedback01.df <- raw.df[raw.df$Q3 %like% "1",c(19,21,83,85)] #will also catch 10,11,12,14  
feedback02.df <- raw.df[raw.df$Q3 %like% "2",c(19,21,147,149)] #will also catch 12  
feedback03.df <- raw.df[raw.df$Q3 %like% "3",c(19,21,211,213)]  
feedback04.df <- raw.df[raw.df$Q3 %like% "4",c(19,21,275,277)] #will also catch 14  
feedback05.df <- raw.df[raw.df$Q3 %like% "5",c(19,21,339,341)]  
feedback06.df <- raw.df[raw.df$Q3 %like% "6",c(19,21,402,403,405)]  
feedback07.df <- raw.df[raw.df$Q3 %like% "7",c(19,21,467,469)]  
feedback08.df <- raw.df[raw.df$Q3 %like% "8",c(19,21,531,533)]  
feedback09.df <- raw.df[raw.df$Q3 %like% "9",c(19,21,595,597)]  
feedback10.df <- raw.df[raw.df$Q3 %like% "10",c(19,21,659,661)]  
feedback11.df <- raw.df[raw.df$Q3 %like% "11",c(19,21,723,725)]  
feedbackws.df <- raw.df[raw.df$Q3 %like% "12",c(19,21,788,789)]  
  
#rename the likelihood session no. for CATA obs & listwise delete missingness  
feedback01.df$Q3[feedback01.df$Q3 != "1"] <- "1" #this is key for CATA  
feedback01.df[feedback01.df==""] <- NA  
  
feedback02.df$Q3[feedback02.df$Q3 != "2"] <- "2"  
feedback02.df[feedback02.df==""] <- NA  
  
feedback03.df$Q3[feedback03.df$Q3 != "3"] <- "3"  
feedback03.df[feedback03.df==""] <- NA  
  
feedback04.df$Q3[feedback04.df$Q3 != "4"] <- "4"  
feedback04.df[feedback04.df==""] <- NA  
  
feedback05.df$Q3[feedback05.df$Q3 != "5"] <- "5"  
feedback05.df[feedback05.df==""] <- NA  
  
feedback06.df$Q3[feedback06.df$Q3 != "6"] <- "6"  
feedback06.df[feedback06.df==""] <- NA  
  
feedback07.df$Q3[feedback07.df$Q3 != "7"] <- "7"  
feedback07.df[feedback07.df==""] <- NA  
  
feedback08.df$Q3[feedback08.df$Q3 != "8"] <- "8"  
feedback08.df[feedback08.df==""] <- NA  
  
feedback09.df$Q3[feedback09.df$Q3 != "9"] <- "9"  
feedback09.df[feedback09.df==""] <- NA  
  
feedback10.df$Q3[feedback10.df$Q3 != "10"] <- "10"  
feedback10.df[feedback10.df==""] <- NA  
  
feedback11.df$Q3[feedback11.df$Q3 != "11"] <- "11"  
feedback11.df[feedback11.df==""] <- NA  
  
feedbackws.df$Q3[feedbackws.df$Q3 != "12"] <- "12"  
feedbackws.df[feedbackws.df==""] <- NA  
  
# rename all the columns in these Likert dataframes  
rename\_feedback <- lapply(list(feedback01.df,  
 feedback02.df,  
 feedback03.df,  
 feedback04.df,  
 feedback05.df,  
# feedback06.df,  
 feedback07.df,  
 feedback08.df,  
 feedback09.df,  
 feedback10.df,  
 feedback11.df,  
 feedbackws.df),function(x) {  
 names(x)<- c("role","session",  
 "session improvements",  
 "session relevance")  
 x})  
   
names(rename\_feedback) <- c("feedback01.df",  
 "feedback02.df",  
 "feedback03.df",  
 "feedback04.df",  
 "feedback05.df",  
# "feedback06.df",  
 "feedback07.df",  
 "feedback08.df",  
 "feedback09.df",  
 "feedback10.df",  
 "feedback11.df",  
 "feedbackws.df")  
list2env(rename\_feedback, envir = .GlobalEnv)

## <environment: R\_GlobalEnv>

names(feedback06.df) <- c("role","session",  
 "use in work",  
 "session improvements",  
 "session relevance")  
   
#bind the dataframes into one  
feedback.df <- rbind(feedback01.df,  
 feedback02.df,  
 feedback03.df,  
 feedback04.df,  
 feedback05.df,  
# feedback06.df,  
 feedback07.df,  
 feedback08.df,  
 feedback09.df,  
 feedback10.df,  
 feedback11.df,  
 feedbackws.df)  
  
#convert variables to factor or numeric type  
feedback.df$session <- factor(feedback.df$session,  
 levels = c(1,2,3,4,5,6,7,8,9,10,11,12),  
 labels = c("Concept of Trust",  
 "Structural Racism",  
 "Participatory Budgeting",  
 "Social Determinants of Health",  
 "Data is More than Numbers",  
 "Equitable Partnerships",  
 "Applying an Equity Lens",  
 "Race & Racism in Healthcare",  
 "Intervention Mapping",  
 "Culturally Responsive Evaluation",  
 "Co-creating Data Visualization",  
 "Co(Laborative) Learning Workshop"))  
feedback06.df$session <- factor(feedback06.df$session,  
 levels = c(6),  
 labels = c("Equitable Partnerships"))  
  
#####################  
# Dumbbell Settings #  
#####################  
gold <- "#d8b365"  
teal <- "#5ab4ab"  
  
xsubtitle <- "Average Ratings (out of 5)"  
caption <- "Source: 2022 HERI Evaluation Survey"  
  
###########################  
# K for all sessions #  
###########################  
prepost.df [prepost.df$id=="knowledge",] %>% #runs code if these cells have a certain value  
 mutate(difference = (post-pre)) %>% #generates a difference score  
 ggplot() +   
 aes(x = pre,  
 xend = post,  
# y = session) + #default: sorted by session name  
 y = reorder(session, difference), group = session) + #sorted by greatest difference  
 geom\_dumbbell(  
 size\_x = 2, colour\_x = gold, #orange pre dot  
 size\_xend = 2, colour\_xend = teal, #green post dot  
 size = 1, #black line  
 dot\_guide = TRUE,  
 dot\_guide\_size = 0.15,  
 dot\_guide\_colour = "grey60") +  
 scale\_x\_continuous(breaks = seq (1,5,1),limits = c(1,5)) + #show x axis from 1 to 5 in intervals of 1, even if no data  
 geom\_text(data=filter(prepost.df, session=="Intervention Mapping"),  
 aes(x=pre, y=session, label="Before"),  
 color = gold, size=2, vjust=-1.5, fontface="bold",family="Lato")+  
 geom\_text(data=filter(prepost.df, session=="Intervention Mapping"),  
 aes(x=post, y=session, label = "After"),  
 color = teal, size=2, vjust=-1.5, fontface="bold",family="Lato")+  
 labs(title = "Knowledge Ratings Before & After Each Session", y = NULL, x = xsubtitle,   
 caption = caption) +  
 theme(panel.border = element\_blank(), #no gap  
 panel.background = element\_rect(fill = 'transparent',color = NA), #transparent bkgd  
 panel.grid.minor = element\_blank(),   
 panel.grid.major.y = element\_blank(),  
 panel.grid.major.x = element\_line(),   
 axis.ticks = element\_blank()) #no ticks

## Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x$label), x$x, x$y, :  
## font family not found in Windows font database

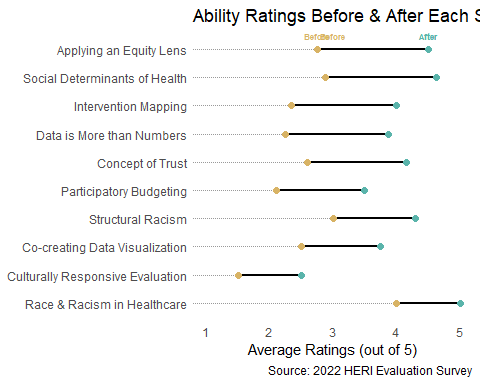


ggsave("All Sessions - Knowledge.png", height = 3.5, width = 7)

## Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x$label), x$x, x$y, :  
## font family not found in Windows font database

###########################  
# A for all sessions #  
###########################  
prepost.df [prepost.df$id == "ability",] %>%   
 mutate(difference = (post-pre)) %>%  
 ggplot() +   
 aes(x = pre,  
 xend = post,  
# y = session) + #default: sorted by session name  
 y = reorder(session, difference), group = session) +   
 geom\_dumbbell(  
 size\_x = 2, colour\_x = gold,  
 size\_xend = 2, colour\_xend = teal,  
 size = 1,  
 dot\_guide = TRUE,  
 dot\_guide\_size = 0.15,  
 dot\_guide\_colour = "grey60") +  
 scale\_x\_continuous(breaks = seq (1,5,1),limits = c(1,5)) +   
 geom\_text(data=filter(prepost.df, session=="Applying an Equity Lens"),  
 aes(x=pre, y=session, label="Before"),  
 color = gold, size=2, vjust=-1.5, fontface="bold",family="Lato") +  
 geom\_text(data=filter(prepost.df, session=="Applying an Equity Lens"),  
 aes(x=post, y=session, label = "After"),  
 color = teal, size=2, vjust=-1.5, fontface="bold",family="Lato") +  
 labs(title = "Ability Ratings Before & After Each Session", y = NULL, x = xsubtitle,   
 caption = caption) +  
 theme(panel.border = element\_blank(),   
 panel.background = element\_rect(fill = 'transparent',color = NA),   
 panel.grid.minor = element\_blank(),   
 panel.grid.major.y = element\_blank(),  
 panel.grid.major.x = element\_line(),   
 axis.ticks = element\_blank())

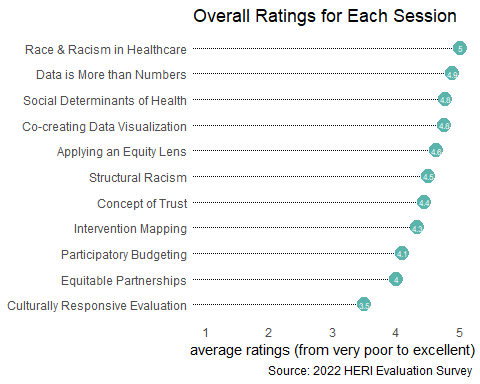
## Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x$label), x$x, x$y, :  
## font family not found in Windows font database



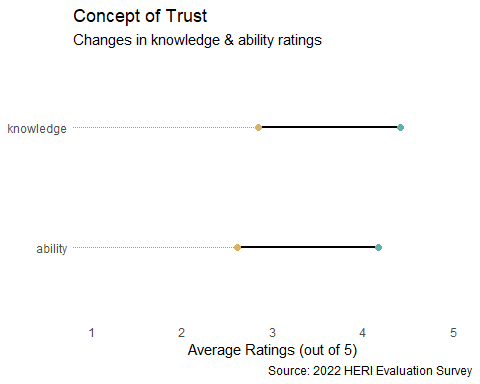
ggsave("All Sessions - Ability.png", height = 3.5, width = 7)

## Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x$label), x$x, x$y, :  
## font family not found in Windows font database

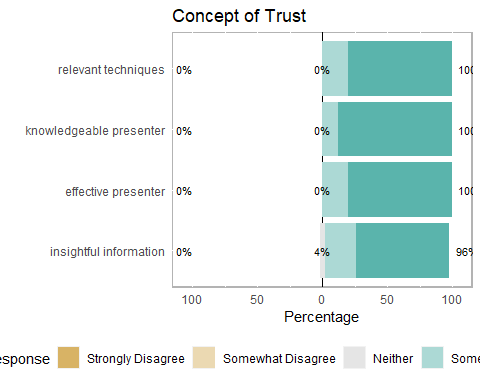
############################  
# Overall for all sessions #  
############################  
 overall.df %>%   
 ggplot() +   
 aes(x = rating,  
 xend = rating,  
 y = reorder(session, rating), group = session) +  
 geom\_dumbbell(  
 size\_x = 4.5, colour\_x = gold,  
 size\_xend = 4.5, colour\_xend = teal,  
 size = 1,  
 dot\_guide = TRUE,  
 dot\_guide\_size = 0.2,  
 dot\_guide\_colour = "black") +  
 scale\_x\_continuous(breaks = seq (1,5,1),limits = c(1,5)) +   
 geom\_text(aes(label = round(rating, digits = 1)), color = "white", size = 2) + #insert rating in lollipop, rounding to 1 decimal place  
 labs(title = "Overall Ratings for Each Session", y = NULL, x = "average ratings (from very poor to excellent)",  
 caption = caption) +  
 theme(panel.border = element\_blank(),   
 panel.background = element\_rect(fill = 'transparent',color = NA),   
 panel.grid.minor = element\_blank(),   
 panel.grid.major.y = element\_blank(),  
 panel.grid.major.x = element\_line(),   
 axis.ticks = element\_blank())



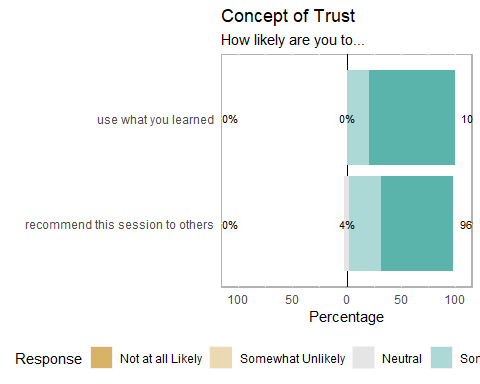
ggsave("All Sessions - Overall Rating.png", height = 3.5, width = 7)  
  
#########################################  
# Session Settings: Dumbbell AND Likert #  
#########################################  
  
# CHANGE thissession, GGSAVEs; RERUN for each session  
  
thissession <- "Concept of Trust"  
#thissession <- "Structural Racism"  
#thissession <- "Participatory Budgeting"  
#thissession <- "Social Determinants of Health"  
#thissession <- "Data is More than Numbers"  
#thissession <- "Equitable Partnerships"  
#thissession <- "Applying an Equity Lens"  
#thissession <- "Race & Racism in Healthcare"  
#thissession <- "Intervention Mapping"  
#thissession <- "Culturally Responsive Evaluation"  
#thissession <- "Co-creating Data Visualization"  
#thissession <- "Co-Lab(orative) Learning Workshop"  
  
######################  
# K and A by session #  
######################  
prepost.df [prepost.df$session==thissession,] %>%   
 ggplot() +   
 aes(x = pre,  
 xend = post,  
 y = id) +   
 geom\_dumbbell(  
 size\_x = 2, colour\_x = gold,  
 size\_xend = 2, colour\_xend = teal,  
 size = 1,  
 dot\_guide = TRUE,  
 dot\_guide\_size = 0.15,  
 dot\_guide\_colour = "grey60") +  
 scale\_x\_continuous(breaks = seq (1,5,1),limits = c(1,5)) +   
 #geom\_text(data=filter(prepost.df, id=="knowledge"),  
 # aes(x=pre, y=session, label="Before"),  
 # color = gold, size=2, vjust=-1.5, fontface="bold",family="Lato") +  
 #geom\_text(data=filter(prepost.df, id=="knowledge"),  
 # aes(x=post, y=session, label = "After"),  
 # color = teal, size=2, vjust=-1.5, fontface="bold",family="Lato") +  
 labs(title = thissession, y = NULL, x = xsubtitle,   
 subtitle = "Changes in knowledge & ability ratings",   
 caption = caption) +  
 theme(  
 panel.grid.minor = element\_blank(),   
 panel.grid.major.y = element\_blank(),  
 panel.grid.major.x = element\_line(),   
 axis.ticks = element\_blank(),   
 panel.border = element\_blank(),   
 panel.background = element\_rect(fill = 'transparent',color = NA)   
 )



ggsave("Session 01 - K and A.png", height = 1.5, width = 5)  
  
#####################  
# Likert by session #  
#####################  
  
quant <- likert(likert.df[likert.df$session==thissession,2:5])  
plot(quant, positive.order = TRUE) +   
 labs(title = thissession)



ggsave("Session 01 - Quant.png", width = 9, height = 2.5)  
  
#########################  
# Likelihood by session #  
#########################  
likely <- likert(likely.df[likely.df$session==thissession,2:3])  
plot(likely, positive.order = TRUE) +   
 labs(title = thissession, subtitle = "How likely are you to...")



ggsave("Session 01 - Likely.png", width = 9, height = 2)