Pattersonetal2016_FMD_Report

2016-09-26

Read in data

Read in the data file and load packages.

```
#required libraries
pacman::p_load(tidyr,
               dplyr,
               ggplot2,
               stringr,
               knitr)

data = read.csv(file="raw_data/data_final2.csv", header=T)
```

Data Processing

First, we need to reshape the data into long format.

```
datalong <- data %>%
  select(ResponseID, starts_with("DR_"), starts_with("IB_")) %>%
  pivot_longer(cols = !ResponseID,
               names_to = c(".value", "variable"),
               names_pattern = "(IB|DR)_(.*)") %>%
  filter(variable %in% c("title", "text", "text2", "text3")==FALSE)
```

Reverse the impact on business (IB) scores, so that higher values = more critical (time until hurts business)

```
datalong$IB_rev <-  5 - datalong$IB
```

Add group for movement type

```
datalong$mvgroup <-  factor(gsub("_[[:digit:]]*", "", as.character(datalong$variable)))
```

Create aggregate data - taking the mean, sd, se by movement.

```
dataagg <- datalong %>%
  group_by(variable) %>%
  summarize(DR_mean = mean(DR, na.rm=T),
            IB_rev_mean = mean(IB_rev, na.rm=T),
            DR_sd = sd(DR, na.rm = T),
            IB_rev_sd = sd(IB_rev, na.rm = T),
            DR_se = DR_sd/sqrt(length(which(!is.na(DR)))),
            IB_rev_se = IB_rev_sd/sqrt(length(which(!is.na(IB)))))
```

Add group for movement type

```
dataagg$mvgroup = factor(gsub("_[[:digit:]]*", "", as.character(dataagg$variable)))
```

Create reversed IB scores in the original data

```
for (i in paste("IB_", dataagg$variable, sep="")) {
    data$a = 5 - data[,which(names(data)==i)]
    names(data)[which(names(data)=="a")] = paste(i, "rev", sep="_")
}
```

Add a column with the number that corresponds to each movement question

```
dataagg$MovmtNumber = as.numeric(gsub(".*_", "", dataagg$variable))

dataagg <- dataagg %>%
  arrange(mvgroup, MovmtNumber) %>%
  mutate(MovmtNumberContin = 1:30)
```

Creating Scores

Based on Reviewer feedback, we will present the average reversed IB score on the graph along with the consensus score for Risk of Disease Spread Presentation of the RDS variable will be changed to indicate the % majority saying the top two categories (some or high risk = above 50) or the bottom two categories (no or low risk = below 50).

The table will include bucketed reverse IB scores based on placement of the means on the rated scale: 3-4 (48 hours to a week); 2.5-3 (7 days to 14 days); 2 - 2.5 (14 days to 21 days); 1-2 (21 days to 60 days).

First, create new variables for each DR rating that codes 1 as "high" (3 or 4) and 0 as "low" risk (1 or 2)

```
for (i in paste("DR_", dataagg$variable, sep="")) {
    data$a = ifelse(data[,which(names(data)==i)] > 2, 1, 0)
    names(data)[which(names(data)=="a")] = paste(i, "phigh", sep="_")
}
```

Then create function to calculate percent of 1's for each variable

```
percentof1s = function(x) {round((sum(x, na.rm=T)/sum(complete.cases(x)))*100, 2)}
```

Use that function to calculate that for each new variable

```
dataagg_p <- data %>%
  select(contains("phigh")) %>%
  summarize(across(.cols = everything(), .fns = percentof1s)) %>%
  t()
```

Add these values to the aggregated dataset

```
dataagg$Percent_DR_High <-  dataagg_p
dataagg$checknames = names(dataagg_p)[1:30]
```

Create a new variable for IB means that codes which above category they fall into:

```
dataagg <- dataagg %>%
  mutate(IB_rev_cat = case_when(IB_rev_mean > 3.01 ~ "2 to 7 days",
                                between(IB_rev_mean, 2.51, 3.01) ~ "7 to 14 days",
                                between(IB_rev_mean, 2.01, 2.51) ~ "14 to 21 days",
                                between(IB_rev_mean, 1, 2.01) ~ "21 to 60 days"))
```

Plot for Figure 1

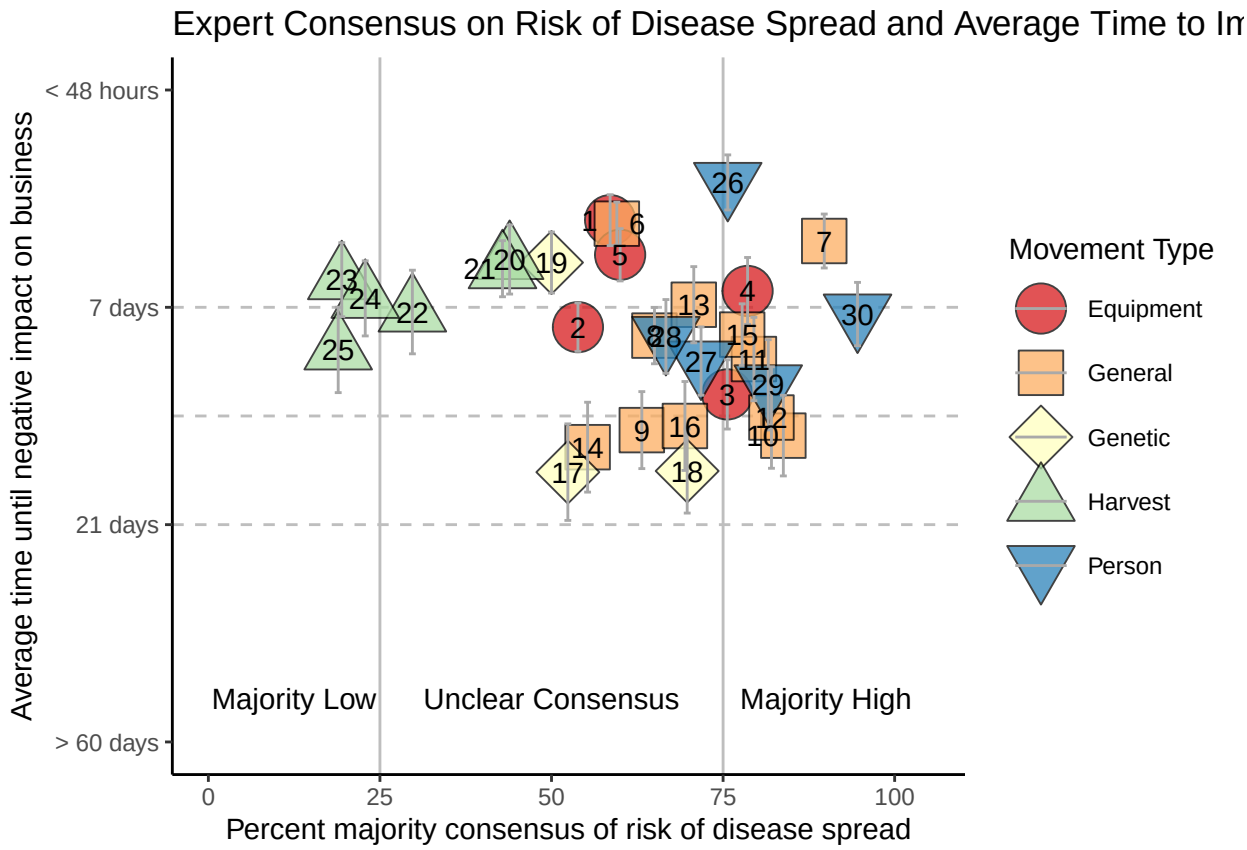
Create the plot for figure 1

Set colors and error bars for plot. Adjust text labels so they don't overlap

```
colors = c("#d7191c","#fdae61", "#ffffbf", "#abdda4", "#2b83ba")

dataagg$eb_min <- dataagg$IB_rev_mean - dataagg$IB_rev_se
dataagg$eb_max <- dataagg$IB_rev_mean + dataagg$IB_rev_se

dataagg$Percent_DR_high_nudge = dataagg$Percent_DR_High
dataagg$Percent_DR_high_nudge[1] = dataagg$Percent_DR_high_nudge[1]-3
dataagg$Percent_DR_high_nudge[21] = dataagg$Percent_DR_high_nudge[21]-3.3
dataagg$Percent_DR_high_nudge[6] = dataagg$Percent_DR_high_nudge[6]+3
dataagg$Percent_DR_high_nudge[10] = dataagg$Percent_DR_high_nudge[10]-3
```

Create the static figure

```
ggplot(dataagg, aes(x=Percent_DR_High, y=IB_rev_mean)) +
  coord_cartesian(ylim=c(1,4), xlim=c(0,105)) +
  labs(title="Expert Consensus on Risk of Disease Spread and Average Time to Impact on Business",
       x="Percent majority consensus of risk of disease spread",
       y="Average time until negative impact on business") +
  geom_vline(xintercept=c(25,75), color="gray") +
  geom_hline(yintercept=c(2, 2.5, 3), lty=2, color="gray") +
  geom_point(aes(x=Percent_DR_High,
                 y=IB_rev_mean,
                 shape=mvgroup,
                 fill=mvgroup), alpha=.75, size=8.5) +
  geom_errorbar(aes(ymin=eb_min,
                    ymax=eb_max,
                    color=mvgroup), width=1) +

  scale_shape_manual(values = c(21:25),
                     name="Movement Type",
                     labels=c("Equipment", "General", "Genetic", "Harvest", "Person"))  +
  scale_fill_manual(name="Movement Type",
                    labels=c("Equipment", "General", "Genetic", "Harvest", "Person"),
                    values=colors)  +
  scale_color_manual(name="Movement Type",
                     labels=c("Equipment", "General", "Genetic", "Harvest", "Person"),
                     values=rep("darkgray", 5)) +
  scale_y_continuous(breaks = c(1, 2, 3, 4),
                     labels=rev(c("< 48 hours",  "7 days",  "21 days", "> 60 days"))) +
  geom_text(aes(x=Percent_DR_high_nudge,
                y=IB_rev_mean,
                label=MovmtNumberContin), size=4) +
  annotate("text", x=12.5, y=1.2, label="Majority Low", size=4) +
  annotate("text", x=50, y=1.2, label="Unclear Consensus", size=4) +
  annotate("text", x=90, y=1.2, label="Majority High", size=4)  +
  theme_classic()
```



Information for Table 1

Make sure all the categories are in formats that will look nice in a table

```
dataagg <- dataagg %>%
  mutate(Category = case_when(mvgroup == "equip" ~ "Equipment",
                              mvgroup == "general" ~ "General",
                              mvgroup == "genetic" ~ "Genetic",
                              mvgroup == "harvest" ~ "Harvest",
                              mvgroup == "person" ~ "Person"),
         NumCat = paste(Category, MovmtNumberContin, sep=" "))
```

Create one column for bin (high, even, low), one for % majority in that bin, then mirror scores to match figure - percent majority low or high

```
dataagg <- dataagg %>%
  mutate(Placement = case_when(between(Percent_DR_High, 25, 75) ~ "Unclear Consensus",
                               Percent_DR_High <= 25 ~ "Low",
                               Percent_DR_High >= 75 ~ "High"))

scores <- dataagg %>%
  select(NumCat, Percent_DR_High, Placement, IB_rev_cat) %>%
  rename(Movement = NumCat,
         "Percent High RDS" = Percent_DR_High,
         "Binned Risk of Disease Spread" = Placement,
         "Time to Negative Business Impact" = IB_rev_cat)
```

Make sure the factors are in the right order

```
scores <- scores %>%
  mutate("Binned Risk of Disease Spread" = factor(`Binned Risk of Disease Spread`, levels=c("Low", "Uncl
         "Time to Negative Business Impact" = factor(`Time to Negative Business Impact`, levels = c("14
```

Then create table

```
scores %>%
  arrange(desc(`Binned Risk of Disease Spread`), desc(`Time to Negative Business Impact`), desc(`Percen
  kable()
```

Movement	Percent High RDS	Binned Risk of Disease Spread	Time to Negative Business Impact
General 7	89.74	High	2 to 7 days
Equipment 4	78.57	High	2 to 7 days
Person 26	75.68	High	2 to 7 days
Person 30	94.59	High	7 to 14 days
Person 29	81.58	High	7 to 14 days
General 11	79.49	High	7 to 14 days
General 15	77.78	High	7 to 14 days
Equipment 3	75.61	High	7 to 14 days
General 10	83.78	High	14 to 21 days
General 12	82.05	High	14 to 21 days
General 13	70.73	Unclear Consensus	2 to 7 days
Equipment 5	60.00	Unclear Consensus	2 to 7 days
General 6	59.52	Unclear Consensus	2 to 7 days
Equipment 1	58.54	Unclear Consensus	2 to 7 days
Genetic 19	50.00	Unclear Consensus	2 to 7 days
Harvest 20	43.90	Unclear Consensus	2 to 7 days
Harvest 21	42.86	Unclear Consensus	2 to 7 days
Person 27	71.79	Unclear Consensus	7 to 14 days
Person 28	66.67	Unclear Consensus	7 to 14 days
General 8	65.00	Unclear Consensus	7 to 14 days
Equipment 2	53.85	Unclear Consensus	7 to 14 days
Harvest 22	29.73	Unclear Consensus	7 to 14 days
Genetic 18	69.77	Unclear Consensus	14 to 21 days
General 16	69.44	Unclear Consensus	14 to 21 days
General 9	63.16	Unclear Consensus	14 to 21 days
General 14	55.26	Unclear Consensus	14 to 21 days
Genetic 17	52.38	Unclear Consensus	14 to 21 days
Harvest 24	22.86	Low	2 to 7 days
Harvest 23	19.44	Low	2 to 7 days
Harvest 25	18.92	Low	7 to 14 days

Session information

```
sessionInfo()
```

```
## R version 4.0.5 (2021-03-31)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Catalina 10.15.7
```

```
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] knitr_1.33    stringr_1.4.0 ggplot2_3.3.5 dplyr_1.0.7   tidyr_1.1.3
##
## loaded via a namespace (and not attached):
##  [1] magrittr_2.0.1    munsell_0.5.0     tidyselect_1.1.1  colorspace_2.0-2
##  [5] R6_2.5.0          rlang_0.4.11      fansi_0.5.0       highr_0.9
##  [9] tools_4.0.5       grid_4.0.5        gtable_0.3.0      xfun_0.24
## [13] pacman_0.5.1      utf8_1.2.2        DBI_1.1.1         withr_2.4.2
## [17] htmltools_0.5.1.1 ellipsis_0.3.2    assertthat_0.2.1  yaml_2.2.1
## [21] digest_0.6.27     tibble_3.1.3      lifecycle_1.0.0   crayon_1.4.1
## [25] farver_2.1.0      purrr_0.3.4       vctrs_0.3.8       glue_1.4.2
## [29] evaluate_0.14     rmarkdown_2.9     labeling_0.4.2    stringi_1.7.3
## [33] compiler_4.0.5    pillar_1.6.1      scales_1.1.1      generics_0.1.0
## [37] pkgconfig_2.0.3
```