Assessing capacity of health centres in Siem Reap, Cambodia Rmarkdown Class Project

Load packages

Load dataset

```
Capacity_survey_dataset <- read_dta("~/Documents/nBox/WHO Cambodia/Task 2 (Survey & Interviews)/Step B/
```

View dataset

```
view(Capacity_survey_dataset)
```

Data cleaning

```
#Add value labels to the `OD name` variable
Capacity_survey_dataset$ODname <-
  factor(Capacity_survey_dataset$ODname,
  levels= c(1,2,3,4),
  labels= c("Krolanh", "Siem Reap", "SotrNikum", "Angkorchum"))</pre>
```

Data analysis

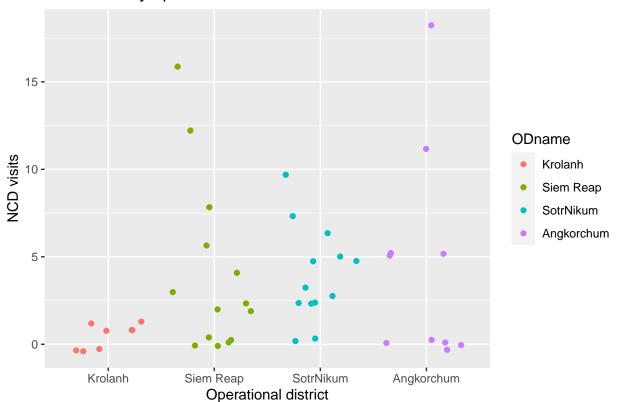
```
#What percent of HCs belong to each of the four operational districts (ODs)?
Capacity_survey_dataset %>%
group_by(ODname) %>%
summarize(n = n()) %>%
mutate(freq = n/sum(n))
```

```
#What is the average distance (in kms) from a HC to a referral hospital
Capacity_survey_dataset["Q18c"][Capacity_survey_dataset["Q18c"] == "0"] <- "0"</pre>
Capacity_survey_dataset$Q18c <- as.numeric(Capacity_survey_dataset$Q18c)</pre>
mean(Capacity_survey_dataset$Q18c)
## [1] 24.54444
#What is the average setting type of each HC?
Capacity_survey_dataset %>%
 group_by(Settingtype) %>%
 summarize(n = n()) \%
 mutate(freq = n/sum(n))
## # A tibble: 3 x 3
    Settingtype
                       n freq
                   <int> <dbl>
##
    <dbl+1b1>
## 1 1 [Rural]
                     38 0.844
## 2 2 [Urban]
                      2 0.0444
## 3 3 [Semi-urban]
                      5 0.111
#What proportion of HCs had been trained in the WHO PEN?
Capacity_survey_dataset %>%
 group_by(WHOtraining) %>%
 summarize(n = n()) \%
 mutate(freq = n/sum(n))
## # A tibble: 2 x 3
    WHOtraining n freq
    <dbl+lbl> <int> <dbl>
## 1 0 [No]
                  30 0.667
## 2 1 [Yes]
                   15 0.333
#Calculate average number of NCD-related visits per day by OD and setting type
Capacity_survey_dataset$NCDvisits <- Capacity_survey_dataset$Q16a + Capacity_survey_dataset$Q16b + Capa
Capacity_survey_dataset %>%
 group_by(ODname, Settingtype) %>%
 summarize(mean_NCD_visits = mean(NCDvisits))
## 'summarise()' has grouped output by 'ODname'. You can override using the
## '.groups' argument.
## # A tibble: 9 x 3
## # Groups: ODname [4]
##
    ODname Settingtype
                              mean_NCD_visits
##
    <fct>
             <dbl+lbl>
                                        <dbl>
## 1 Krolanh 1 [Rural]
                                        0.714
## 2 Krolanh 3 [Semi-urban]
                                        0
## 3 Siem Reap 1 [Rural]
                                       2.64
## 4 Siem Reap 2 [Urban]
                                       12
## 5 Siem Reap 3 [Semi-urban]
                                        2
```

```
## 6 SotrNikum 1 [Rural] 4.09
## 7 SotrNikum 3 [Semi-urban] 2.5
## 8 Angkorchum 1 [Rural] 4.89
## 9 Angkorchum 3 [Semi-urban] 0
```

```
#Create a scatter plot of average NCD visits by district
ggplot(data=Capacity_survey_dataset, aes(x=ODname, y=NCDvisits)) + geom_jitter(aes(colour=ODname)) + 1
```

NCD visits by operational district



```
#Create a boxplot of average NCD visits by district
ggplot(data=Capacity_survey_dataset, aes(x=NCDvisits, y=ODname)) + geom_boxplot(aes(colour=ODname)) +
```



