

final_scratch

Ahmed Eldeeb

2022-11-30

Spinal Injury Data

```
analysis <- read.csv("./CSpine/CSV datasets/analysisvariables.csv")
```

```
# remove holdout sites
```

```
analysis_all <- analysis
```

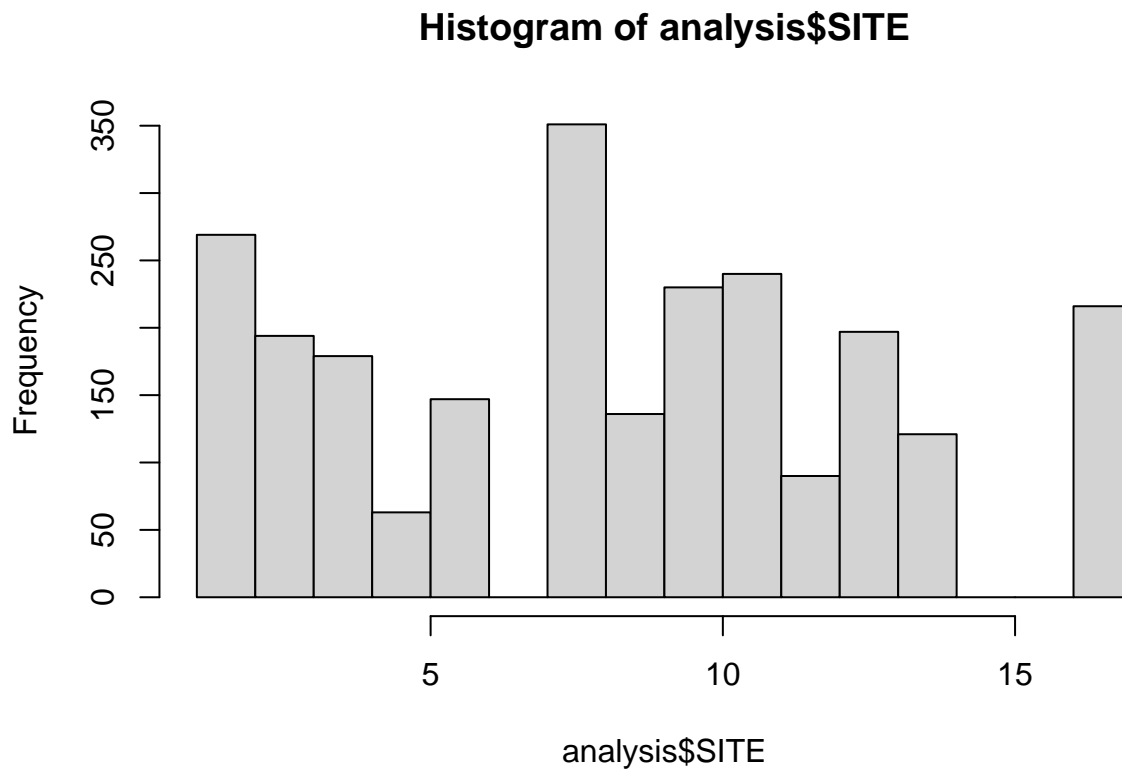
```
summary(as.factor(analysis$SITE))
```

```
##      1      2      3      4      5      6      7      8      9     10     11     12     13     14     15     16     17
## 148 121 194 179   63 147 267 351 136 230 240   90 197 121 285 329 216
```

```
analysis <- analysis[analysis$SITE != 15 & analysis$SITE != 16 & analysis$SITE != 7,]
summary(as.factor(analysis$SITE))
```

```
##      1      2      3      4      5      6      8      9     10     11     12     13     14     17
## 148 121 194 179   63 147 351 136 230 240   90 197 121 216
```

```
hist(analysis$SITE)
```



```
summary(as.factor(analysis$ControlType))
```

```
## case  ems  moi  ran
## 399  521  734  779
```

```
#cross.tabulation <- table(Q65, Q66)
#p.table <- prop.table(cross.tabulation, 1)
#ptdf<- as.data.frame(p.table)
#ptdf_f <- ptdf |> filter(Q65!="no response", Q66!="no response", Q65!="other", Q66!="other", Q66!="no
#ggplot(ptdf_f, aes(x=Q65, y=Freq,
#                   color=Q66, group=Q66)) +
#  geom_line(size=1)# +
#  #scale_y_continuous(label=percent)

#ggplot(ptdf_f, aes(x=Q65, y=Freq, fill=Q66)) + geom_col()

#dv1 <- dummyVars(~ Q65, data = qdf_f)
#dummyCats1 <- predict(dv1, qdf_f)

library(reshape2)

grp1 <- analysis[,c(13,14,15,16)]
grp1[is.na(grp1)] <- 0
grp2 <- analysis[,c(18,19,20,21,22,23)]
```

```

grp2[is.na(grp2)] <- 0
# correlation matrix for the two sets of 1-hot encoded variables
cormat <- cor(grp1, grp2)
melted_cormat <- melt(cormat)

# Heatmap of the correlation matrix
melted_cormat |> filter(!is.na(value)) |>
  ggplot(aes(x=Var1, y=Var2, fill=value)) +
  geom_tile() +
  geom_text(aes(Var1, Var2, label = round(value,2)),
            color = "white", size = 4) +
  labs(x="", y="")

```



```

grp1 <- analysis[,c(5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23)]
grp1[is.na(grp1)] <- 0
grp2 <- analysis[,c(5,6)]
grp2$EMS <- analysis$ControlType == "ems"
grp2$RAN <- analysis$ControlType == "ran"
grp2$MOI <- analysis$ControlType == "moi"
grp2$PositiveCase <- analysis$ControlType == "case"
# correlation matrix for the two sets of 1-hot encoded variables
cormat <- cor(grp1, grp2)
melted_cormat <- melt(cormat)

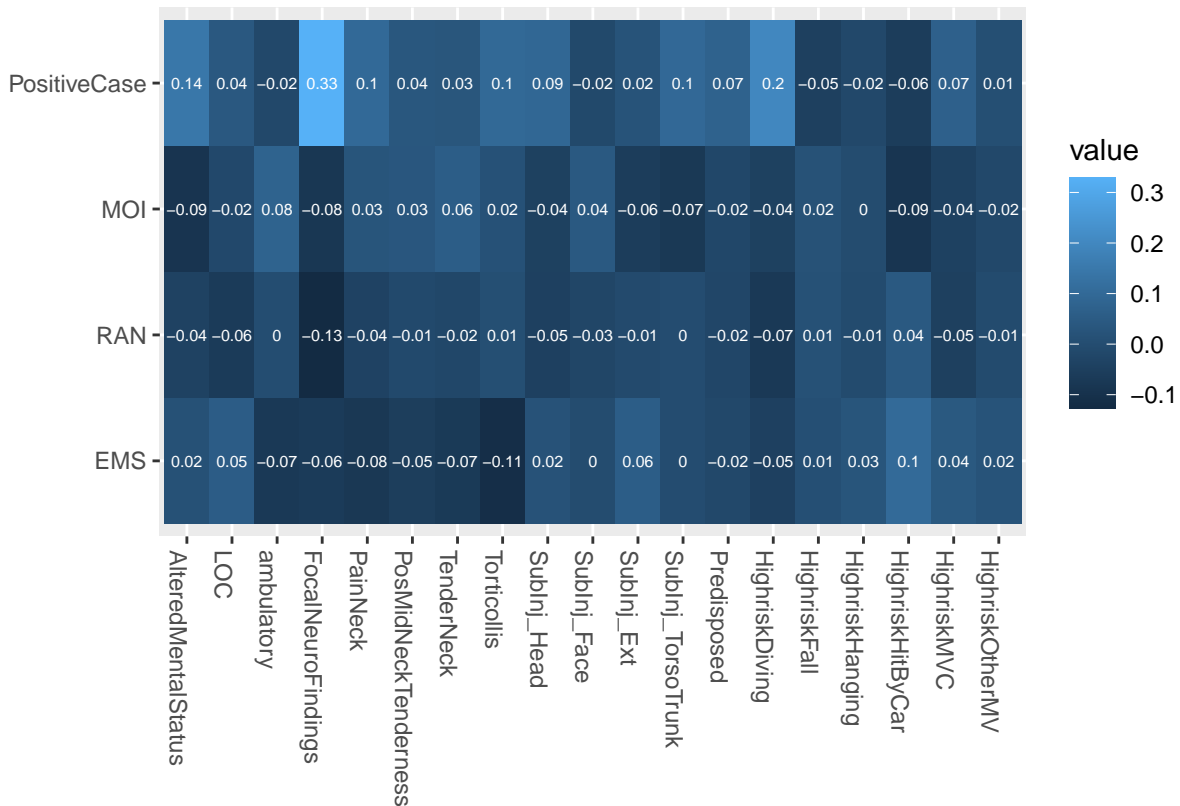
# Heatmap of the correlation matrix

```

```

melted_cormat |> filter(!is.na(value)) |>
  ggplot(aes(x=Var1, y=Var2, fill=value)) +
  geom_tile() +
  geom_text(aes(Var1, Var2, label = round(value,2)),
    color = "white", size = 2) +
  labs(x="", y="") +
  theme(axis.text.x=element_text(angle = -90, hjust = 0))

```



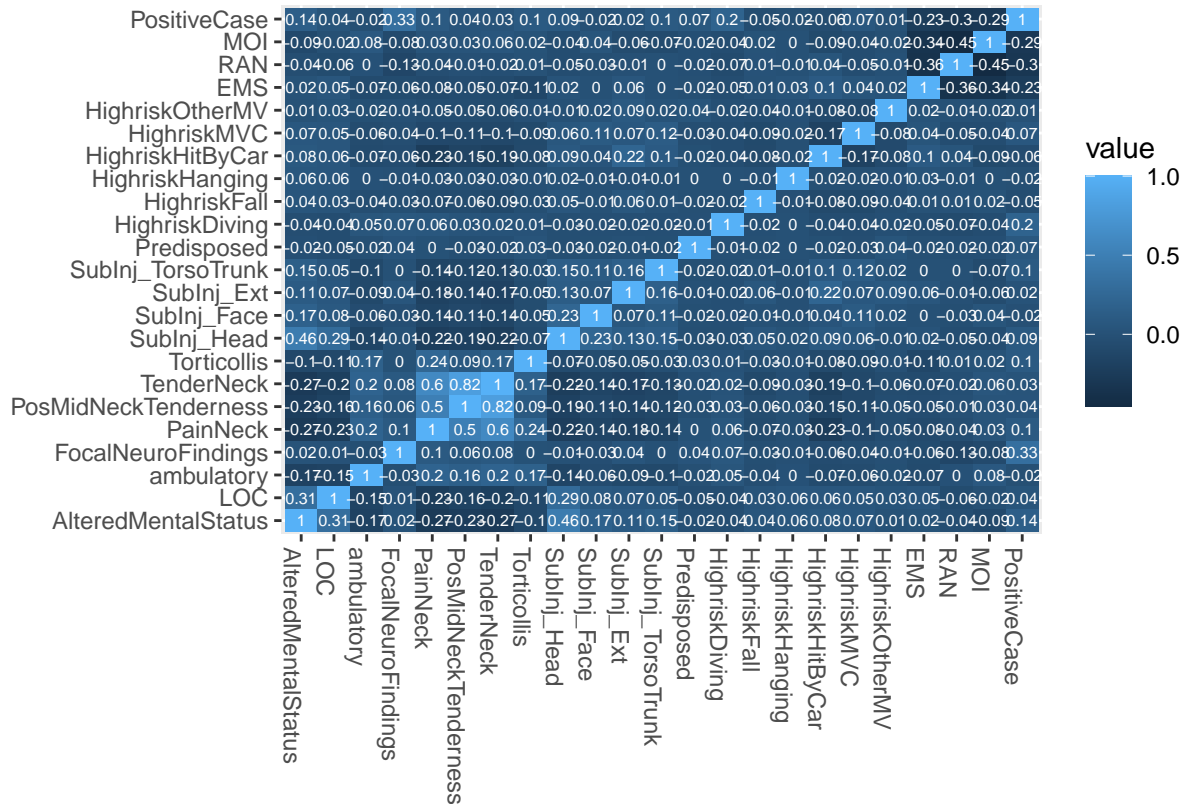
```

grp1 <- analysis[,c(5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23)]
grp1[is.na(grp1)] <- 0
grp1$EMS <- analysis$ControlType == "ems"
grp1$RAN <- analysis$ControlType == "ran"
grp1$MOI <- analysis$ControlType == "moi"
grp1$PositiveCase <- analysis$ControlType == "case"
# correlation matrix for the two sets of 1-hot encoded variables
cormat <- cor(grp1, grp1)
melted_cormat <- melt(cormat)

# Heatmap of the correlation matrix
melted_cormat |> filter(!is.na(value)) |>
  ggplot(aes(x=Var1, y=Var2, fill=value)) +
  geom_tile() +
  geom_text(aes(Var1, Var2, label = round(value,2)),
    color = "white", size = 2) +
  labs(x="", y="") +

```

```
theme(axis.text.x=element_text(angle = -90, hjust = 0))
```

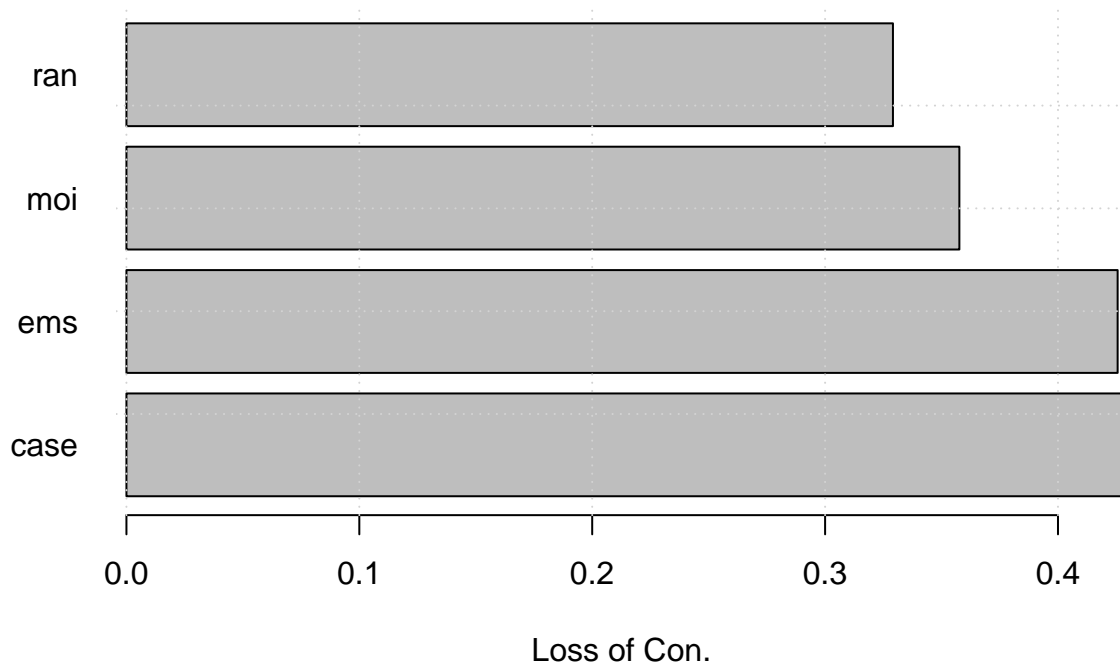


```
analysis$ControlType <- as.factor(analysis$ControlType)
```

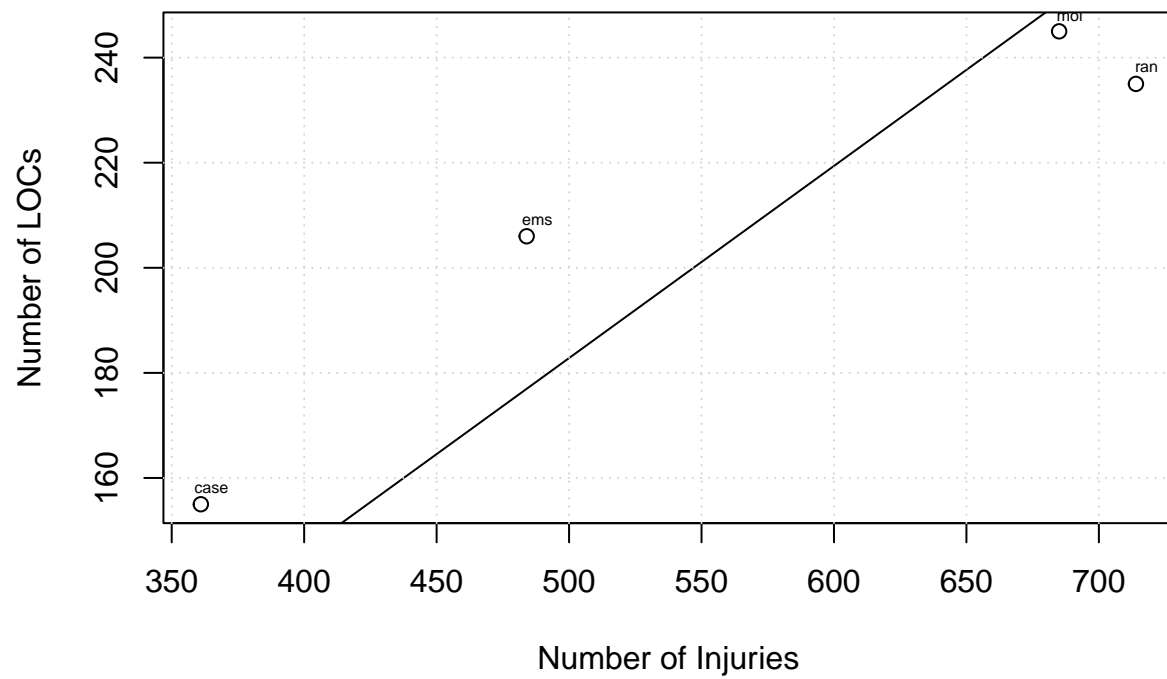
```
CtrlType <- analysis |> select(ControlType, LOC) |> filter(!is.na(LOC)) |>
  group_by(ControlType) |> summarize(tot=n(), LOC=sum(LOC))
```

```
barplot(CtrlType$LOC/CtrlType$tot, names.arg=levels(analysis$ControlType), main="Loss of consciousness by Control Type",
  grid())
```

Loss of concioussness rate for different control groups



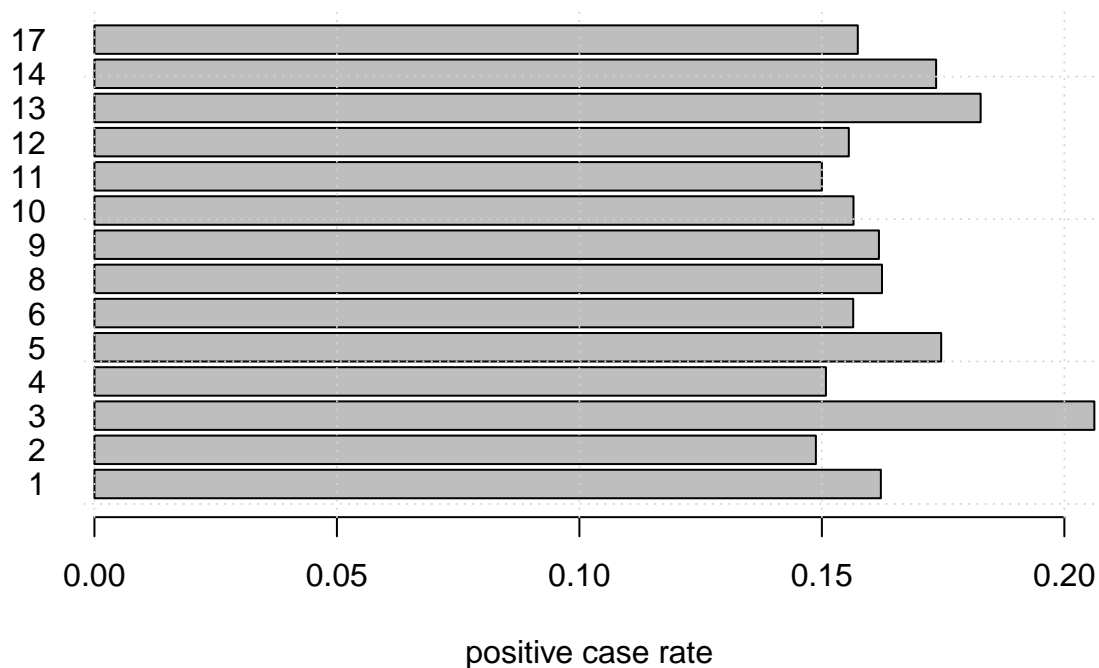
```
plot(CtrlType$tot,CtrlType$LOC, main="", xlab="Number of Injuries", ylab="Number of LOCs")
text(CtrlType$tot+4, CtrlType$LOC+3, labels=CtrlType$ControlType, cex=0.5)
abline(lm(CtrlType$LOC~CtrlType$tot-1))
grid()
```



```
SiteType <- analysis |> select(SITE, ControlType) |> filter(!is.na(ControlType)) |>
  group_by(SITE) |> summarize(tot=n(), TypeCase=sum(ControlType=="case"), TypeRan=sum(ControlType=="ran"))

barplot(SiteType$TypeCase/SiteType$tot, names.arg=SiteType$SITE, main="Case Rate for different Sites",
  grid())
```

Case Rate for different Sites



```
plot(SiteType$tot,SiteType$TypeCase, main="", xlab="Number of Patients", ylab="Number of Positive Cases",
text(SiteType$tot+2, SiteType$TypeCase+1, labels=SiteType$SITE, cex=0.5)
abline(lm(SiteType$TypeCase~SiteType$tot -1))
grid()
```