Anzahl Resistenzen und einfache plots

11.03.2022

# Bibliotheken laden, Hilfsfunktion

#library(xlsx) # für Excel files  
#library(stringr) # String-verarbeitung  
library(ggplot2) # moderne plots  
  
debug <- T # debug printout  
debug <- F # kein debug printout  
Log <- function(string) {  
 if(debug){print(string)}   
}

# Resistenzen.Rmd erzeugte Resistenzen.csv, dieses einlesen

Und evtl. ansehen

Resistenzen <- read.csv("Resistenzen.csv")  
  
# csv schreiben fügt vorne Index-Spalte an; diese entfernen :  
Resistenzen[,1] <- NULL   
  
View(Resistenzen)

# Resistenzen pro Betrieb

Resistenzen pro Betrieb in neuer Tabelle “NResistenzen” zählen, Multirestenz dokumentieren und als NResistenzen.csv ausschreiben

ResRow <- nrow(Resistenzen) # Zeilen Resistenzen : 4 pro Betrieb  
NResRow <- ResRow/4 # Zeilen NResistenzen : 1 pro Betrieb  
NAntib <- 15 # wir untersuchen 15 Antibiotika (wird von Resistenzen.Rmd so aus 2 Excel files eingelesen)  
  
NResistenzen <- Resistenzen[0,] # header wie"Resistenzen"  
for(line in 1:NResRow){ # 1 bis 60, aber 30 fehlt  
 i <- (line - 1)\*4 + 1  
 NResistenzen[line,] <- Resistenzen[(line - 1)\*4 + 1,] # WM.group etc. kopieren  
 NResistenzen[line,2:(NAntib+1)] <- 0 # aber Antibiotika auf 0 setzen : hier später Resistenzen zählen  
}  
for(col in 2:(NAntib+1)){  
 NResistenzen[,col] <- as.numeric(NResistenzen[,col]) # muss immer noch in type double konvertieren  
}  
#View(NResistenzen)   
  
# für jedes Antibiotikum Resistenzen über die 4 Proben zählen, also mögliche Werte 0-4 :  
for(i in 1:ResRow){ # Liniennummer (Betriebe in 4er Gruppen) für dataframe Resistenzen  
 Log(paste("i=",i))  
   
 line <- floor((i-1)/4)+1 # Liniennummer für dataframe NResistenzen  
   
 for(j in 2:(NAntib+1)){ # Spaltennummer: Antibiotikum  
  
 if(substr(Resistenzen[i,j],1,1)==">"){ # wenn Resistenz  
 Log(paste(" NResistenzen[",line,j,"]=",NResistenzen[line,j],typeof(NResistenzen[line,j]) ))  
 NResistenzen[line,j] <- NResistenzen[line,j] + 1 # gef. Resistenz zählen  
} } }   
  
NResistenzen$NRes <- rep(0,NResRow) # neue Spalte, zählt für jeden Betrieb Resistenzen über Antibiotika; erstmal 0  
NResistenzen$MultiR <- rep(F,NResRow) # neue Spalte, dokumentiert für jeden Betrieb Multiresistenz; erstmal False  
for(line in 1:NResRow){ # 1 bis 60, aber 30 fehlt  
 for(col in 2:(NAntib+1)){  
 if(NResistenzen[line,col] > 0){  
 NResistenzen[line,"NRes"] <- NResistenzen[line,"NRes"]+1 # Resistenz zählen  
 }  
 }  
 if(NResistenzen[line,"NRes"] >= 3){ # Multiresistenz heisst mind. 3 Resistenzen  
 NResistenzen[line,"MultiR"] <- T  
 }  
}  
View(NResistenzen)   
write.csv(NResistenzen,"NResistenzen.csv")

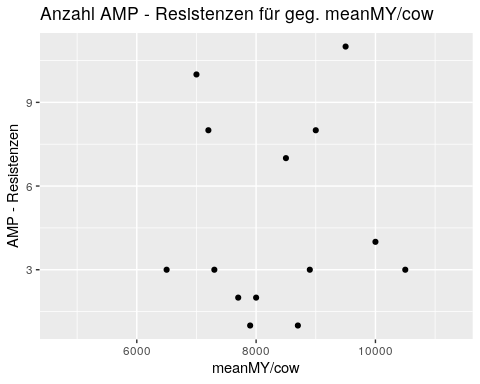
# Numerische und Ordinale Unabhängige Variablen

graphisch2 <- function(gruppe, join, antibiotikum) {  
 group <- Resistenzen[,gruppe ]   
 antib <- Resistenzen[,antibiotikum ]  
  
 X <- c()  
 Y <- c()  
 for(i in 1:ResRow){ # Liniennummer für dataframe Resistenzen  
 x <- as.numeric(group[i]) # ,na.rm=TRUE) hilft nicht weil's "NA" ist, nicht NA  
 if(substr(antib[i],1,1)==">"){ # wenn   
   
 pos <- match(x,X)   
 if(is.na(pos)){  
 X <- c(X,x) # faster: pre-allocate+assign,  
 Y <- c(Y,1) # in this way vector copied in every iteration  
 } else {  
 Y[pos] <- Y[pos] + 1  
 }  
 }  
 }   
  
 df <- data.frame(X,Y)  
 ylab <- paste(antibiotikum,"- Resistenzen")  
   
 #print(paste("gruppe=",gruppe))  
 if( gruppe == "WM.group" ){xlab <- "Wastemilk-Gruppe"}  
 if( gruppe == "OLS.group"){xlab <- "Other LiveStock-Gruppe"}   
 if( gruppe == "IAC.group"){xlab <- "Ill Animals in Calving box-Gruppe"}  
 if( gruppe == "HSC.group"){xlab <- "Husbandry System Calves-Gruppe"}   
   
 if( gruppe == "MY.group" ){xlab <- "meanMY/cow"}  
 if( gruppe == "SCC.group"){xlab <- "mean SCC/11mo"}   
 if( gruppe == "CBC.group"){xlab <- "calvingbox\_clean"}   
 if( gruppe == "DIA.group"){xlab <- "IN\_diarrhea<30d"}   
 #print(paste("xlab=",xlab))  
  
 min <- min(as.numeric(Resistenzen[,gruppe]), na.rm=T)  
 max <- max(as.numeric(Resistenzen[,gruppe]), na.rm=T)  
 #print(paste("gruppe=",gruppe,": max=",max))  
 puffer <- (max - min)/20   
 min <- min - puffer # links und rechts 5% freier Platz  
 max <- max + puffer  
   
 print( ggplot(df, aes(X, Y)) +   
 geom\_point() +  
 xlim(min,max) +  
 xlab(xlab) + ylab(ylab) +   
 ggtitle(paste("Anzahl", ylab, join,xlab))   
 )  
}

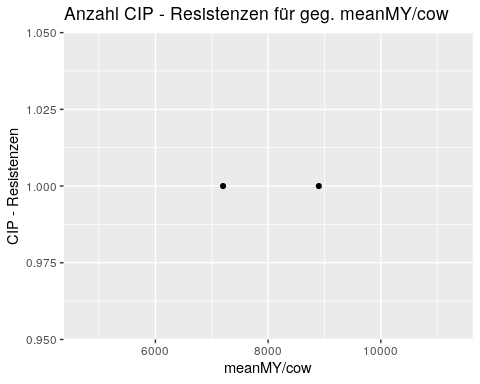
# Plot Anzahl der Resistenzen für verschiedene Antibiotika

* MERO, AMI, TGC, TAZ COL, keine Resistenzen
* FOT , AZI nur eine (die AZI-CBC und AZI-IAC plots sind korrekterweise leer: Diese Resistenz hat NA für CBC und IAC)

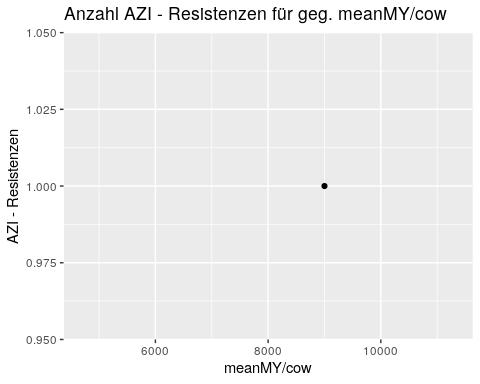
# NA warnings interessieren nicht  
  
groups <- c() # nur Verteilungen plotten  
groups <- c("MY.group","SCC.group","CBC.group","DIA.group")  
for( group in groups) {   
  
 for( antib in c("AMP","CIP","AZI","GEN","FOT","CHL","NAL","TET","TMP","SMX") ){   
   
 graphisch2(group,"für geg.",antib)   
 print("")  
 }   
 print("--------------------------------------------------------")  
}



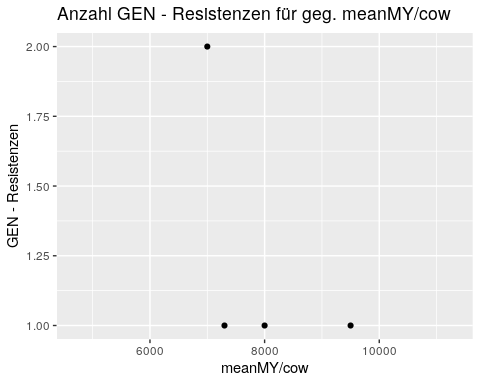
## [1] ""



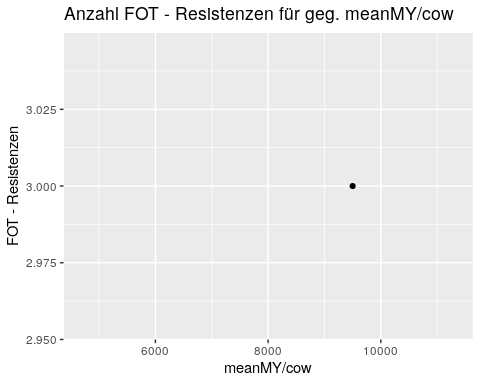
## [1] ""



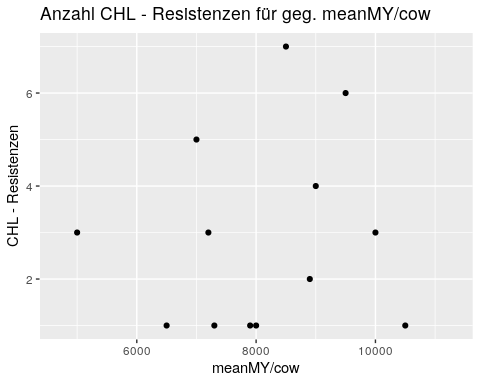
## [1] ""



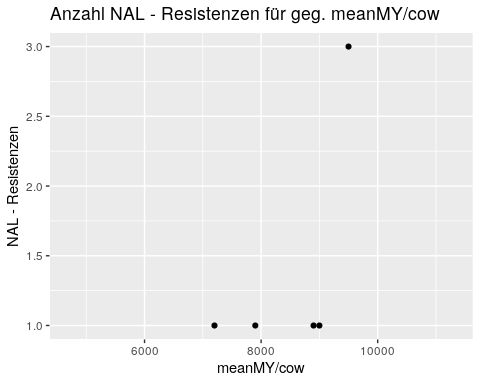
## [1] ""



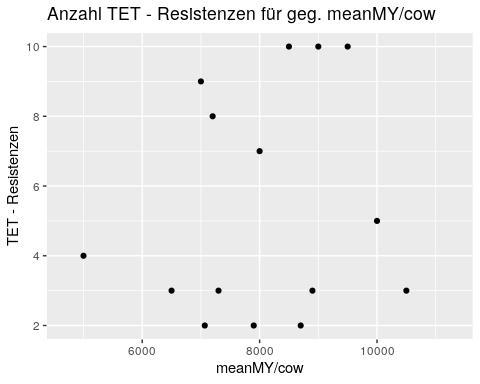
## [1] ""



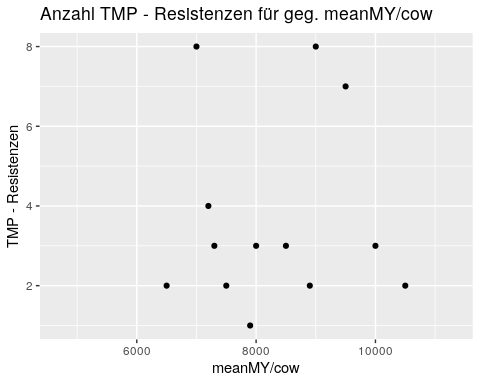
## [1] ""



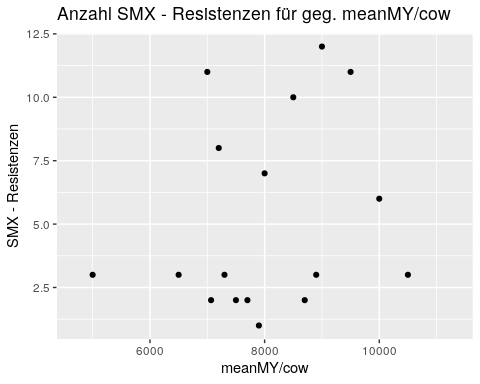
## [1] ""



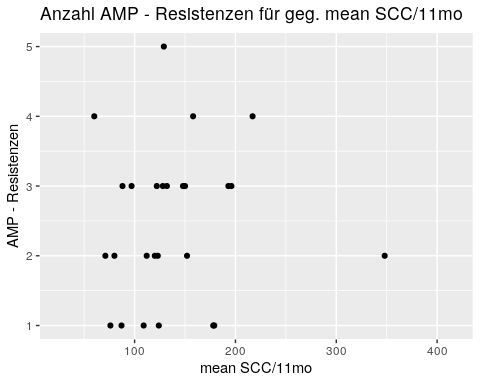
## [1] ""



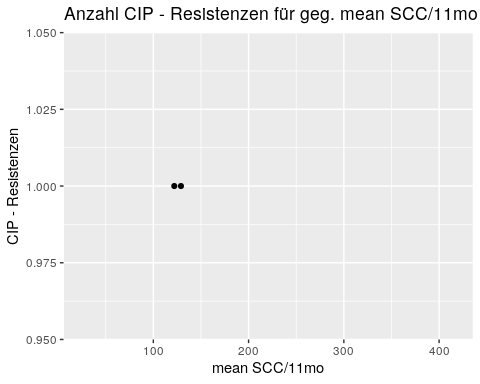
## [1] ""



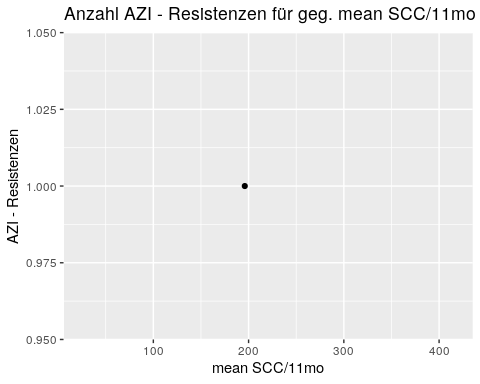
## [1] ""  
## [1] "--------------------------------------------------------"



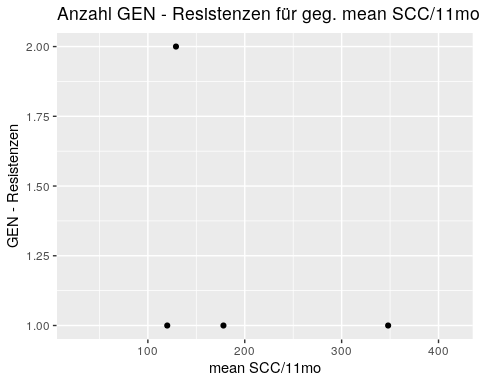
## [1] ""



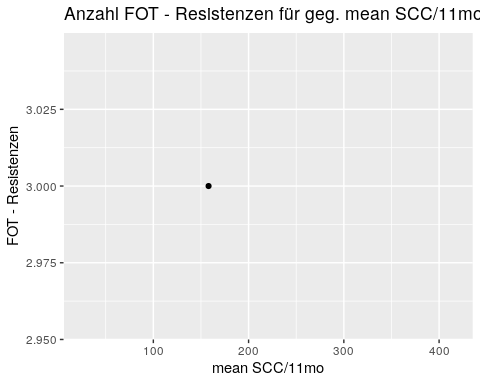
## [1] ""



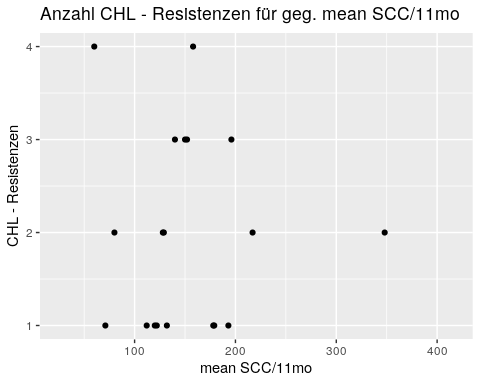
## [1] ""



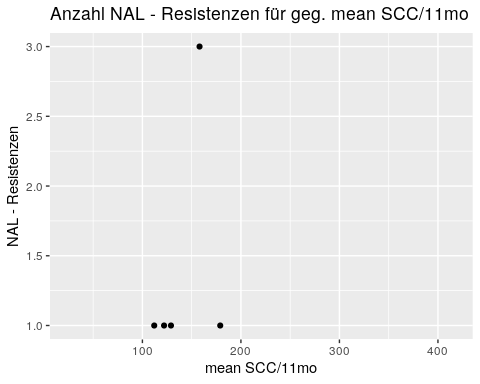
## [1] ""



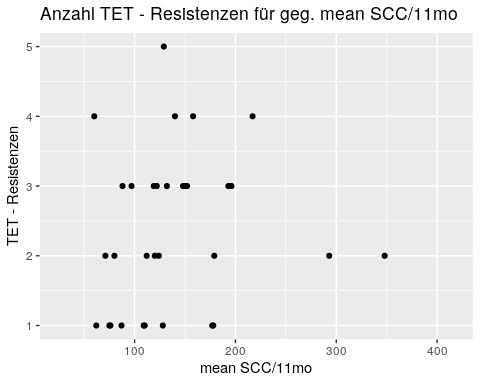
## [1] ""



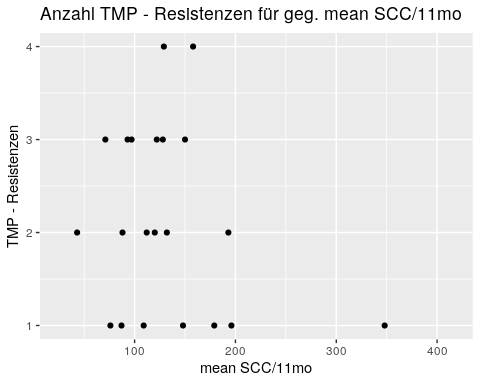
## [1] ""



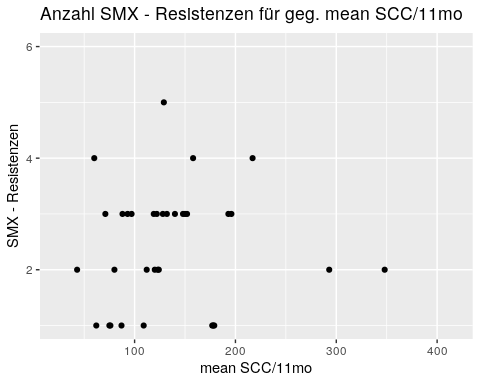
## [1] ""



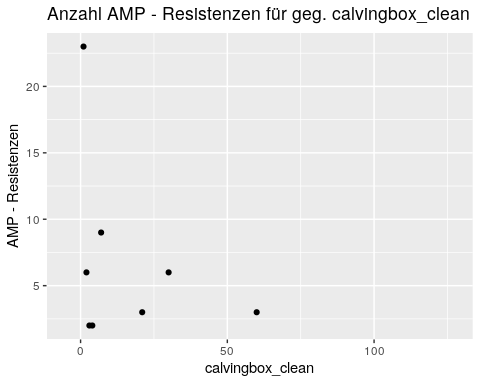
## [1] ""



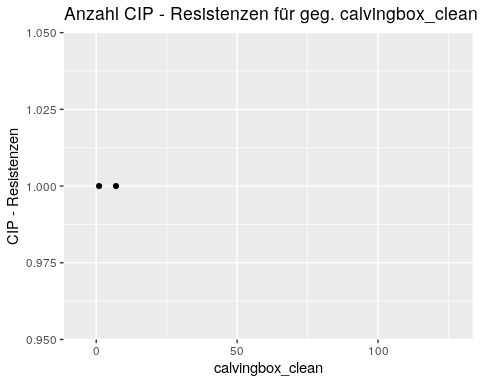
## [1] ""



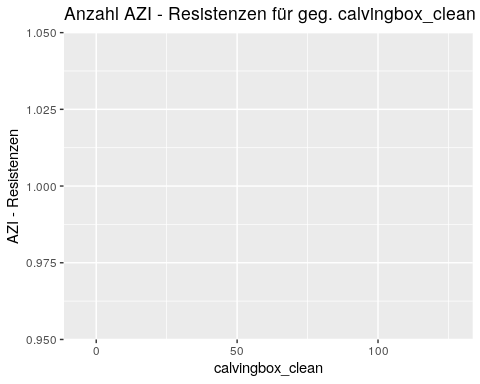
## [1] ""  
## [1] "--------------------------------------------------------"



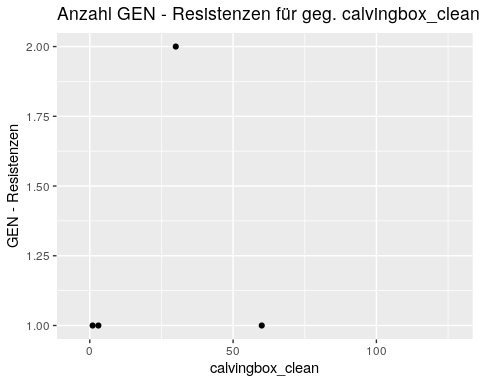
## [1] ""



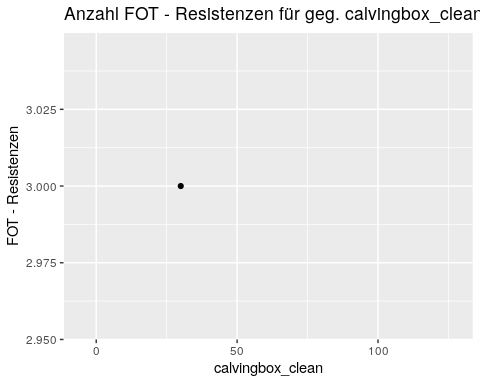
## [1] ""



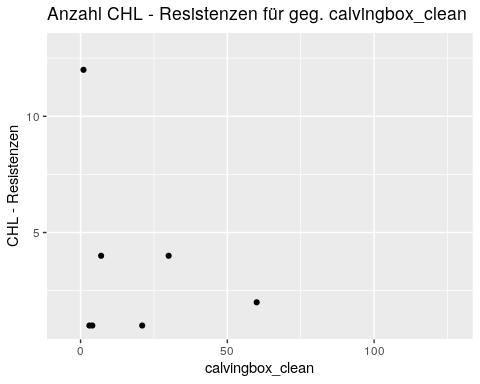
## [1] ""



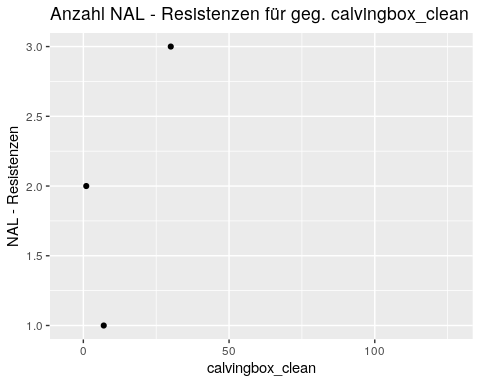
## [1] ""



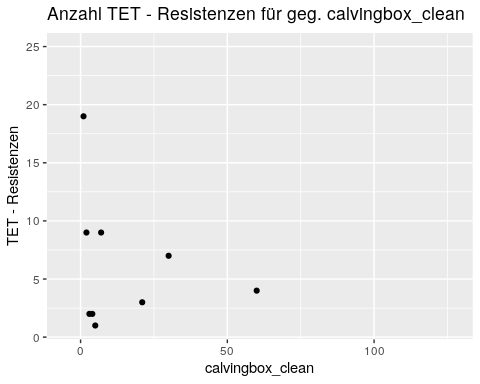
## [1] ""



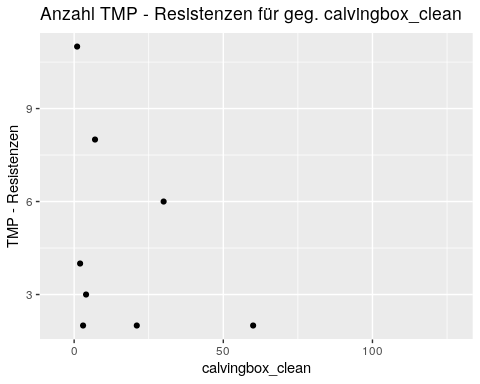
## [1] ""



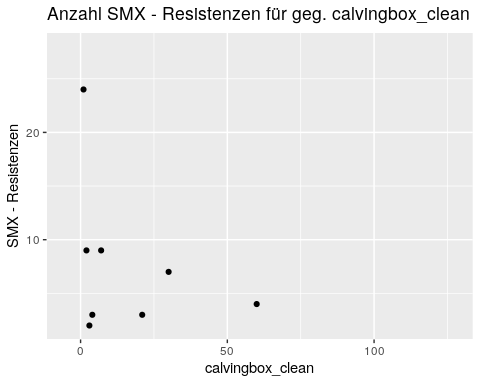
## [1] ""



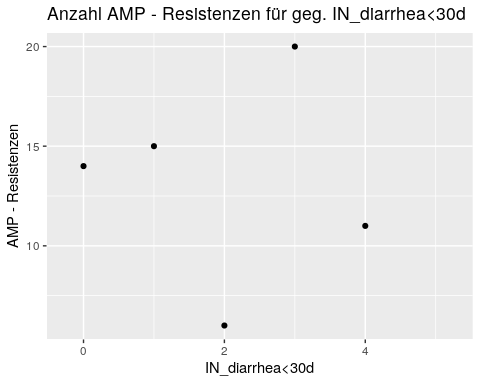
## [1] ""



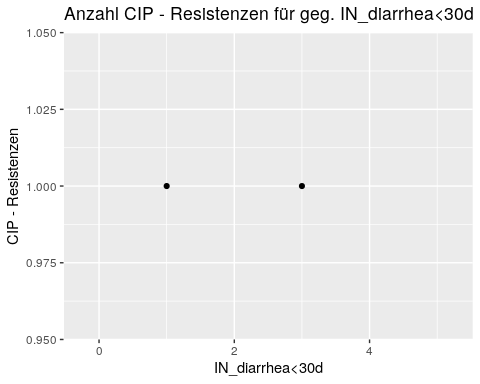
## [1] ""



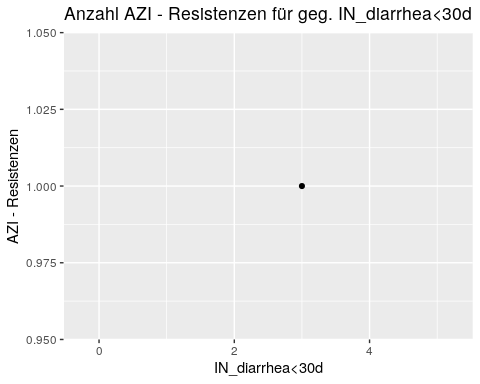
## [1] ""  
## [1] "--------------------------------------------------------"



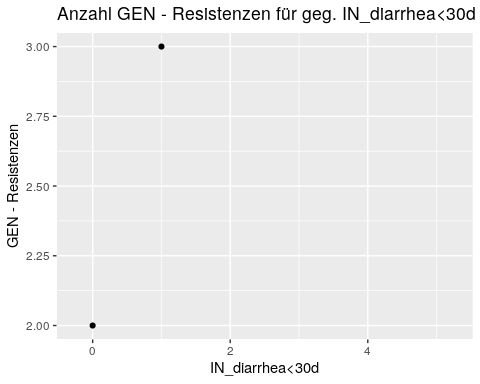
## [1] ""



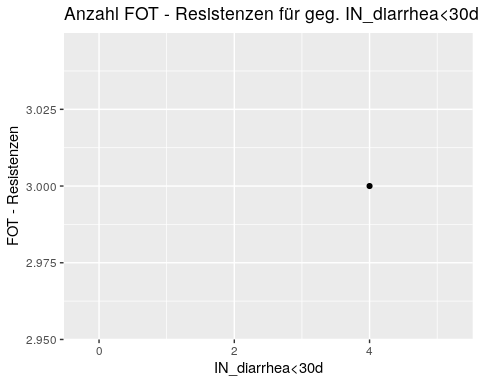
## [1] ""



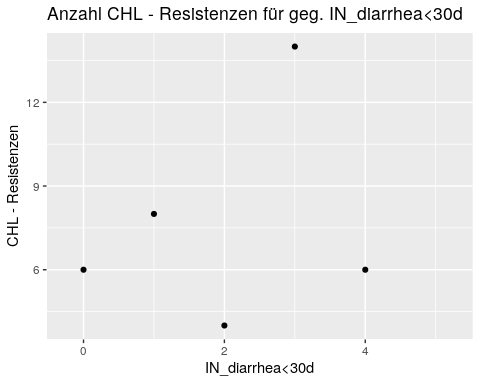
## [1] ""



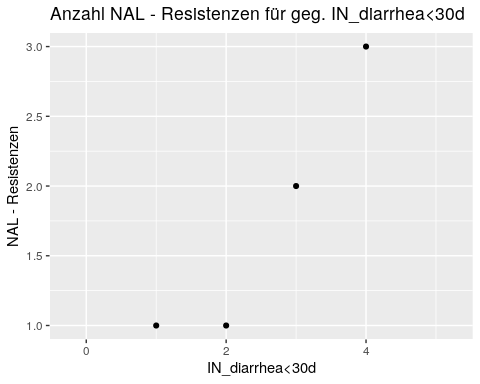
## [1] ""



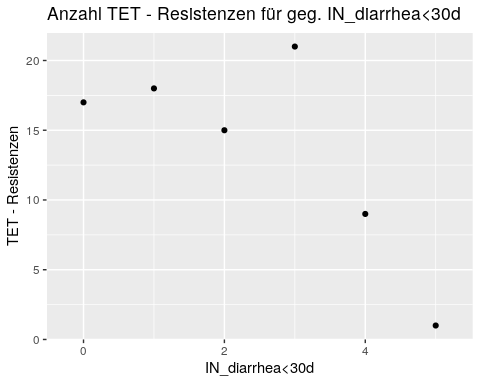
## [1] ""



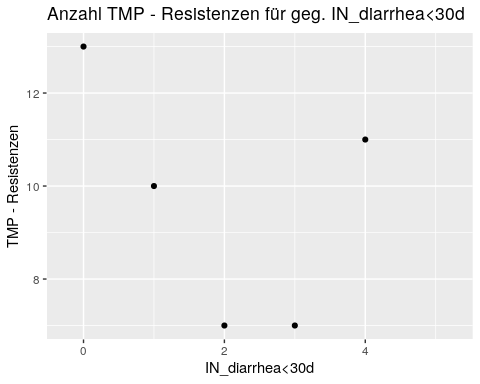
## [1] ""



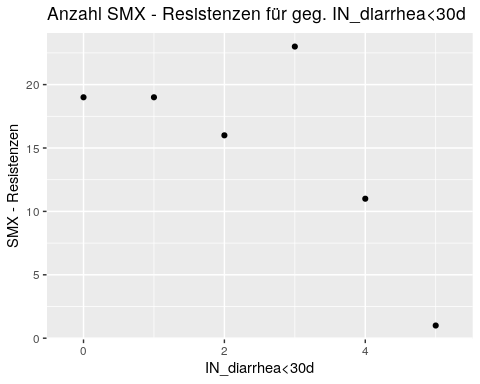
## [1] ""



## [1] ""



## [1] ""



## [1] ""  
## [1] "--------------------------------------------------------"

Resistenzen scheinen tendenziell zu

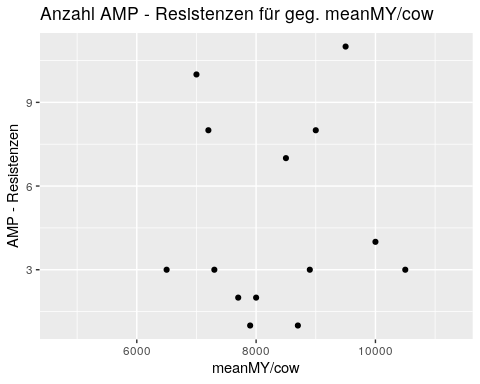
* steigen mit MY.group
* fallen mit SCC.group, CBC.group
* ? mit DIA.group

Eine Regression sagt mehr.

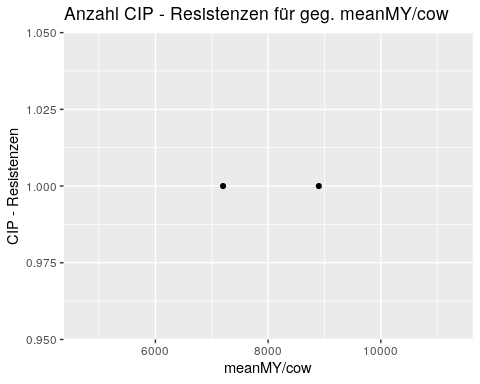
# Binäre und Nominale Unabhängige Variablen

## Anzahl Resistenzen

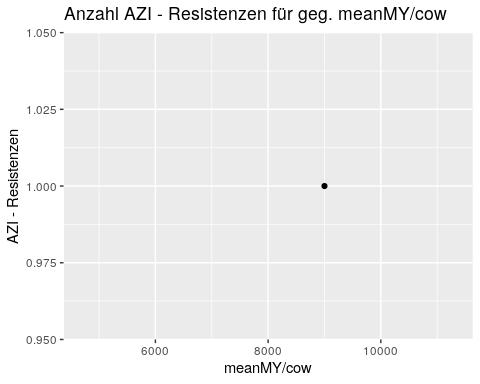
# NA warnings interessieren nicht  
  
for( group in groups ){  
 for( antib in c("AMP","CIP","AZI","GEN","FOT","CHL","NAL","TET","TMP","SMX") ){  
 graphisch2(group,"für geg.",antib)   
 print("")  
 }   
 print("--------------------------------------------------------")  
}



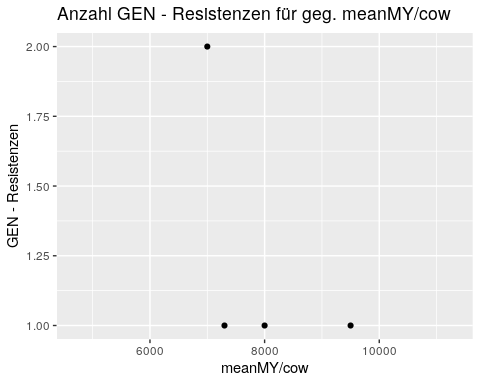
## [1] ""



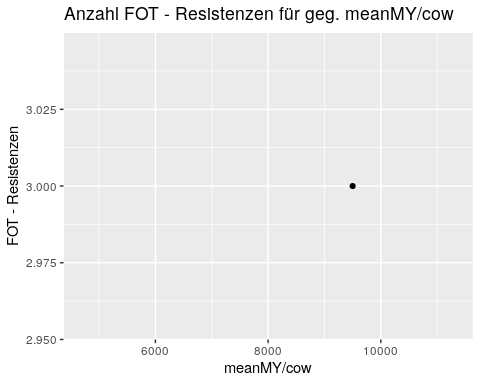
## [1] ""



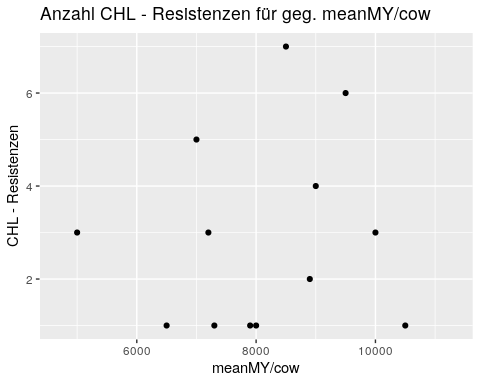
## [1] ""



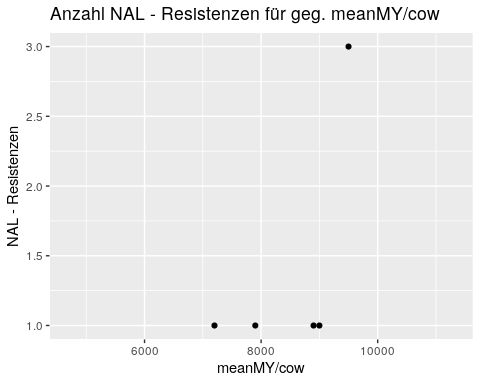
## [1] ""



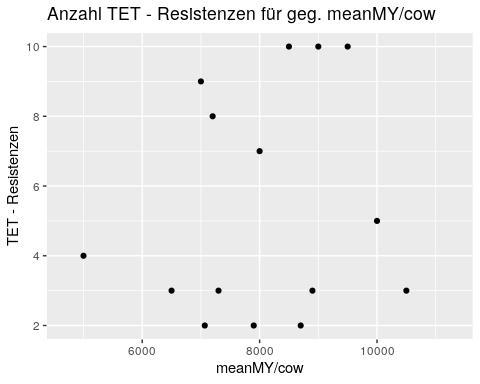
## [1] ""



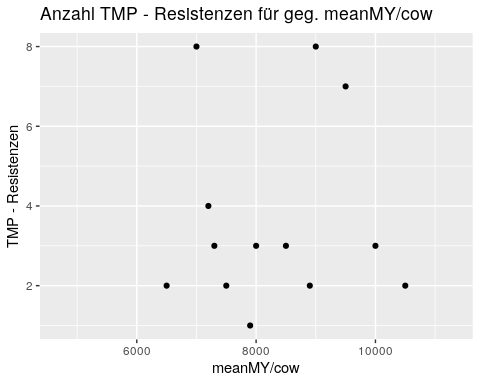
## [1] ""



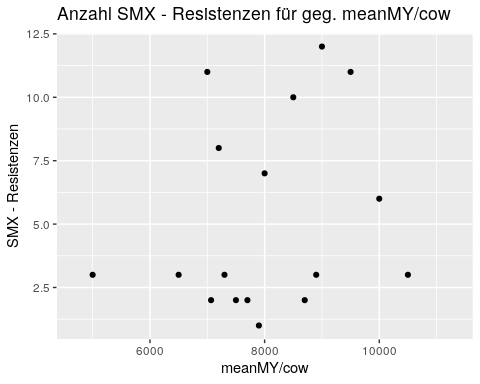
## [1] ""



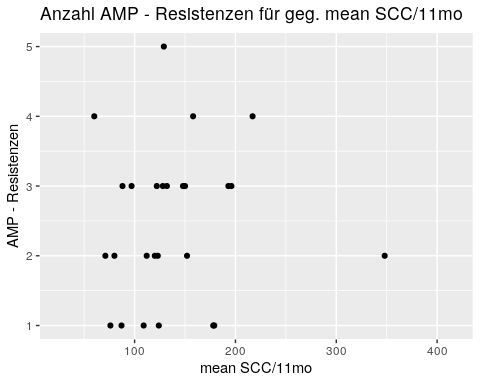
## [1] ""



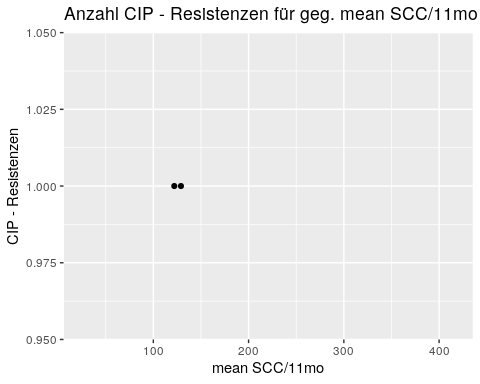
## [1] ""



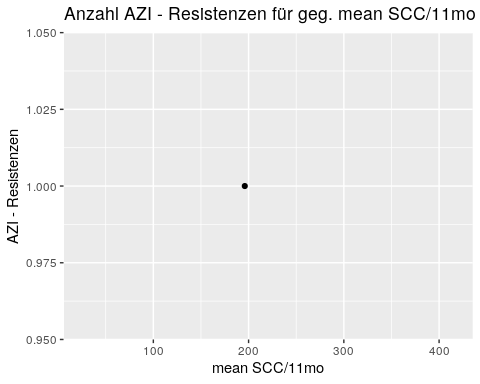
## [1] ""  
## [1] "--------------------------------------------------------"



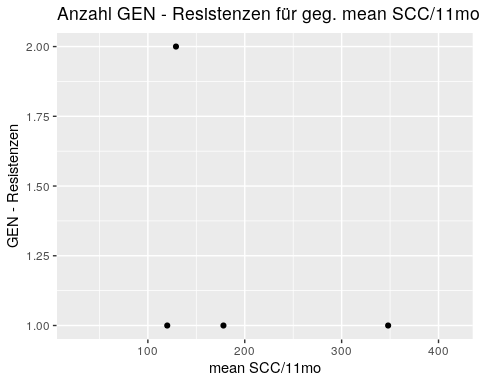
## [1] ""



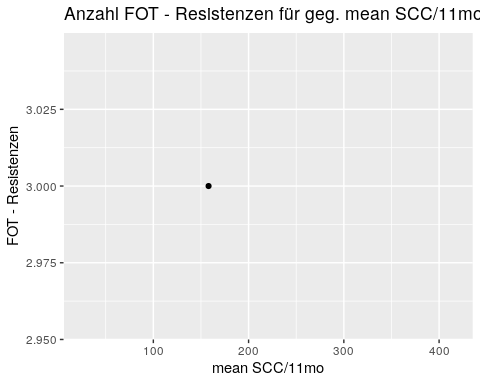
## [1] ""



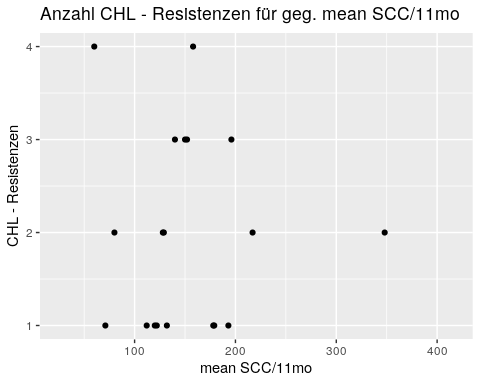
## [1] ""



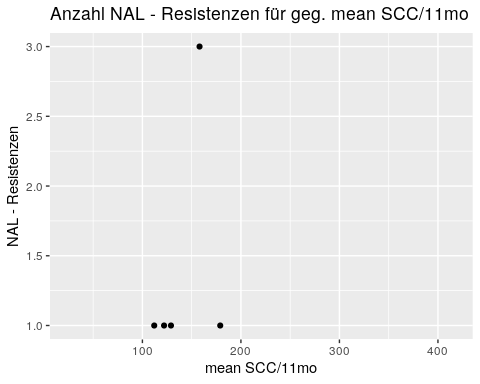
## [1] ""



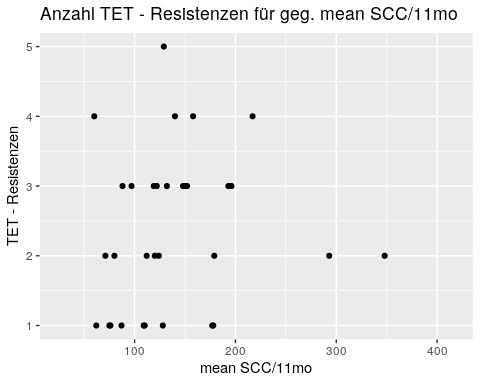
## [1] ""



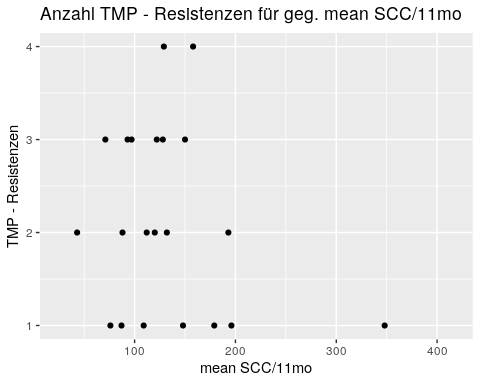
## [1] ""



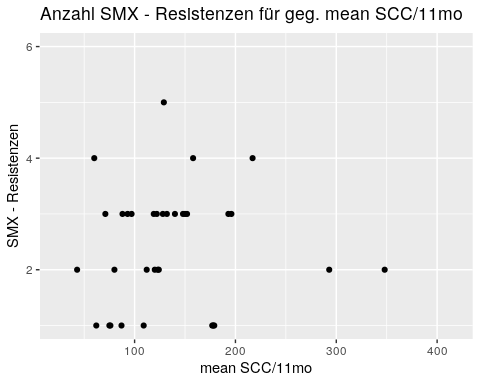
## [1] ""



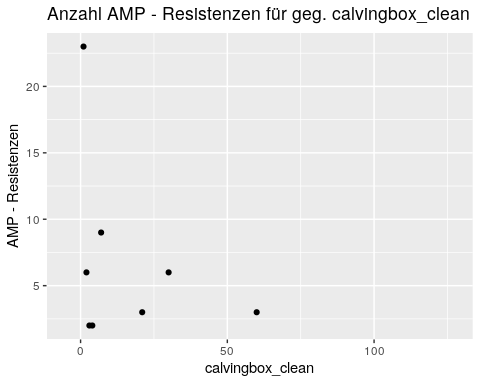
## [1] ""



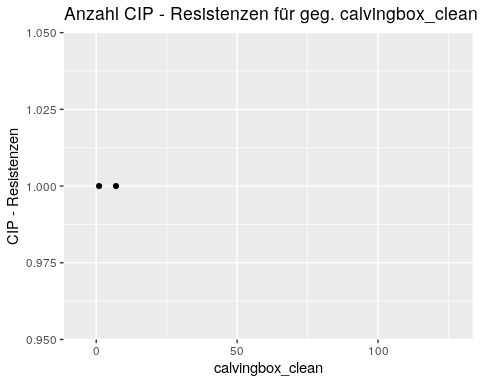
## [1] ""



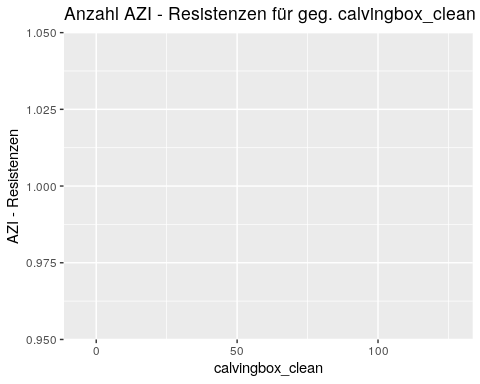
## [1] ""  
## [1] "--------------------------------------------------------"



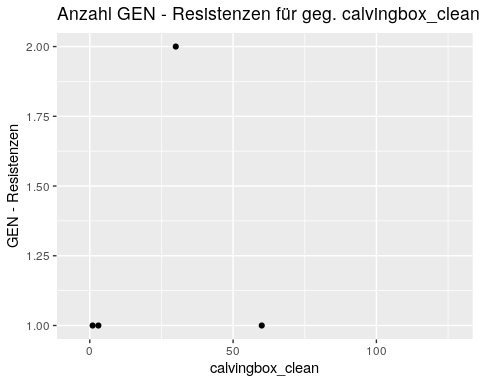
## [1] ""



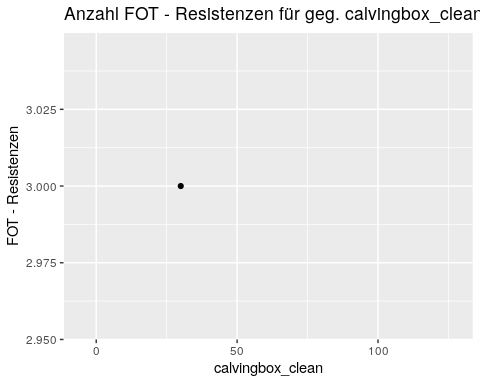
## [1] ""



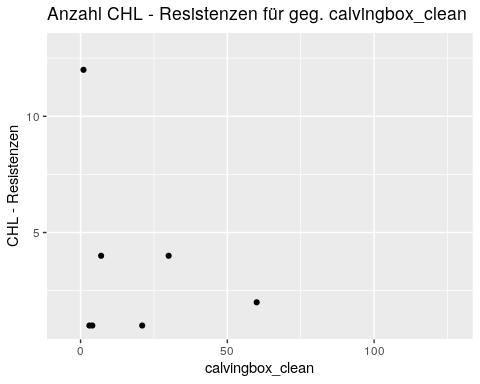
## [1] ""



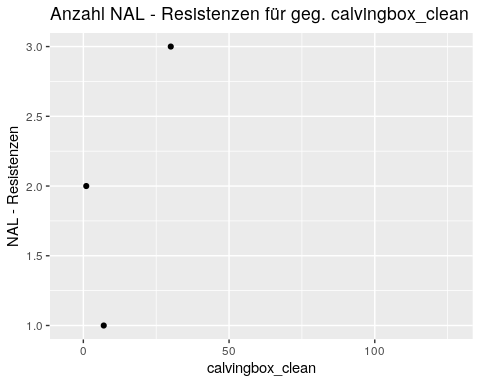
## [1] ""



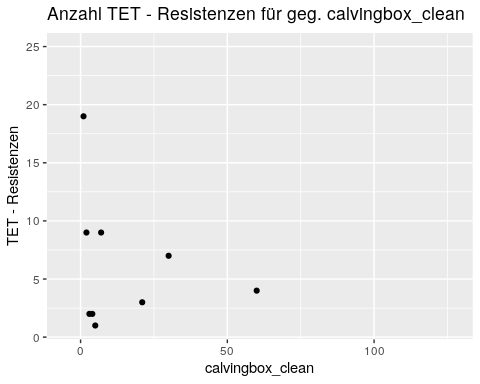
## [1] ""



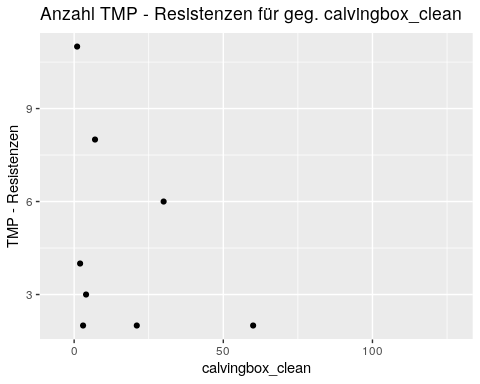
## [1] ""



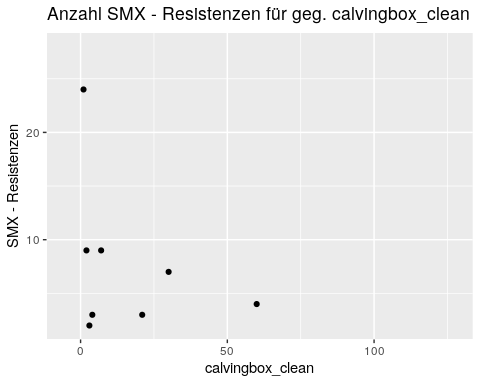
## [1] ""



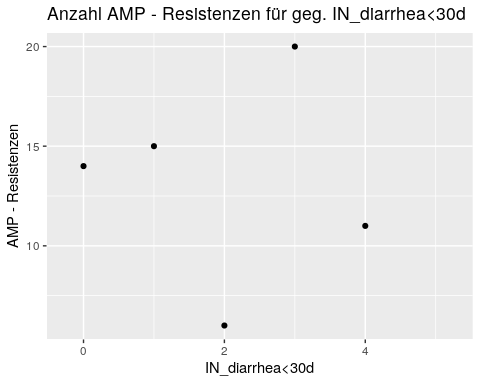
## [1] ""



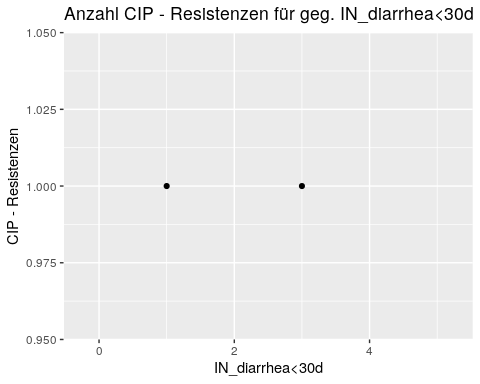
## [1] ""



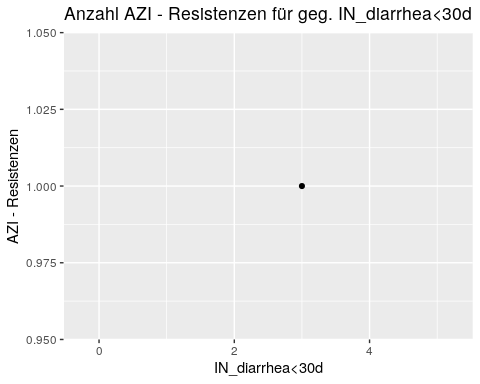
## [1] ""  
## [1] "--------------------------------------------------------"



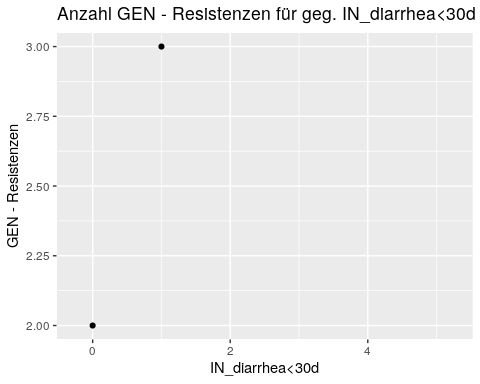
## [1] ""



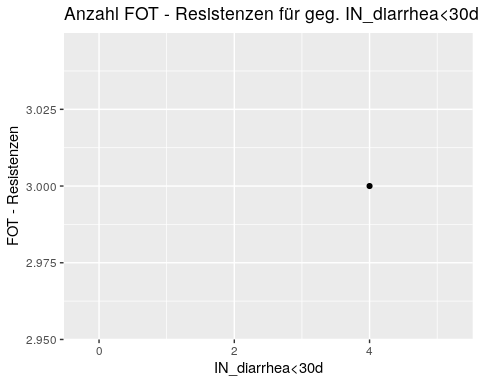
## [1] ""



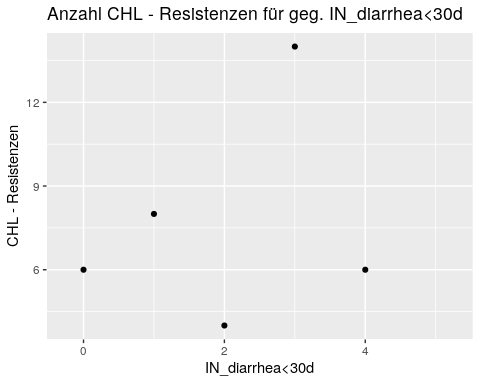
## [1] ""



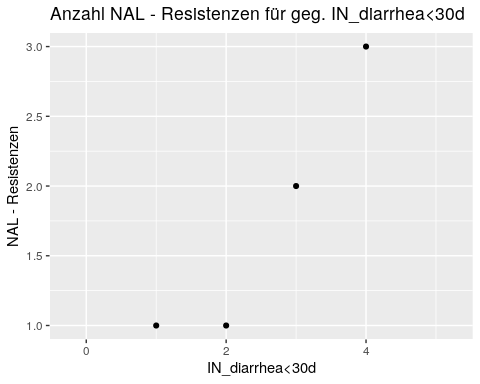
## [1] ""



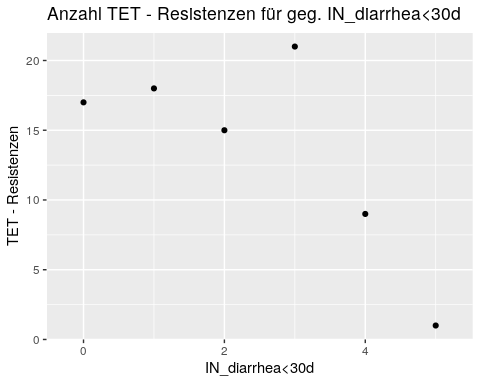
## [1] ""



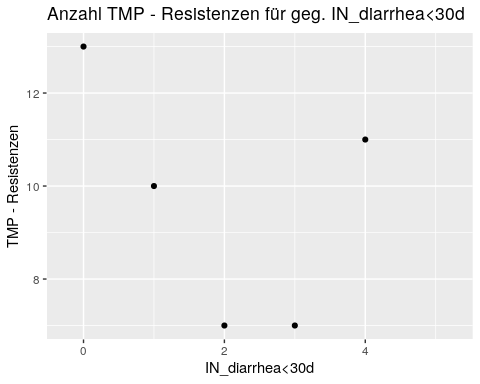
## [1] ""



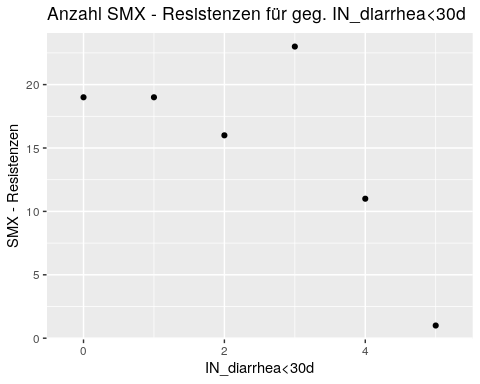
## [1] ""



## [1] ""



## [1] ""



## [1] ""  
## [1] "--------------------------------------------------------"

Resistenzen scheinen zu

* steigen mit MY.group (das sahen wir schon aus den Verteilungen), OLS.group, tendenziell auch IAC.group
* fallen bis HSC.group = 3, dann wieder etwas zu steigen (die Steigung von scheint einleuchtend, da 5=0+2 und 4=1+2; man könnte im plot vertauschen)
* jedenfalls sind die Trends klarer als aus den Verteilungen. Eine Regression sagt nochmal mehr