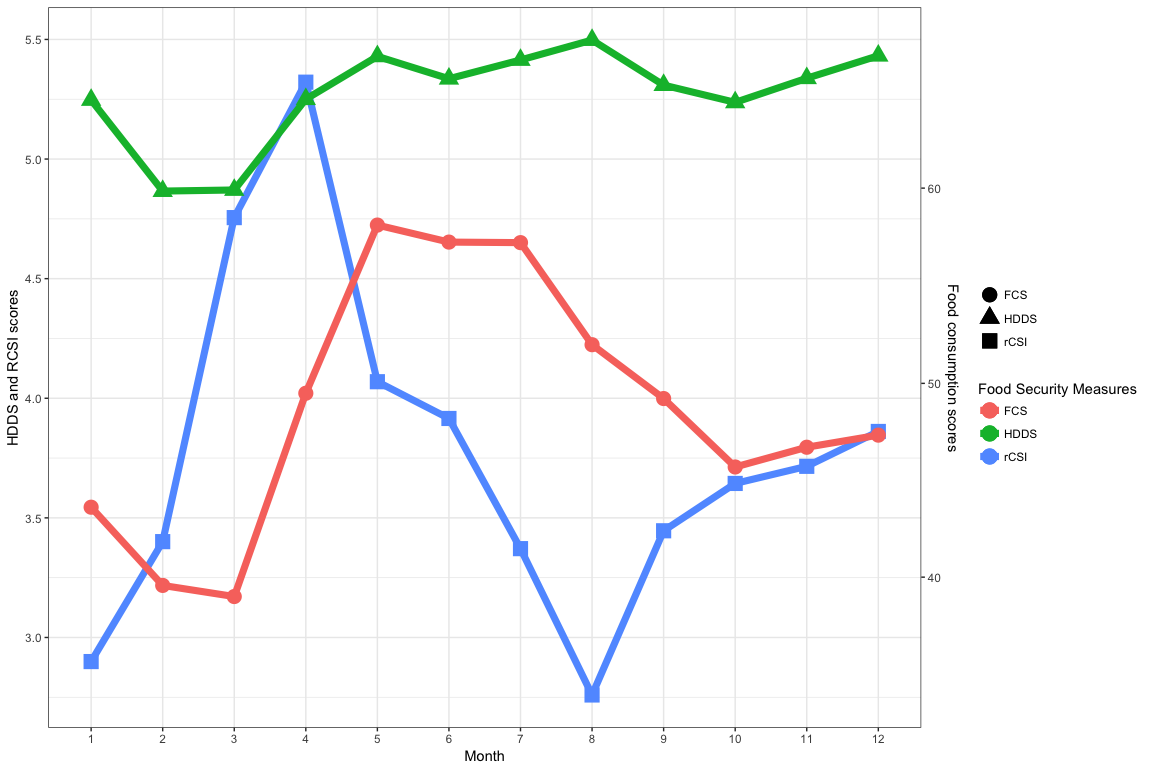
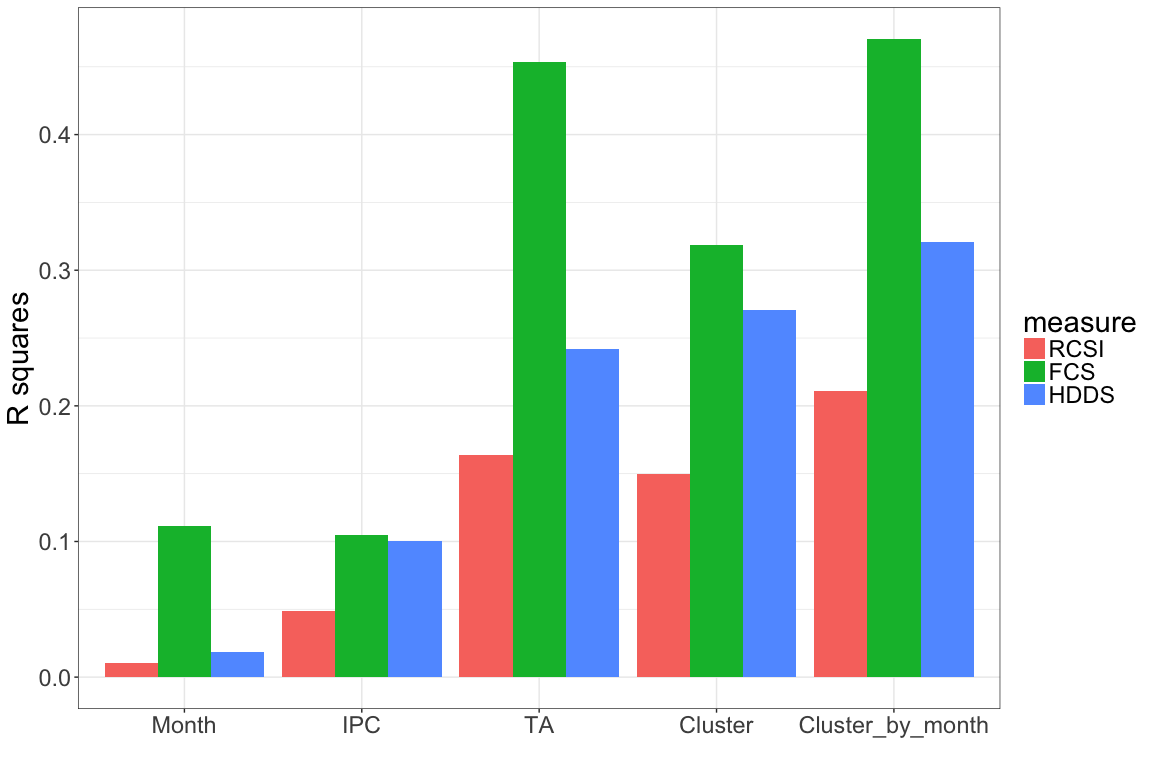
Result\_writeup

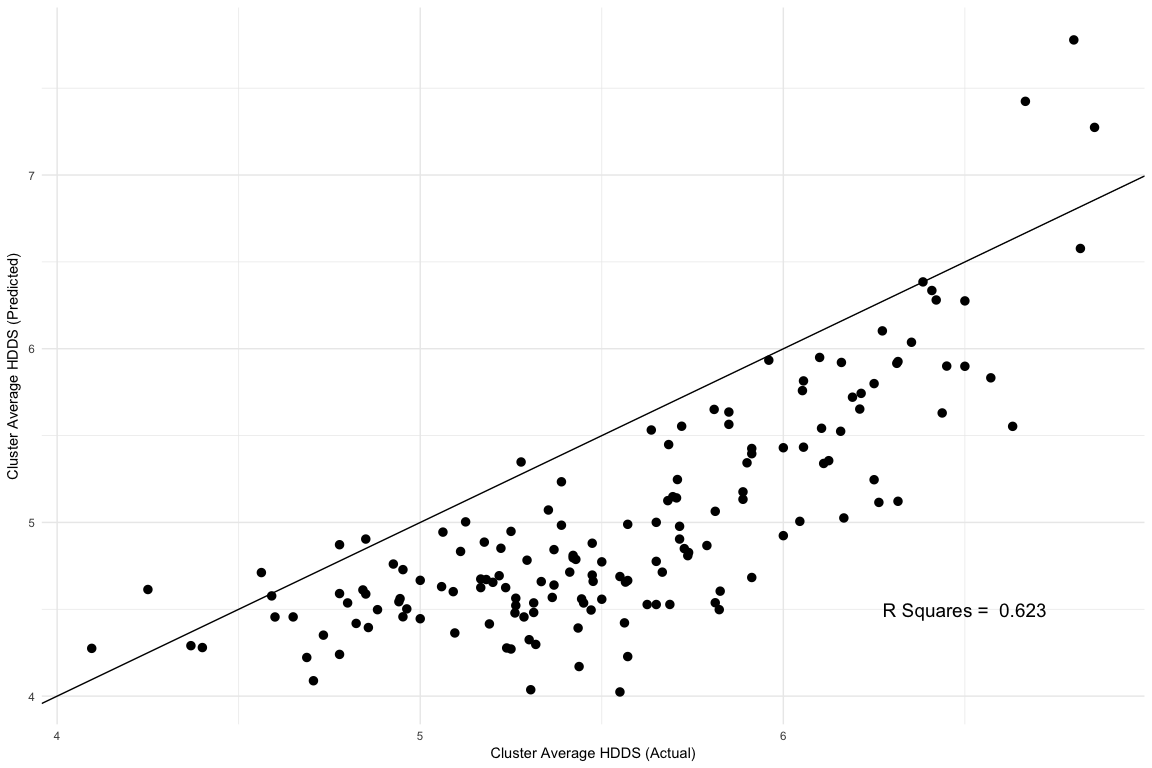
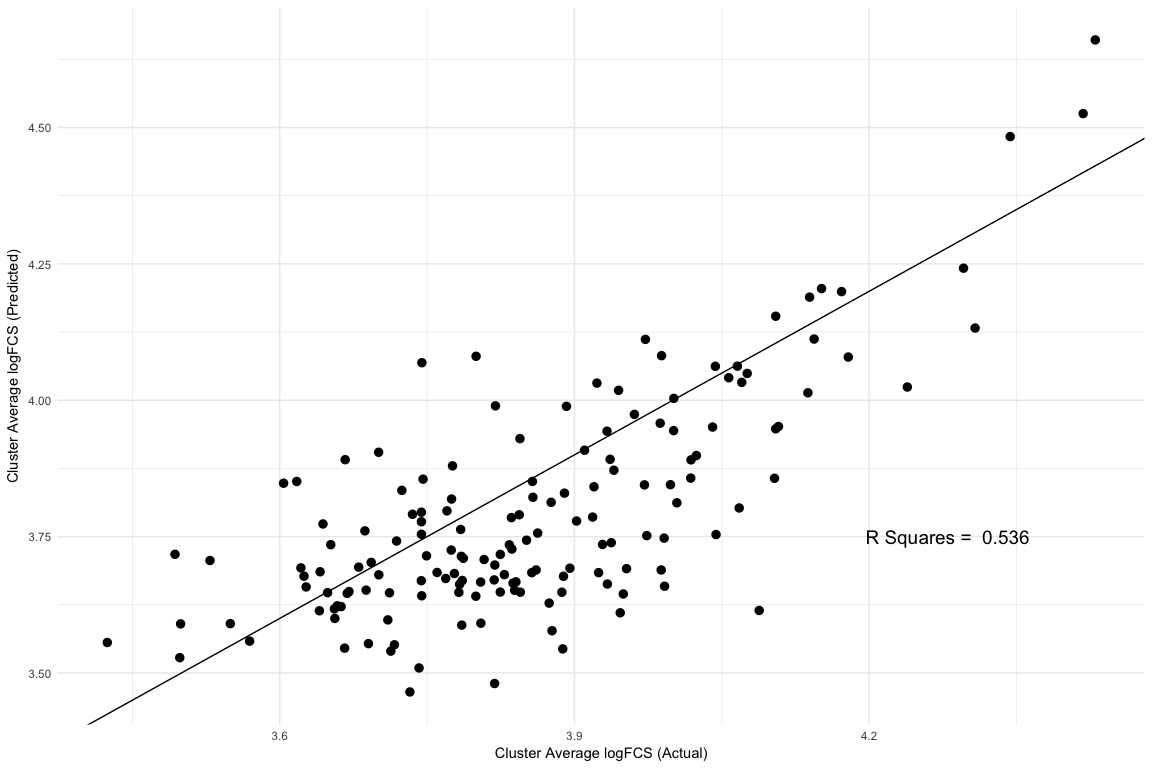
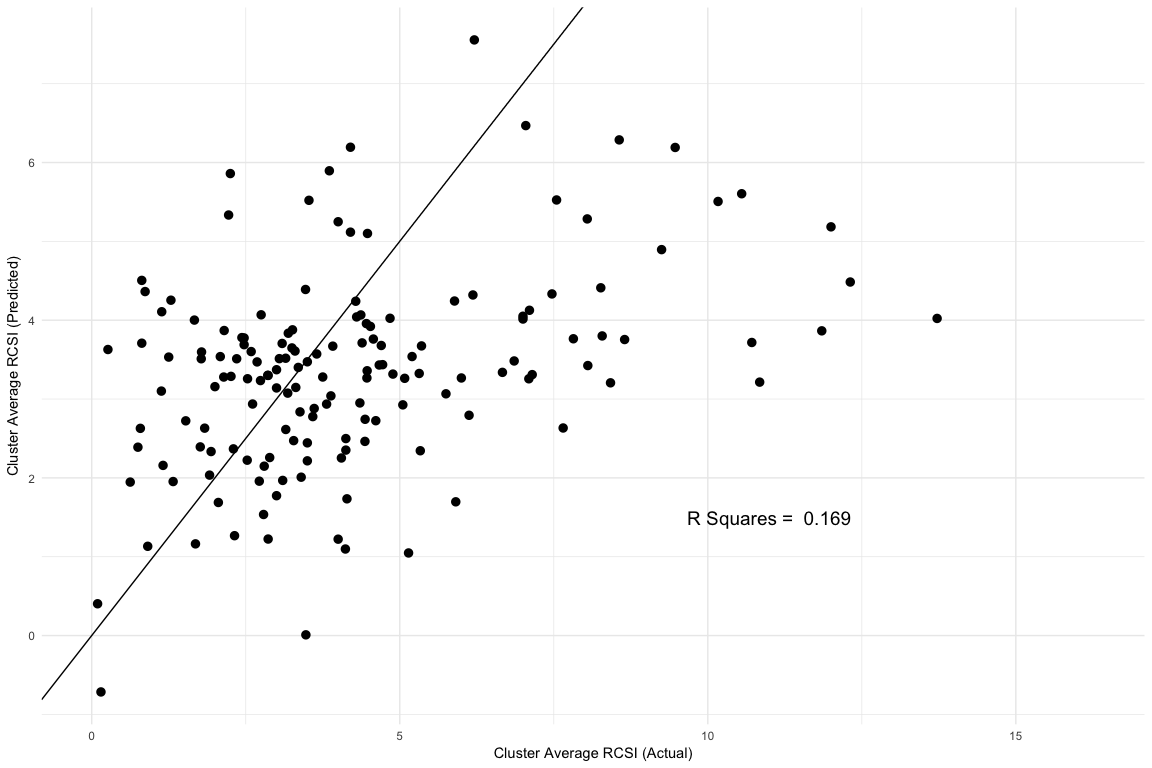
1. summary stats (pooled data of 2010 and 2013)
2. FS by month plot



1. bar chart of variation

data = malawi\_join %>%   
 select(rCSI,FCS,HDDS,FS\_month,fnid,TA\_names,ea\_id)  
  
  
data$month\_clust <-interaction(data$FS\_month,data$ea\_id)  
data$rCSI[data$rCSI>42]<-42  
data<-data[data$HDDS!=0,]  
  
colnames(data)[4] <-"Month"  
colnames(data)[5] <-"IPC"  
colnames(data)[6] <-"TA"  
colnames(data)[7] <-"Cluster"  
colnames(data)[8] <-"month\_clust"  
  
   
  
# Maybe a bar chart of variation in FS security measures (month, IPCzone, TA, cluster, month x cluster)  
  
r2 = matrix(NA,3,5)  
my\_lms <- lapply(1:3, function(x) lm(data[,x] ~ as.factor(Month),data=data))  
summaries <- lapply(my\_lms, summary)  
r2[,1]<-sapply(summaries, function(x) c(r\_sq = x$r.squared))  
  
  
my\_lms <- lapply(1:3, function(x) lm(data[,x] ~ as.factor(IPC),data=data))  
summaries <- lapply(my\_lms, summary)  
r2[,2]<-sapply(summaries, function(x) c(r\_sq = x$r.squared))  
  
  
my\_lms <- lapply(1:3, function(x) lm(data[,x] ~ as.factor(TA),data=data))  
summaries <- lapply(my\_lms, summary)  
r2[,3]<-sapply(summaries, function(x) c(r\_sq = x$r.squared))  
  
  
my\_lms <- lapply(1:3, function(x) lm(data[,x] ~ as.factor(Cluster),data=data))  
summaries <- lapply(my\_lms, summary)  
r2[,4]<-sapply(summaries, function(x) c(r\_sq = x$r.squared))  
  
  
  
my\_lms <- lapply(1:3, function(x) lm(data[,x] ~ as.factor(month\_clust),data=data))  
summaries <- lapply(my\_lms, summary)  
r2[,5]<-sapply(summaries, function(x) c(r\_sq = x$r.squared))  
  
  
r2 <-as.data.frame(r2)  
  
colnames(r2)<-c("Month","IPC","TA","Cluster","Cluster\_by\_month")  
rownames(r2)<-c("RCSI","FCS","HDDS")  
#r2$measure<-c("RCSI","FCS","HDDS")  
  
  
long\_r2<- melt(t(r2),na.rm=TRUE,measure.vars =c("RCSI","FCS","HDDS"))  
long\_r2<-long\_r2[long\_r2$Var1!="measure",]  
  
colnames(long\_r2)<-c("variation","measure","rsquares")  
  
ar<-as.character(long\_r2$rsquares)  
long\_r2$ar<-as.numeric(ar)  
  
  
# RCSI\_r2\_fe = .22320556  
# HDDS\_r2\_fe = .30987621  
# logFCS\_r2\_fe = .33045673  
  
  
p<-ggplot(long\_r2, aes(variation,ar)) +geom\_bar(stat = "identity",position = "dodge",aes(fill = measure))  
  
p <- p + labs(y = "R squares",  
 x = " ",  
 colour = "Food Security Measures",  
 shape= " ")  
  
p <- p+ theme\_bw()   
p <- p + theme(text = element\_text(size=22))   
p



1. Result of 2013 prediction
2. Scatter plots (predict vs. actual) 
3. Density plot (predication using different scales + household)
   1. Unexplored variation of household level

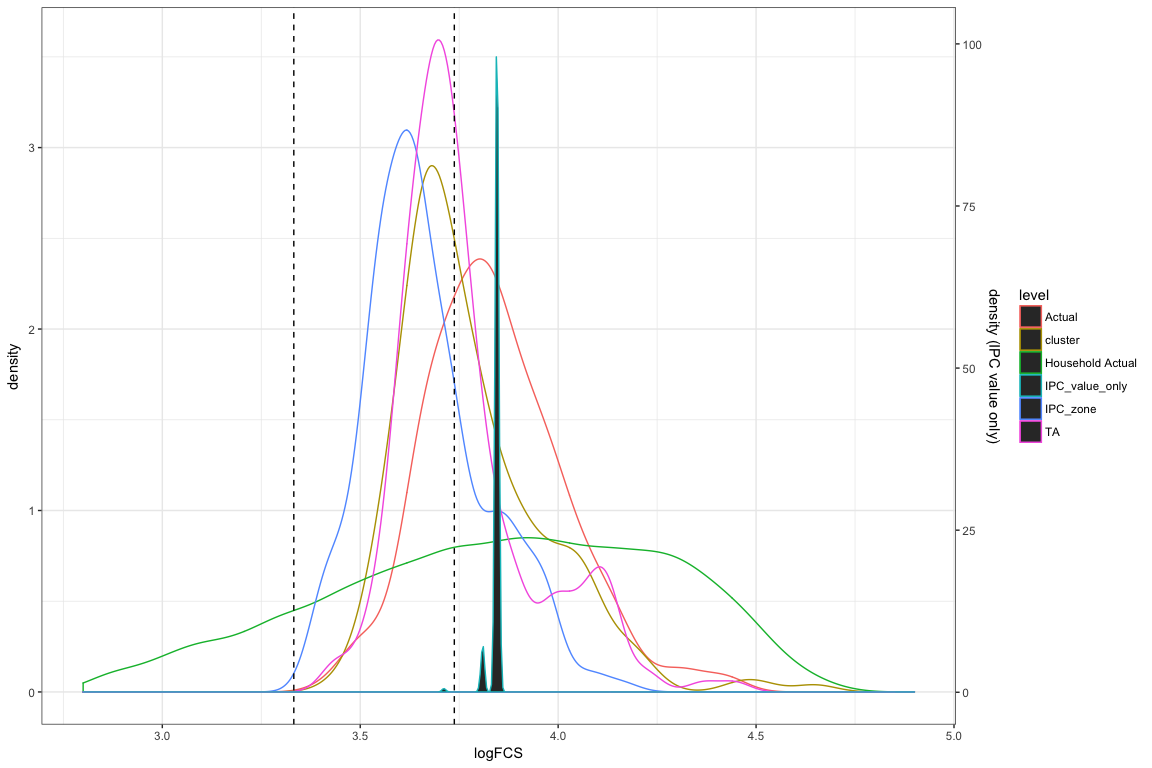
## <environment: R\_GlobalEnv>

## Warning in cbind(predict\_df$clust\_logFCS\_ipczone\_predict\_m3, predict\_df  
## $clust\_logFCS\_TA\_predict\_m3, : number of rows of result is not a multiple  
## of vector length (arg 1)

## Warning in bind\_rows\_(x, .id): binding factor and character vector,  
## coercing into character vector

## Warning in bind\_rows\_(x, .id): binding character and factor vector,  
## coercing into character vector

## Warning: Removed 423 rows containing non-finite values (stat\_density).

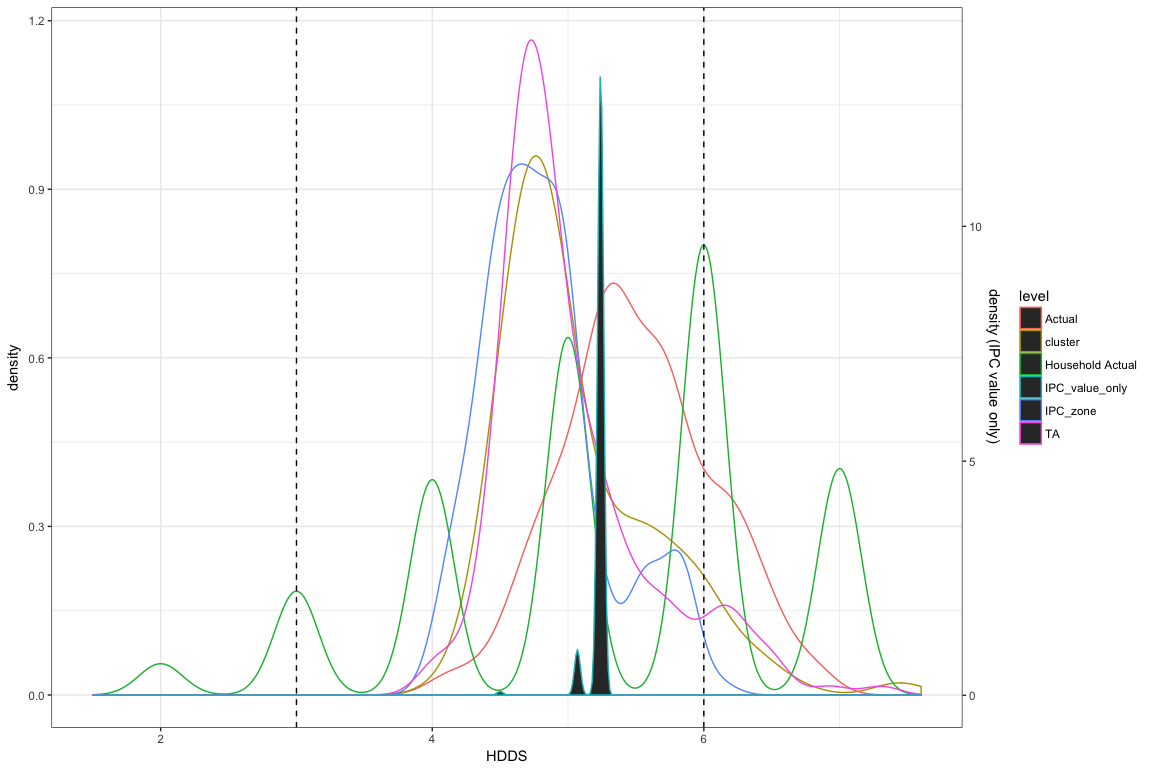


## Warning in cbind(predict\_df$clust\_HDDS\_ipczone\_predict\_m3, predict\_df  
## $clust\_HDDS\_TA\_predict\_m3, : number of rows of result is not a multiple of  
## vector length (arg 1)

## Warning in bind\_rows\_(x, .id): binding factor and character vector,  
## coercing into character vector

## Warning in bind\_rows\_(x, .id): binding character and factor vector,  
## coercing into character vector

## Warning: Removed 7 rows containing non-finite values (stat\_density).

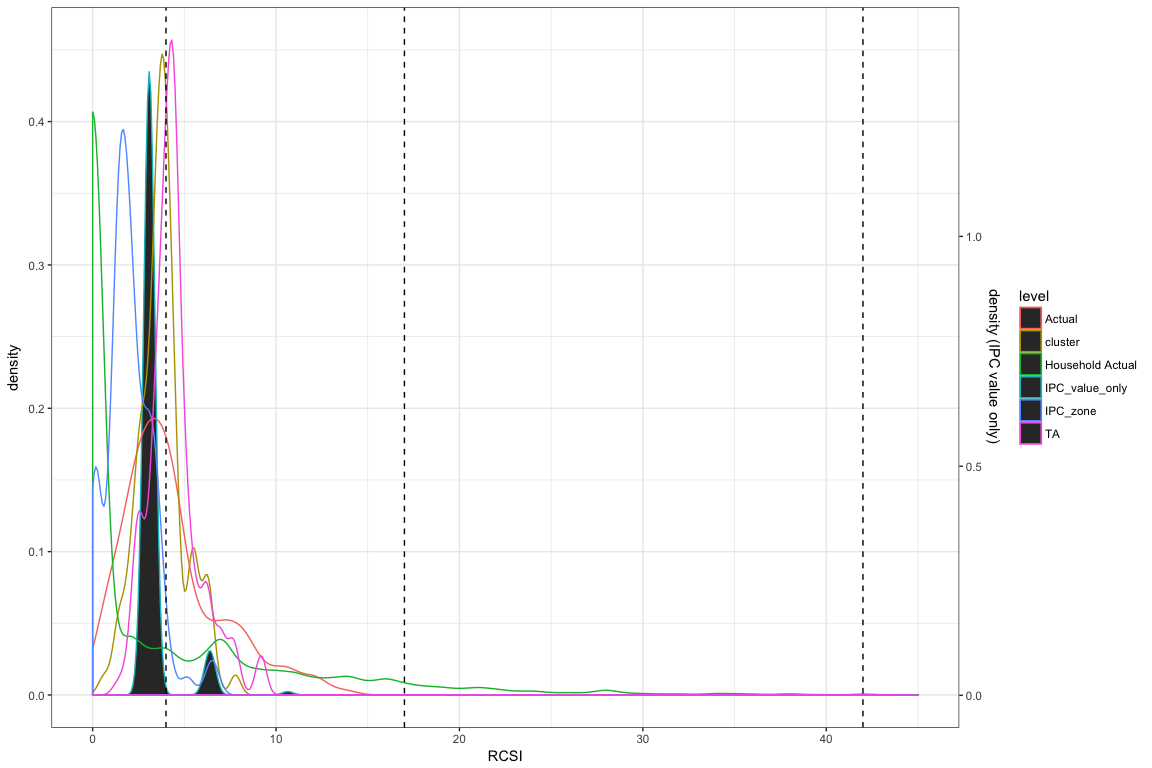


## Warning in cbind(predict\_df$clust\_RCSI\_ipczone\_predict\_m3, predict\_df  
## $clust\_RCSI\_TA\_predict\_m3, : number of rows of result is not a multiple of  
## vector length (arg 1)

## Warning in bind\_rows\_(x, .id): binding factor and character vector,  
## coercing into character vector

## Warning in bind\_rows\_(x, .id): binding character and factor vector,  
## coercing into character vector

## Warning: Removed 40 rows containing non-finite values (stat\_density).



1. R squares of 2013 predication

## Warning in cbind(predict\_df$clust\_logFCS\_ipczone\_predict\_m2, predict\_df  
## $clust\_logFCS\_TA\_predict\_m2, : number of rows of result is not a multiple  
## of vector length (arg 1)

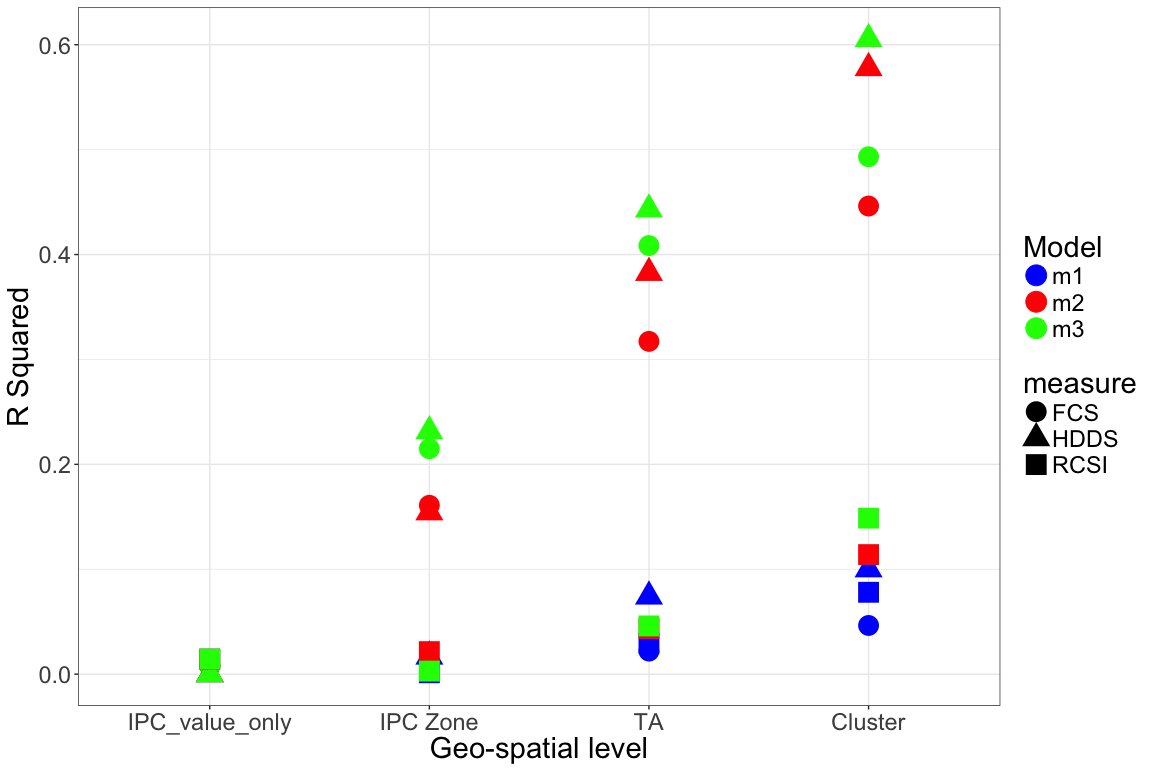
## Warning in cbind(predict\_df$clust\_logFCS\_ipczone\_predict\_m1, predict\_df  
## $clust\_logFCS\_TA\_predict\_m1, : number of rows of result is not a multiple  
## of vector length (arg 1)

## Warning in cbind(predict\_df$clust\_HDDS\_ipczone\_predict\_m2, predict\_df  
## $clust\_HDDS\_TA\_predict\_m2, : number of rows of result is not a multiple of  
## vector length (arg 1)

## Warning in cbind(predict\_df$clust\_HDDS\_ipczone\_predict\_m1, predict\_df  
## $clust\_HDDS\_TA\_predict\_m1, : number of rows of result is not a multiple of  
## vector length (arg 1)

## Warning in cbind(predict\_df$clust\_RCSI\_ipczone\_predict\_m2, predict\_df  
## $clust\_RCSI\_TA\_predict\_m2, : number of rows of result is not a multiple of  
## vector length (arg 1)

## Warning in cbind(predict\_df$clust\_RCSI\_ipczone\_predict\_m1, predict\_df  
## $clust\_RCSI\_TA\_predict\_m1, : number of rows of result is not a multiple of  
## vector length (arg 1)



1. Discussion of classification
2. hit and miss tables (for the predications)
   1. one is cluster to actual cluster outcomes
   2. one is cluster predication to actual household level outcomes
   3. how bad it can be to just target the cluster level ?
   4. put it in SI ?
3. 2010 data (only what matters for 2010 is the tables (coefficients and variables))
4. regression results
5. discussion on the coefficients

# library("ggpubr")  
# ggscatter(logFCS\_pair, x = "actual", y = "predict",  
# add = "reg.line", conf.int = TRUE,   
# cor.coef = FALSE, cor.method = "pearson",  
# xlab = "actual", ylab = "predict")  
#   
# ggsave("logFCS\_scatter.png", plot = last\_plot(),device = "png",path = "output/figures/",  
# dpi = 1000, limitsize = TRUE)  
#   
# ggscatter(d, x ="RCSI\_cluster" , y = "RCSI\_pred",  
# add = "reg.line", conf.int = TRUE,   
# cor.coef = TRUE, cor.method = "pearson",  
# xlab = "actual", ylab = "predict")  
#   
# ggsave("RCSI\_scatter.png", plot = last\_plot(),device = "png",path = "output/figures/",  
# dpi = 1000, limitsize = TRUE)  
#   
# ggscatter(d, x ="HDDS\_cluster" , y = "HDDS\_pred",  
# add = "reg.line", conf.int = TRUE,   
# cor.coef = FALSE, cor.method = "pearson",  
# xlab = "actual", ylab = "predict")  
#   
# ggsave("HDDS\_scatter.png", plot = last\_plot(),device = "png",path = "output/figures/",  
# dpi = 1000, limitsize = TRUE)