Counterfactual Testing Associational Model

```
# clear the working space
rm(list = ls())

library(here)
library(stats) # need this to calculate Mahalanobis Distance
library(parallel) # parallelize
library(dplyr)
library(FNN)
library(cluster)
library(ggplot2)
library(rpart)
library(caret)
```

Counterfactual Data Input

```
# we need the renaming function for cleaning
melor_2015 <- read.csv(here("data", "melor15_CF_data.csv"))
```

```
# this is going to load mun_clusters which is the list with the clusters
load(here("data", "clusters.RData"))
```

Counterfactual

Importing trained models

```
# Read the .rds models
base_reg <- readRDS(here("associational XGBOOST", "damage_fit_reg_base.rds"))
trunc_reg <- readRDS(here("associational XGBOOST", "trunc_damage_fit_reg.rds"))
clas_model  <- readRDS(here("associational XGBOOST", "ass_XGBOOST_class.rds"))
```

Counterfactual predictions

```
source(here("R", "ass_hurdle_function.R"))

# setting threshold for classification step
threshold = 0.30

preds <- ass_hurdle_function(df = melor_2015, ass_clas_model = clas_model,
  ass_base_model = base_reg, ass_trunc_model = trunc_reg ,threshold = threshold)
```

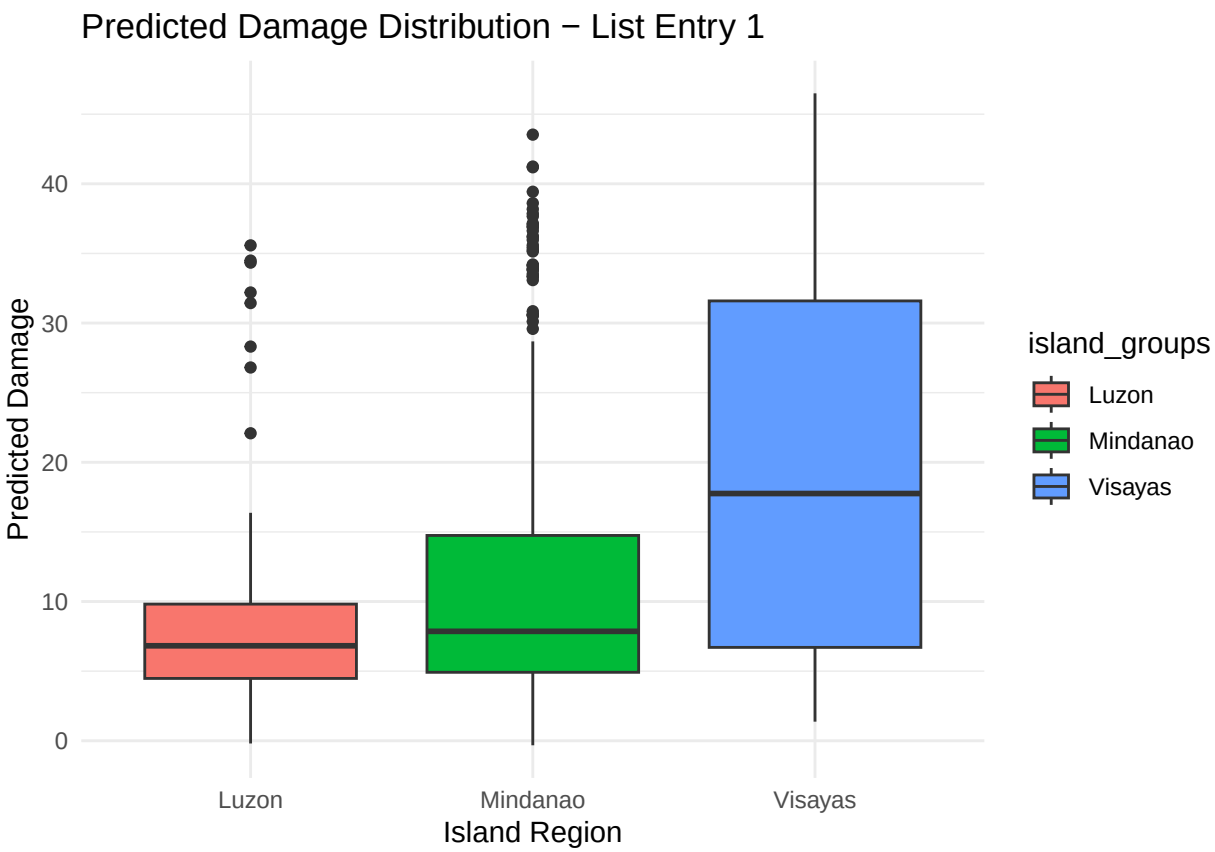
```r
# append the results to the counterfactual dataset
melor_2015  <- melor_2015 %>%
    mutate(damage_preds = preds)
```

Counterfactual clusters

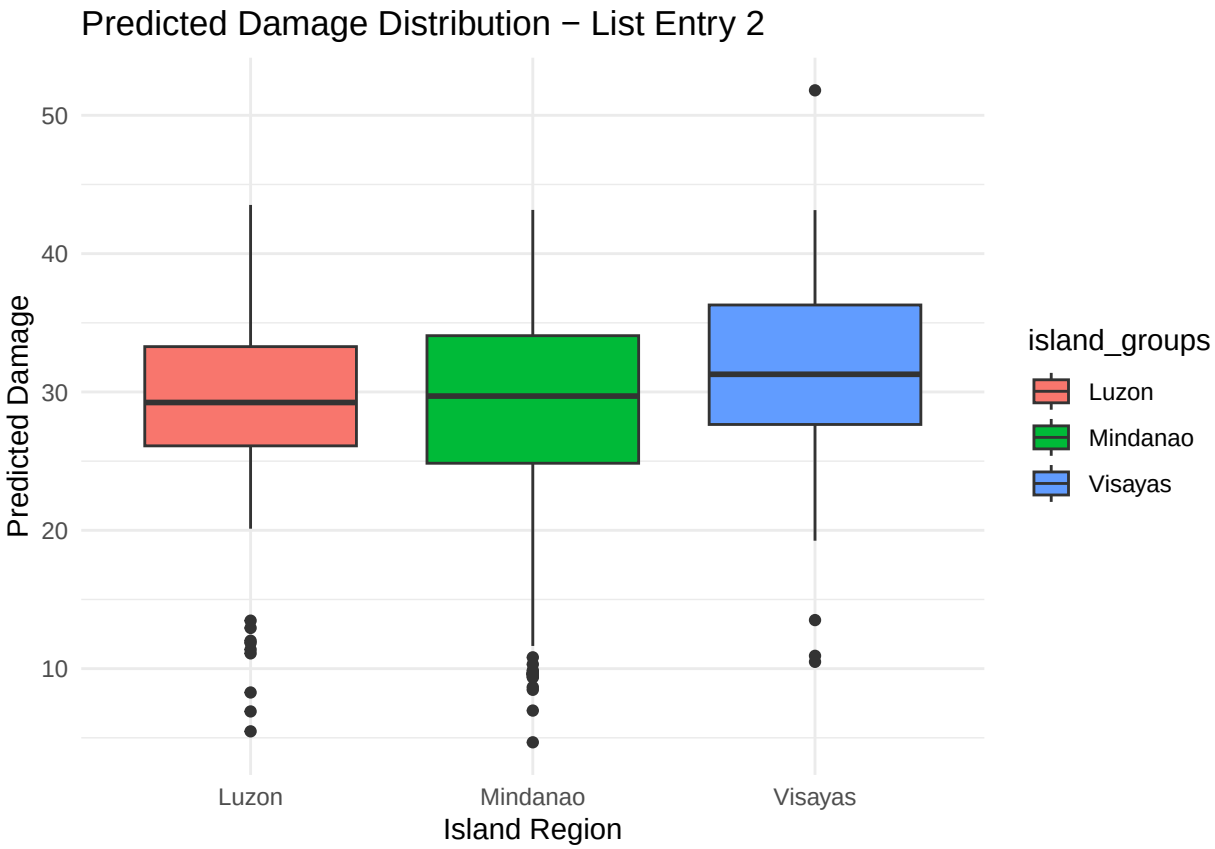
```r
plots_list  <- list()
means_list  <- list()
median_list  <- list()

for (i in seq_along(mun_clusters)) {
  # Get the current entry
  current_entry <- mun_clusters[[i]]

  # Convert the nested list entry to a data frame format
  plot_data <- bind_rows(lapply(names(current_entry), function(region) {
    data.frame(Mun_Code = unlist(current_entry[[region]]), island_regions = region, stringsAsFactors = F
  }))

  # Merge with original data to get predicted damage
  merged_data <- melor_2015 %>%
    inner_join(plot_data, by = "Mun_Code")

  # Create boxplot
  p <- ggplot(merged_data, aes(x = island_groups, y = damage_preds, fill = island_groups)) +
    geom_boxplot() +
    labs(title = paste("Predicted Damage Distribution - List Entry", i),
         x = "Island Region",
         y = "Predicted Damage") +
    theme_minimal()

 # Save the plot in the list
  plots_list[[i]] <- p

  # Calculate the mean of damage_preds for each island_groups
  mean_values <- merged_data %>%
    group_by(island_groups) %>%
    summarise(mean_damage = mean(damage_preds, na.rm = TRUE))

  # Save the means in the list
  means_list[[i]] <- mean_values

 # Calculate median of the damage_preds for each island groups
  median_values  <- merged_data %>%
    group_by(island_groups) %>%
    summarise(median_damage = median(damage_preds, na.rm = TRUE))

  # save the medians in the list
   median_list[[i]]  <- median_values

}
```

```
# Check the list to confirm plots are stored
print(plots_list)
```

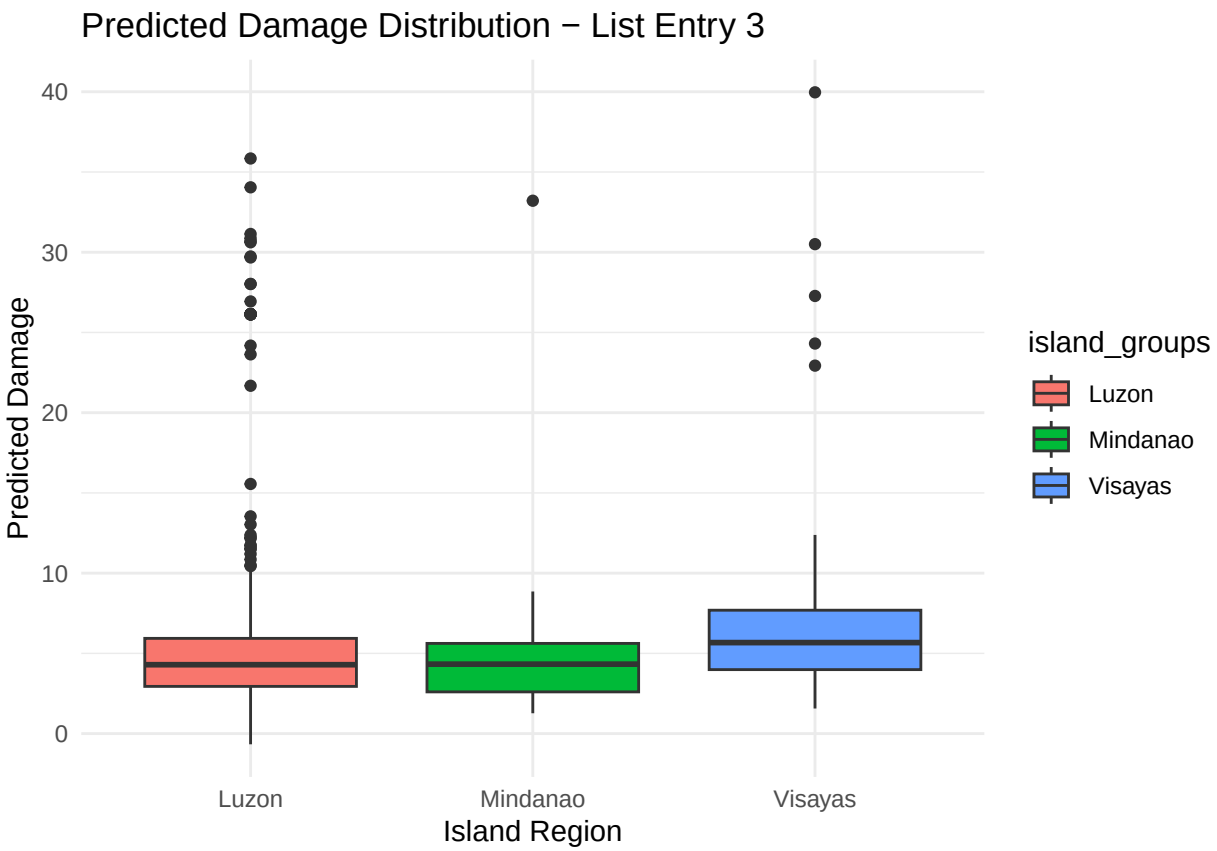
```
## [[1]]
```



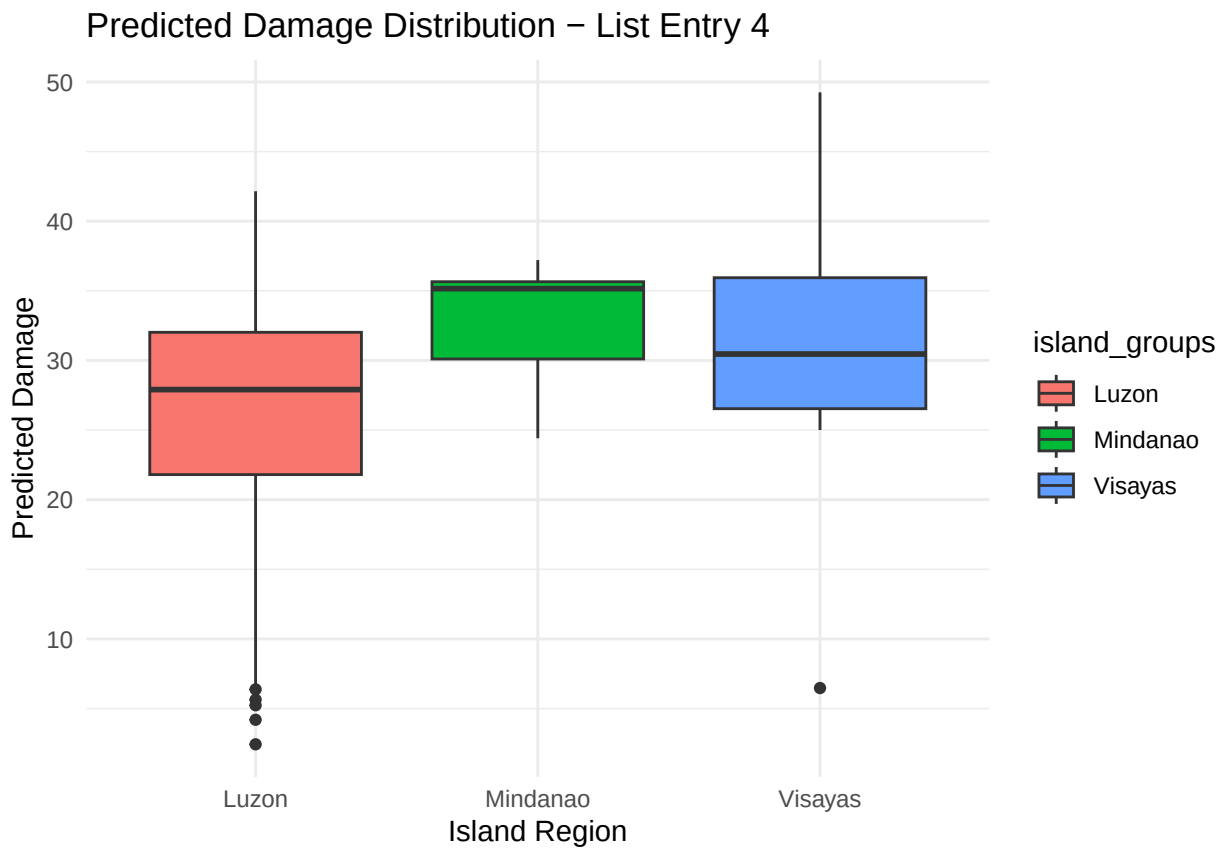

```
##
## [[2]]
```


Predicted Damage Distribution − List Entry 2
Predicted Damage
50
40
30
20
10
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

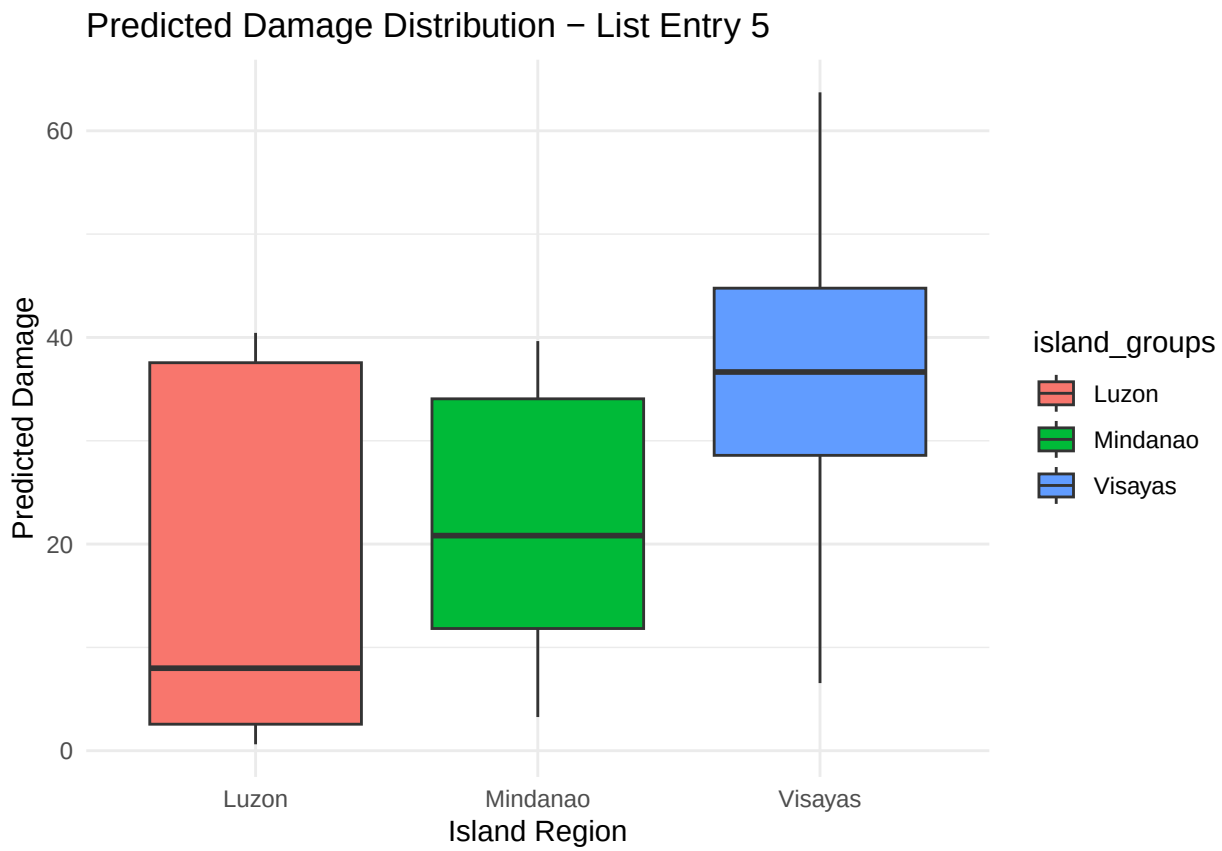
```
##
## [[3]]
```


Predicted Damage Distribution − List Entry 3
Predicted Damage
40
30
20
10
0
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

```
##
## [[4]]
```


Predicted Damage Distribution − List Entry 4
Predicted Damage
50
40
30
20
10
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

```
##
## [[5]]
```




```
print(median_list)
```

```
## [[1]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  6.82
## 2 Mindanao               7.86
## 3 Visayas               17.8
##
## [[2]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  29.2
## 2 Mindanao               29.7
## 3 Visayas                31.3
##
## [[3]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  4.29
## 2 Mindanao               4.33
## 3 Visayas                5.67
```

```
##
## [[4]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  27.9
## 2 Mindanao               35.2
## 3 Visayas                30.5
##
## [[5]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  7.99
## 2 Mindanao              20.8
## 3 Visayas               36.7
```

Output the counterfactual predictions

Saving the counterfactual predictions for mapping differences between this associational model with in QGIS.

```
# prep

CF_output <- melor_2015 %>%
  select(Mun_Code, damage_preds) %>%
  rename(ass_CF_M15 = damage_preds)

write.csv2(CF_output, file = here("associational XGBOOST", "ass_CF_M15.csv"))
```