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", "clustered_M15_CF_data2.csv"))
```

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)
```

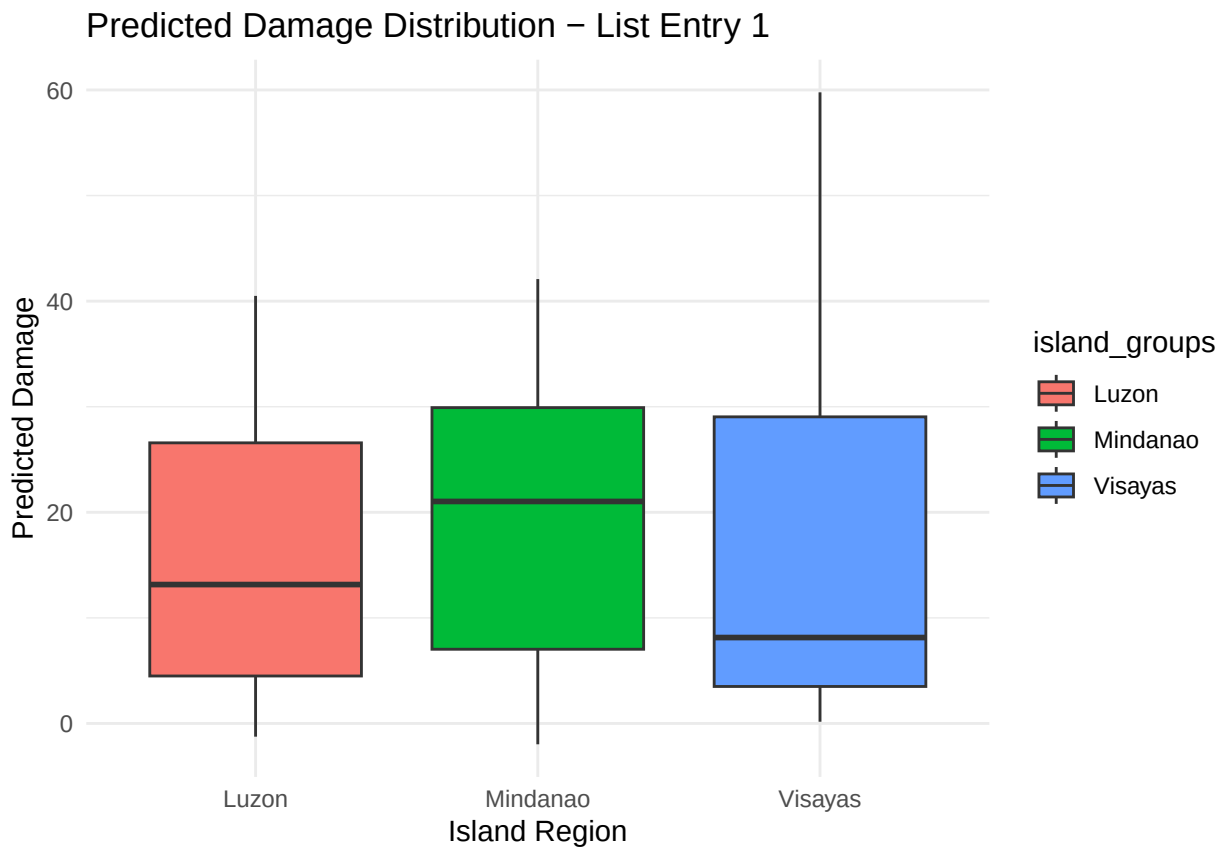
Counterfactual results

```r
# convert the Cluster column to factor
melor_2015$Cluster <- as.factor(melor_2015$Cluster)

# extract cluster_levels
cluster_levels <- levels(melor_2015$Cluster)

# Source function: counterfactual_results
source(here("R", "counterfactual_results.R"))


cf_results <- counterfactual_results(cf_data = melor_2015,
                                     cluster_levels = cluster_levels)
```

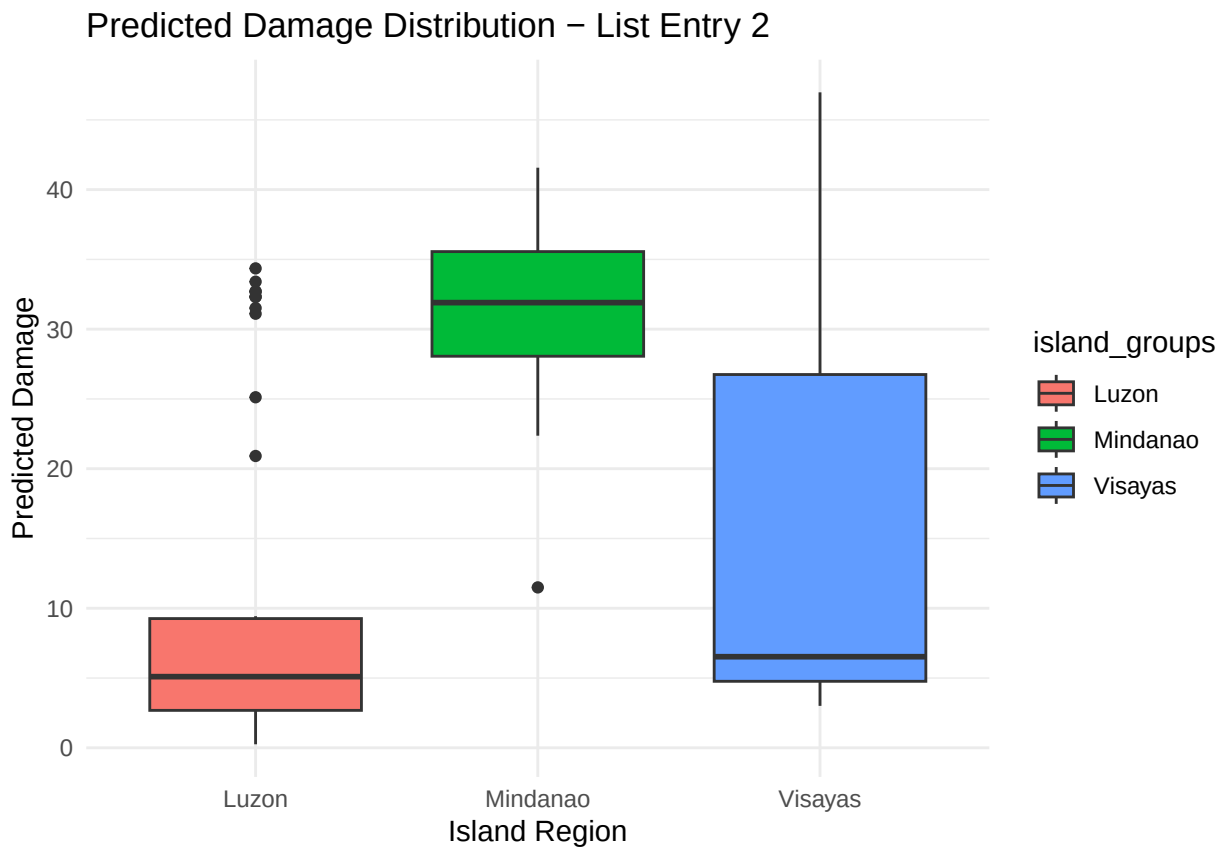
```
## currently evaluating cluster: 1currently evaluating cluster: 2currently evaluating cluster: 3current
```

```r
# Check the list to confirm plots are stored
print(cf_results$plots)
```

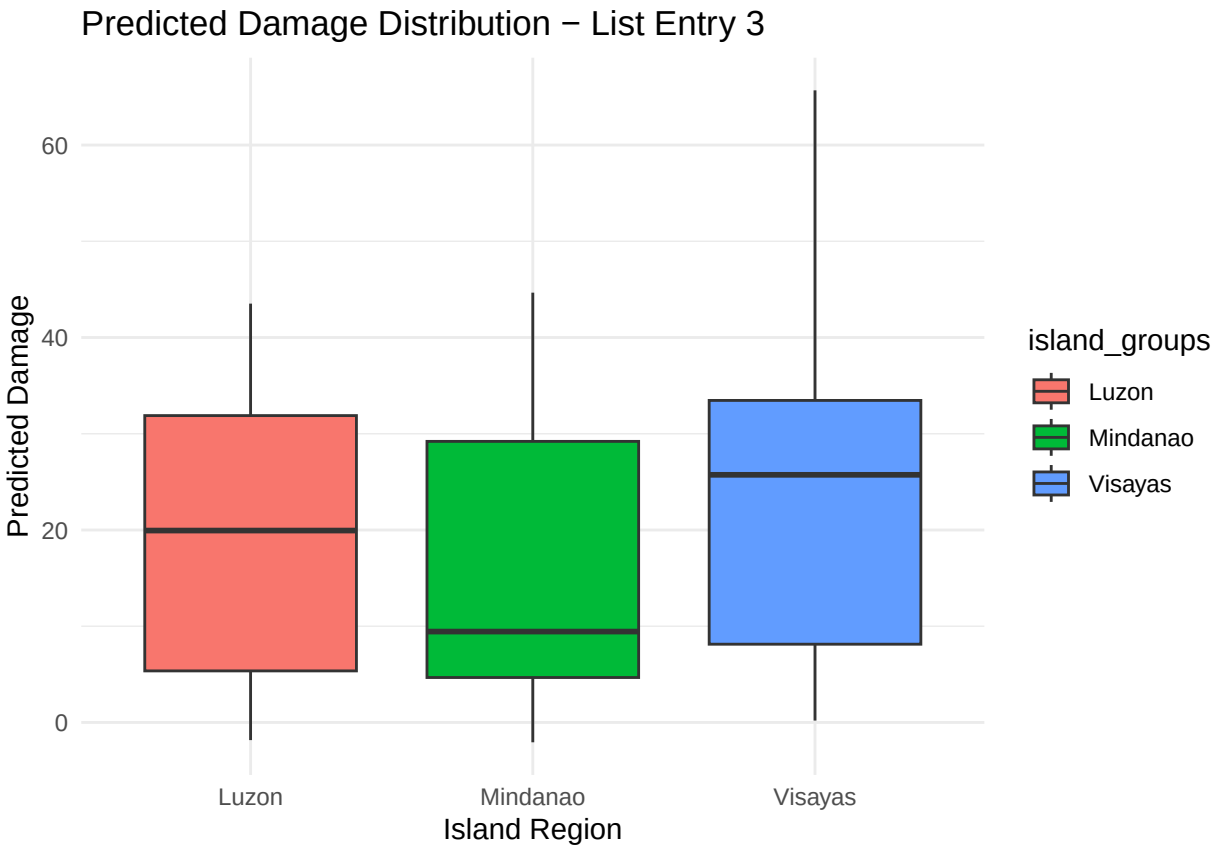
```
## [[1]]
```


Predicted Damage Distribution − List Entry 1
60
40
20
0
Predicted Damage
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

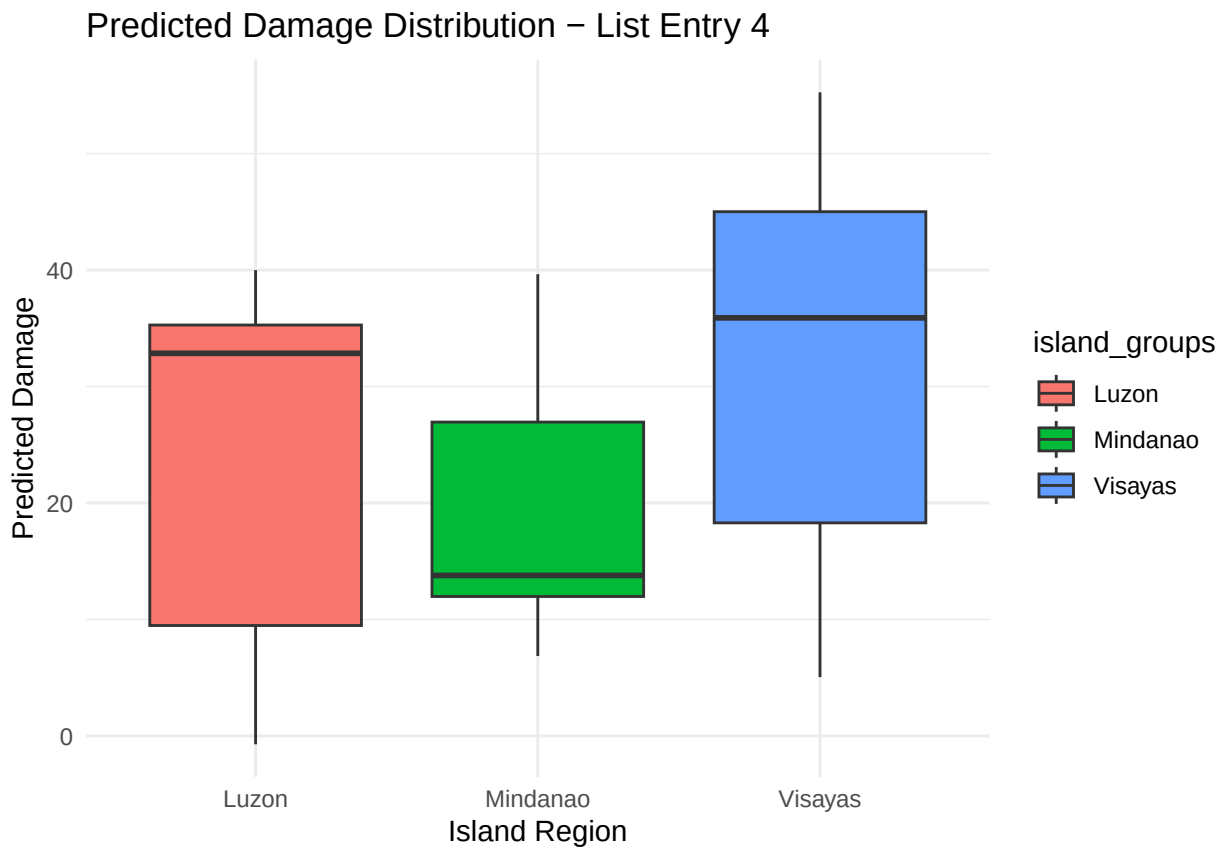
```
##
## [[2]]
```


Predicted Damage Distribution − List Entry 2
Predicted Damage
0
10
20
30
40
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

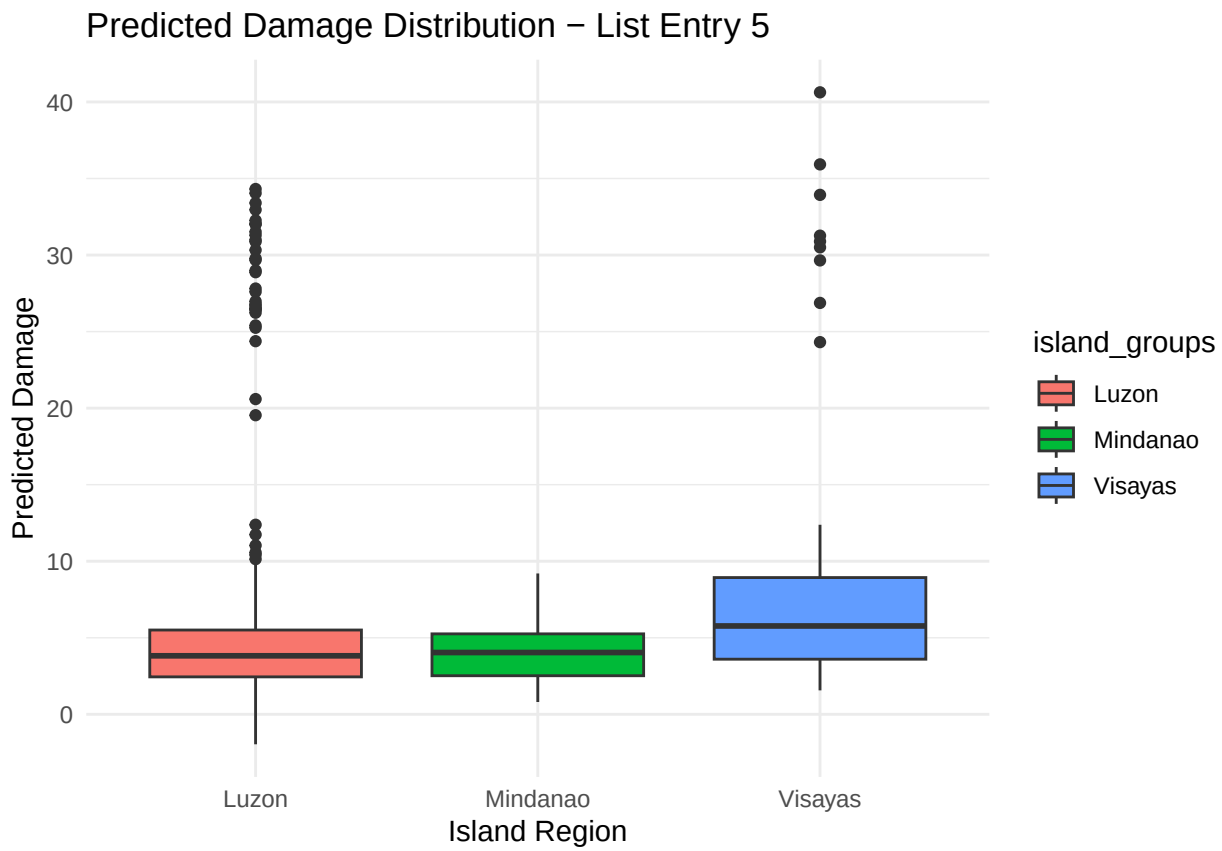
```
##
## [[3]]
```


Predicted Damage Distribution − List Entry 3
Predicted Damage
60
40
20
0
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

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


Predicted Damage Distribution − List Entry 4
Predicted Damage
40
20
0
Luzon
Mindanao
Visayas
Island Region
island_groups
Luzon
Mindanao
Visayas

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




```
print(cf_results$median)
```

```
## [[1]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                 13.2
## 2 Mindanao              21.0
## 3 Visayas                8.14
##
## [[2]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  5.10
## 2 Mindanao              31.9
## 3 Visayas                6.53
##
## [[3]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                 19.9
## 2 Mindanao               9.45
## 3 Visayas               25.7
```

```
##
## [[4]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  32.8
## 2 Mindanao               13.8
## 3 Visayas                35.9
##
## [[5]]
## # A tibble: 3 x 2
##   island_groups median_damage
##   <chr>                 <dbl>
## 1 Luzon                  3.82
## 2 Mindanao               4.04
## 3 Visayas                5.77
```

```
print(cf_results$averages)
```

```
## [[1]]
## # A tibble: 3 x 2
##   island_groups mean_damage
##   <chr>               <dbl>
## 1 Luzon                15.9
## 2 Mindanao             19.4
## 3 Visayas              17.2
##
## [[2]]
## # A tibble: 3 x 2
##   island_groups mean_damage
##   <chr>               <dbl>
## 1 Luzon                10.4
## 2 Mindanao             30.5
## 3 Visayas              18.8
##
## [[3]]
## # A tibble: 3 x 2
##   island_groups mean_damage
##   <chr>               <dbl>
## 1 Luzon                18.8
## 2 Mindanao             15.4
## 3 Visayas              22.4
##
## [[4]]
## # A tibble: 3 x 2
##   island_groups mean_damage
##   <chr>               <dbl>
## 1 Luzon                23.8
## 2 Mindanao             19.3
## 3 Visayas              33.0
##
## [[5]]
## # A tibble: 3 x 2
##   island_groups mean_damage
```

```
##   <chr>               <dbl>
## 1 Luzon                5.45
## 2 Mindanao             4.13
## 3 Visayas             10.2
```

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"))
```

OLD CODE