Management Actions Preliminary Analysis

Jessica French

2023-06-25

Description

This document contains histograms showing the number and frequency of management classifications both overall and or actions that are for climate change. The data from the Google Sheet was read in and combined into one data frame. A column containing the plan each action belonged to was added. The data were then pivoted into long form and filtered to values of 1. This resulted in a data frame where each row represents a positive classification.

Bar charts were created to visualize the total number of categorizations in each category and the proportion of actions that each category applied to. These were further broken down into non-climate change actions and climate change actions.

Set-up

```
# Read in classification data. There are 5 csvs, one for
# each plan create a plan column to make it easier to track
# after combing them in a later step.
sundarbans <- read_csv(here("../../Data/ci_data/Management_Plan_Classification - Sundarbans.csv")) |>
    mutate(plan = "sundarbans")

kisite <- read_csv(here("../../Data/ci_data/Management_Plan_Classification - Kisite-Mpunguti.csv")) |>
    mutate(plan = "kisite")

gladden <- read_csv(here("../../Data/ci_data/Management_Plan_Classification - gladden_spit_silk_cayes.cs
    mutate(plan = "gladden_spit")

lorenzo <- read_csv(here("../../Data/ci_data/Management_Plan_Classification - San_Lorenzo.csv")) |>
    mutate(plan = "san_lorenzo")

heard <- read_csv(here("../../Data/ci_data/Management_Plan_Classification - Hear_McDonalds.csv")) |>
    mutate(plan = "heard_mcdonalds")

# create list of data frames.
actions_vec <- list(sundarbans, kisite, gladden, lorenzo, heard)
# combine csvs using row bind

actions <- bind_rows(actions_vec) |>
    dplyr::select(-text, -summary, -notes)
```

```
# Create long form data so that histograms can be created

columns <- colnames(actions)  # get column names
columns_to_exclude <- c("plan")  # set up for removing plan column
columns_filtered <- setdiff(columns, columns_to_exclude)  # pivot longer.

# pivot actions wider so can use binaries as cateogries
actions_longer <- pivot_longer(actions, cols = columns_filtered,
    names_to = "binaries") |>
    filter(value == 1)
```

Plots

Category plots

These plots show the total and proportion of categorizations across all actions. In other words the number and proportion of times categories were applied.

```
# create histogram to see the counts in each category
# across all plans

# create plot of total counts
ggplot(actions_longer, aes(x = fct_infreq(binaries))) + geom_bar() +
    ylim(0, 450) + labs(x = "", title = "Count of categorizations",
    caption = "Histogram showing the number of times each category was applied. Multiple categories coul
    theme_minimal() + geom_text(stat = "count", aes(label = after_stat(count)),
    vjust = -1) + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1), plot.caption = element_text(hjust = 0))
```
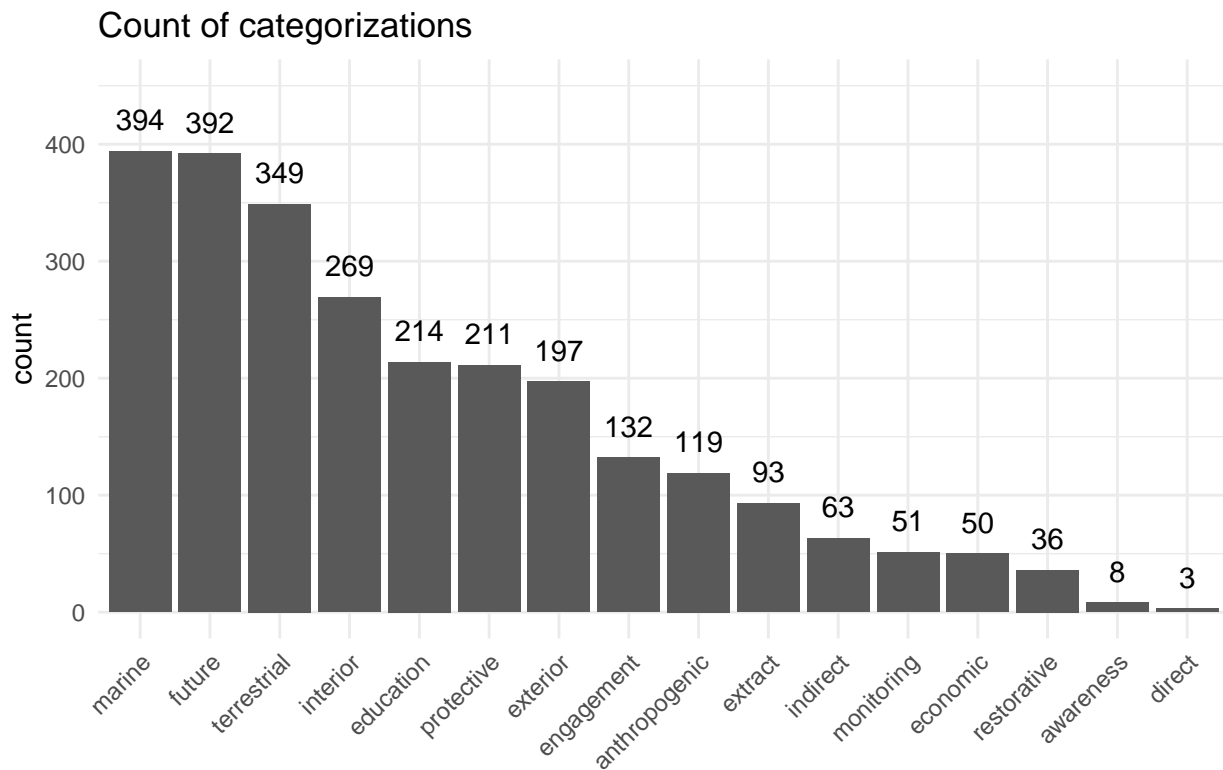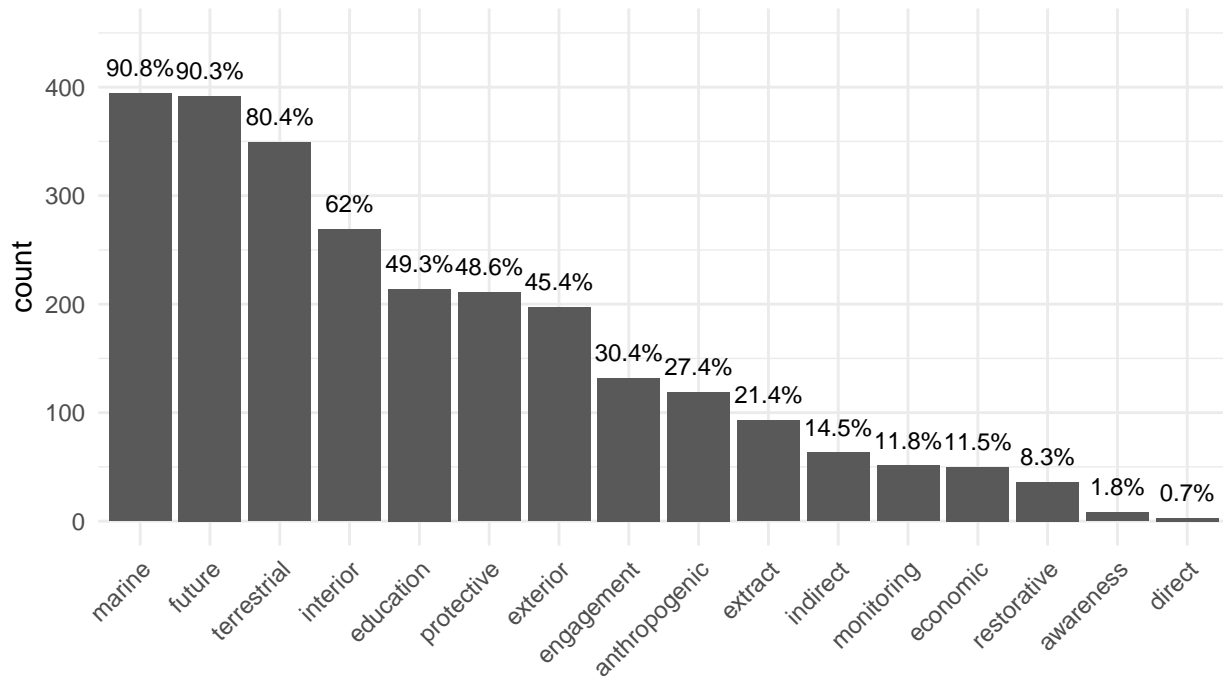
Count of categorizations

Histogram showing the number of times each category was applied. Multiple categories could be apploed to

```r
# alter the above plot to show the proportion
ggplot(actions_longer, aes(x = fct_infreq(binaries))) + geom_bar() +
    ylim(0, 450) + labs(x = "", title = "Proportion of actions in each category",
    subtitle = "Total number of actions is 434", caption = "Histogram showing the proportion of actions
    geom_text(stat = "count", aes(label = paste0(round(after_stat(count)/nrow(actions) *
        100, 1), "%")), vjust = -1, size = 3) + theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        plot.caption = element_text(hjust = 0))
```

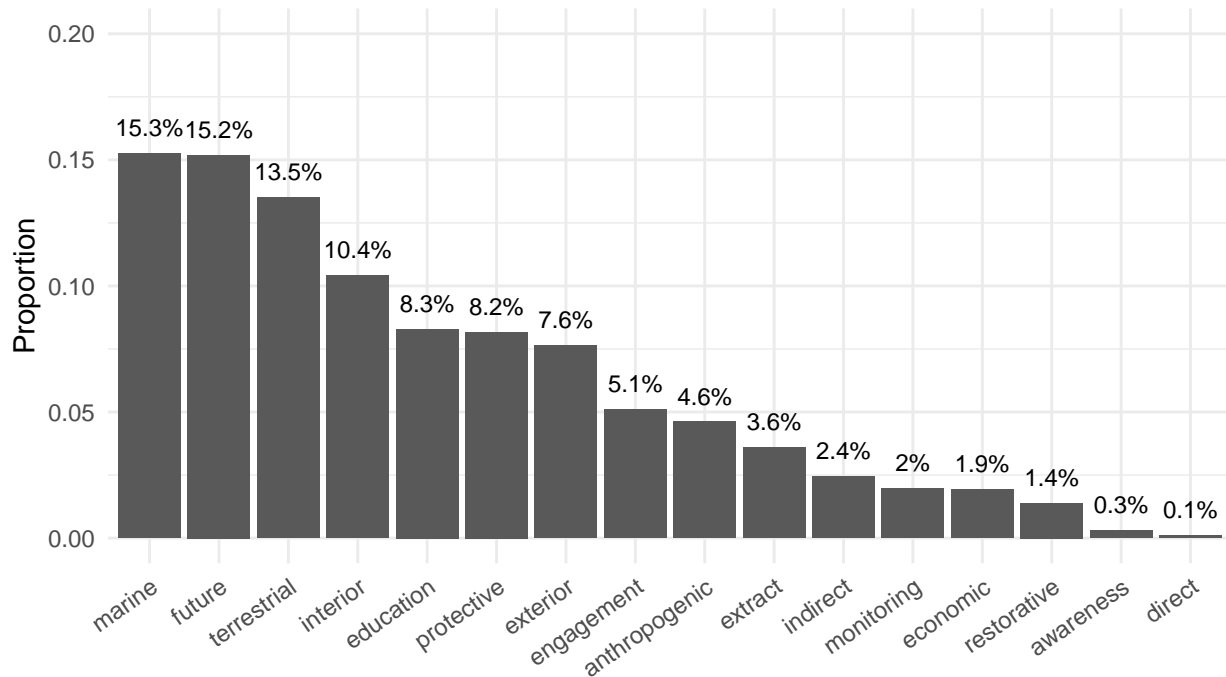# Proportion of actions in each category

Total number of actions is 434

Histogram showing the proportion of actions categorized with each binary.

```
# create plot of proportion of categorizations, uses that
# total number of categorizations rather than the number of
# actions to determine the proportion.
prop_table <- prop.table(table(actions_longer$binaries))
prop_df <- data.frame(binaries = names(prop_table), proportion = prop_table)


ggplot(prop_df, aes(x = reorder(binaries, -proportion.Freq),
    y = proportion.Freq)) + geom_bar(stat = "identity") + ylim(0,
    0.2) + labs(x = "", y = "Proportion", title = "Proportion of classifications in each category",
    subtitle = "Total number of classifications 2581", caption = "Histogram showing the proportion of a
    geom_text(aes(label = paste0(round(proportion.Freq * 100,
        1), "%")), vjust = -1, size = 3) + theme_minimal() +
    theme(axis.text.x = element_text(angle = 35, vjust = 1, hjust = 1),
        plot.caption = element_text(hjust = 0))
```

Proportion of classifications in each category
Total number of classifications 2581

Histogram showing the proportion of all classifications that fall into each category.

```r
# split data into those with climate change specific action
# and those without.

# Filtered to actions that are for CC
actions_cc <- actions |>
    filter(direct == 1 | indirect == 1)

# pivot actions cc longer
actions_cc_long <- pivot_longer(actions_cc, cols = columns_filtered,
    names_to = "binaries") |>
    filter(value == 1 & binaries != "direct" & binaries != "indirect")


# Filtered to actions that were not specifically for CC
actions_no_cc <- actions |>
    filter(direct == 0 & indirect == 0)

actions_no_cc_long <- pivot_longer(actions_no_cc, cols = columns_filtered,
    names_to = "binaries") |>
    filter(value == 1)

# ggplot of category frequency in cc actions
ggplot(actions_cc_long, aes(x = fct_infreq(binaries))) + geom_bar() +
    geom_text(stat = "count", aes(label = after_stat(count)),
        vjust = -1) + ylim(0, 70) + labs(title = "Frequency of Categories in CC Actions",
    x = "") + theme_minimal() + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1))
```
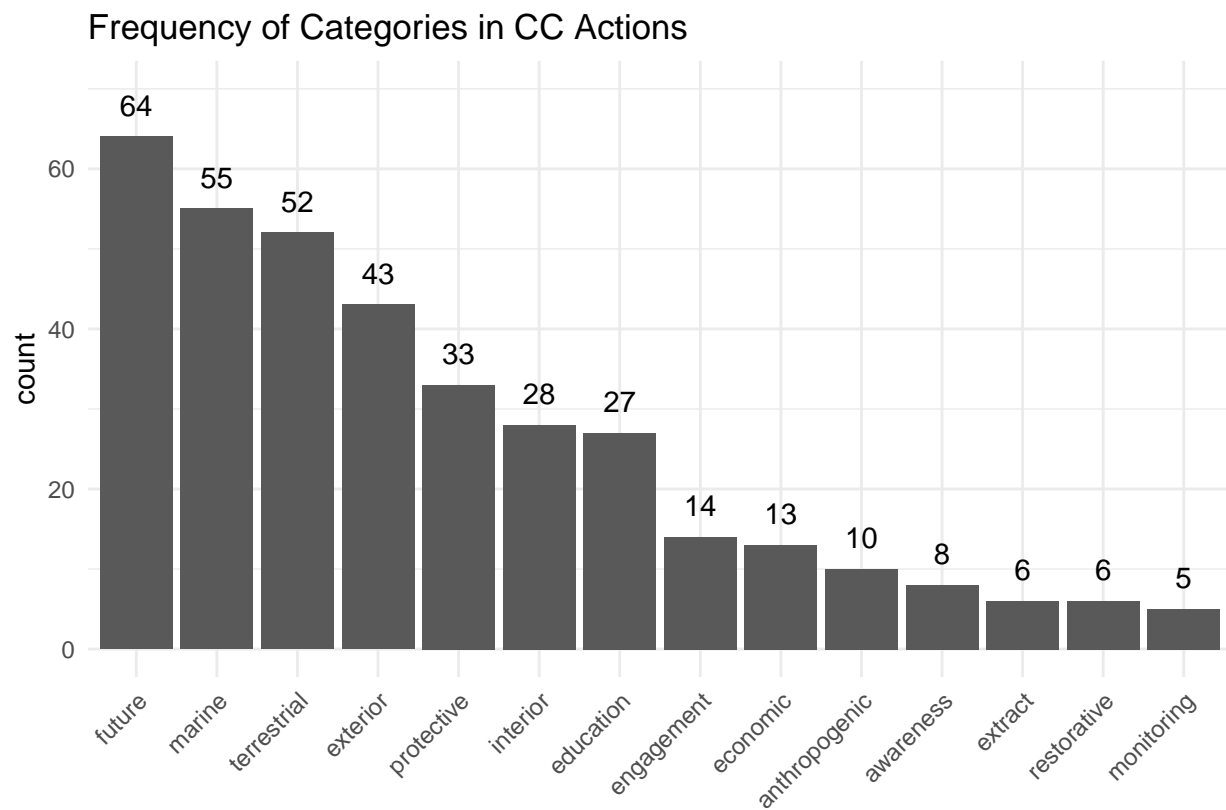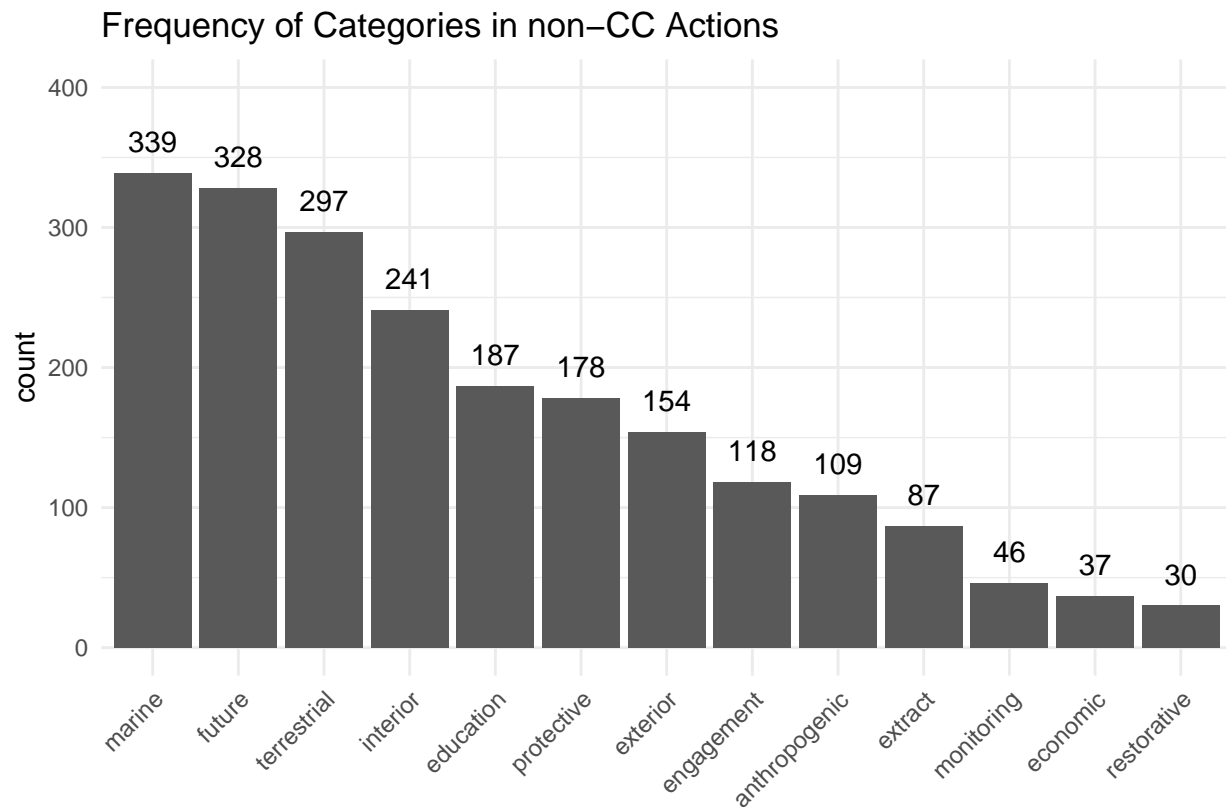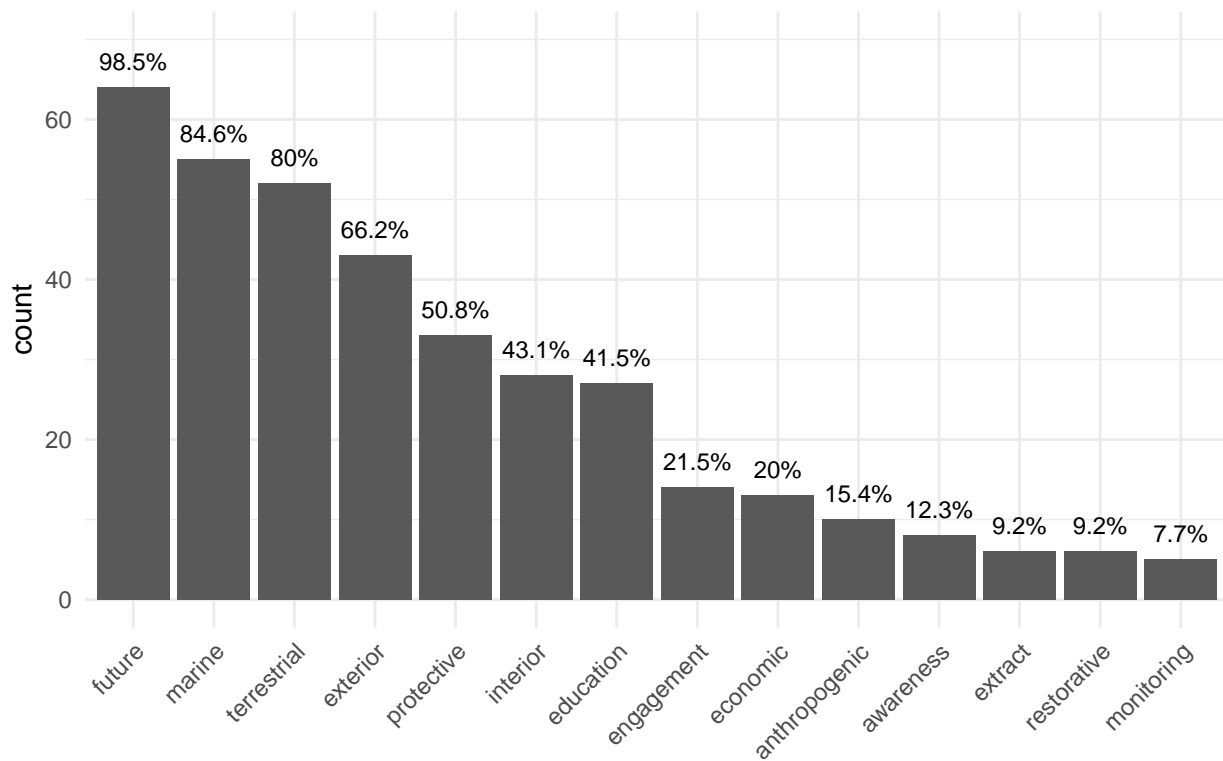
Frequency of Categories in CC Actions


```r
# ggplot of category frequency in non-cc actions.
ggplot(actions_no_cc_long, aes(x = fct_infreq(binaries))) + geom_bar() +
    geom_text(stat = "count", aes(label = after_stat(count)),
        vjust = -1) + ylim(0, 400) + labs(title = "Frequency of Categories in non-CC Actions",
    x = "") + theme_minimal() + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1))
```

## Frequency of Categories in non−CC Actions


```
# ggplot of category frequency in cc actions
ggplot(actions_cc_long, aes(x = fct_infreq(binaries))) + geom_bar() +
    ylim(0, 70) + labs(title = "Frequency of categries in actions for climate change",
    x = "") + geom_text(stat = "count", aes(label = paste0(round(after_stat(count)/nrow(actions_cc) *
    100, 1), "%")), vjust = -1, size = 3) + theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1))
```

Frequency of categries in actions for climate change


```r
# ggplot of category frequency in non-cc actions.
ggplot(actions_no_cc_long, aes(x = fct_infreq(binaries))) + geom_bar() +
    geom_text(stat = "count", aes(label = paste0(round(after_stat(count)/nrow(actions_no_cc) *
        100, 1), "%")), vjust = -1, size = 3) + ylim(0, 375) +
    labs(title = "Frequency of categories in actions not for climate change",
        x = "") + theme_minimal() + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1))
```

Frequency of categories in actions not for climate change
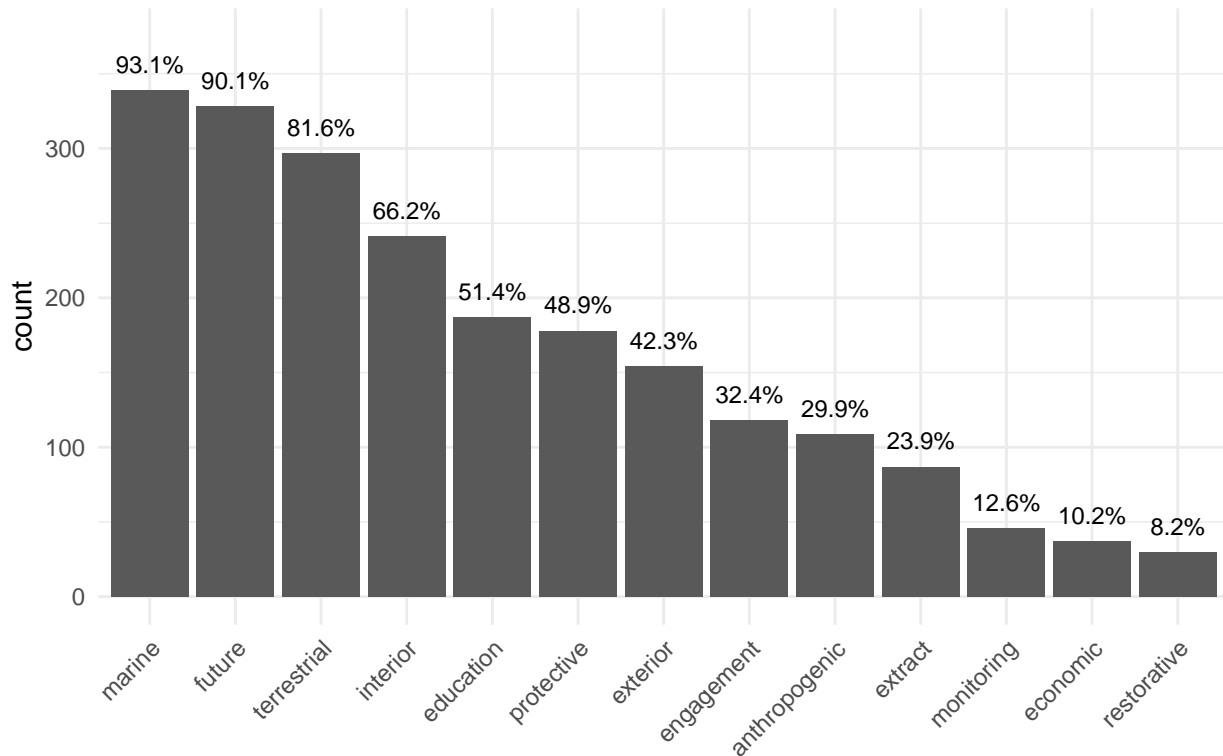
Number of category combinations.

```r
# Want to find all of the unique combinations of
# categorizations to get a sense of how the different
# categories occur together.

# use distinct to find the number of unique combinations.
# Will use this as a check on future operations.
unique_combinations <- actions |>
    dplyr::select(-plan) |>
    distinct()

# count the number of times each combinations occurs
count_combinations <- actions |>
    dplyr::select(-plan) |>
    group_by(across(everything())) |>
    summarise(count = n())
```

```
## `summarise()` has grouped output by 'direct', 'indirect', 'interior',
## 'exterior', 'protective', 'restorative', 'awareness', 'engagement', 'extract',
## 'anthropogenic', 'terrestrial', 'marine', 'economic', 'future', 'education'.
## You can override using the `.groups` argument.
```

There are 201 unique combinations of categorizations. Lets use this as a base to see how the frequency changes with whether or not the action was specifically for climate change.

```r
# count the number of times each combination occurs for
# actions that were specifically for cc.
count_combinations_cc <- actions |>
    dplyr::select(-plan) |>
    filter(direct == 1 | indirect == 1) |>
    select(-direct, -indirect) |>
    group_by(across(everything())) |>
    summarise(count = n())


# see combinations that occur when the action is for cc and
# classified as economic. And count each one.
count_combinations_cc_econ <- actions |>
    dplyr::select(-plan) |>
    filter(direct == 1 | indirect == 1 & economic == 1) |>
    select(-direct, -indirect, -economic) |>
    group_by(across(everything())) |>
    summarise(count = n())

# to plot don't want to collapse the data frame and need to
# pivot longer

count_combinations_cc_econ <- actions |>
    dplyr::select(-plan) |>
    filter(direct == 1 | indirect == 1 & economic == 1) |>
    select(-direct, -indirect, -economic) |>
    pivot_longer(cols = everything(), names_to = "binaries") |>
    filter(value == 1)

# now can make a histogram
ggplot(count_combinations_cc_econ, aes(x = fct_infreq(binaries))) +
    geom_bar() + geom_text(stat = "count", aes(label = after_stat(count)),
    vjust = -1) + ylim(0, 20) + labs(title = "Frequency of Categories in cc and economic",
    x = "") + theme_minimal() + theme(axis.text.x = element_text(angle = 45,
    vjust = 1, hjust = 1))
```

Frequency of Categories in cc and economic