Untitled

Data Carpentry contributors & Montana State University R Workshops Team

10/29/2020

R Markdown

```
surveys_w_days <- surveys %>%
  mutate(date = ymd(paste(year,
                          month,
                          day,
                          sep = "-")
                    ),
         day_of_week = wday(date, label = TRUE)
         ## Creating a day of the week variable
         ## label = TRUE prints the name, not the level!
         )
```

```
surveys_days_full <- surveys_w_days %>%
  mutate(day_of_week = case_when(day_of_week == "Mon" ~ "Monday",
                                 day_of_week == "Tue" ~ "Tuesday",
                                 day_of_week == "Wed" ~ "Wednesday",
                                 day_of_week == "Thu" ~ "Thursday",
                                 day_of_week == "Fri" ~ "Friday",
                                 day_of_week == "Sat" ~ "Saturday",
                                 day_of_week == "Sun" ~ "Sunday")
         )
#glimpse(surveys_days_full$day_of_week)
```

```
surveys_edited <- surveys_days_full %>%
  mutate(day_of_week = fct_relevel(day_of_week,
                                   "Monday",
                                   "Tuesday",
                                   "Wednesday",
                                   "Thursday",
                                   "Friday",
                                   "Saturday",
                                   "Sunday")
  )
glimpse(surveys_edited$day_of_week)
```

```
combined <- surveys_edited %>%
              left_join(plots, by = "plot_id") %>%  # adding the type of plot
              left_join(species, by = "species_id") # adding the genus, species, and taxa
```

```
surveys_gw <- combined %>%
  filter(!is.na(weight)) %>%
  group_by(plot_id, genus) %>%
```

```
  summarize(mean_weight = mean(weight))
```

```
surveys_wide <- surveys_gw %>%
  pivot_wider(names_from = genus, values_from = mean_weight)
```

```
surveys_wide_genera <- combined %>%
  group_by(plot_id, year) %>%
  summarize(num_genera = n_distinct(genus)) %>%
  pivot_wider(names_from = year, values_from = num_genera)
```

```
combined_longer <- combined %>%
  pivot_longer(cols = c(hindfoot_length, weight), names_to = "measurement",
               values_to = "values")
```

```
combined_longer <- combined %>%
  pivot_longer(cols = c(hindfoot_length, weight), names_to = "measurement",
               values_to = "values")

measurement_averages <- combined_longer %>%
  group_by(year, plot_type, measurement) %>%
  summarize(avg_measure = mean(values))

measurement_avg_wide <- measurement_averages %>%
  pivot_wider(names_from = measurement, values_from = avg_measure)
```

```
## Extract the most common species_id
species_counts <- surveys_complete %>%
    count(species_id) %>%
    filter(n >= 50)

## Only keep the most common species
surveys_complete_subset <- surveys_complete %>%
  filter(species_id %in% species_counts$species_id)
  ## using the relational operator %in%
```

```
surveys_complete <- surveys_edited %>%
  filter(!is.na(weight),           # remove missing weight
         !is.na(hindfoot_length),  # remove missing hindfoot_length
         !is.na(sex))                # remove missing sex
```