Andrew Martinez

10/9/2022

Stat 331

Week 6 Reflection

Table Of Contents

Reflection Question 1:

Week 6 Reflection:

1) What learning targets have you demonstrated an understanding of?

Working with Data

- ❑ **WD-1:** I can import data from a variety of formats (e.g., csv, xlsx, txt, etc.).
- ❑ **WD-2:** I can select necessary columns from a dataset.
- ❑ **WD-3:** I can filter rows from a dataframe for a variety of data types (e.g., numeric, integer, character, factor, date).
- ❑ **WD-4:** I can modify existing variables and create new variables in a dataframe for a variety of data types (e.g., numeric, integer, character, factor, date).
- ❑ **WD-5:** I can use mutating joins to combine multiple dataframes.
- ❑ **WD-6:** I can use filtering joins to filter rows from a dataframe.

WD-1:

Lab 4: Loading In Datasets

```r
{r}
#| message: FALSE

library(tidyverse)
library(here)
library(readxl)
avocado <- read_csv(here::here("supporting_artifacts", "Lab & Challenge 4",
                               "avocado.csv"))
housing_data <- read_xlsx(here::here("supporting_artifacts", "Lab & Challenge 4",
                                     "Housing DataSet.xlsx"))
```

WD-2:

Challenge 3: Part A

Part A

```r
{r}
gender_hiphop <- hiphop |>
  group_by(sex) |>
  distinct(subj, .keep_all = TRUE) |>
  summarize(across(intl:unclassifiable, mean)) |>
  summarize(across(intl:unclassifiable, diff))


gender_hiphop
```

WD-3:

Lab 3: Finding Justin Bieber

Finding Justin Bieber

```r
justin_bieber_data <- hiphop |>
  distinct(subj, .keep_all = TRUE) |>
  filter(sex == "Male", age >= 17, age <= 23, ethnic == "white", city >= 10000,
         city <= 60000) |>
  slice_max(bieber)
justin_bieber_data$subj
```

WD-4:.

Lab 3: Re-categorizing The Ethnic Variable

Re-categorizing The Ethnic Variable

```r
clean_hiphop <- clean_hiphop |>
  mutate(ethnic_categorical = ifelse(ethnic == "white",
                                     "white",
                                     "non-white")) |>
  relocate(ethnic_categorical, .after = ethnic)
```

Lab 3: Familiar Words

Familiar Words

People Under 20

```r
young_clean_hiphop_max <- clean_hiphop |>
  filter(age < 20) |>
  select(word, familiarity) |>
  group_by(word) |>
  summarise(mean = mean(familiarity), n = n()) |>
  slice_max(mean)
```

Lab 5: Time Series Plot 2

Time Series Plot 2

```r
time_series_surveys <- surveys |>
  separate(col=date, into=c('Year', 'Month', "Day"), sep='-') |>
  group_by(year, genus) |>
  mutate(avg_weight = mean(weight))
```

WD-5:
Mutating Join
Challenge 5: Altering / Joining Dataset

Altering / Joining Dataset

```{r}
housing_data <- housing_data |>
  group_by(region) |>
  mutate( mean_housing_price = mean(`Housing Price`))

avocado_housing_data <- avocado |>
  inner_join(housing_data, by = "region") |>
  group_by(region, type) |>
  mutate(avg_avg_price = mean(AveragePrice)) |>
  rename(small = `4046`, large = `4225`, extralarge = `4770`)
```

WD-6:
Filtering Join
Lab 4: Identifying Important Regions

Identifying Important Regions

```{r}
library(dplyr)

avocado_regions_dataframe <- data.frame(
  region = c("Midsouth", "Northeast", "Plains", "Southeast", "West", "SouthCentral"
, "GreatLakes" )
)

avocado_clean_regions <- avocado |>
  semi_join(y = avocado_regions_dataframe, by = "region") |>
  separate(col=Date, into=c('Year', 'Month', "Day"), sep='-') |>
  rename(small = `4046`, large = `4225`, extralarge = `4770`)

avocado_clean_cities <- avocado |>
  anti_join(y = avocado_clean_regions, by = "region") |>
  separate(col=Date, into=c('Year', 'Month', "Day"), sep='-') |>
  rename(small = `4046`, large = `4225`, extralarge = `4770`)
```

Reproducibility

- ❑ **R-1:** I can create a professional looking, reproducible analysis using RStudio projects, Quarto documents, and the here package.
- ❑ **R-2:** I can write well documented and tidy code.
- ❑ **R-3:** I can write robust programs that are resistant to changes in inputs.

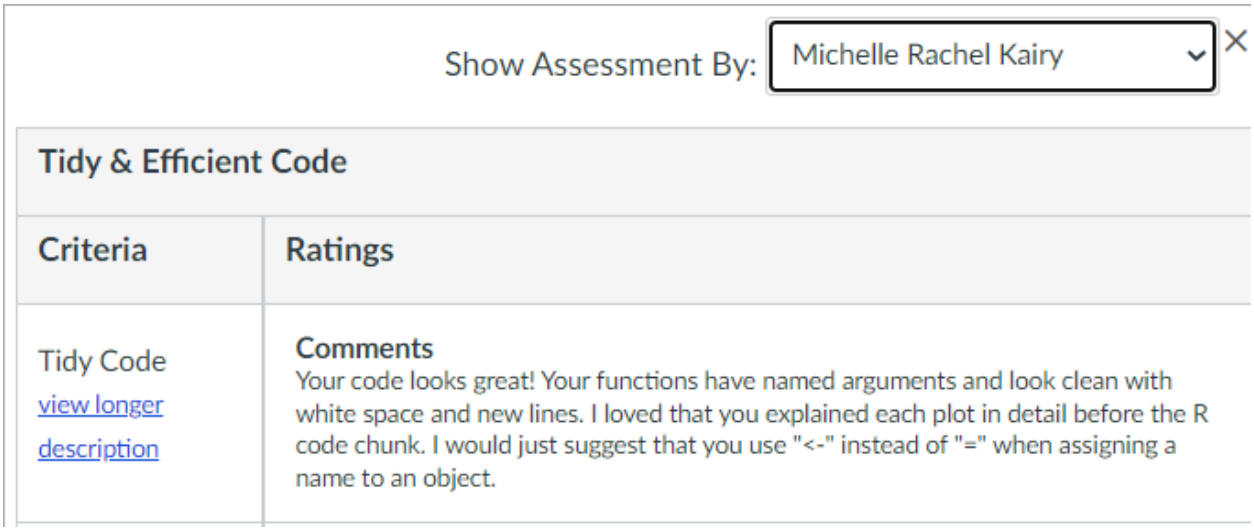
R-1:
Here:
See WD-1
Professional Looking:
Please See Formatting Quarto Document For Challenges 1-4

R-2
Lab 2: Question 10

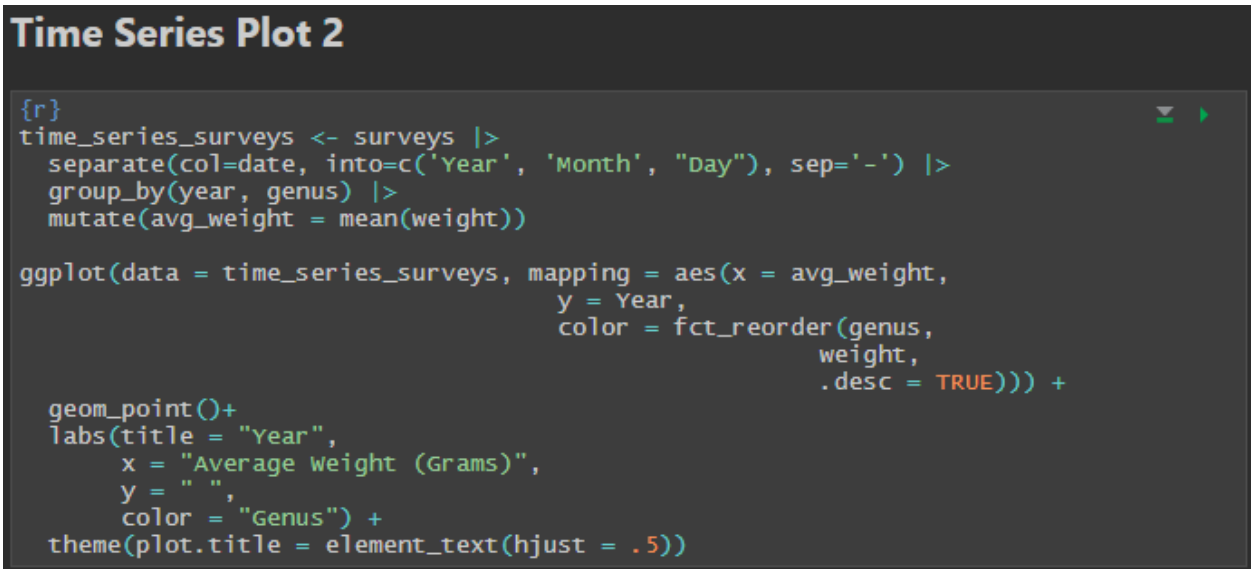
```{r}
# Code for question 10!
ggplot(data = surveys, mapping = aes( x = weight,
                                      y = species)) +
  geom_boxplot(outlier.shape = NA) +
  geom_jitter(color = "tomato", alpha = 1/50) +
  labs(title = "Distribution Of Weight (In Grams) By Species",
       x = "Weight (Grams)",
       y = "Species") +
  theme(plot.title = element_text(hjust = .5))
```

R-3

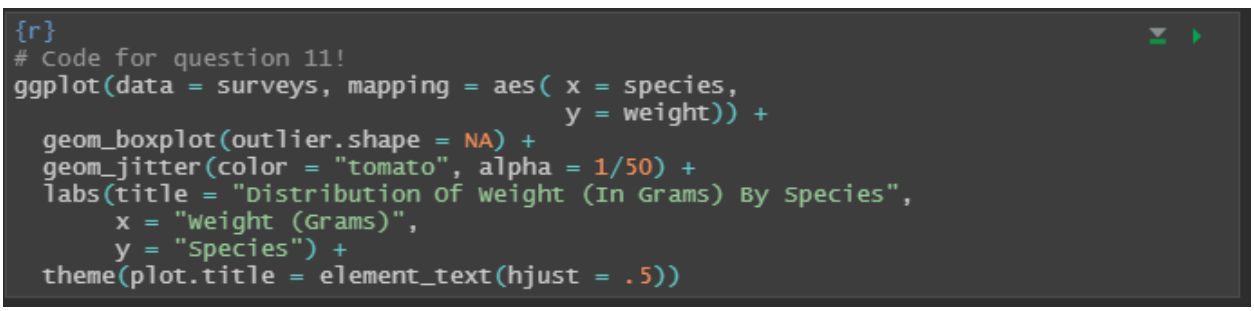
Lab 1 Peer Review



Lab 5: Time Series Plot 2



Lab 2: Question 11



Challenge 2: Challenge 1 Code

Challenge Code 1

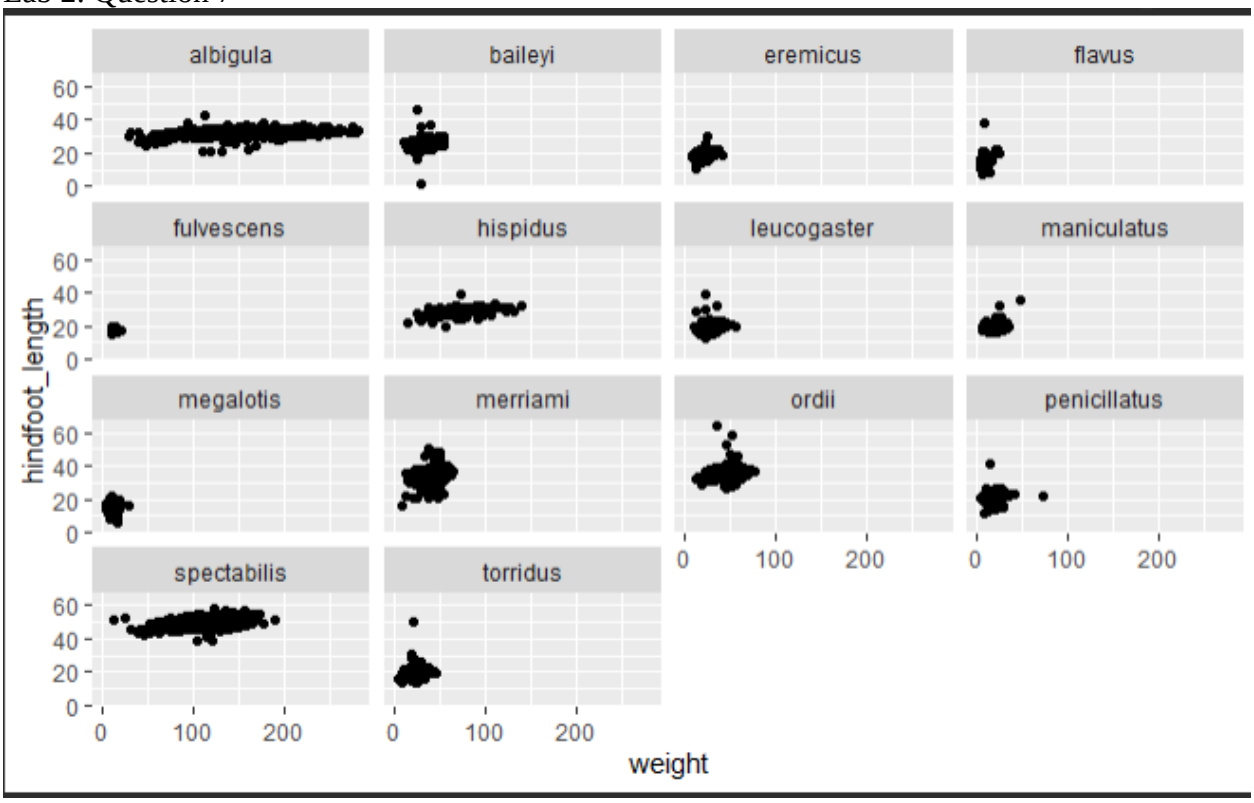
```{r}
ggplot(data = surveys, mapping = aes( x = weight,
                                      y = species,
                                      color = genus)) +
  geom_density_ridges() +
  geom_jitter(color = "tomato", alpha = 1/50) +
  labs(title = "Distribution Of Weight (In Grams) By Species",
       x = "Weight (Grams)",
       y = "Species") +
  theme(plot.title = element_text(hjust = .5))
```

Data Visualization & Summarization

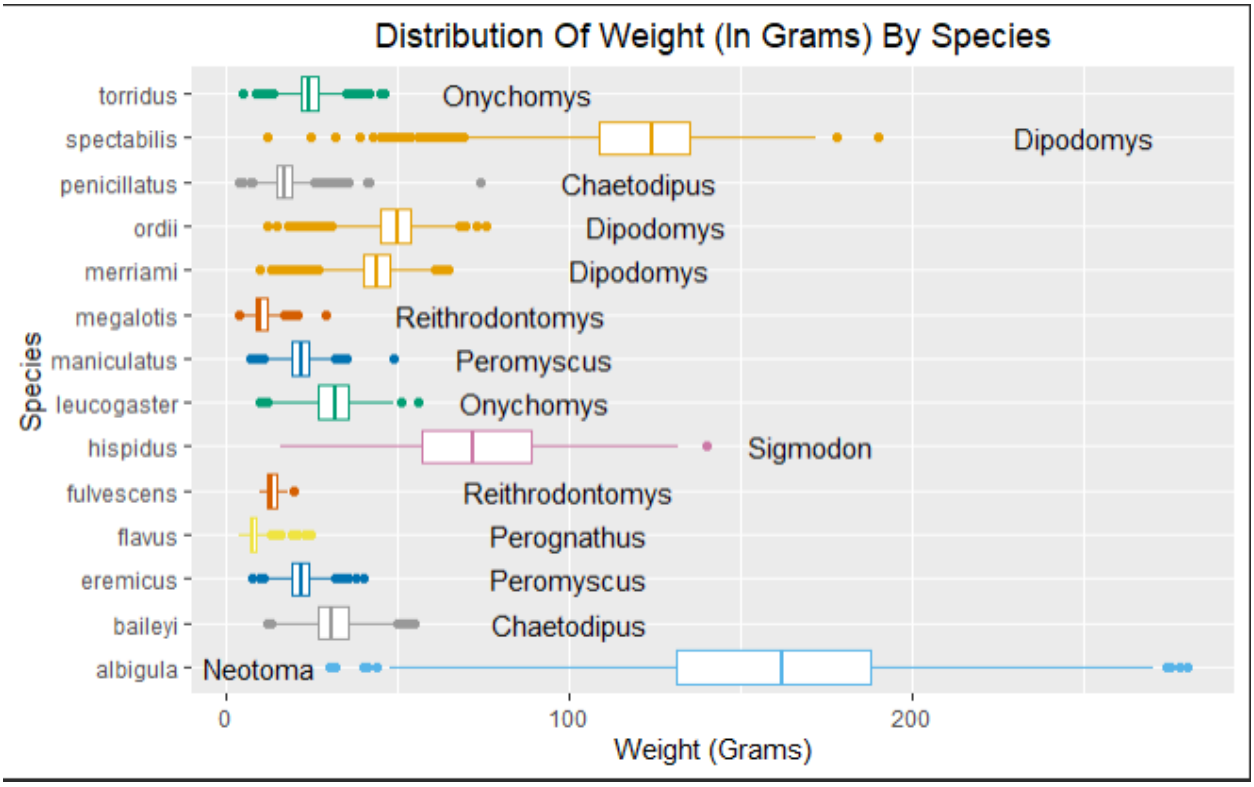
- ❑ **DVS-1:** I can create visualizations for a variety of variable types.
- ❑ **DVS-2:** I use plot modifications to make my visualization clear to the reader.

DVS-1:

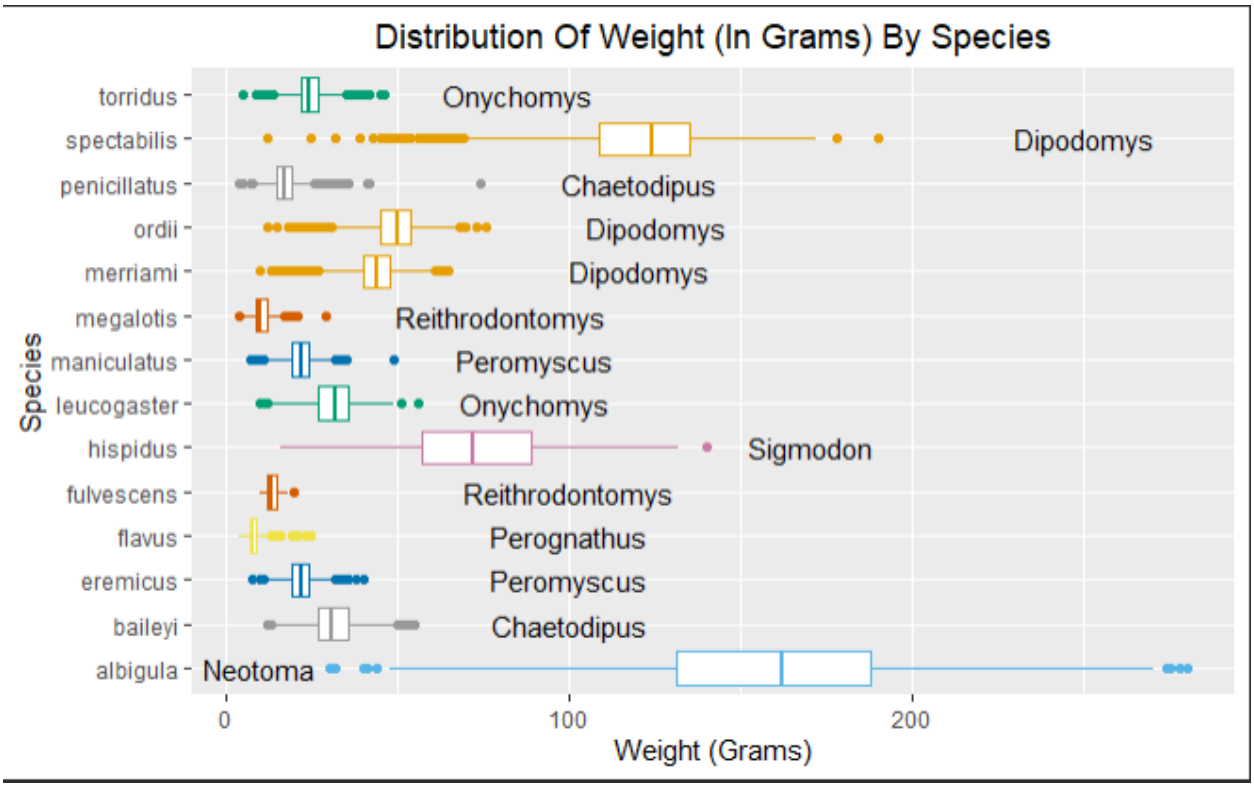
Lab 2: Question 7



Challenge 2: Challenge 3 Code



DVS-2:
Challenge 2: Challenge 3 Code

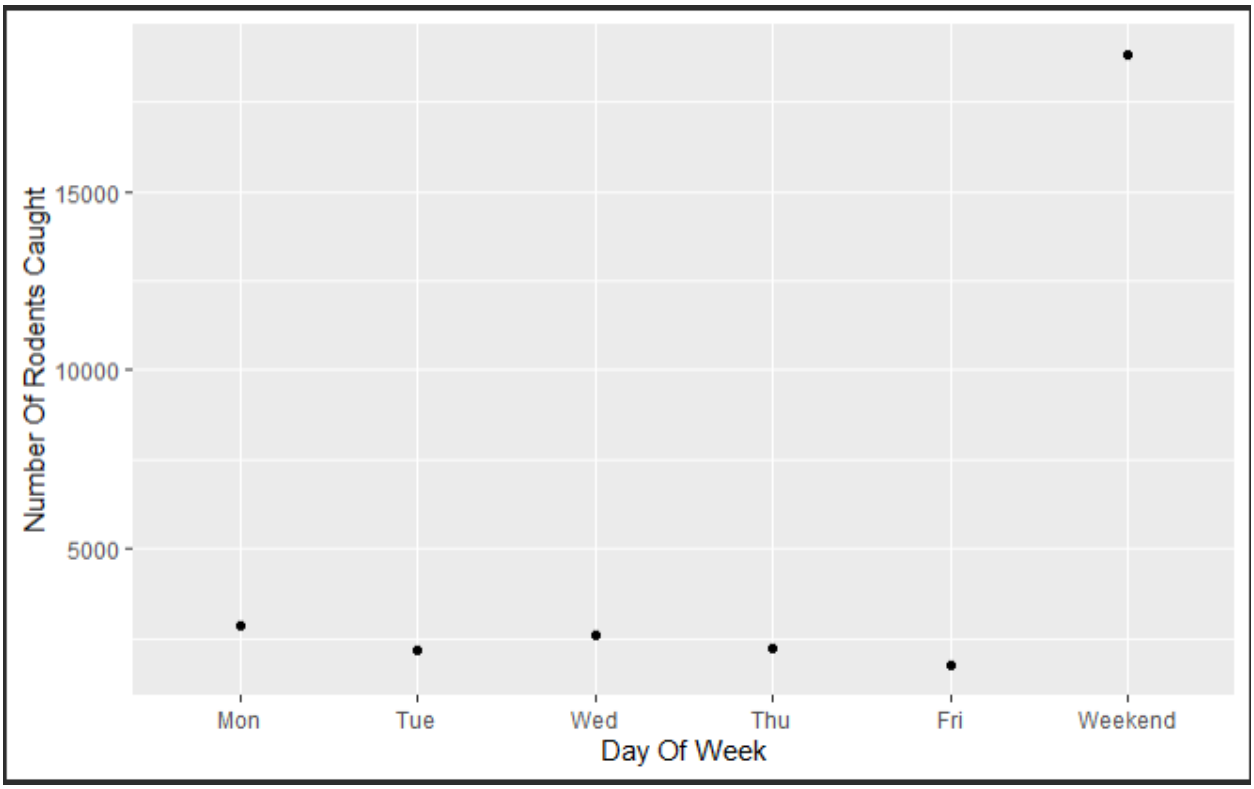


Lab 5: Time Series #3 Plot

```{r}

time_series_surveys_by_week_count <- surveys |>
  count(day_of_week) |>
  mutate(day_of_week =
           fct_collapse(day_of_week,
             Weekend = c("Sat", "Sun"))) |>
  group_by(day_of_week) |>
  summarise(day_of_week, sum(n)) |>
  distinct(day_of_week, .keep_all = TRUE) |>
  filter(day_of_week != "NA")

ggplot(data = time_series_surveys_by_week_count, mapping = aes(
  fct_relevel(day_of_week, c("Mon","Tue","Wed","Thu","Fri","Weekend"))
  , `sum(n)`))+
  geom_point()+
  labs(y = "Number Of Rodents Caught",
       x = "Day Of Week")
```



2) How have you demonstrated a commitment to continued learning?

On all of my challenges, 1, 2, 3, and 4, I have gone above and beyond the outlined challenge tasks. For example, for challenge 2 I completed all three of the challenge options instead of just settling for one. Additionally, for challenge 3 I went online to research how to merge separate datasets in an attempt to plot the intl:unclassifiable data rather than just typing out an explanation. While I was unsuccessful in the plotting portion of my prior statement, I do plan on coming into office hours to try and figure that portion out. I believe that my commitment to go above and beyond the required portion of the challenges demonstrates a commitment to continued learning. Additionally, you personally have reinforced my belief in this opinion as your comments on my challenges have suggested that they are more than “satisfactory,” and that they display continued learning.

3) How have you grown as a team member and leader?

Since the beginning of the quarter, I believe that I have grown as both a team member and leader. As you can see in class, as we sit right next to you, I am more than willing to assist my team members, and often will check in on them and their progress to ensure that everyone is on the same page. With the group project coming up soon, I look to take a leadership role in the group, to ensure that no one is left behind, and ensure that every group member has a part and understands the group project. While the quarter is on the tail end, I believe the group project is a great opportunity for me to show my leadership capabilities.

4) How have you contributed to creating a respectful classroom learning community?

As of this reflection, I still have had no experiences of a disrespectful classroom community, nor would I expect anything to arise in the future. While I have not had the opportunity to prevent a disrespectful environment from occurring, I have continued to be a quality group member, which perpetuates a positive learning community. I have not been as active as before on the discord, and I believe this reflection serves as a great reminder for me to not only assist my classmates on discord, but also ask my questions on there as well. Although I have not been as active on discord, I have continued to create a respectful classroom learning community within my group.

5) Are you accomplishing the goals you outlined for yourself at the start of the course?

I do not have a traditional goal-setting system. I believe the best path to success is to fall in love with the journey, rather than fixate on a goal or "top of the mountain moment." My thought process is that if you have an ultimate goal, such as gaining proficiency in R, at the end of the quarter you will lose motivation to have "continued learning" as you will have already reached the top of the mountain. As a result, I think my biggest goal this quarter is to learn to enjoy the coding process, and continue to apply my learning in the future. With this goal in mind, I do believe I have done that so far, as I have enjoyed the class content and assignments thoroughly thus far.

Hypothesized Grade: A

Explanation:

I believe that of the five questions you posed of us, I have had superior attainment of all of them. Since the week three reflection, I have continued obtaining superior attainment of the learning objectives and course goals. After reviewing the red annotations you wrote within the syllabus, I also feel that I coincide with those remarks and goals towards an A as well. I look to finish out the quarter strong, to ensure that I maintain an A level of completion in this class!