Mini Project III - Tial and Jessica

```
#load libraries for extracting inat data
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   4.0.0     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```
library(rinat)
```

```
#extract raw data of all observations
raw_obs <- get_inat_obs_project("st-olaf-natural-lands", type = "observations")
```

```
9483 records
```

```
Getting records 0-200
```

```
Getting records up to 400
```

```
Getting records up to 600
```

```
Getting records up to 800
```

```
Getting records up to 1000

Getting records up to 1200

Getting records up to 1400

Getting records up to 1600

Getting records up to 1800

Getting records up to 2000

Getting records up to 2200

Getting records up to 2400

Getting records up to 2600

Getting records up to 2800

Getting records up to 3000

Getting records up to 3200

Getting records up to 3400

Getting records up to 3600

Getting records up to 3800

Getting records up to 4000

Getting records up to 4200

Getting records up to 4400

Getting records up to 4600
```

```
Getting records up to 4800

Getting records up to 5000

Getting records up to 5200

Getting records up to 5400

Getting records up to 5600

Getting records up to 5800

Getting records up to 6000

Getting records up to 6200

Getting records up to 6400

Getting records up to 6600

Getting records up to 6800

Getting records up to 7000

Getting records up to 7200

Getting records up to 7400

Getting records up to 7600

Getting records up to 7800

Getting records up to 8000

Getting records up to 8200

Getting records up to 8400
```

```
Getting records up to 8600

Getting records up to 8800

Getting records up to 9000

Getting records up to 9200

Getting records up to 9400

Getting records up to 9600

Done.
```

```
#Created a new data frame called "obs_clean" from the previous data frame "raw_obs" which co
obs_clean <- raw_obs |>

#Using the select() function, we selected specific columns we want to include in our new data
   select(observed_on, description, latitude, longitude, species_guess, user_id, created_at,

#We used the mutate() function to modify specific columns.
#The ymd() function converts date columns to follow this format:  year-month-day
#For the "created_at" and "updated_at" columns, we used the str_extract() function to isolate
#The parse_number() function extracts numeric values from "latitude" and "longitude" columns
#Finally, the replace_na() function for the "species_guess" column to replace any "NAs" with

  mutate(observed_on = ymd(observed_on),
         latitude = parse_number(latitude),
         longitude = parse_number(longitude),
         created_at = ymd(str_extract(created_at, "\\d{4}-\\d{2}-\\d{2}")),
         updated_at = ymd(str_extract(updated_at, "\\d{4}-\\d{2}-\\d{2}")),
         species_guess = replace_na(species_guess, "unknown"))|>

#Here, we used the rename() function to assign new names to specific columns.
          rename (species = species_guess,
          location = place_guess,
          quality = quality_grade,
          URL = uri)
```

```
write_csv(obs_clean, "TJ_mini3")
```

Variable Description Table :

“observed_on”	“description”	“latitude”	“Tomato”
“Orange”	26.6%	rng	138521