Project Proposal

due October 15, 2021 by 11:59 PM

Thomas Barker, Phoebe Ducote, Ebenezer Teshome, Alexander Du

October 8 2021

Load Packages

```
library(tidyverse)
```

Load Data

```
ceodepartures <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/
nyc_squirrels <- readr::read_csv("https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/
athletes <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/
```

Introduction and Data

CEO Departures https://github.com/rfordatascience/tidytuesday/blob/master/data/2021/2021-04-27/readme.md -This dataset came from TidyTuesday, and the research was conducted by Richard J. Gentry, Joseph S. Harrison,Timothy J. Quigley, and Steven Boivie. In the research paper they claim to have researched CEO departures from 2000-2018, but the data set has entries beginning in 1987 and ending in 2020, so the original data (which they got from Execucomp, a Wharton database which collects data on top CEOs and companies), must have begun being collected in 1987. Each case is a single CEO departure/CEO. Relevant variables include the company name, the CEOs name, a code number for the reason of departure, the year of the event which caused the departure, the year and date they actually departed, and how many times they served as CEO.

NYC Squirrel Census: -This dataset came from TidyTuesday, by way of the Squirrel Census. The Squirrel Census is an organization that tracks Eastern gray squirrels, and used volunteers in (for this data set) New York City to count and observe squirrels. This data was originally collected in October of 2018. Each case represents a single squirrel. Some of the variables are the unique squirrel id, “hectare” (which describes the location on a grid map of Central Park), the date they were observed, and serveral variables like “running”, “climbing”, and “eating” to describe the squirrell’s actions.

Paralympics https://github.com/rfordatascience/tidytuesday/blob/master/data/2021/2021-08-03/readme.md -The data came from TidyTuesday, which was sourced from the International Paralympic Committee. It was webscraped and details data from 1980 - 2016. Each datapoint represents a given athlete in the Paralympics. Relevant variables include the gender, name, and type of medal that an athlete received in a specific event. If the athlete was vision-impaired, the guides and pilots were noted as well. And of course, each athlete was also accompanied by their country, its abbreviation, and the year of the Paralympics. It was originally collected to be utilized for an article published around August 2021.

Research Questions

-Does the reason for CEO departure differ significantly among different S&P 1500 companies and does it differ significantly by year of departure as well?

-Does the proportion of squirrels seen eating differ significantly among juvenile and adult squirrels?

-In any given year, how does the type of medal countries receive compare across all events?

Glimpse

Please use `glimpse` for your datasets here.

```
glimpse(ceodepartures)
```

```
## Rows: 9,423
## Columns: 19
## $ dismissal_dataset_id <dbl> 559043, 12, 13, 31, 43, 51, 61, 63, 62, 65, 75, 7~
## $ coname               <chr> "SONICBLUE INC", "AMERICAN AIRLINES GROUP INC", "~
## $ gvkey                <dbl> 27903, 1045, 1045, 1078, 1161, 1177, 1194, 1194, ~
## $ fyear                <dbl> 2002, 1997, 2002, 1998, 2001, 1997, 1993, 1997, 1~
## $ co_per_rol           <dbl> -1, 1, 3, 6, 11, 16, 21, 22, 24, 28, 33, 34, 38, ~
## $ exec_fullname        <chr> "L. Gregory Ballard", "Robert L. Crandall", "Dona~
## $ departure_code       <dbl> 7, 5, 3, 5, 5, 5, 5, 7, 9, 5, 5, 5, 3, 5, 5, 3, 3~
## $ ceo_dismissal        <dbl> 0, 0, 1, 0, 0, 0, 0, 0, NA, 0, 0, 0, 1, 0, 0, 1, ~
## $ interim_coceo        <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ tenure_no_ceodb      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ max_tenure_ceodb     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ fyear_gone           <dbl> 2003, 1998, 2003, 1998, 2002, 1997, 1993, 1998, 1~
## $ leftofc              <dttm> 2003-03-21, 1998-05-20, 2003-04-24, 1998-12-31, ~
## $ still_there          <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ notes                <chr> "Ballard took over when the outgoing CEO said tha~
## $ sources              <chr> "https://www.wsj.com/articles/SB10288576921909334~
## $ eight_ks             <chr> "https://www.sec.gov/Archives/edgar/data/850519/0~
## $ cik                  <dbl> 850519, 6201, 6201, 1800, 2488, 1122304, 771667, ~
## $ `_merge`             <chr> "matched (3)", "matched (3)", "matched (3)", "mat~
```

```
glimpse(nyc_squirrels)
```

```
## Rows: 3,023
## Columns: 36
## $ long                                       <dbl> -73.95613, -73.95704, -73.9~
## $ lat                                        <dbl> 40.79408, 40.79485, 40.7667~
## $ unique_squirrel_id                         <chr> "37F-PM-1014-03", "37E-PM-1~
## $ hectare                                    <chr> "37F", "37E", "02E", "05D",~
## $ shift                                      <chr> "PM", "PM", "AM", "PM", "AM~
## $ date                                       <dbl> 10142018, 10062018, 1010201~
## $ hectare_squirrel_number                    <dbl> 3, 3, 3, 5, 1, 2, 2, 3, 9, ~
## $ age                                        <chr> NA, "Adult", "Adult", "Juve~
## $ primary_fur_color                          <chr> NA, "Gray", "Cinnamon", "Gr~
## $ highlight_fur_color                        <chr> NA, "Cinnamon", NA, NA, NA,~
## $ combination_of_primary_and_highlight_color <chr> "+", "Gray+Cinnamon", "Cinn~
```

```
## $ color_notes                                <chr> NA, NA, NA, NA, NA, NA, NA,~
## $ location                                   <chr> NA, "Ground Plane", "Above ~
## $ above_ground_sighter_measurement           <chr> NA, "FALSE", "4", "3", NA, ~
## $ specific_location                          <chr> NA, NA, NA, NA, NA, NA, NA,~
## $ running                                    <lgl> FALSE, TRUE, FALSE, FALSE, ~
## $ chasing                                    <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ climbing                                   <lgl> FALSE, FALSE, TRUE, TRUE, F~
## $ eating                                     <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ foraging                                   <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ other_activities                           <chr> NA, NA, NA, NA, "unknown", ~
## $ kuks                                       <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ quaas                                      <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ moans                                      <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ tail_flags                                 <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ tail_twitches                              <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ approaches                                 <lgl> FALSE, FALSE, FALSE, FALSE,~
## $ indifferent                                <lgl> FALSE, FALSE, TRUE, FALSE, ~
## $ runs_from                                  <lgl> FALSE, TRUE, FALSE, TRUE, F~
## $ other_interactions                         <chr> NA, "me", NA, NA, NA, NA, N~
## $ lat_long                                   <chr> "POINT (-73.9561344937861 4~
## $ zip_codes                                  <dbl> NA, NA, NA, NA, NA, NA, NA,~
## $ community_districts                        <dbl> 19, 19, 19, 19, 19, 19, 19,~
## $ borough_boundaries                         <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, ~
## $ city_council_districts                     <dbl> 19, 19, 19, 19, 19, 19, 19,~
## $ police_precincts                           <dbl> 13, 13, 13, 13, 13, 13, 13,~
```

```
glimpse(athletes)
```

```
## Rows: 19,547
## Columns: 10
## $ gender  <chr> "Men", "Men", "Men", "Men", "Men", "Men", "Men", "Men", "Men",~
## $ event   <chr> "Double FITA Round Amputee", "Double FITA Round Amputee", "Dou~
## $ medal   <chr> "Gold", "Silver", "Bronze", "Gold", "Silver", "Bronze", "Gold"~
## $ athlete <chr> "LARSEN Finn", "BRENNE Manfred", "SATO Masao", "GEISS H.", "GR~
## $ abb     <chr> "DEN", "FRG", "JPN", "FRG", "BEL", "GBR", "CAN", "NOR", "FRA",~
## $ country <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ grp_id  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ type    <chr> "Archery", "Archery", "Archery", "Archery", "Archery", "Archer~
## $ year    <dbl> 1980, 1980, 1980, 1980, 1980, 1980, 1980, 1980, 1980, 1980, 19~
## $ guide   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
```