

Past STAT 184 Activities

Brady Seyler

Sunday, November 9, 2025

Activity 8/10: Armed Forces Data

Activity 8 asked us to take a data frame containing information on armed forces and make the data table tidy where each case was an individual, with the attributes being their sex, branch, and rank. Then in Activity 10 we took that tidy data and created a frequency table to show the relationship between the rank and sex of each officer based on each branch.

Table 1: Army Frequency Table

Group	Male	Female
Army E1	7429	1326
Army E2	22338	4336
Army E3	43775	10229
Army E4	79234	15143
Army E5	54803	10954
Army E6	49502	7363
Army E7	30264	4410
Army E8	9482	1472
Army E9	2865	394
Army W1	3727	460
Army W2	6024	692
Army W3	2794	346
Army W4	1378	137
Army W5	494	43
Army O1	7122	2400
Army O2	9550	3006
Army O3	20986	6053
Army O4	12350	3044

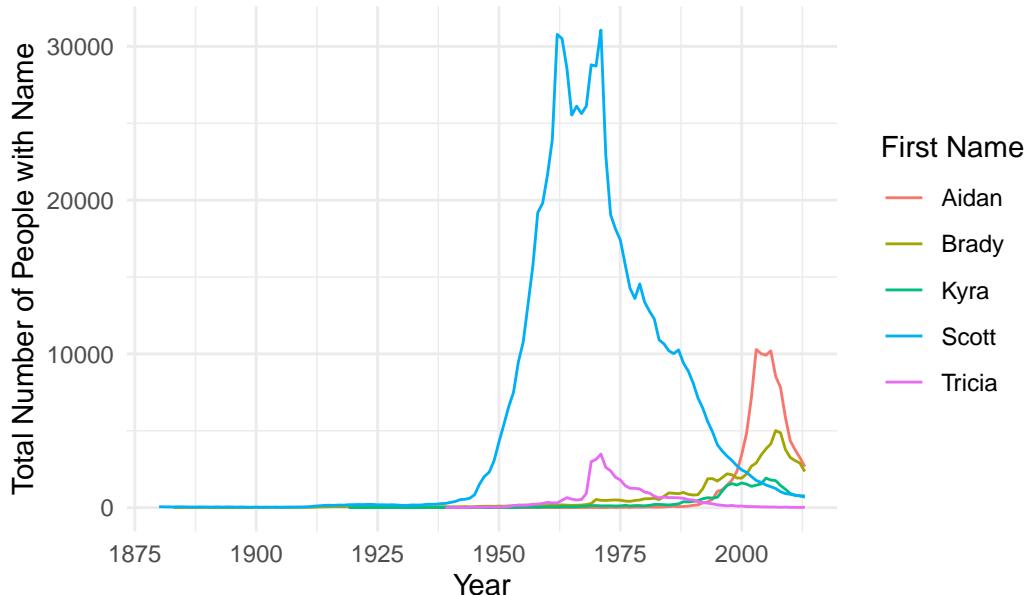
Army O5	6939	1531
Army O6	3161	452
Army O7	100	18
Army O8	80	8
Army O9	46	5
Army O10	11	0

In this frequency table we are looking to see if rank and sex are independent from each other in the Army. Each case is a different rank in the Army, with the columns indicating how many men and women are enrolled in each rank. When looking at the table, the men vastly outnumber the women for every rank. That means in the Army sex and rank are independent, since there will be more men than women regardless of rank.

Activity 13: Popularity of Baby Names

Activity 13 asked us to create a plot showing the popularity of different names over a span of time.

Popularity of Seyler Family First Names Over Time



The path plot shows the popularity of the names Scott, Tricia, Aidan, Kyra, and Brady from 1880 to 2013. I picked these names since they are the five first names in my family. The x-axis represents the different years, with the y-axis representing how many babies were named each

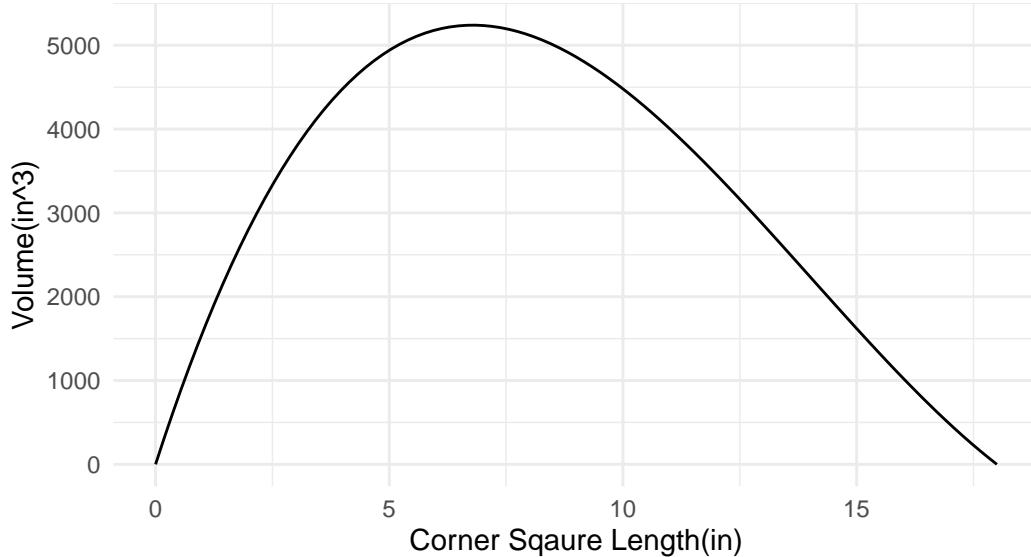
of the five names every year. Each different name is represented by a different line, with each line having a different color. Scott has the biggest peak around 1960, with a peak over 30,000. Scott remains the most popular name till around the year 2000, where Aidan takes over as the most popular name, with that name peaking around 10,000. A ranking of the peaks of all five names would be Scott, Aidan, Brady, Tricia, and Kyra with the lowest peak.

Activity 4: Plotting a Mathematical Function

The goal of this activity was to create a function of the volume of an open-faced box using a 36 by 48 inch piece of paper.

Volume of an Open-Faced Box

From a 36 by 48 inch Paper



In the box problem, you are given a piece of paper with a given set of measurements, in this case 36 by 48 inches. If you were to cut four identical sized squares out of each corner, you would create an open-faced box. The problem asks you to create a function where you can insert the side length of the square cutouts and it gives you the volume of the open-face box with that length. This graph plots that function, and when looking at the plot you can see the max volume is over 5000 in³. To be specific, to get the max volume the side length would need to be about 6.79 inches long, which would give you a volume of 5239.82 in³.

What I've Learned So Far

These assignments have taught me a lot about R. I've learned how to load a data frame into R and how to modify it to my desire. I've used this skill in multiple assignments like assignment 8, along with assignments in another class that also uses R. I've also learned how to create multiple different plots in R, like the one's in assignments 13 and 4. I have also learned more general things, like how important it is to plan when doing anything. Writing how thorough plans not only helps me complete tasks, but it can also help my future self or someone else trying to recreate what I did.

Code Appendix

```
#Loading in the necessary packages
library(googlesheets4)
library(tidyverse)
library(knitr)
library(kableExtra)

# Loading the Data Frame into R
gs4_deauth()
militaryraw<-read_sheet(
  ss = 'https://docs.google.com/spreadsheets/d/19xQnI1cBh6Jkw7eP8YQuuicMlVDF7Gr-nXCb5qbwb_E/'
)

# Removing the Unnecessary Rows
militaryfix<-militaryraw[-c(1,2,12,18,29,30,31), ]

# Renaming the Columns
militaryrename <- militaryfix %>% rename(
  GroupofSoldiers = "Active-Duty Personnel by Service Branch, Sex, and Pay Grade",
  ArmyMale = ...2,
  ArmyFemale = ...3,
  Army = ...4,
  NavyMale = ...5,
  NavyFemale = ...6,
  Navy = ...7,
  MCMale = ...8,
  MCFemale = ...9,
  MC = ...10,
  AFMale = ...11,
```

```

AFFemale = ...12,
AF = ...13,
SFMale = ...14,
SFFemale = ...15,
SF = ...16,
MaleTotal = ...17,
FemaleTotal = ...18,
TotalTotal = ...19
)

#Creating a Data Frame for the Army
army<-militaryrename %>%
  dplyr::select(1,2,3,4) %>%
  pivot_longer(
    cols = Army,
    names_to = "Branch",
    values_to = "Total"
  ) %>%
  unite(
    col = "Group",
    Branch,
    GroupofSoldiers,
    sep=" "
  ) %>%
  dplyr::select(!Total) %>%
  rename(Male=ArmyMale,Female=ArmyFemale)

#Creating a Data Frame for the Navy
navy<-militaryrename %>%
  pivot_longer(
    cols = Navy,
    names_to = "Branch",
    values_to = "Total"
  ) %>%
  unite(
    col = "Group",
    Branch,
    GroupofSoldiers,
    sep=" "
  ) %>%
  rename(Male=NavyMale,Female=NavyFemale) %>%
  dplyr::select(1,5,6)

```

```

#Creating a Data Frame for the Marine Corps
marinecorps<-militaryrename %>%
  pivot_longer(
    cols = MC,
    names_to = "Branch",
    values_to = "Total"
  ) %>%
  unite(
    col = "Group",
    Branch,
    GroupofSoldiers,
    sep=" "
  ) %>%
  rename(Male=MCMale,Female=MCFemale) %>%
  dplyr::select(1,8,9)

#Creating a Data Frame for the Air Force
airforce<-militaryrename %>%
  pivot_longer(
    cols = AF,
    names_to = "Branch",
    values_to = "Total"
  ) %>%
  unite(
    col = "Group",
    Branch,
    GroupofSoldiers,
    sep=" "
  ) %>%
  rename(Male=AFMale,Female=AFFemale) %>%
  dplyr::select(1,11,12)

#Creating a Data Frame for the Space Force
spaceforce<-militaryrename %>%
  pivot_longer(
    cols = SF,
    names_to = "Branch",
    values_to = "Total"
  ) %>%
  unite(
    col = "Group",
    Branch,

```

```

  GroupofSoldiers,
  sep=" "
) %>%
rename(Male=SFMale,Female=SFFemale) %>%
dplyr::select(1,14,15)

#Combining the five Branch Data Frames
militarygroups<-bind_rows(army,navy,marinecorps,airforce,spaceforce)
militarygroups<-militarygroups%>%filter(Male >= 0)

#Making a Frequency Table for the Army Data
army %>%
  kable(caption = "Army Frequency Table") %>%
  kableExtra::kable_classic()

#Loading the necessary packages
library(dcData)
library(ggplot2)

#Shortening the "Baby Names" data frame to only include the desired first names.
seyler<-BabyNames %>%
  filter(name=="Scott" |
         name=="Tricia" |
         name=="Aidan" |
         name=="Kyra" |
         name=="Brady") %>%
  pivot_wider(
    id_cols = c(name,year),
    names_from = sex,
    values_from = count
) %>%
  mutate_all(~replace(.,is.na(.),0)) %>%
  mutate(count = M + F)

#Creating a plot that shows the popularity of the five Seyler first names over time
ggplot(seyler) +
  aes(
    x = year,
    y = count,
    color = name,
    group = name
) +

```

```

geom_path() +
labs(
  title = "Popularity of Seyler Family First Names Over Time",
  x = "Year",
  y = "Total Number of People with Name",
  color = "First Name",
  alt = "A plot that shows the popularity of different baby names over time."
) +
theme_minimal()

#Creating the function of the volume of an open-faced box
Max_Volume<-function(SquareLength){
  volume<- (36-2*SquareLength)*(48-2*SquareLength)*SquareLength
  return(volume)
}

# Creating a plot that shows the relationship between the size of the square and the size of
ggplot()+
  xlim(0, 18) +
  stat_function(
    fun=Max_Volume
  ) +
  labs(
    title = "Volume of an Open-Faced Box",
    subtitle = "From a 36 by 48 inch Paper",
    x = "Corner Sqaure Length(in)",
    y = "Volume(in^3)",
    alt = "Plot showing volume of open-faced box based on the length of the corner square."
  ) +
  theme_minimal()

```