Strengths

Sam Newman

2023-10-21

Background: TThe data comes from one qualitative question asked in a survey that was given to a universities (named disclosed for confidentiality) 2nd year MPP students (n = 269) from 2019-2023. The question was free-response and asked students to list their strengths about the UCLA MPP program. This script preforms thematic analysis by tallying the total number of mentions a theme is mentioned. A bar chart is created showing the 5 most common themes and a table is produced displaying all the themes with their respective number of mentions.

Outline: 1. Clean Data 2. Analyze Data 3. Visualize Data

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.2.3

## Warning: package 'ggplot2' was built under R version 4.2.3

## Warning: package 'tibble' was built under R version 4.2.3

## Warning: package 'tidyr' was built under R version 4.2.3

## Warning: package 'readr' was built under R version 4.2.3

## Warning: package 'purrr' was built under R version 4.2.3

## Warning: package 'dplyr' was built under R version 4.2.3

## Warning: package 'stringr' was built under R version 4.2.3

## Warning: package 'forcats' was built under R version 4.2.3

## Warning: package 'lubridate' was built under R version 4.2.3

## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- 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 become errors
```

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.2.3
```

```
library(writexl)
```

```
## Warning: package 'writexl' was built under R version 4.2.3
```

```
library(stringr)
library(gt)
```

```
## Warning: package 'gt' was built under R version 4.2.3
```

```
library(webshot2)
```

```
## Warning: package 'webshot2' was built under R version 4.2.3
```

```
setwd("C:/Grad School/Student Affairs/Exit Survey")
```

load clean exit survey data 2019-2023 (created in another script)

```
sub_e2019_2023 <- readRDS("sub_e2019_2023.rds")
```

Clean Variable

subset data and investigate subset

```
#create subset of Themes_Strengths
subset_strengths<- sub_e2019_2023 %>%
  select(Themes_Strengths)

#269 observations
nrow(subset_strengths)
```

```
## [1] 269
```

```
#11 NA values
sum(is.na(subset_strengths))
```

```
## [1] 11
```

```
#reduces observations to 258 (practice)
subset_strengths %>%
  select(Themes_Strengths) %>%
  filter(!is.na(Themes_Strengths)) %>%
  nrow()
```

```
## [1] 258
```

```
#remove NAs
clean_subset_strengths <- subset_strengths %>%
  select(Themes_Strengths) %>%
  filter(!is.na(Themes_Strengths))

#observations are 258. Code worked.
nrow(clean_subset_strengths)
```

```
## [1] 258
```

```
#124 unqiue answers. Although this will change once I seperate entries by commas
unique(clean_subset_strengths$Themes_Strengths)
```

```
##   [1] "Qualitative, Quantitative"
##   [2] "courses, Core Classes, Department, APP"
##   [3] "Academic Resources, Career Resources"
##   [4] "courses, APP, Career Resources"
##   [5] "courses"
##   [6] "APP, Career Resources"
##   [7] "Alumni Connections, Relationships"
##   [8] "Relationships"
##   [9] "Core Classes"
##  [10] "Quantitative"
##  [11] "APP, Career Resources, Relationships,"
##  [12] "APP"
##  [13] "Core classes, Relationships"
##  [14] "Quantitative, core classes"
##  [15] "Core classes"
##  [16] "Personal Growth"
##  [17] "Courses, Relationships"
##  [18] "Classes"
##  [19] "DEI"
##  [20] "DEI, Core Classes"
##  [21] "DEI, Core Classes, Faculty"
##  [22] "DEI, Career Resources, Relationship, Alumni Networking"
##  [23] "Core Classes, courses"
##  [24] "Relationships, courses"
##  [25] "Relationships, core classes"
##  [26] "Core Classes, Quantitative, courses"
##  [27] "courses, Core classes"
##  [28] "Flexibility through Covid"
##  [29] "courses, Faculty"
##  [30] "Practical Skills, faculty"
##  [31] "Applied Learning"
```

```
##  [32] "Courses"
##  [33] "courses, Relationships"
##  [34] "Relationships, Courses"
##  [35] "Quantitative, courses, APP"
##  [36] "Quantitative, Core Classes"
##  [37] "Core Classes, Relationships, Classes, DEI, Flexibility through Covid"
##  [38] "Relationships, APP"
##  [39] "Quantitative, APP"
##  [40] "Academic Resources, courses, Relationships"
##  [41] "Work Experience"
##  [42] "Career Resources"
##  [43] "Core classes, Practical Skills, Practioners"
##  [44] "Relationships, Courses, Work Experience"
##  [45] "Quantiative, Qualitative, Applied Learning"
##  [46] "Work Experience, Career Resources, courses"
##  [47] "Quantitative, Relationships"
##  [48] "Relationships, courses, APP"
##  [49] "Relationships, Qualitative"
##  [50] "Career Resources, Quantitative,"
##  [51] "courses, Core classes, work experience"
##  [52] "Core Classes, Career Resources"
##  [53] "Core Classes, APP, Courses"
##  [54] "Quntitative, Core Classes"
##  [55] "Quantitative, Courses"
##  [56] "courses, Quantitative, APP, Leanring from Professors,"
##  [57] "Quantitative, Career Resources"
##  [58] "Quantiative, APP, Relationships, Work Experience"
##  [59] "Quantiative, courses"
##  [60] "Quantitative, Practical skills"
##  [61] "Quantitative, Career Resources, Practical skills, APP"
##  [62] "Quantitative, Core Classes, courses"
##  [63] "Core Classes, Relationships, Courses"
##  [64] "Practical Skills"
##  [65] "APP, courses"
##  [66] "Core classes, courses"
##  [67] "Relationships, Academic Resources"
##  [68] "DEI, courses"
##  [69] "Relationships, courses, Courses"
##  [70] "Career Resources, APP, Relationships"
##  [71] "Core Classes, Relationships, Funding Opportunities"
##  [72] "Quantitative, core classes, courses"
##  [73] "Relationships, Flexibility through COVID"
##  [74] "Relationships, core classes, courses, work experience, APP"
##  [75] "Quantitative, Core classes"
##  [76] "Quantitative, courses"
##  [77] "core classes"
##  [78] "practical skills"
##  [79] "relationships, work experience"
##  [80] "relationships"
##  [81] "relationships, career resources, work experience,"
##  [82] "Quantitative, courses, core classes"
##  [83] "Practical skills, Quantiative, Core Classes"
##  [84] "core classes, applied learning"
##  [85] "core classes, courses"
```

```
##  [86] "Core Classes, relationships, DEI"
##  [87] "applied learning"
##  [88] "Quantitative, Relationships, courses"
##  [89] "Classses outside of Luskin, Relationships"
##  [90] "Classses outside of Luskin, Relationships, Flexibility through COVID"
##  [91] "Work experience, core classes, quantitative, applied learning, Practical skills"
##  [92] "Courses, core classes, relationships"
##  [93] "courses, Core Classes"
##  [94] "APP, Courses, Realtionships"
##  [95] "Career Resourses, Work Experience, Relationships"
##  [96] "Courses, Funding Opportunities, Work Experience"
##  [97] "APP, Personal Growth, DEI, Career Resources"
##  [98] "NA"
##  [99] "Funding Opportunities,"
## [100] "Relationships, DEI"
## [101] "Quantitative, Practical skills, Courses"
## [102] "Quantiative"
## [103] "APP, Relationships"
## [104] "Quantitative, Qualitative"
## [105] "Courses, Work Experience, Relationships"
## [106] "Courses, Career Resources"
## [107] "Personal Growth, Relationships"
## [108] "Quntitative, Relationships,"
## [109] "Work experience"
## [110] "Practical skills"
## [111] "Career Resources, Courses,"
## [112] "Relationships, pratical skills, career resources"
## [113] "Relationships, Career Resources"
## [114] "Relationships, Courses, Career Resources"
## [115] "Funding Opportunities, Relationships, Personal Growth"
## [116] "Courses, core classses, work experience, APP"
## [117] "Relationships, Academic Resources, Courses"
## [118] "Relationships, work experience"
## [119] "Relationships, Funding Opportunities, APP, Core classes, Courses"
## [120] "Practical Skills, Quantitative"
## [121] "Career Networking"
## [122] "Quantitative, DEI, Relationships"
## [123] "Courses, Funding Opportunities, Relationships"
## [124] "Career Networking, Courses, Work Experience, APP, Relationships"
```

```
#Class is character
class(clean_subset_strengths$Themes_Strengths)
```

```
## [1] "character"
```

seperate values by commas

```
comma_seperated_strengths <- clean_subset_strengths %>%
  separate_rows(Themes_Strengths,
                sep = ",")
```

```
#object now has 451 observations
length(comma_seperated_strengths$Themes_Strengths)
```

```
## [1] 451
```

clean comma_seperated_strengths

```
#remove white space before and after values
clean_comma_seperated_strengths <- comma_seperated_strengths %>%
  mutate(Themes_Strengths =
           str_squish(Themes_Strengths))

#convert all letters to lower case
clean_comma_seperated_strengths <- clean_comma_seperated_strengths %>%
  mutate(Themes_Strengths =
           tolower(Themes_Strengths))

#33 unique values
unique(clean_comma_seperated_strengths$Themes_Strengths)
```

```
##  [1] "qualitative"                "quantitative"
##  [3] "courses"                    "core classes"
##  [5] "department"                 "app"
##  [7] "academic resources"         "career resources"
##  [9] "alumni connections"         "relationships"
## [11] ""                           "personal growth"
## [13] "classes"                    "dei"
## [15] "faculty"                    "relationship"
## [17] "alumni networking"          "flexibility through covid"
## [19] "practical skills"           "applied learning"
## [21] "work experience"            "practioners"
## [23] "quantiative"                "quntitative"
## [25] "leanring from professors"   "funding opportunities"
## [27] "classses outside of luskin" "realtionships"
## [29] "career resourses"           "na"
## [31] "pratical skills"            "core classses"
## [33] "career networking"
```

```
#recode misspelled values
clean_comma_seperated_strengths <- clean_comma_seperated_strengths %>%
  mutate(Themes_Strengths =
           recode(Themes_Strengths,

                                  " " = 'NA',
                                  "alumni connections" = "career resources",
                                  "alumni networking" = "career resources",
                                  "career networking" = "career resources",
                                  "career resourses" = "career resources",
                                  "classes" = "courses",
                                  "classses outside of luskin" = "courses",
```

```
                                  "core classses" = "core classes",
                                  "na" = 'NA',
                                  "quantiative" = "quantitative",
                                  "quntitative" = "quantitative",
                                  "realtionships" = "relationships",
                                  "relationship" = "relationships",
                                  "department" = "relationships",
                                  "faculty" = "relationships",
                                  "leanring from professors" = "courses",
                                  "practioners" = "practical skills",
                                  "pratical skills" = "practical skills",
                                  "applied learning" = "practical skills"))

#17 unique values
unique(clean_comma_seperated_strengths$Themes_Strengths)
```

```
##  [1] "qualitative"               "quantitative"
##  [3] "courses"                   "core classes"
##  [5] "relationships"             "app"
##  [7] "academic resources"        "career resources"
##  [9] ""                          "personal growth"
## [11] "dei"                       "flexibility through covid"
## [13] "practical skills"          "work experience"
## [15] "funding opportunities"     "NA"
```

```
#check code to remove NAs
clean_comma_seperated_strengths %>%
  filter(Themes_Strengths != "NA") %>%
  unique()
```

```
## # A tibble: 15 x 1
##    Themes_Strengths
##    <chr>
##  1 "qualitative"
##  2 "quantitative"
##  3 "courses"
##  4 "core classes"
##  5 "relationships"
##  6 "app"
##  7 "academic resources"
##  8 "career resources"
##  9 ""
## 10 "personal growth"
## 11 "dei"
## 12 "flexibility through covid"
## 13 "practical skills"
## 14 "work experience"
## 15 "funding opportunities"
```

```
#check code to remove blank values
clean_comma_seperated_strengths %>%
  filter(Themes_Strengths != "") %>%
  unique()
```

```
## # A tibble: 15 x 1
##    Themes_Strengths
##    <chr>
##  1 qualitative
##  2 quantitative
##  3 courses
##  4 core classes
##  5 relationships
##  6 app
##  7 academic resources
##  8 career resources
##  9 personal growth
## 10 dei
## 11 flexibility through covid
## 12 practical skills
## 13 work experience
## 14 funding opportunities
## 15 NA
```

```
#remove NAs and blank values
clean_comma_seperated_strengths <-  clean_comma_seperated_strengths %>%
  filter(Themes_Strengths != "NA",
         Themes_Strengths != "",
         Themes_Strengths != " ")


#14 unique values
unique(clean_comma_seperated_strengths$Themes_Strengths)
```

```
##  [1] "qualitative"               "quantitative"
##  [3] "courses"                   "core classes"
##  [5] "relationships"             "app"
##  [7] "academic resources"        "career resources"
##  [9] "personal growth"           "dei"
## [11] "flexibility through covid" "practical skills"
## [13] "work experience"           "funding opportunities"
```

Analyze variable

add count and percent to the number of times a value is repeated

```
#create object containing count
count_comma_seperated_strengths <- clean_comma_seperated_strengths %>%
  group_by(Themes_Strengths) %>%
  summarize(count = n())

#create object containing percent
count_comma_seperated_strengths <- count_comma_seperated_strengths %>%
  mutate(percent =
           round(count / sum(count), 4))
```

```
#order objects by numerically
count_comma_seperated_strengths <- count_comma_seperated_strengths %>%  arrange(-count)

#number of observations is 443
sum(count_comma_seperated_strengths$count)
```

```
## [1] 443
```

keep top 5 counts

```
top5_strenghts <- count_comma_seperated_strengths[1:5,]
```

change to uppercase and change the names of the variables

```
#Capitalize the first letter of each word
top5_strenghts <- top5_strenghts %>%
  mutate(Themes_Strengths =
           str_to_title(Themes_Strengths))

#change the name of courses
top5_strenghts <- top5_strenghts %>%
  mutate(Themes_Strengths =
           recode(Themes_Strengths,
                  "Courses" = "Electives",
                  "Quantitative" = "Quantitative Skills"))
```

factor bars (preparing for visualization)

```
factor_top5_strengths <- top5_strenghts

factor_top5_strengths$Themes_Strengths <- factor_top5_strengths$Themes_Strengths %>%
  factor(
  factor_top5_strengths$Themes_Strengths,
  levels = c(
             "Electives",
             "Relationships",
             "Quantitative Skills",
             "Core Classes",
             "Career Resources"
                               ))

#check class type (it's factor - good)
class(factor_top5_strengths$Themes_Strengths)
```

```
## [1] "factor"
```

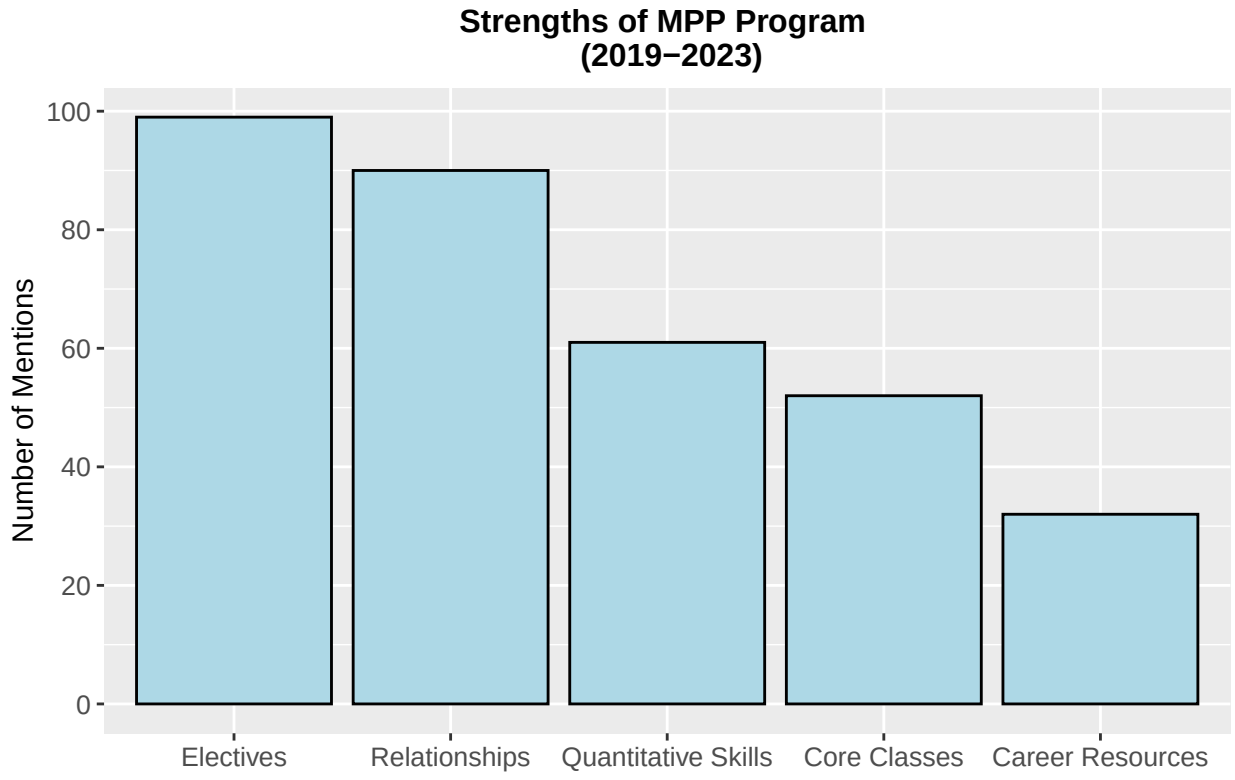
Visualize Data

create horizontal bar chart

```
ggplot(factor_top5_strengths,
       aes(x = Themes_Strengths,
           y = count)) +

  geom_bar(
         stat = 'identity',
         fill = "lightblue",
         position = 'dodge', #makes bar chart side by side
         color = "black", #adds outline color to bins
         na.rm = TRUE)+

  scale_y_continuous(
                    breaks = seq(0,100, 20)) +
  labs(
        x = "",
        y = "Number of Mentions",
        title = "Strengths of MPP Program \n (2019-2023)")+
  theme(
        axis.text = element_text(hjust = .5, size =10),
        plot.caption = element_text(hjust = 1,),
        plot.title = element_text(hjust = .5, face = "bold", size = 12))+
        labs(caption ="443 mentions from 258 students")
```



443 mentions from 258 students

create table showing all strengths for count_sep_strengths

```
#create row for total count
new_row <- data.frame(
  Themes_Strengths = "Total",
  count = sum(count_comma_seperated_strengths$count),
  percent = NA)

#add row to dataframe
table_full_strengths <- add_row(count_comma_seperated_strengths, new_row)
```

clean table_strengths

```
#Format names of Themes Strength and capitalized variables name
clean_table_full_strengths <- table_full_strengths %>%
  mutate(
    Themes_Strengths = str_to_title(Themes_Strengths),
    Themes_Strengths = recode(Themes_Strengths,
                              "App" = "APP",
                              "Dei" = "DEI")) %>%
```

```
  rename(
    Mentions = count,
    Percent = percent,
    Strengths = Themes_Strengths)

#remove percent
clean_table_full_strengths <- clean_table_full_strengths %>%  select(-Percent)
```

create table

```
strengths_table <- gt(clean_table_full_strengths)
```

create file directory and save table in directory as .png

```
#create object that points to folder where I want table to be saved
table_dir <- "C:/Grad School/Student Affairs/Exit Survey/Visualizations"

#save table to that directory
gtsave(strengths_table, filename = file.path(table_dir, "strengths_table.png"))
```