

Librarian Census Data Code

Dan Anthony Dorado

2023-11-30

Table of contents

Abstract	2
About this document	3
R packages and dataset	3
Data Cleaning, Wrangling, and Visualization	7
Question: Are you currently a licensed librarian (R.A. 9246) ?	7
Question: Are you currently residing in the Philippines?	8
Question: Kindly indicate which gender you identify the most	9
Question: Kindly indicate your age	10
Question: Are you currently employed and/or working?	13
Question: Kindly indicate the setting, field or industry you are primarily working in	14
Question: How long have you been working as a professional after passing the licen- sure exam for librarians (cumulative years)	15
Question: Kindly indicate the type of institution (in terms of ownership) you are primarily working in	18
Question: Kindly indicate your current tenure status	19
Question: Kindly indicate your current monthly gross salary or income range (before deductions)	20
Kindly indicate your current monthly net salary or income range (after deductions) .	23
Question: Kindly indicate additional benefits, if any (please check all that apply) . .	26
Question: Other benefits not mentioned in the previous page	28
Question: What kind of materials are you currently working with? Kindly check all that applies.	29
Question: Kindly indicate your current work position	30
Question: Within the next five (5) years, do you see your self working in the same industry?	31

Question: If your answer was NO or MAYBE in the previous page, indicate the field you wish to pursue	32
Question: Kindly indicate your highest educational attainment	33
Question: Are you currently enrolled in formal studies? (Short Courses, Certificate Program, Diploma Program, Master Program, PhD)	35
Question: If your answer was YES, indicate the degree you are currently completing	35
Question: Are you planning to study again in a formal university setting within the next five (5) years?	37
Question: If your answer was YES, kindly provide the degree you wish to study within the next five (5) years	38
Question: Where do you usually acquire your CPD trainings in LIS? Kindly check all that applies	39
Where do you usually get funding or support for your trainings in LIS?	40
Question: Are you satisfied with the amount of CPD trainings you have attended for the last two (2) years?	41
Question: Kindly indicate any challenges you encounter when fulfilling the required CPD trainings	43
Kindly indicate any topics you wish to explore in future trainings	44
Kindly indicate your primary work location	45
Question: Have you moved out of your original / permanent residence in order to work on your current work place?	48
Is Filipino / Tagalog your native language?	48
Kindly check the local languages or dialects you can speak on a conversational level other than Filipino / Tagalog	49
Kindly indicate any foreign languages or dialects you can speak on a conversational level other than English (if any)	50
Kindly indicate the type of library you are currently working in	51
Other than being a librarian, what other roles do you serve in your institution/ organization?	56
Kindly indicate your current work position	57
Kindly indicate your primary job / designation inside the library	58
Kindly indicate your current / existing professional affiliations	60
Kindly indicate what factors influence your decision when it comes to joining a professional organization?	69

Abstract

The University of the Philippines School of Library and Information Studies (UPSLIS) conducted a study of Philippine Librarians from November 2018 to October 2019. This endeavor was the first-ever wide-scale study undertaken to collect the occupational profile of Librarians in the Philippines. The survey was designed to be an online questionnaire via Mentimeter and

was presented for dissemination during the 2018 Philippine Librarians Association Inc. (PLAI) Congress at Novotel in Quezon City. To further disseminate, the link was posted on the UP-SLIS Facebook account and was shared to library association group pages. A total of 685 responses were recorded which is only 7.5% of the almost 9000 registered librarians in the Philippines but can be used in making statistical inference with a 4% margin of error.

About this document

This paper aims to discuss the methods and process used in cleaning, visualizing, and processing of the **Philippine Librarian Census Dataset** (see <https://zenodo.org/record/6450452>). It details the:

- R packages used
- Data cleaning methods
- Data wrangling methods
- Data summarization methods
- Data computation methods
- Data visualization methods

The researchers hope that this would provide a framework for future LIS research to be more open by sharing their datasets and documentation. This would allow research verification and provide an opportunity for other researchers to use different theories and approaches to shared datasets in order to promote the culture of research in the LIS field.

R packages and dataset

```
# Load needed package  
library(tidyverse)
```

Warning: package 'tidyverse' was built under R version 4.3.1

Warning: package 'ggplot2' was built under R version 4.3.2

Warning: package 'tidyr' was built under R version 4.3.1

Warning: package 'dplyr' was built under R version 4.3.2

Warning: package 'forcats' was built under R version 4.3.1

Warning: package 'lubridate' was built under R version 4.3.2

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.3      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      1.0.1
-- 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
```

```
library(tm)
```

Warning: package 'tm' was built under R version 4.3.1

Loading required package: NLP

Attaching package: 'NLP'

The following object is masked from 'package:ggplot2':

annotate

```
library(infer)
```

Warning: package 'infer' was built under R version 4.3.2

```
#Load the raw "librariancensus.csv" dataset
dataset <- read_csv("librariancensus.csv")
```

Warning: One or more parsing issues, call `problems()` on your data frame for details,
e.g.:

```
dat <- vroom(...)
problems(dat)
```

Rows: 684 Columns: 67

-- Column specification -----

Delimiter: ","

chr (62): licensed, philippines, gender, working, industry, institution, tenure...

dbl (5): voter, age, years_service, gross_salary, net_salary

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

```
#Inspect the loaded dataset
```

```
glimpse(dataset)
```

Rows: 684

Columns: 67

```
$ voter      <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ~
$ licensed   <chr> "Yes, I have a valid and up-to-date PRC license fo~
$ philippines <chr> NA, "No, I am in the Philippines right now but onl~
$ gender      <chr> NA, "Prefer not to say", "Woman", "Man", "Man", "M~
$ age         <dbl> 59, 100, 40, 46, 32, 46, 48, 44, 25, 24, 34, 39, 3~
$ working     <chr> NA, "No, I am a retiree.", "Yes, I am currently em~
$ industry    <chr> NA, "information technology", "academic institutio~
$ years_service <dbl> NA, 60, 14, 23, 10, 24, 23, 12, 5, 1, 12, 18, 14, ~
$ institution <chr> NA, "private (BPO)", "government", "government", "~
$ tenure      <chr> NA, "self-employed, freelancing or working with mu~
$ gross_salary <dbl> NA, 100, 51, 53, 40, 56, 36, 38, 19, 11, 95, 58, 3~
$ net_salary  <dbl> NA, 100, 30, 38, 30, 42, 29, 26, NA, 10, 95, 36, 2~
$ benefits1   <chr> NA, "night differential pay", "paid vacation and/o~
$ benefits2   <chr> NA, NA, "monetary allowances", "monetary allowance~
$ benefits3   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, "h~
$ benefits4   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ benefits5   <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ otherbenefits1 <chr> NA, "123", "Rice allowance\r\nClothing allowance\r~
$ otherbenefits2 <chr> NA, "123", NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ otherbenefits3 <chr> NA, "Hello Philippines and hello world", NA, NA, N~
$ otherbenefits4 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ materials1  <chr> NA, "law materials and resources", NA, "archives (~
$ materials2  <chr> NA, NA, NA, "hardcopy theses and dissertations", N~
$ materials3  <chr> NA, NA, NA, "hardcopy books / pamphlets / etc", NA~
$ materials4  <chr> NA, NA, NA, "institutional files and records", NA,~
$ materials5  <chr> NA, NA, NA, "data sets", NA, "archives (physical)"~
$ materials6  <chr> NA, NA, NA, "electronic material (digital titles, ~
```

\$ materials7	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ othermaterials1	<chr> NA, "Huehs", "Grading sheets\r\nExam papers", "Boo~
\$ othermaterials2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ othermaterials3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ othermaterials4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ position	<chr> NA, "supervisory / mid-management level", "supervi~
\$ fiveyears	<chr> NA, "No (planning to shift to other fields or indu~
\$ pursue	<chr> NA, "Hfeudnr", NA, NA, NA, NA, NA, NA, NA, NA, NA,~
\$ education	<chr> NA, "Doctor of Philosophy (Ph.D)", "Master's degre~
\$ enrolled	<chr> NA, "Yes", "Yes", "Yes", "No.", "Yes", "Yes", "No.~
\$ completing	<chr> NA, "Post-graduate Certificate or Diploma (non-LIS~
\$ study_again	<chr> NA, NA, "No", "Yes", "Yes", "Yes", "Yes", "Yes", "~
\$ `5yearsdegree`	<chr> NA, NA, NA, "Short Certificate Diploma Program~
\$ cpd1	<chr> NA, NA, "trainings and symposiums facilitated by p~
\$ cpd2	<chr> NA, NA, "government facilitated trainings and symp~
\$ cpd3	<chr> NA, NA, NA, NA, "within the work place (in-house t~
\$ cpd4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ cpdsupport1	<chr> NA, NA, "from the government (scholarships, free t~
\$ cpdsupport2	<chr> NA, NA, NA, "from the government (scholarships, fr~
\$ cpdsupport3	<chr> NA, NA, NA, "from the work place (private sponsors~
\$ cpdsatisfaction	<chr> NA, NA, "Yes, the amount of trainings I received w~
\$ cpdbarriers1	<chr> NA, NA, "CPD program sources that I really like", ~
\$ cpdbarriers2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ cpdbarriers3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ topics1	<chr> NA, NA, "Ethics\r\nCommunication skills\r\nCooking~
\$ topics2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ topics3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ topics4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ topics5	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ location	<chr> NA, NA, "NCR", "NCR", "NCR", "NCR", "NCR", "NCR", ~
\$ worktravel	<chr> NA, NA, "Yes", "Yes", "Yes", "Yes", "No", "No", "N~
\$ filipino	<chr> NA, NA, "No", "Yes", "Yes", "Yes", "No", "Yes", "Y~
\$ dialect	<chr> NA, NA, "Ilocano", NA, "Others", NA, "Kapampangan"~
\$ foreignlang	<chr> NA, NA, "None", NA, "Chinese", NA, NA, "English", ~
\$ type	<chr> NA, NA, "Slis", "Library school", NA, "Library sch~
\$ roles	<chr> NA, NA, "Cleaner\r\nCook", NA, "Faculty", "Asst pr~
\$ institutionposition	<chr> NA, NA, "Head", "Asst prof 5", "Asst. Prof.", "Ass~
\$ primaryjob	<chr> NA, NA, NA, "Education", "Research, Extention, Tea~
\$ affiliation	<chr> NA, NA, "Plai\r\nCodhlis\r\nPatls", "Plai, patls, ~
\$ factor	<chr> NA, NA, "Friends", "Similar thrusts", "colleagues"~

```
#controls how many digits will be printed
options(digits=4)

#Initialize a pseudorandom number generator to obtain the same output during simulation
set.seed(12)
```

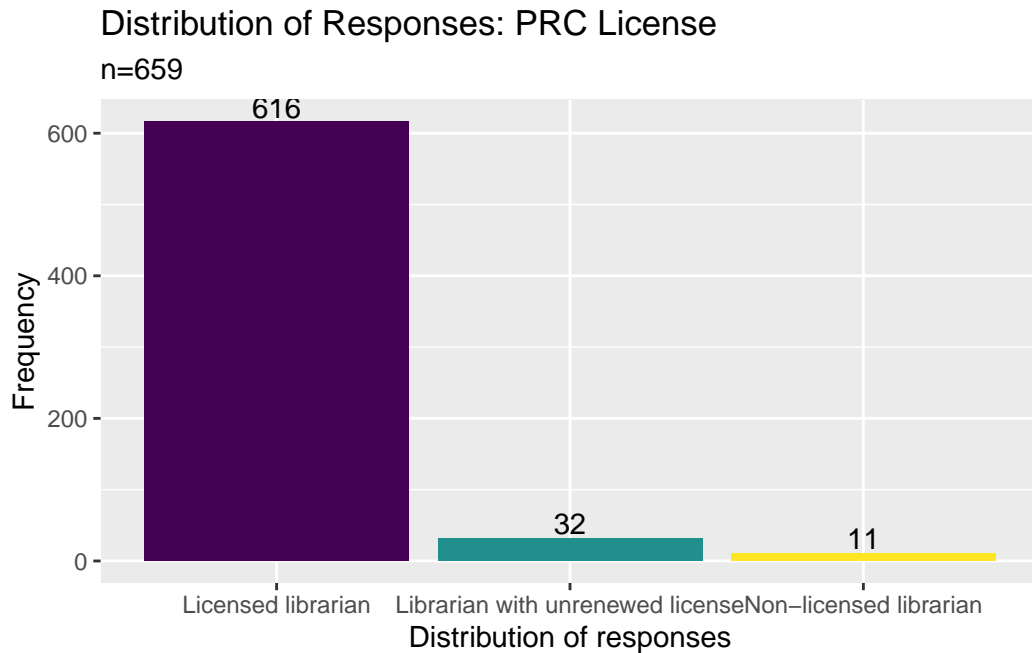
Data Cleaning, Wrangling, and Visualization

Question: Are you currently a licensed librarian (R.A. 9246) ?

- Yes, I have a valid and up-to-date PRC license for librarians
- No, I have not renewed my license
- No, I have never taken any licensure examination for librarians and/or never applied for a PRC license for librarians

```
# Change the factor type
dataset$licensed <- as.factor(dataset$licensed)
# Recode responses
dataset$licensed <- fct_recode(dataset$licensed, "Licensed librarian" = "Yes, I have a val
# Rearrange the responses
dataset$licensed <- factor(dataset$licensed, levels = c("Licensed librarian", "Librarian w
# Graph the data
dataset %>%
filter(!is.na(licensed)) %>%
ggplot(aes(x = licensed, fill = licensed)) + geom_bar() + geom_text(aes(label = ..count..))
```

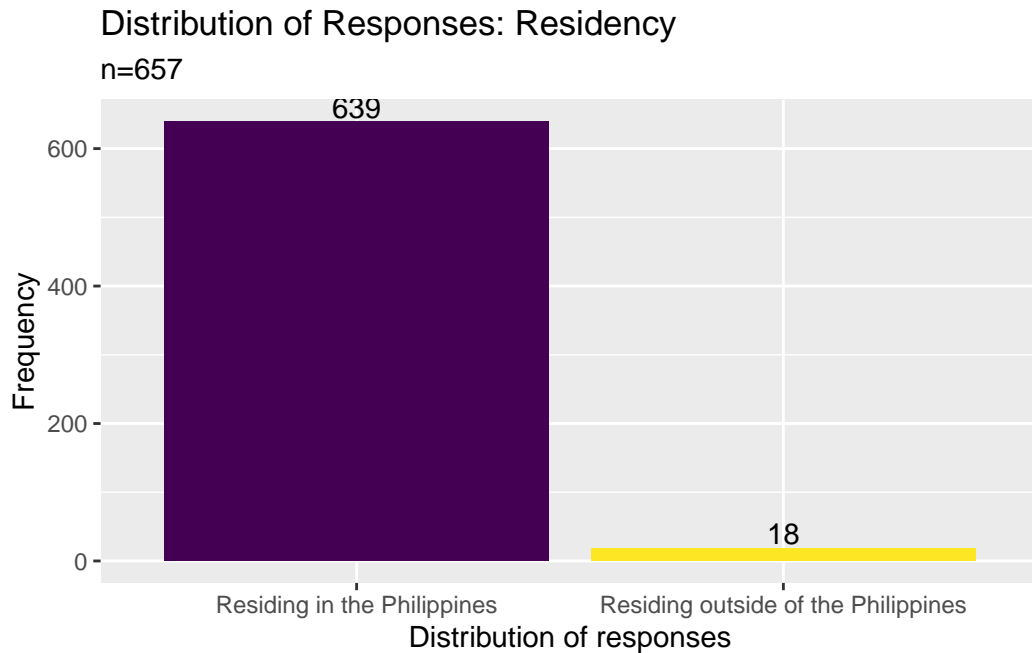
Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
i Please use `after_stat(count)` instead.



Question: Are you currently residing in the Philippines?

- Yes, I am currently residing and staying in the Philippines
- No, I am in the Philippines right now but only for vacation
- No, I am currently residing outside the Philippines

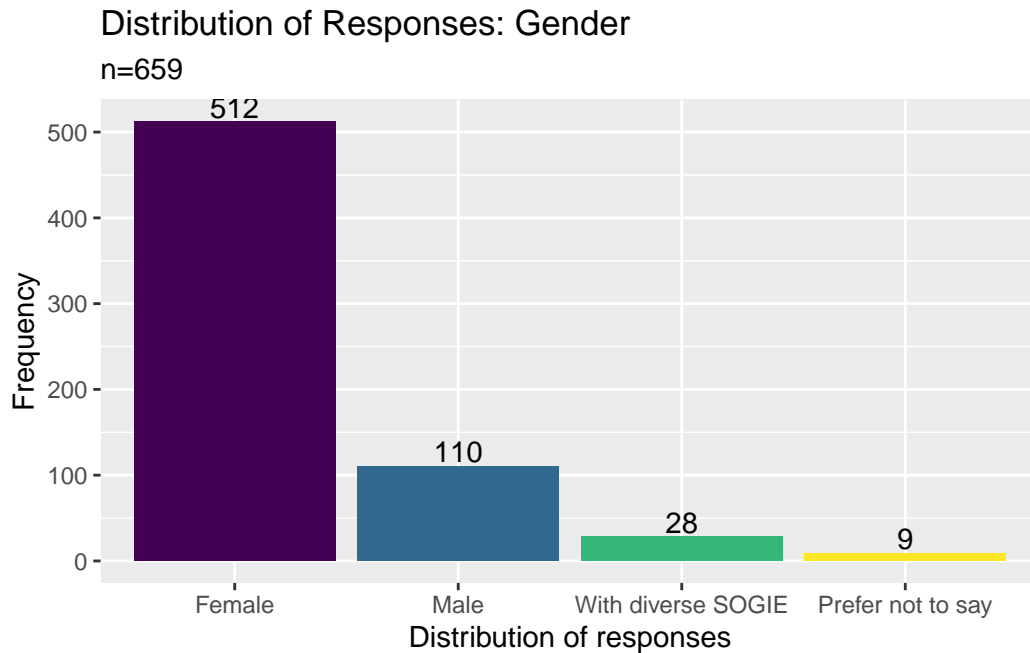
```
# Change the factor type
dataset$philippines <- as.factor(dataset$philippines)
# Recode responses
dataset$philippines <- fct_recode(dataset$philippines, "Residing in the Philippines" = "Yes")
# Rearrange the responses
dataset$philippines <- factor(dataset$philippines, levels = c("Residing in the Philippines", "No"))
# Graph the data
dataset %>%
  filter(!is.na(philippines)) %>%
  ggplot(aes(x = philippines, fill = philippines)) + geom_bar() + geom_text(aes(label = ..count..))
```

Question: Kindly indicate which gender you identify the most

- Man
- Woman
- LGBTQ+
- Prefer not to say

```
# Change the factor type
dataset$gender <- as.factor(dataset$gender)
# Recode responses
dataset$gender <- fct_recode(dataset$gender, "With diverse SOGIE" = "LGBTQ+", "Male" = "Ma
# Rearrange the responses
dataset$gender <- factor(dataset$gender, levels = c("Female", "Male", "With diverse SOGIE")
# Graph the data
dataset %>%
filter(!is.na(gender)) %>%
ggplot(aes(x = gender, fill = gender)) + geom_bar() + geom_text(aes(label = ..count..), st
```

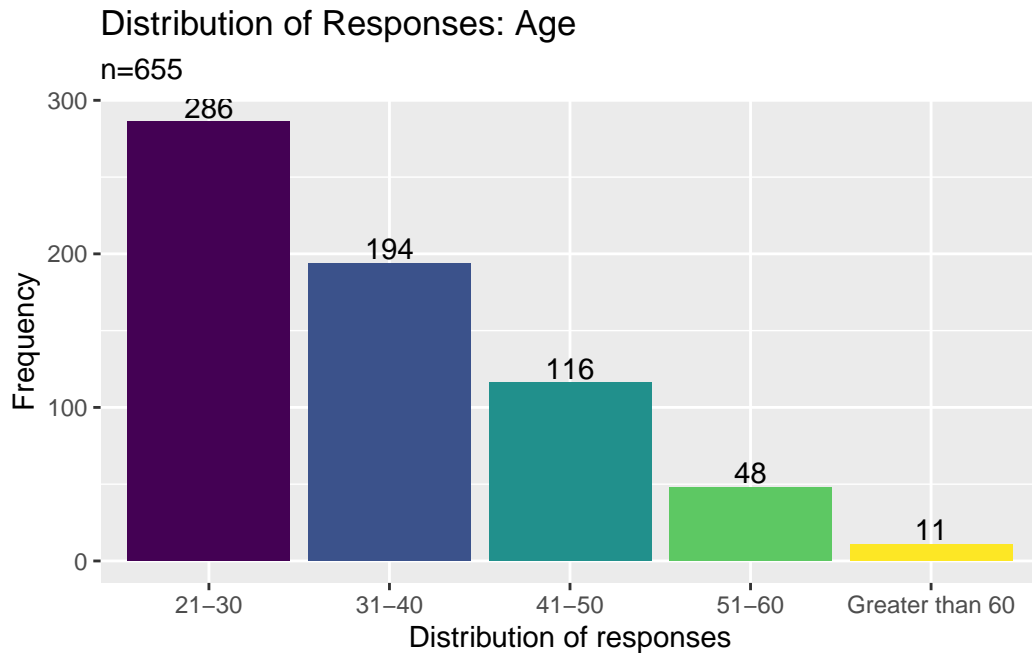


Question: Kindly indicate your age

```
# Inspect the responses
summary(dataset$age)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
18.0	25.0	31.0	33.8	40.0	100.0	24

```
# Create a new column: age_group
age_group <- cut(dataset$age, c(20,30,40,50,60,70,80), right = FALSE)
# Attach age_group to dataset
dataset <- cbind(dataset, age_group)
# Recode responses
dataset$age_group <- fct_recode(dataset$age_group, "21-30" = "[20,30)", "31-40" = "[30,40)
# Graph the data: bar graph
dataset %>%
filter(!is.na(age_group)) %>%
ggplot(aes(x = age_group, fill = age_group)) + geom_bar() + geom_text(aes(label = ..count..
```



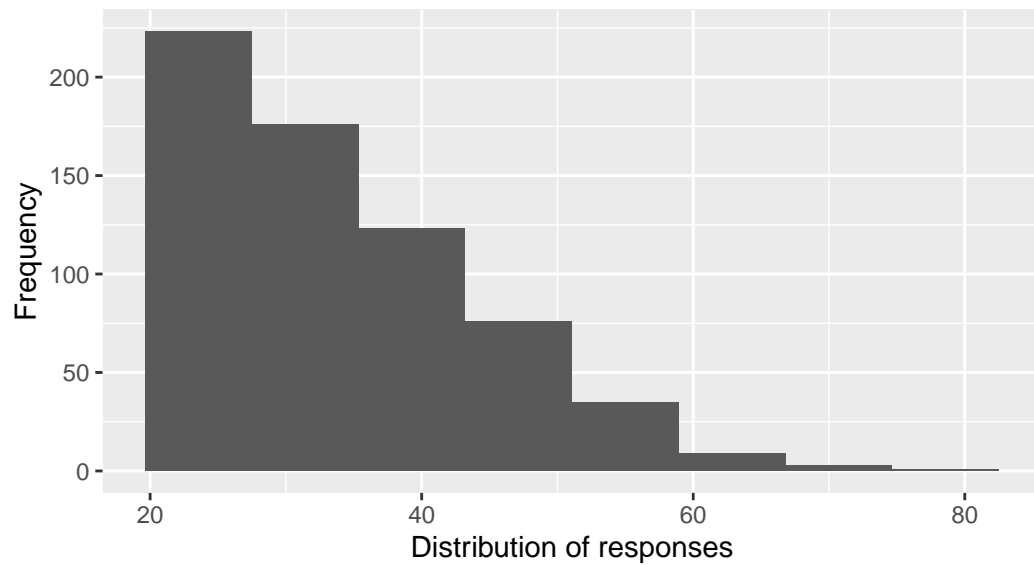
```
# Graph the data: histogram
dataset %>%
  filter(!is.na(age)) %>%
  filter(age > 20) %>%
  filter(age < 90) %>%
  ggplot(aes(x = age, fill = age)) + geom_histogram(bins = 8) + xlab("Distribution of responses")
```

Warning: The following aesthetics were dropped during statistical transformation: fill

- i This can happen when ggplot fails to infer the correct grouping structure in the data.
- i Did you forget to specify a `group` aesthetic or to convert a numerical variable into a factor?

Distribution of Responses: Age

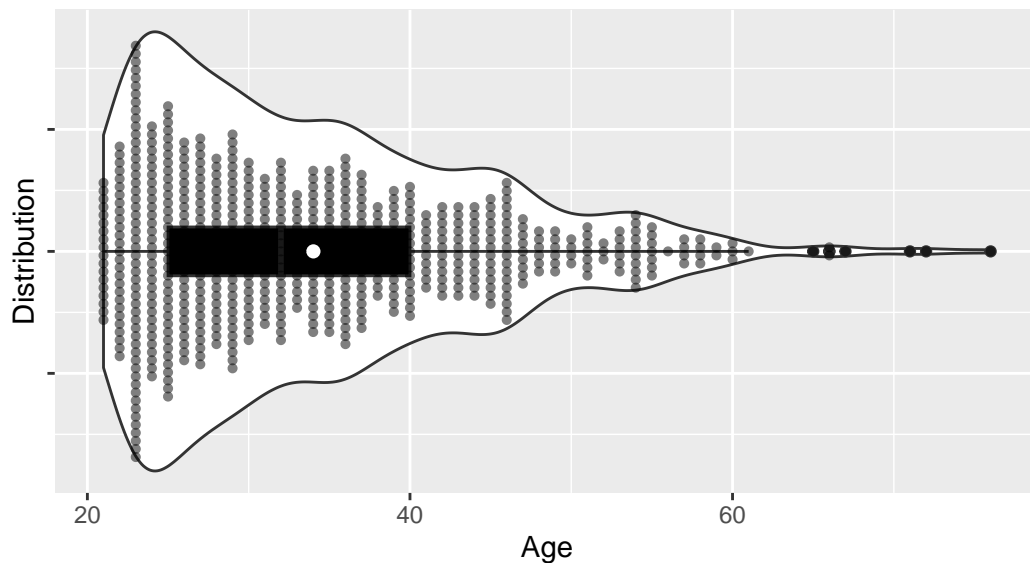
n=645



```
# Graph the data: boxplot + violin plot + dot plot
dataset %>%
  filter(!is.na(age)) %>%
  filter(age > 20) %>%
  filter(age < 90) %>%
  ggplot(aes(x = 1, y = age)) + geom_violin(adjust=.7) + geom_boxplot(width = .1, fill = "bl
```

Distribution of Responses: Age(Boxplot)

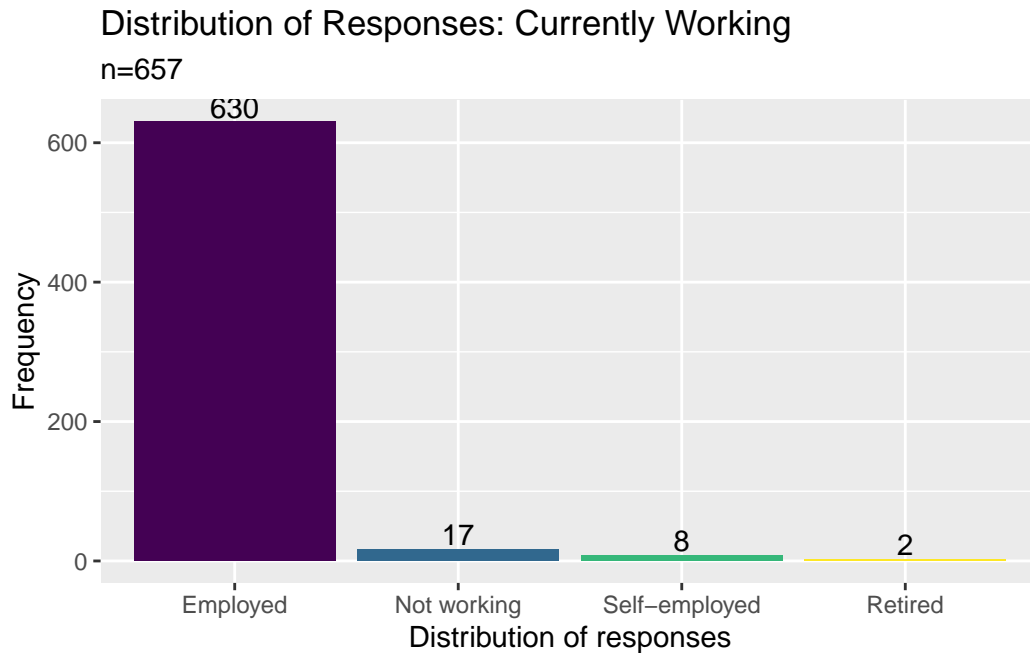
n=646



Question: Are you currently employed and/or working?

- Yes, I am currently employed by an institution, agency, individual or similar.
- Yes, I am currently self-employed and/or doing freelance and consultancy work.
- No, I am currently not working for reasons other than retirement.
- No, I am a retiree.

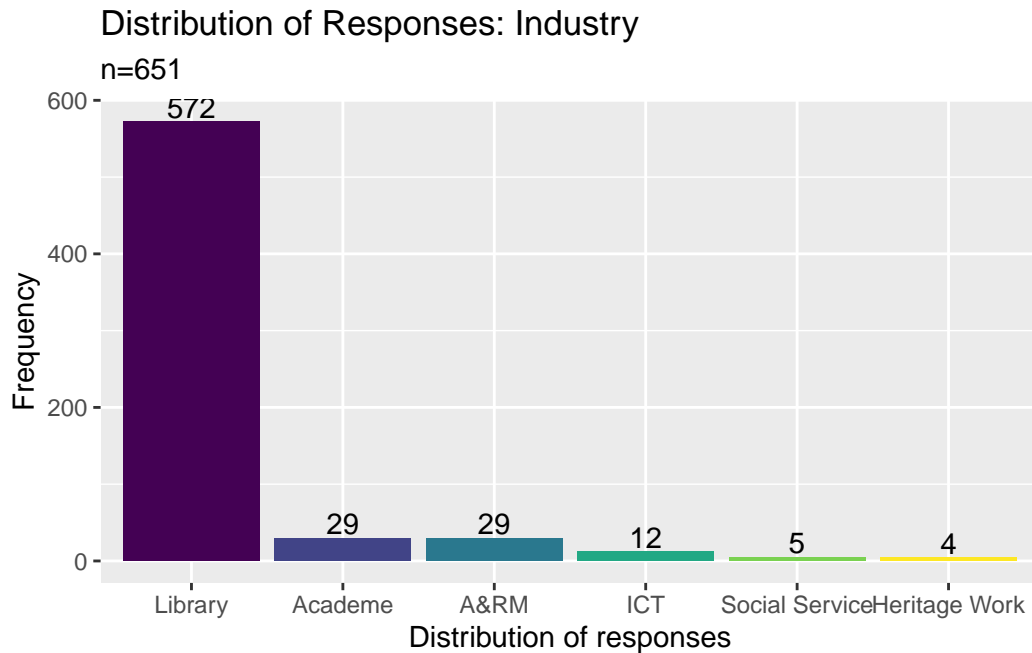
```
# Change the factor type
dataset$working <- as.factor(dataset$working)
# Recode responses
dataset$working <- fct_recode(dataset$working, "Retired" = 'No, I am a retiree.', "Not working" = 'No, I am currently not working for reasons other than retirement.')
# Rearrange the responses
dataset$working <- factor(dataset$working, levels = c('Employed', 'Not working', 'Self-employed'))
# Graph the data
dataset %>%
  filter(!is.na(working)) %>%
  ggplot(aes(x = working, fill = working)) + geom_bar() + geom_text(aes(label = ..count..),
```



Question: Kindly indicate the setting, field or industry you are primarily working in

- library
- archives and/or records management
- academic institution (non-librarian position)
- information technology, computer science, software / database engineering and other similar line of work
- museum and heritage work
- social services

```
# Change the factor type
dataset$industry <- as.factor(dataset$industry)
# Recode responses
dataset$industry <- fct_recode(dataset$industry, "Academe" = 'academic institution (non-li
# Rearrange the responses
dataset$industry <- factor(dataset$industry, levels = c('Library', 'Academe', 'A&RM', 'ICT
# Graph the data
dataset %>%
filter(!is.na(industry)) %>%
ggplot(aes(x = industry, fill = industry)) + geom_bar() + geom_text(aes(label = ..count..))
```



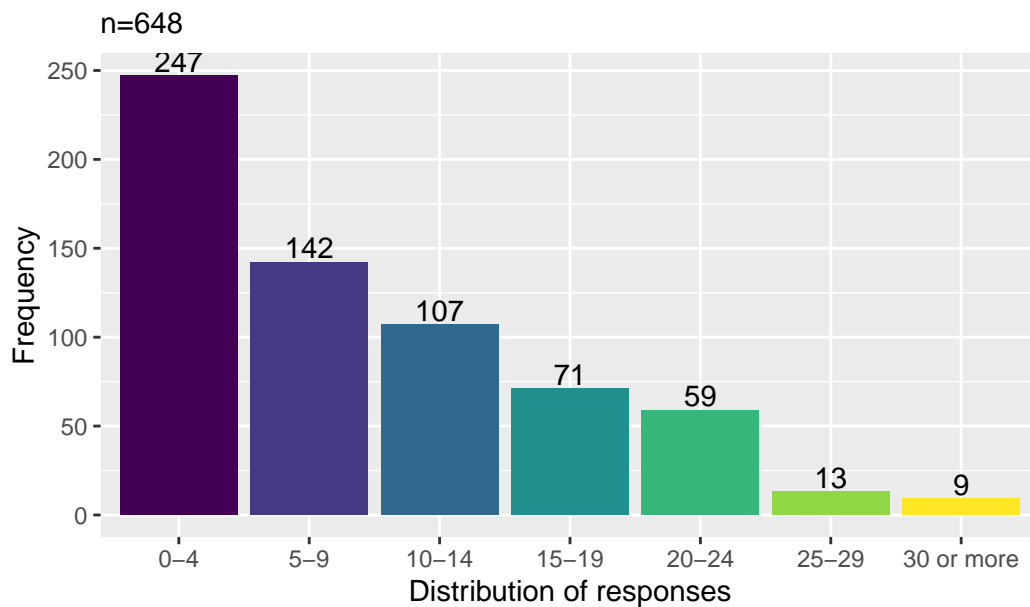
Question: How long have you been working as a professional after passing the licensure exam for librarians (cumulative years)

```
# Inspect the responses
summary(dataset$years_service)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.00	3.00	7.00	9.21	14.00	60.00	34

```
# Create a new column: years_service_group
years_service_group <- cut(dataset$years_service, c(0,5,10,15,20,25,30,35,40), right = FALSE)
# Attach years_service_group to dataset
dataset <- cbind(dataset, years_service_group)
# Recode responses
dataset$years_service_group <- fct_recode(dataset$years_service_group, "0-4" = "[0,5)", "5-9" = "[5,10)", "10-14" = "[10,15)", "15-19" = "[15,20)", "20-24" = "[20,25)", "25-29" = "[25,30)", "30-34" = "[30,35)", "35-39" = "[35,40)", "40+" = "[40,61)")
# Graph the data: bar graph
dataset %>%
  filter(!is.na(years_service_group)) %>%
  ggplot(aes(x = years_service_group, fill = years_service_group)) + geom_bar() + geom_text(aes(x = years_service_group, y = count, label = count))
```

Distribution of Responses: Years of Service



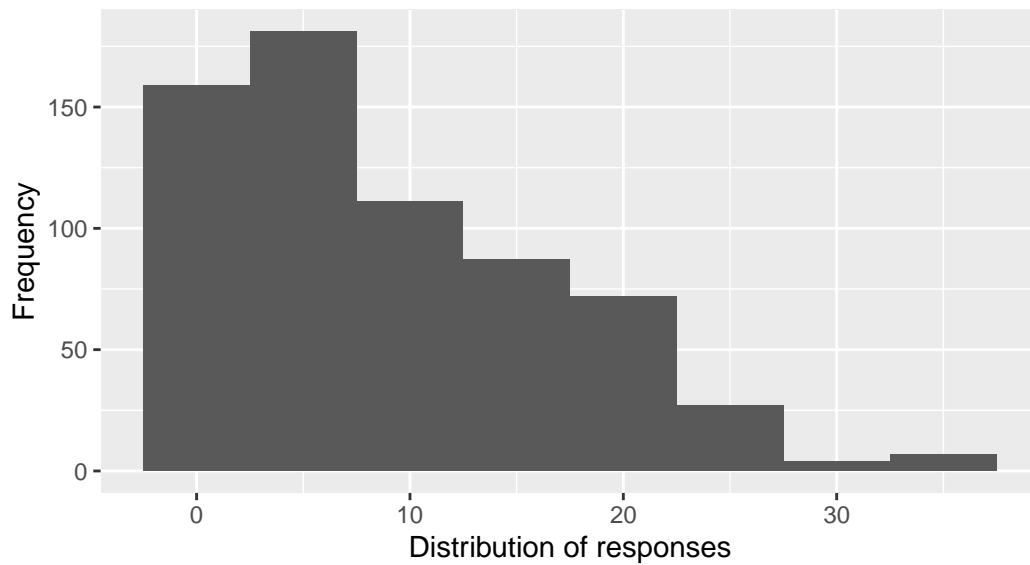
```
# Graph the data: histogram
dataset %>%
  filter(!is.na(years_service)) %>%
  filter(years_service < 40) %>%
  ggplot(aes(x = years_service, , fill = years_service)) + geom_histogram(bins = 8) + xlab("Years of Service")
```

Warning: The following aesthetics were dropped during statistical transformation: fill

- i This can happen when ggplot fails to infer the correct grouping structure in the data.
- i Did you forget to specify a `group` aesthetic or to convert a numerical variable into a factor?

Distribution of Responses: Years of Service

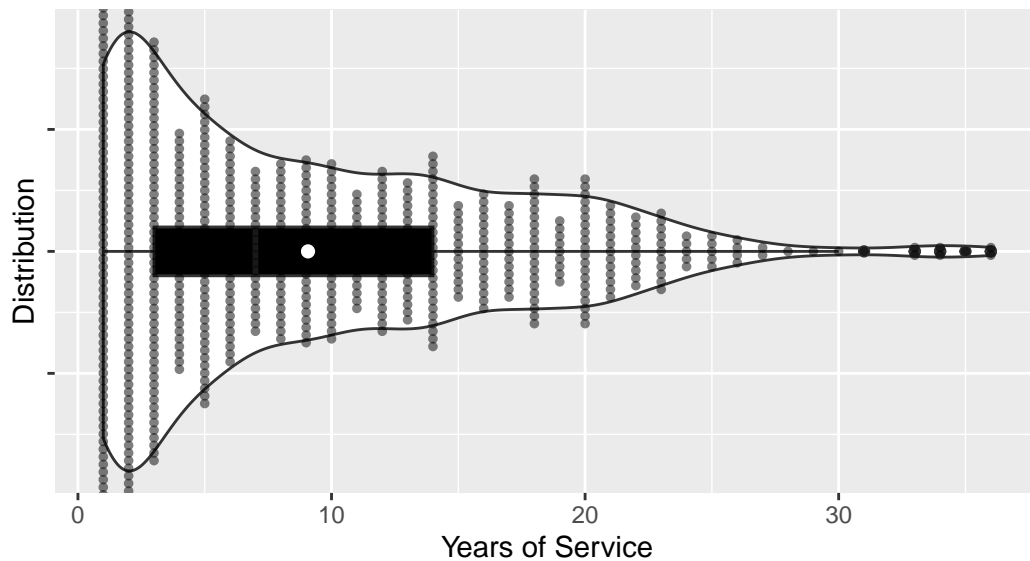
n=648



```
# Graph the data: boxplot + violin plot + dot plot
dataset %>%
  filter(!is.na(years_service)) %>%
  filter(years_service < 40) %>%
  ggplot(aes(x = 1, y = years_service)) + geom_violin(adjust=.7) + geom_boxplot(width = .1,
```

Distribution of Responses: Years of Service(Boxplot)

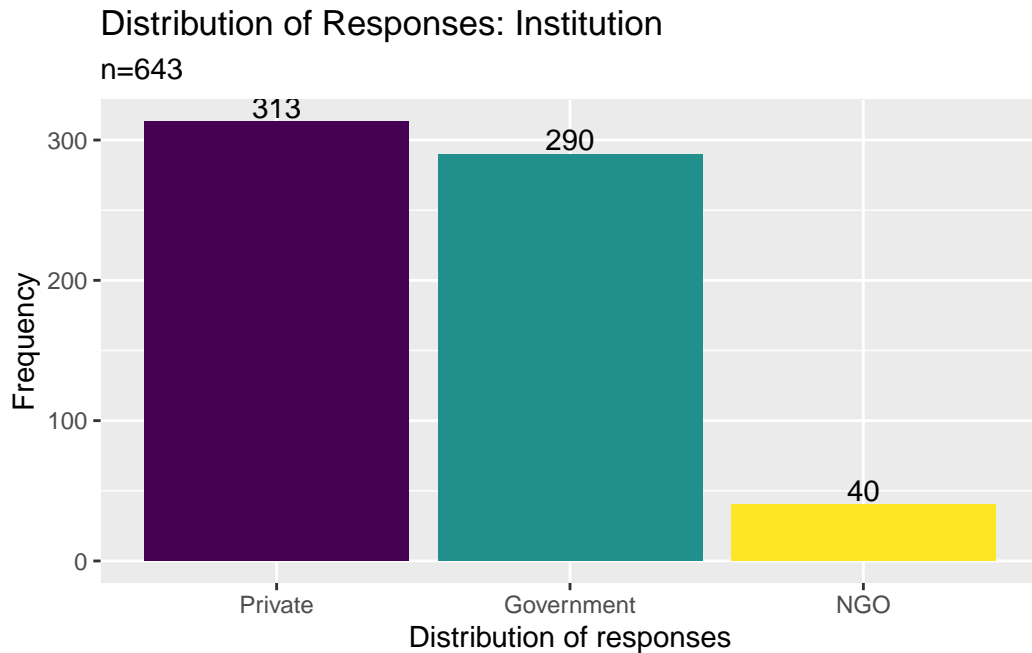
n=648



Question: Kindly indicate the type of institution (in terms of ownership) you are primarily working in

- government
- private (in-house)
- private (BPO)
- non-profit

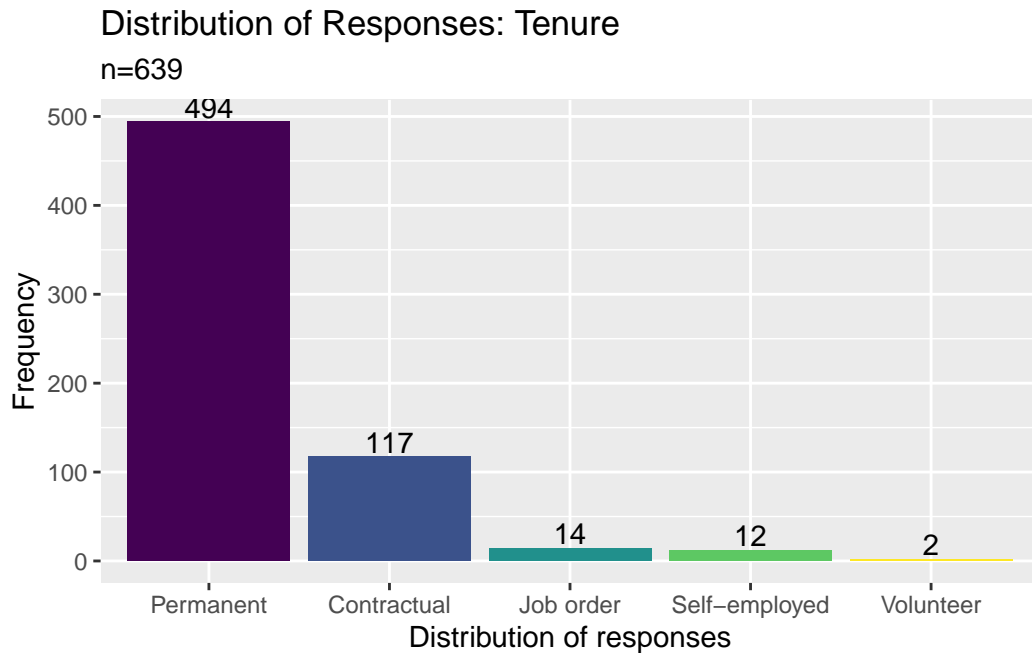
```
# Change the factor type
dataset$institution <- as.factor(dataset$institution)
# Recode responses
dataset$institution <- fct_recode(dataset$institution, "Government" = 'government', "NGO"
# Rearrange the responses
dataset$institution <- factor(dataset$institution, levels = c('Private', 'Government', 'NGO')
# Graph the data
dataset %>%
filter(!is.na(institution)) %>%
ggplot(aes(x = institution, fill = institution)) + geom_bar() + geom_text(aes(label = ..co
```



Question: Kindly indicate your current tenure status

- permanent
- contractual
- voluntary work
- job order (J.O.)
- self-employed, freelancing or working with multiple institutions

```
# Change the factor type
dataset$tenure <- as.factor(dataset$tenure)
# Recode responses
dataset$tenure <- fct_recode(dataset$tenure, "Contractual" = 'contractual', "Job order" = 
# Rearrange the responses
dataset$tenure <- factor(dataset$tenure, levels = c('Permanent', 'Contractual', 'Job order
# Graph the data
dataset %>%
filter(!is.na(tenure)) %>%
ggplot(aes(x = tenure, fill = tenure)) + geom_bar() + geom_text(aes(label = ..count..), st
```



Question: Kindly indicate your current monthly gross salary or income range (before deductions)

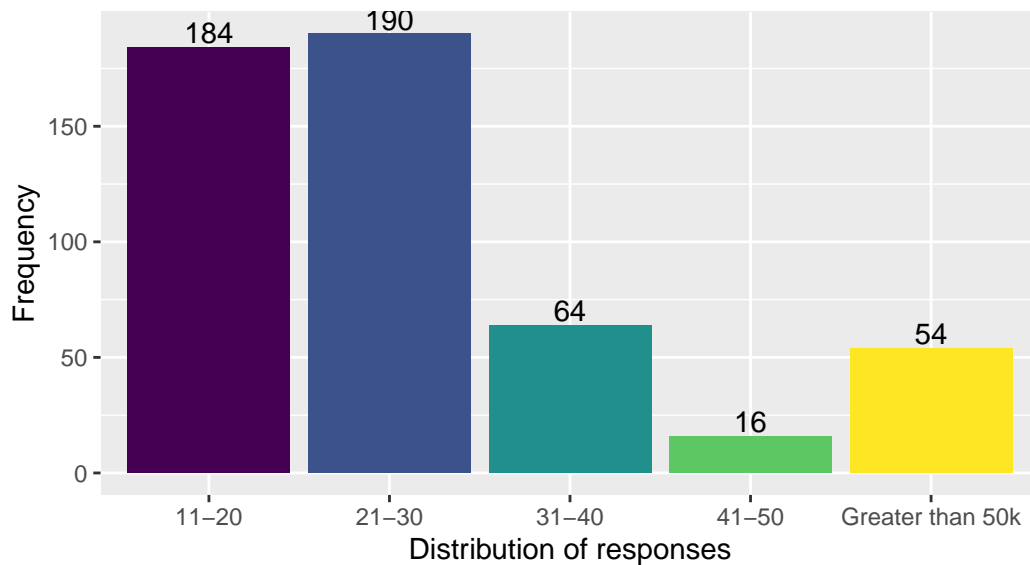
```
# Inspect the responses
summary(dataset$gross_salary)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.0	17.0	22.0	26.9	30.0	100.0	122

```
# Create a new column: gross_salary_group
gross_salary_group <- cut(dataset$gross_salary, c(10,20,30,40,50,60,70,80,90,100), right =
# Attach gross_salary_group to dataset
dataset <- cbind(dataset, gross_salary_group)
# Recode responses
dataset$gross_salary_group <- fct_recode(dataset$gross_salary_group, "11-20" = '(10,20]',
# Graph the data: bar graph
dataset %>%
filter(!is.na(gross_salary_group)) %>%
ggplot(aes(x = gross_salary_group, fill = gross_salary_group)) + geom_bar() + geom_text(aes
```

Distribution of Responses: Gross Salary

n=562



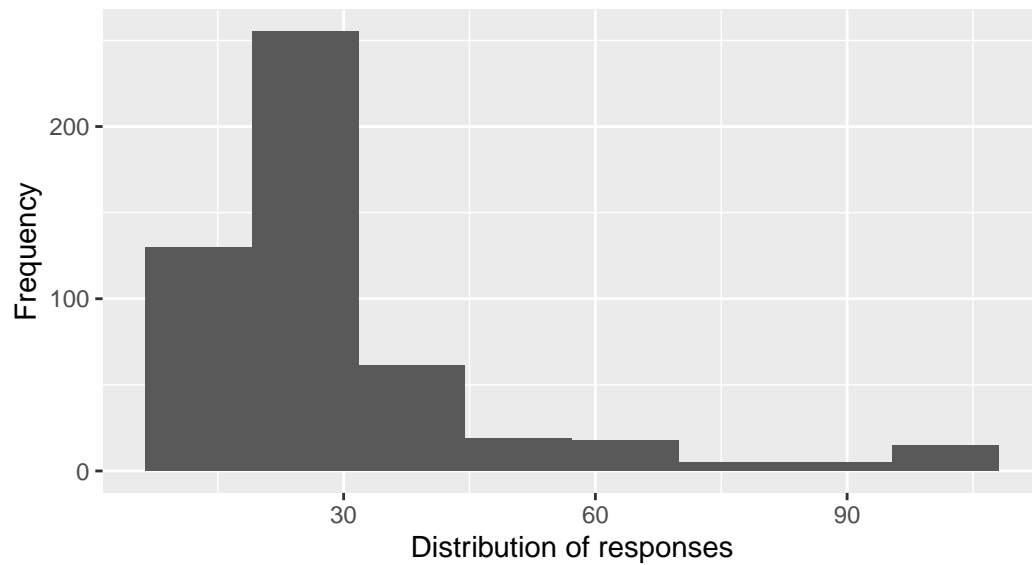
```
# Graph the data: histogram
dataset %>%
  filter(!is.na(gross_salary)) %>%
  filter(gross_salary > 10) %>%
  ggplot(aes(x = gross_salary, fill = gross_salary)) + geom_histogram(bins = 8) + xlab("Dist
```

Warning: The following aesthetics were dropped during statistical transformation: fill

- i This can happen when ggplot fails to infer the correct grouping structure in the data.
- i Did you forget to specify a `group` aesthetic or to convert a numerical variable into a factor?

Distribution of Responses: Gross Salary

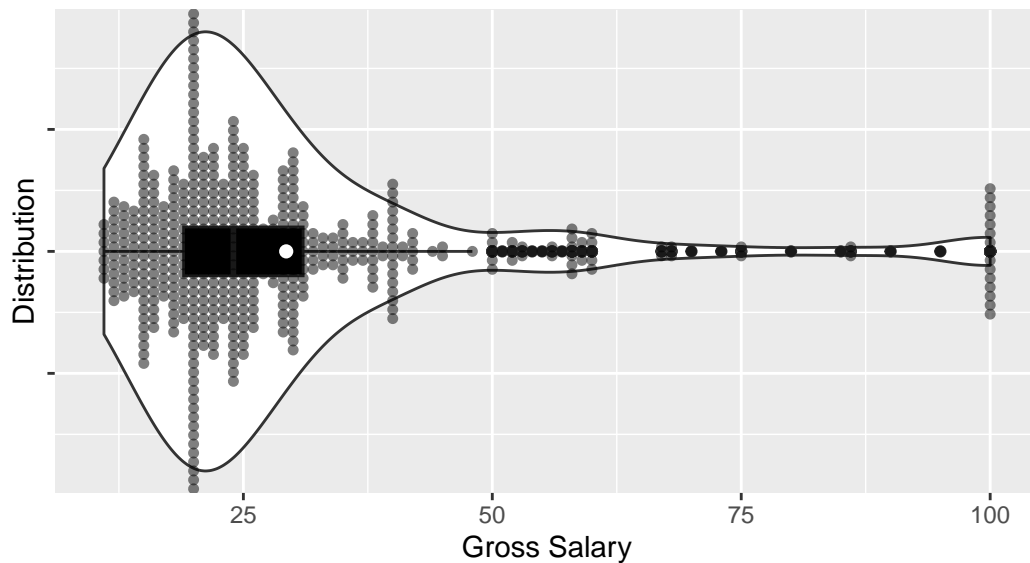
n=508



```
# Graph the data: boxplot + violin plot + dot plot
dataset %>%
  filter(!is.na(gross_salary)) %>%
  filter(gross_salary > 10) %>%
  ggplot(aes(x = 1, y = gross_salary)) + geom_violin(adjust=2) + geom_boxplot(width = .1, fi
```

Distribution of Responses: Gross Salary(Boxplot)

n=508



Kindly indicate your current monthly net salary or income range (after deductions)

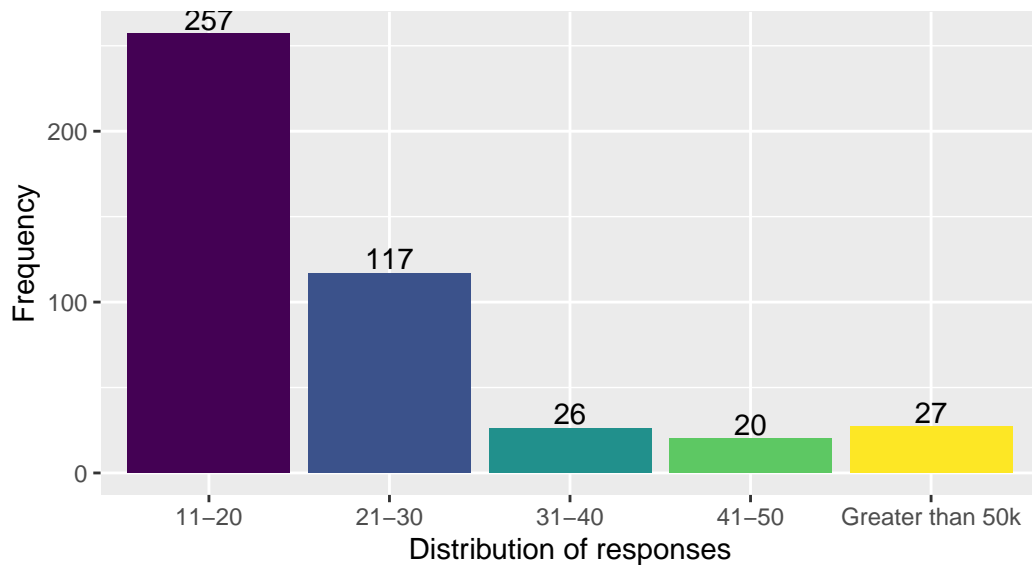
```
# Inspect the responses
summary(dataset$net_salary)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.0	12.0	18.0	21.1	24.0	100.0	134

```
# Create a new column: gross_salary_group
net_salary_group <- cut(dataset$net_salary, c(10,20,30,40,50,60,70,80,90,100), right = TRUE)
# Attach net_salary_group to dataset
dataset <- cbind(dataset, net_salary_group)
# Recode responses
dataset$net_salary_group <- fct_recode(dataset$net_salary_group, "11-20" = "(10,20]", "21-30" = "(20,30]", "31-40" = "(30,40]", "41-50" = "(40,50]", "51-60" = "(50,60]", "61-70" = "(60,70]", "71-80" = "(70,80]", "81-90" = "(80,90]", "91-100" = "(90,100]", "101+" = "(100,1000]")
# Graph the data: bar graph
dataset %>%
  filter(!is.na(net_salary_group)) %>%
  ggplot(aes(x = net_salary_group, fill = net_salary_group)) + geom_bar() + geom_text(aes(label = paste0("n = ", sum(dataset$net_salary_group == net_salary_group)), y = 10))
```

Distribution of Responses: Net Salary

n=550



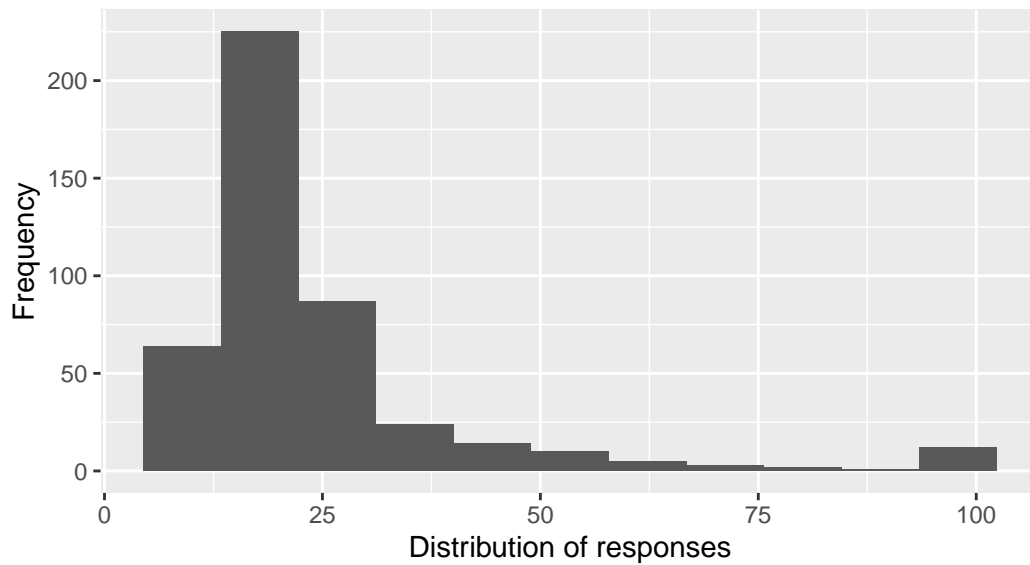
```
# Graph the data: bar graph
dataset %>%
  filter(!is.na(net_salary)) %>%
  filter(net_salary > 10) %>%
  ggplot(aes(x = net_salary, fill = net_salary)) + geom_histogram(bins = 11) + xlab("Distrib
```

Warning: The following aesthetics were dropped during statistical transformation: fill

- i This can happen when ggplot fails to infer the correct grouping structure in the data.
- i Did you forget to specify a `group` aesthetic or to convert a numerical variable into a factor?

Distribution of Responses: Net Salary

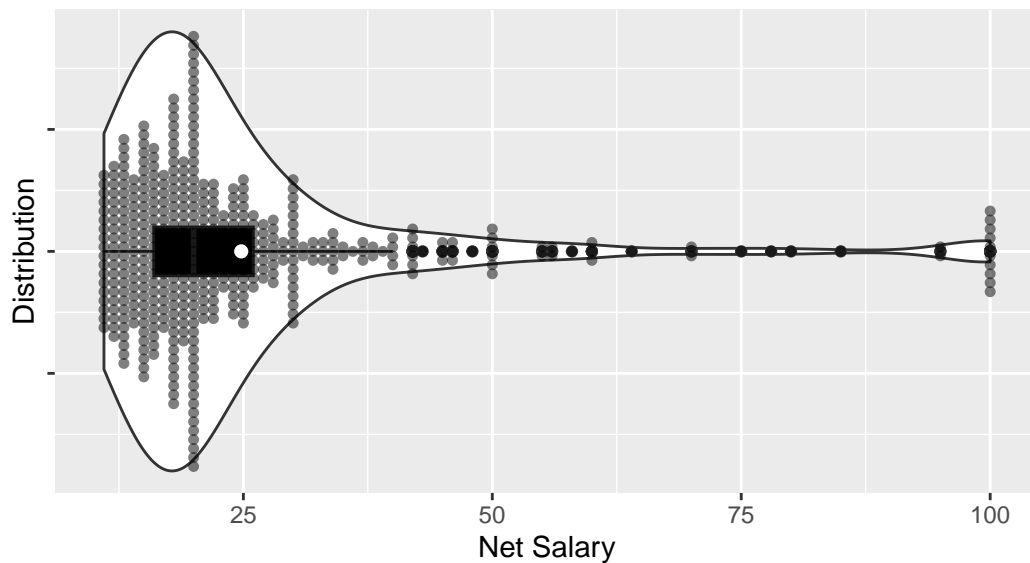
n=446



```
# Graph the data: boxplot + violin plot + dot plot
dataset %>%
  filter(!is.na(net_salary)) %>%
  filter(net_salary > 10) %>%
  ggplot(aes(x = 1, y = net_salary)) + geom_violin(adjust=2) + geom_boxplot(width = .1, fill
```

Distribution of Responses: Net Salary(Boxplot)

n=446

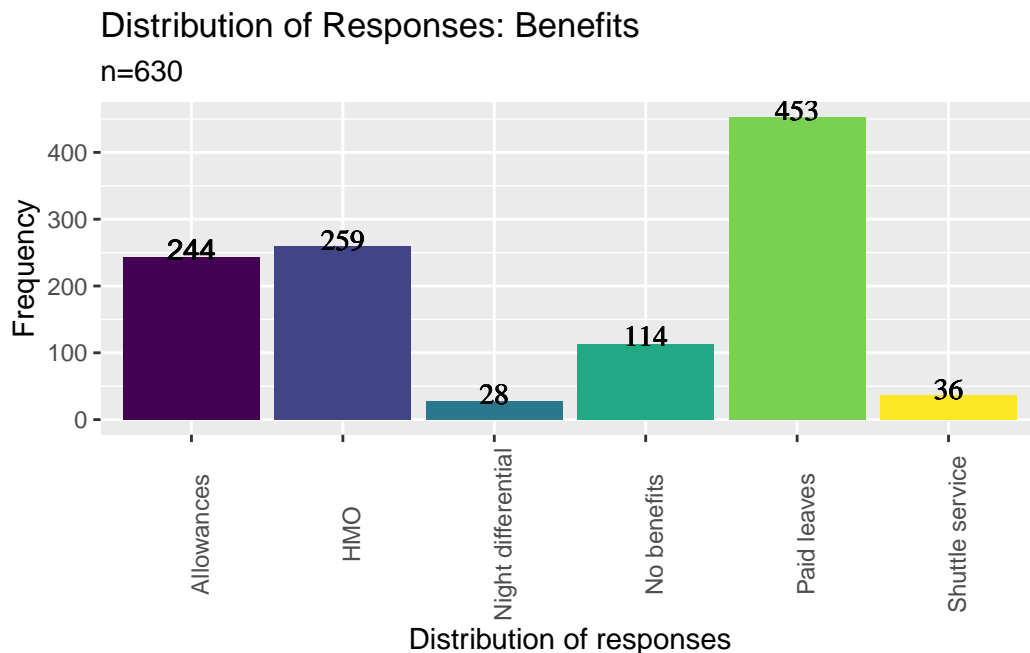


Question: Kindly indicate additional benefits, if any (please check all that apply)

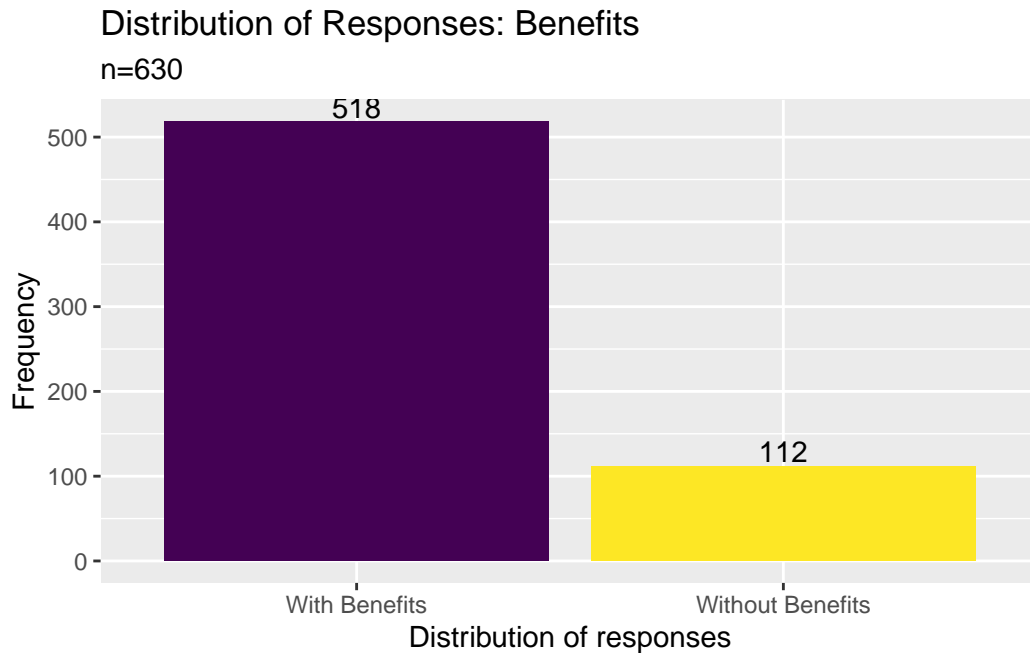
- free shuttle service
- night differential pay
- paid vacation and/or sick leaves
- health card and/or insurance other than PhilHealth
- monetary allowances
- NONE / Not Applicable

```
# Combine columns: benefits1, benefits2, benefits3, benefits4, benefits5
dataset$benefits <- paste(dataset$benefits1, dataset$benefits2, dataset$benefits3, dataset$benefits4, dataset$benefits5)
#Create new columns: shuttle, night_differential, paid_leaves, hmo, allowances, no_benefit
dataset <- transform(dataset, shuttle = grepl("shuttle", benefits), night_differential = grepl("night_differential", benefits),
# Summarize by counting the TRUE values per benefits
dataset %>%
summarise("Shuttle service" = sum(shuttle == TRUE), "Night differential" = sum(night_differential == TRUE),
gather() %>%
ggplot(aes(x = key, y = value, fill = key)) + geom_col(position = "dodge") + xlab("Distribution of Responses: Net Salary")
geom_text(x = 1, y = 254, label = "244", colour = "black", size = 4) +
geom_text(x = 2, y = 269, label = "259", family = "serif", colour = "black", size = 4) +
geom_text(x = 3, y = 38, label = "28", family = "serif", colour = "black", size = 4) +
```

```
geom_text(x = 4, y = 124, label = "114", family = "serif", colour = "black", size = 4) +
geom_text(x = 5, y = 463, label = "453", family = "serif", colour = "black", size = 4) +
geom_text(x = 6, y = 46, label = "36", family = "serif", colour = "black", size = 4)
```



```
# Change the factor type
dataset$benefits1 <- as.factor(dataset$benefits1)
# Recode responses
dataset$benefits6 <- fct_recode(dataset$benefits1, "With Benefits" = 'free shuttle service
# Rearrange the responses
dataset$benefits6 <- factor(dataset$benefits6, levels = c('With Benefits', 'Without Benefi
# Graph the data
dataset %>%
filter(!is.na(benefits6)) %>%
ggplot(aes(x = benefits6, fill = benefits6)) + geom_bar() + geom_text(aes(label = ..count..
```



Question: Other benefits not mentioned in the previous page

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en")))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}

# Create a text dataset
benefits.text <- paste(dataset$otherbenefits1, dataset$otherbenefits2, dataset$otherbenefits3)

# Create a text corpus
benefit.corpus <- benefits.text %>%
  VectorSource() %>%
  VCorpus()

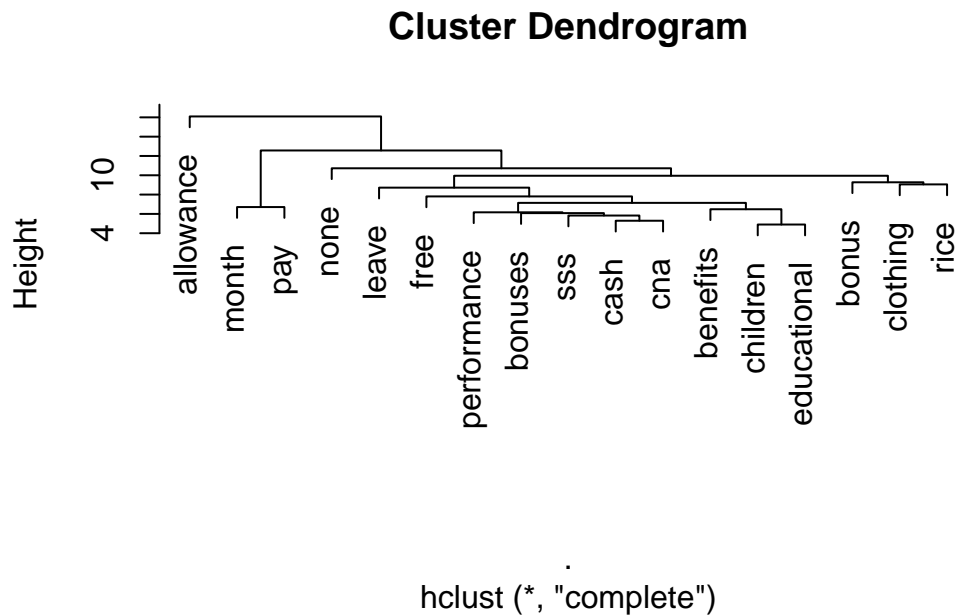
# Clean the created corpus
benefit.corpus <- clean_corpus(benefit.corpus)

# Create a dendrogram
benefit.corpus %>%
```

```

TermDocumentMatrix() %>%
removeSparseTerms(sparse = 0.98) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()

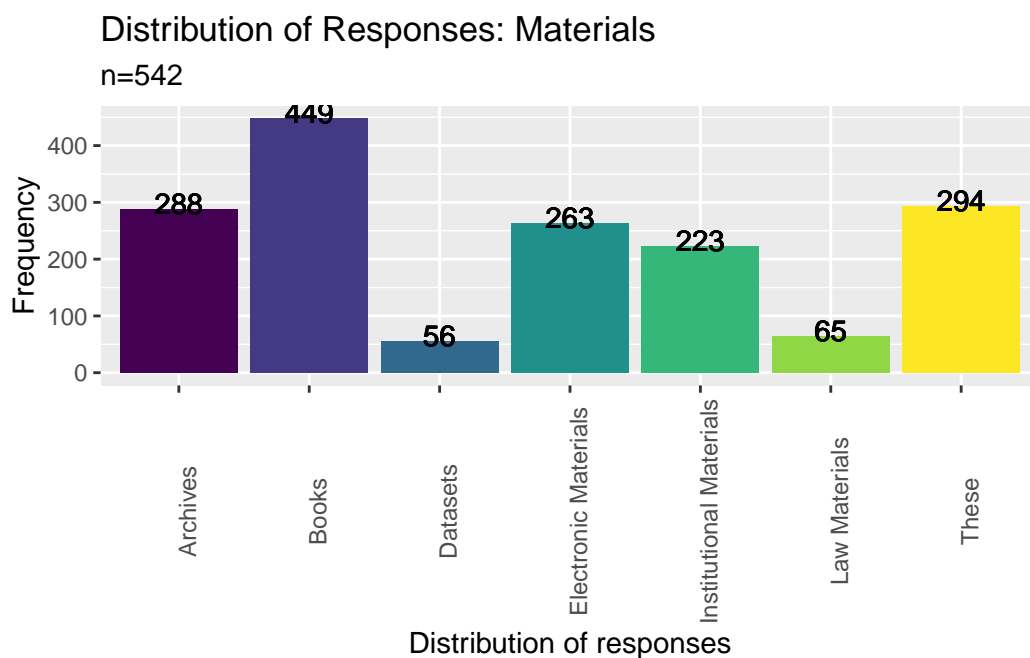
```



Question: What kind of materials are you currently working with? Kindly check all that applies.

- hardcopy theses and dissertations
- hardcopy books / pamphlets / etc
- electronic material (digital titles, digital archives, etc)
- law materials and resources
- archives (physical)
- data sets
- institutional files and records
- NA

```
# Combine columns: materials1, materials2, materials3, materials4, materials5, materials6,
dataset$materials <- paste(dataset$materials1, dataset$materials2, dataset$materials3, dat
# Create new columns: theses, books, electronic_materials, law_materials, datasets, instit
dataset <- transform(dataset, theses = grepl("theses", materials), books = grepl("books",
# Summarize by counting the TRUE values per materials
dataset %>%
summarise("These" = sum(theses == TRUE), "Books" = sum(books == TRUE), "Electronic Material
gather() %>%
ggplot(aes(x = key, y = value, fill = key)) + geom_col(position = "dodge") + xlab("Distrib
geom_text(x = 1, y = 298, label = "288", colour = "black", size = 4) +
geom_text(x = 2, y = 459, label = "449", colour = "black", size = 4) +
geom_text(x = 3, y = 66, label = "56", colour = "black", size = 4) +
geom_text(x = 4, y = 273, label = "263", colour = "black", size = 4) +
geom_text(x = 5, y = 233, label = "223", colour = "black", size = 4) +
geom_text(x = 6, y = 75, label = "65", colour = "black", size = 4) +
geom_text(x = 7, y = 304, label = "294", colour = "black", size = 4)
```

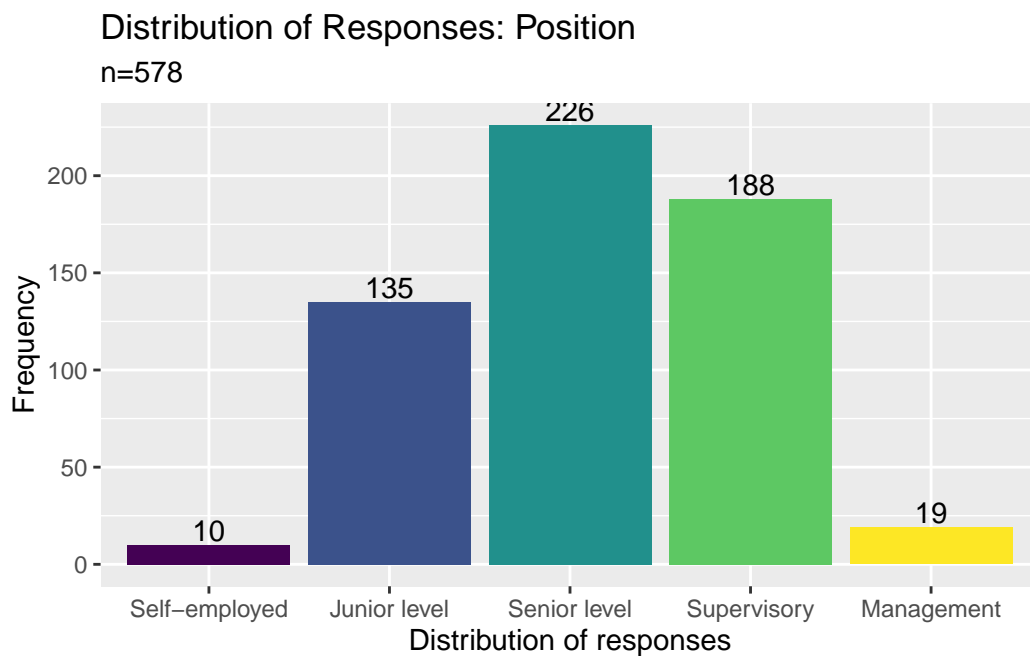


Question: Kindly indicate your current work position

- entry level / junior staff
- regular / senior staff (non-supervisory)

- supervisory / mid-management level
- upper management level / top tier executive
- consultant / freelancer / self-employed

```
# Change the factor type
dataset$position <- as.factor(dataset$position)
# Recode responses
dataset$position <- fct_recode(dataset$position, "Self-employed" = 'consultant / freelancer')
# Graph the data
dataset %>%
  filter(!is.na(position)) %>%
  ggplot(aes(x = position, fill = position)) + geom_bar() + geom_text(aes(label = ..count..))
```



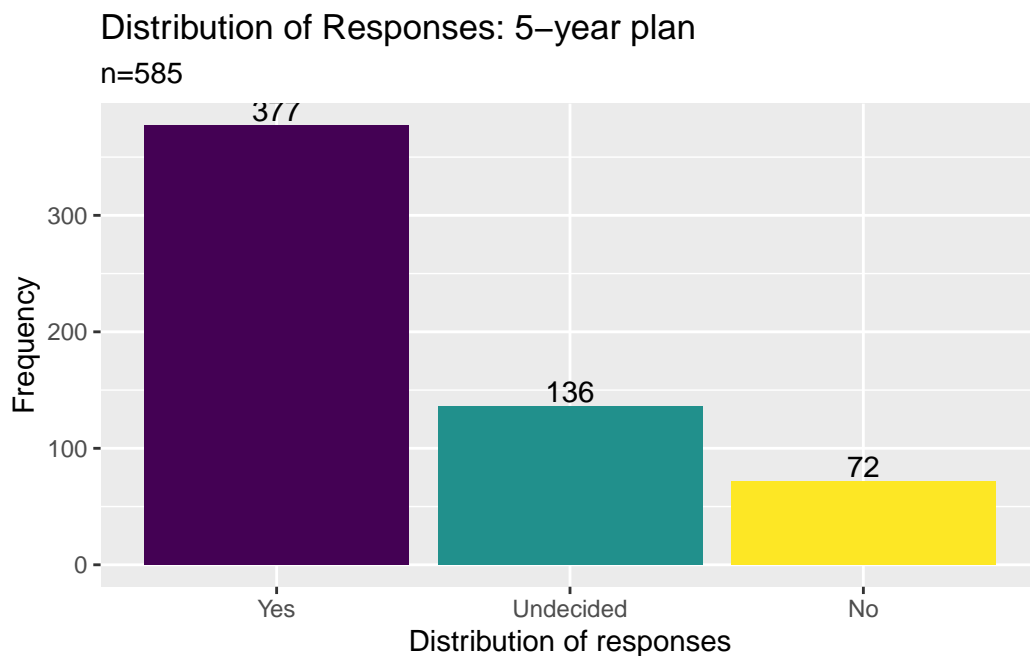
Question: Within the next five (5) years, do you see your self working in the same industry?

- Yes
- No (planning to shift to other fields or industry)
- Maybe / Undecided

```

# Change the factor type
dataset$fiveyears <- as.factor(dataset$fiveyears)
# Recode responses
dataset$fiveyears <- fct_recode(dataset$fiveyears, "Undecided" = 'Maybe / Undecided', "No")
# Rearrange the responses
dataset$fiveyears <- factor(dataset$fiveyears, levels = c('Yes', 'Undecided', 'No'))
# Graph the data
dataset %>%
  filter(!is.na(fiveyears)) %>%
  ggplot(aes(x = fiveyears, fill = fiveyears)) + geom_bar() + geom_text(aes(label = ..count..))

```



Question: If your answer was NO or MAYBE in the previous page, indicate the field you wish to pursue

```

# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "none", "yes", "like", "libra

```

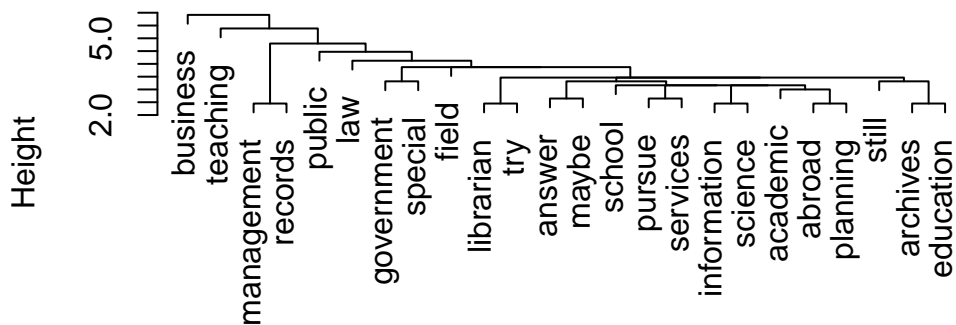


```

    corpus <- tm_map(corpus, stripWhitespace)
    return(corpus)
}
# Create a text corpus
fiveyearplan.corpus <- dataset$pursue %>%
  VectorSource() %>%
  VCorpus()
# Clean the created corpus
fiveyearplan.corpus <- clean_corpus(fiveyearplan.corpus)
# Create a dendrogram
fiveyearplan.corpus %>%
  TermDocumentMatrix() %>%
  removeSparseTerms(sparse = 0.995) %>%
  as.matrix() %>%
  dist() %>%
  hclust() %>%
  plot()

```

Cluster Dendrogram



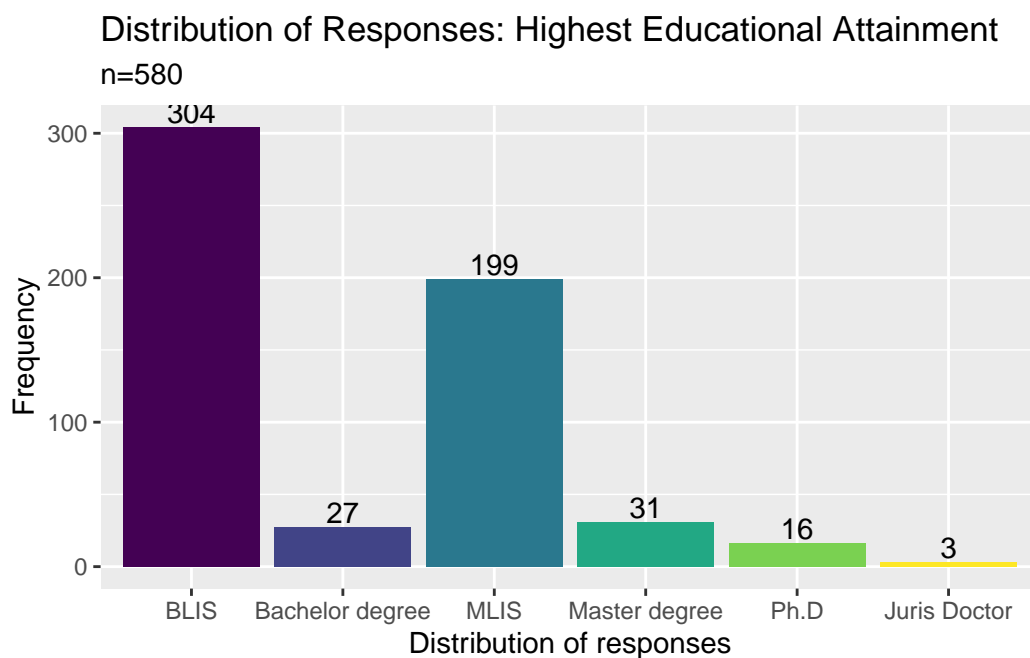
hclust (*, "complete")

Question: Kindly indicate your highest educational attainment

- High School

- Bachelor's degree (BLIS)
- Bachelor's degree (non-BLIS)
- Master's degree (MLIS)
- Master's degree (non-MLIS)
- Juris Doctor / Bachelor of Laws
- Doctor of Philosophy (Ph.D)

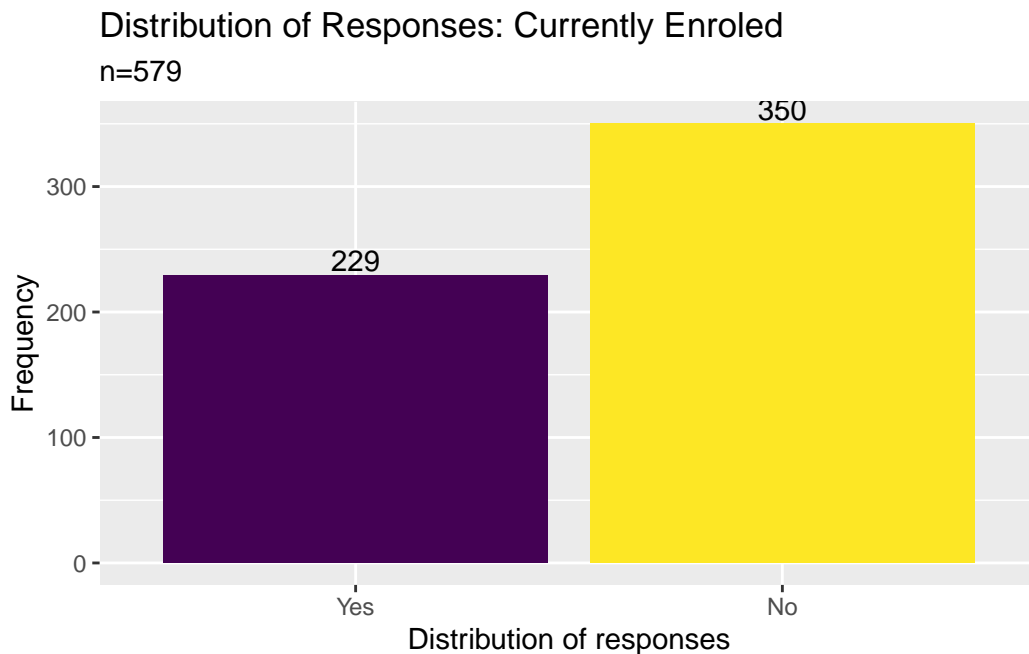
```
# Change the factor type
dataset$education <- as.factor(dataset$education)
# Recode responses
levels(dataset$education)[1] <- "BLIS"
levels(dataset$education)[2] <- "Bachelor degree"
levels(dataset$education)[3] <- "Ph.D"
levels(dataset$education)[4] <- "Juris Doctor"
levels(dataset$education)[5] <- "MLIS"
levels(dataset$education)[6] <- "Master degree"
# Rearrange the responses
dataset$education <- fct_relevel(dataset$education, 'BLIS', 'Bachelor degree', 'MLIS', 'Ma
# Graph the data
dataset %>%
  filter(!is.na(education)) %>%
  ggplot(aes(x = education, fill = education)) + geom_bar() + geom_text(aes(label = ..count..
```



Question: Are you currently enrolled in formal studies? (Short Courses, Certificate Program, Diploma Program, Master Program, PhD)

- Yes
- No

```
# Change the factor type
dataset$enrolled <- as.factor(dataset$enrolled)
# Recode responses
levels(dataset$enrolled)[1] <- "No"
# Rearrange the responses
dataset$enrolled <- fct_relevel(dataset$enrolled, 'Yes', 'No')
# Graph the data
dataset %>%
  filter(!is.na(enrolled)) %>%
  ggplot(aes(x = enrolled, fill = enrolled)) + geom_bar() + geom_text(aes(label = ..count..))
```

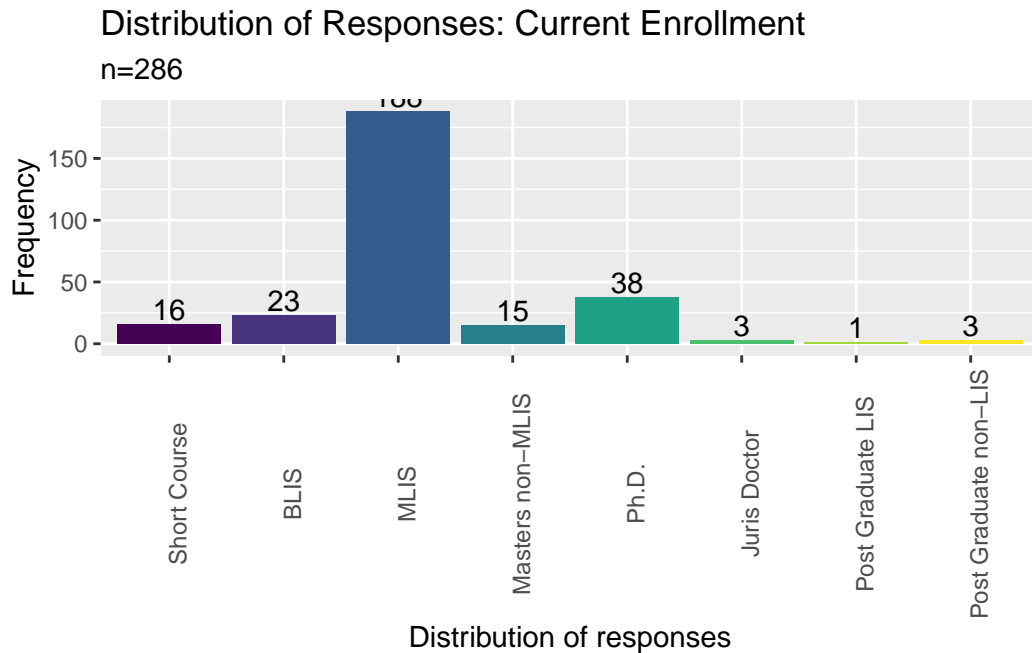


Question: If your answer was YES, indicate the degree you are currently completing

- Short | Certificate | Diploma Program
- Bachelor's degree (BLIS)

- Master's degree (MLIS)
- Master's degree (non-MLIS)
- Juris Doctor / Bachelor of Laws
- Doctor of Philosophy (Ph.D)
- Post-graduate Certificate or Diploma (LIS)
- Post-graduate Certificate or Diploma (non-LIS)

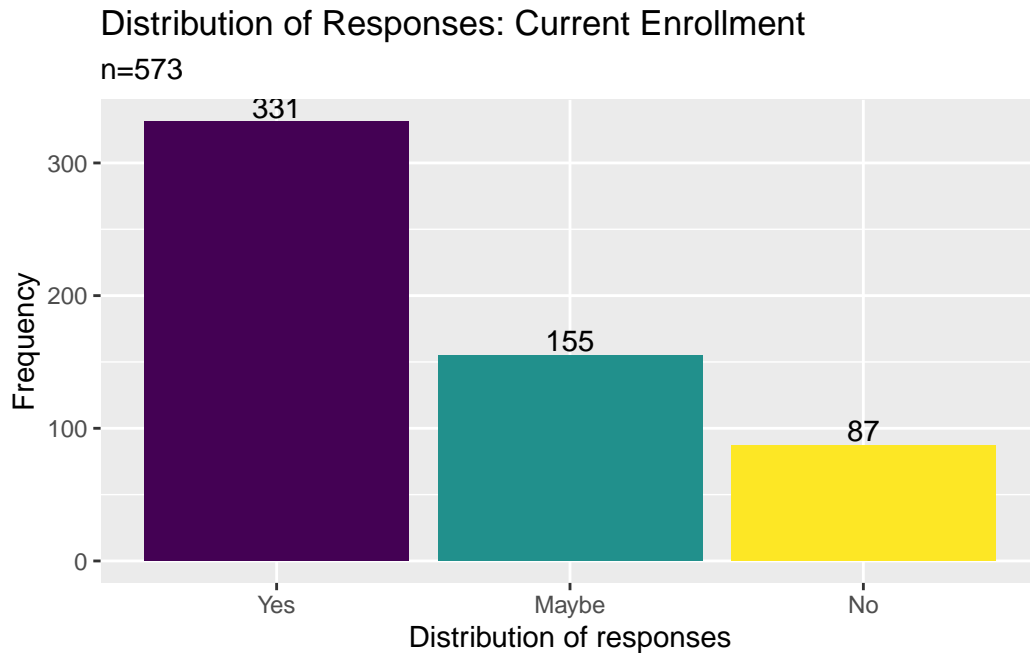
```
# Change the factor type
dataset$completing <- as.factor(dataset$completing)
# Recode responses
levels(dataset$completing)[1] <- "BLIS"
levels(dataset$completing)[2] <- "Ph.D."
levels(dataset$completing)[3] <- "Juris Doctor"
levels(dataset$completing)[4] <- "MLIS"
levels(dataset$completing)[5] <- "Masters non-MLIS"
levels(dataset$completing)[6] <- "Post Graduate LIS"
levels(dataset$completing)[7] <- "Post Graduate non-LIS"
levels(dataset$completing)[8] <- "Short Course"
# Rearrange the responses
dataset$completing <- factor(dataset$completing, levels = c("Short Course", "BLIS", "MLIS"
# Graph the data
dataset %>%
filter(!is.na(completing)) %>%
ggplot(aes(x = completing, fill = completing)) + geom_bar() + geom_text(aes(label = ..count..))
```



Question: Are you planning to study again in a formal university setting within the next five (5) years?

- Yes
- No
- Maybe

```
# Change the factor type
dataset$study_again <- as.factor(dataset$study_again)
# Rearrange the responses
dataset$study_again <- factor(dataset$study_again, levels = c("Yes", "Maybe", "No"))
# Graph the data
dataset %>%
  filter(!is.na(study_again)) %>%
  ggplot(aes(x = study_again, fill = study_again)) + geom_bar() + geom_text(aes(label = ..co
```



Question: If your answer was YES, kindly provide the degree you wish to study within the next five (5) years

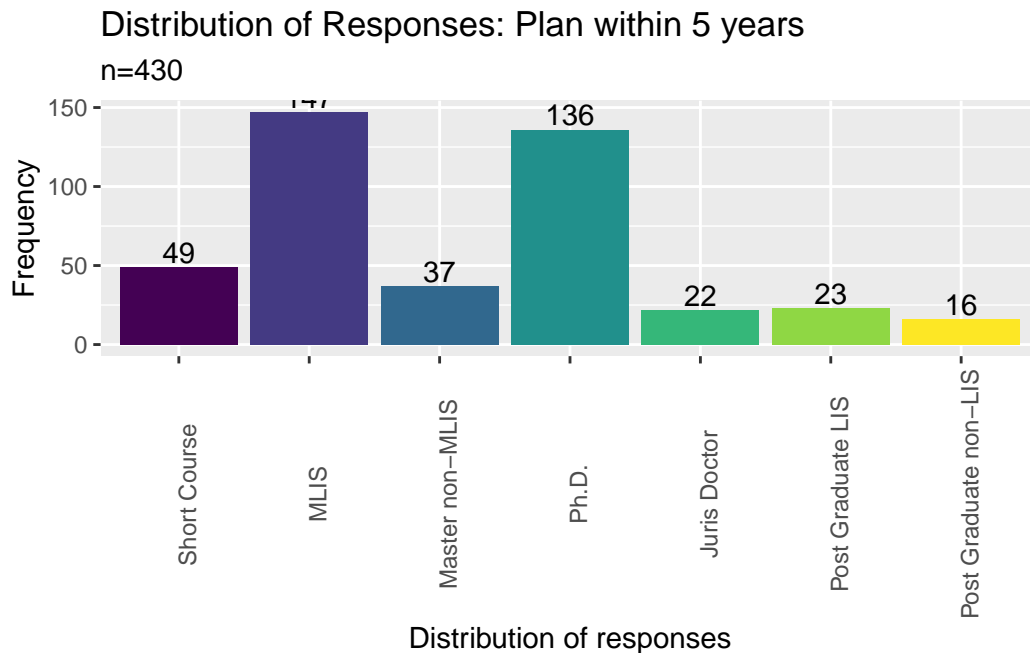
- Short | Certificate | Diploma Program
- Bachelor's degree (BLIS)
- Master's degree (MLIS)
- Master's degree (non-MLIS)
- Juris Doctor / Bachelor of Laws
- Doctor of Philosophy (Ph.D)
- Post-graduate Certificate or Diploma (LIS)
- Post-graduate Certificate or Diploma (non-LIS)

```
# Change the column name of '5yearsdegree'  
colnames(dataset)[40] <- "after5years"  
# Change the factor type  
dataset$after5years <- as.factor(dataset$after5years)  
# Recode responses  
levels(dataset$after5years)[1] <- "Ph.D."  
levels(dataset$after5years)[2] <- "Juris Doctor"  
levels(dataset$after5years)[3] <- "MLIS"  
levels(dataset$after5years)[4] <- "Master non-MLIS"
```

```

levels(dataset$after5years)[5] <- "Post Graduate LIS"
levels(dataset$after5years)[6] <- "Post Graduate non-LIS"
levels(dataset$after5years)[7] <- "Short Course"
# Rearrange the responses
dataset$after5years <- factor(dataset$after5years, levels = c("Short Course", "MLIS", "Mas
# Graph the data
dataset %>%
filter(!is.na(after5years)) %>%
ggplot(aes(x = after5years, fill = after5years)) + geom_bar() + geom_text(aes(label = ..co

```



Question: Where do you usually acquire your CPD trainings in LIS? Kindly check all that applies

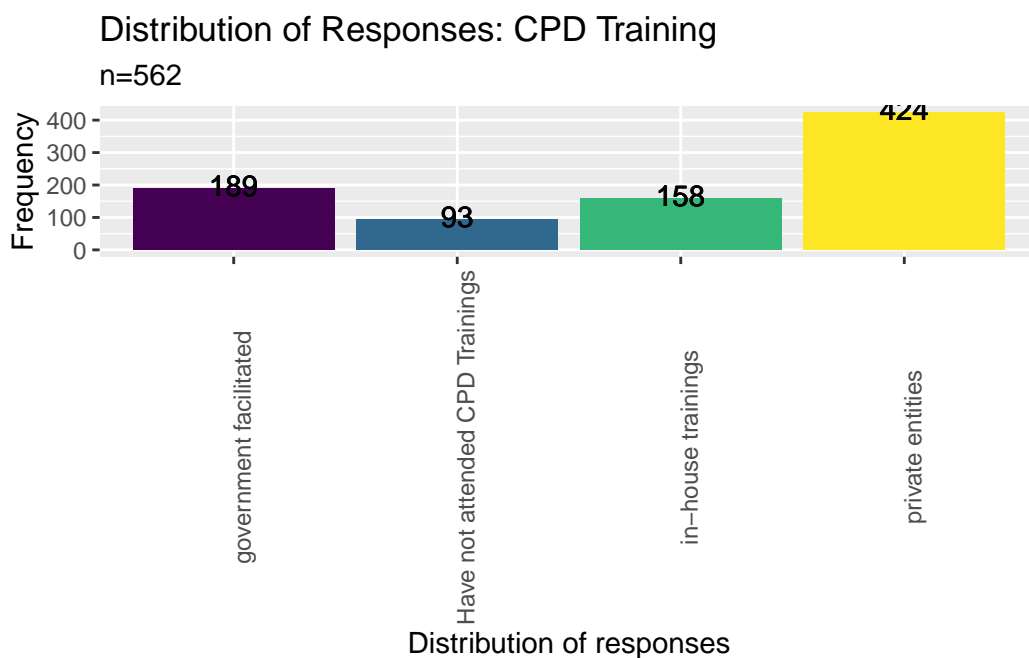
- within the work place (in-house trainings)
- government facilitated trainings and symposiums
- trainings and symposiums facilitated by private entities, individuals and/or corporations
- NA (Have not attended CPD Trainings yet)

```

# Combine columns: cpd1, cpd2, cpd3, cpd4
dataset$cpd <- paste(dataset$cpd1, dataset$cpd2, dataset$cpd3, dataset$cpd4)

```

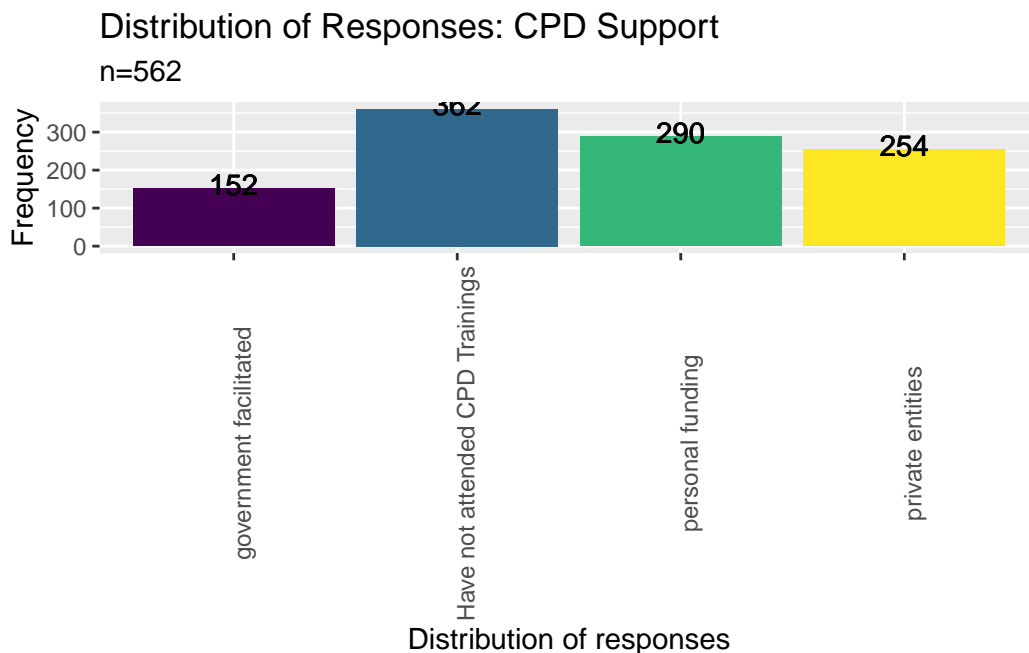
```
# Create new columns: inhouse, government, private, xattended
dataset <- transform(dataset, inhouse = grepl("within the work place", cpd), government =
# Summarize by counting the TRUE values per materials
dataset %>%
summarise("in-house trainings" = sum(inhouse == TRUE), "government facilitated" = sum(gove
gather() %>%
ggplot(aes(x = key, y= value, fill = key)) + geom_col(position = "dodge") + xlab("Distribu
geom_text(x = 1, y = 199, label = "189", colour = "black", size = 4) +
geom_text(x = 2, y = 103, label = "93", colour = "black", size = 4) +
geom_text(x = 3, y = 168, label = "158", colour = "black", size = 4) +
geom_text(x = 4, y = 424, label = "424", colour = "black", size = 4)
```



Where do you usually get funding or support for your trainings in LIS?

- from the work place (private sponsorship)
- from the government (scholarships, free trainings and seminars, etc)
- personal funding (personally attended trainings and symposiums and shouldered most of the expenses)
- NA (Have not attended CPD Trainings yet)


```
# Combine columns: cpdsupport1, cpdsupport2, cpdsupport3
dataset$cpdsupport <- paste(dataset$cpdsupport1, dataset$cpdsupport2, dataset$cpdsupport3)
# Create new columns: support_inhouse, sgovernment, sprivate, sxattended
dataset <- transform(dataset, support_self = grepl("personal", cpdsupport), support_govern
# Summarize by counting the TRUE values per materials
dataset %>%
summarise("personal funding" = sum(support_self == TRUE), "government facilitated" = sum(s
gather() %>%
ggplot(aes(x = key, y= value, fill = key)) + geom_col(position = "dodge") + xlab("Distribu
geom_text(x = 1, y = 162, label = "152", colour = "black", size = 4) +
geom_text(x = 2, y = 372, label = "362", colour = "black", size = 4) +
geom_text(x = 3, y = 300, label = "290", colour = "black", size = 4) +
geom_text(x = 4, y = 264, label = "254", colour = "black", size = 4)
```

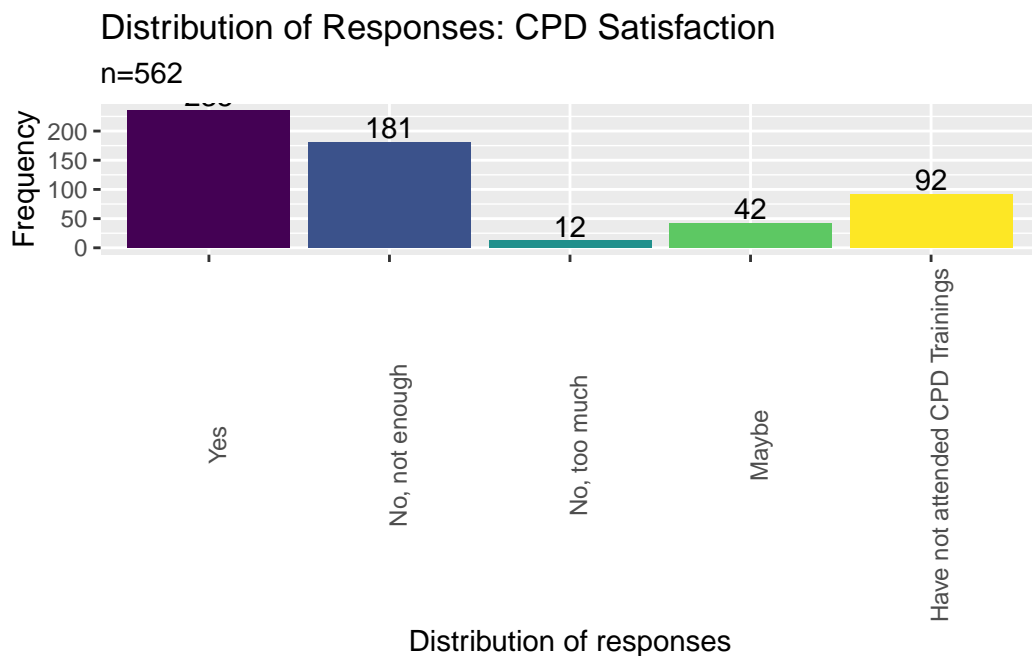


Question: Are you satisfied with the amount of CPD trainings you have attended for the last two (2) years?

- Yes, the amount of trainings I received within the last two (2) years is enough.
- No, the amount of trainings I received within the last two (2) years is not enough.
- No, the amount of trainings received within the last two (2) years is too much.
- Maybe

- NA (Have not attended CPD Trainings yet)

```
# Change the factor type
dataset$cpdsatisfaction <- as.factor(dataset$cpdsatisfaction)
# Recode responses
levels(dataset$cpdsatisfaction)[1] <- "Maybe"
levels(dataset$cpdsatisfaction)[2] <- "Have not attended CPD Trainings"
levels(dataset$cpdsatisfaction)[3] <- "No, not enough"
levels(dataset$cpdsatisfaction)[4] <- "No, too much"
levels(dataset$cpdsatisfaction)[5] <- "Yes"
# Rearrange the responses
dataset$cpdsatisfaction <- factor(dataset$cpdsatisfaction, levels = c("Yes", "No, not enough", "No, too much", "Maybe", "Have not attended CPD Trainings"))
# Graph the data
dataset %>%
filter(!is.na(cpdsatisfaction)) %>%
ggplot(aes(x = cpdsatisfaction, fill = cpdsatisfaction)) + geom_bar() + geom_text(aes(label = frequency))
```

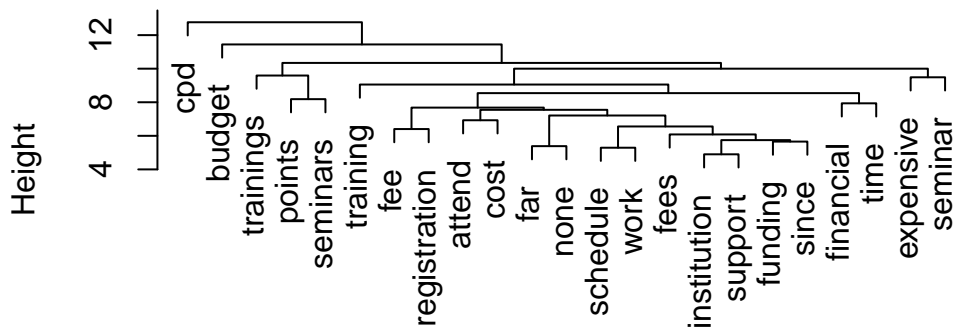


Question: Kindly indicate any challenges you encounter when fulfilling the required CPD trainings

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en")))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}

# Create a text dataset
cpdbarriers.text <- paste(dataset$cpdbarriers1, dataset$cpdbarriers2, dataset$cpdbarriers3)
# Create a text corpus
cpdbarriers.corpus <- cpdbarriers.text %>%
  VectorSource() %>%
  VCorpus()
# Clean the created corpus
cpdbarriers.corpus <- clean_corpus(cpdbarriers.corpus)
# Create a dendrogram
cpdbarriers.corpus %>%
  TermDocumentMatrix() %>%
  removeSparseTerms(sparse = 0.98) %>%
  as.matrix() %>%
  dist() %>%
  hclust() %>%
  plot()
```

Cluster Dendrogram



`hclust (*, "complete")`

Kindly indicate any topics you wish to explore in future trainings

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "library", "librarianship"))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}

# Create a text dataset
topics.text <- paste(dataset$topics1, dataset$topics2, dataset$topics3, dataset$topics4, d

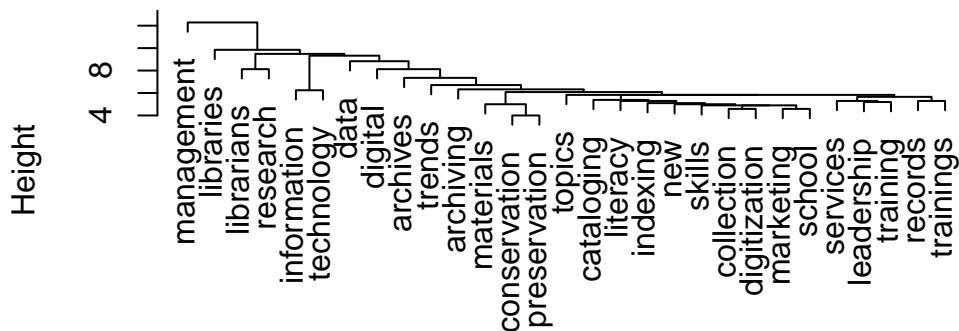
# Create a text corpus
topics.corpus <- topics.text %>%
  VectorSource() %>%
  VCorpus()

# Clean the created corpus
topics.corpus <- clean_corpus(topics.corpus)

# Create a dendrogram
topics.corpus %>%
```

```
TermDocumentMatrix() %>%
removeSparseTerms(sparse = 0.985) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()
```

Cluster Dendrogram



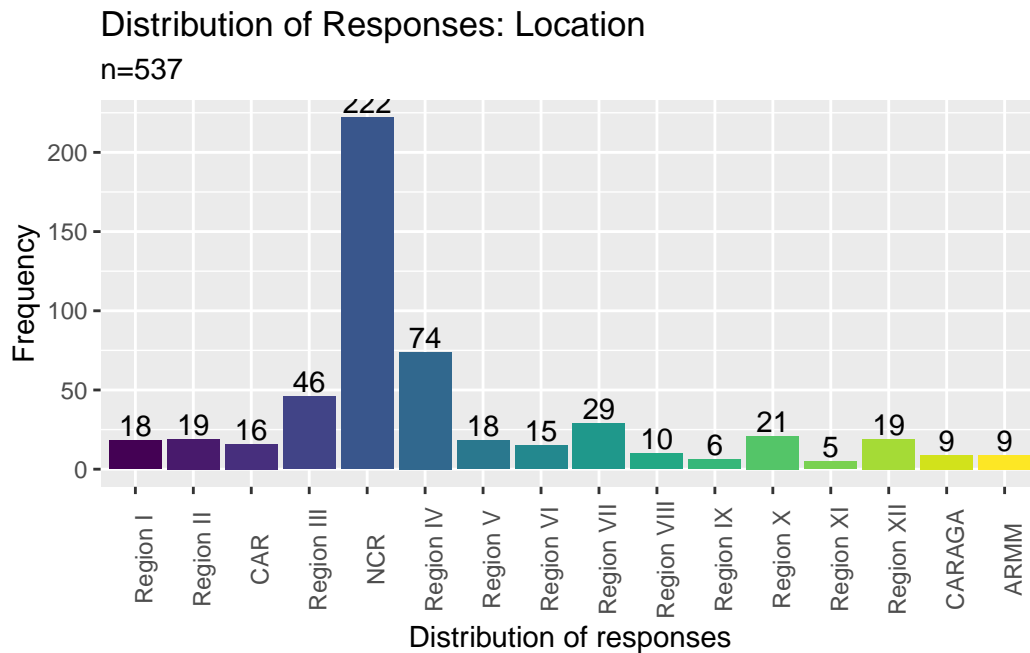
hclust (*, "complete")

Kindly indicate your primary work location

- NCR
- Region I
- Region II
- CAR
- Region III
- Region IV
- Region V
- Region VI
- Region VII
- Region VIII
- Region IX
- Region X

- Region XI
- Region XII
- CARAGA
- ARMM

```
# Change the factor type
dataset$location <- as.factor(dataset$location)
# Rearrange the responses
dataset$location <- factor(dataset$location, levels = c( 'Region I', 'Region II', 'CAR', '
# Graph the data
dataset %>%
filter(!is.na(location)) %>%
ggplot(aes(x = location, fill = location)) + geom_bar() + geom_text(aes(label = ..count..))
```

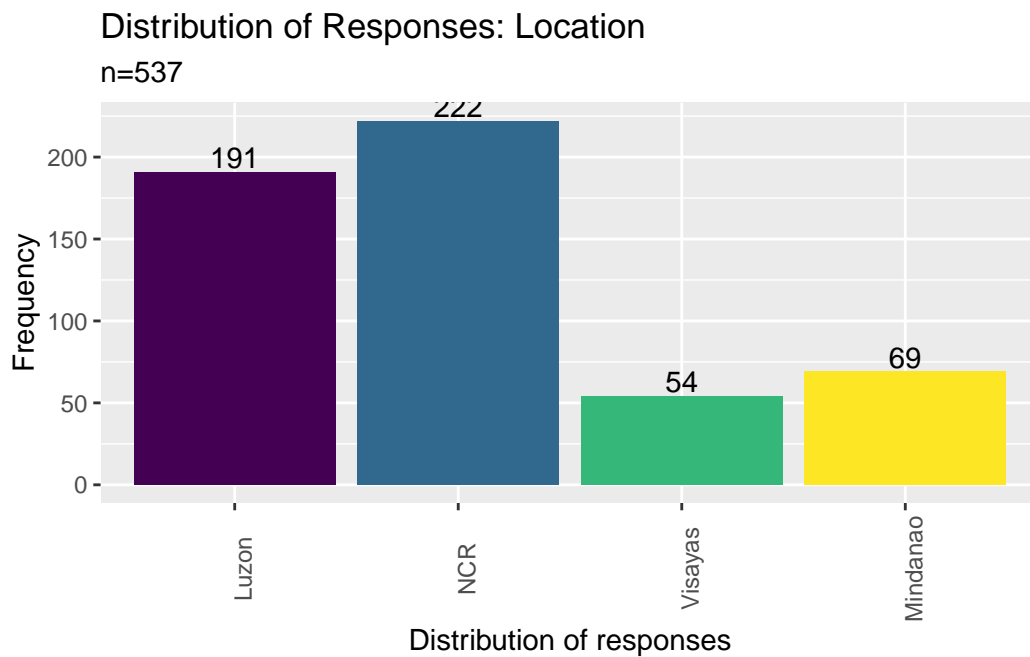


```
# Create lvm column
dataset <- dataset %>%
mutate(lvm = location)
# Recode responses
dataset$lvm <- as.factor(dataset$lvm)
levels(dataset$lvm)[16] <- "Mindanao"
levels(dataset$lvm)[15] <- "Mindanao"
```

```

levels(dataset$lvm)[14] <- "Mindanao"
levels(dataset$lvm)[13] <- "Mindanao"
levels(dataset$lvm)[12] <- "Mindanao"
levels(dataset$lvm)[11] <- "Mindanao"
levels(dataset$lvm)[10] <- "Visayas"
levels(dataset$lvm)[9] <- "Visayas"
levels(dataset$lvm)[8] <- "Visayas"
levels(dataset$lvm)[7] <- "Luzon"
levels(dataset$lvm)[6] <- "Luzon"
levels(dataset$lvm)[5] <- "NCR"
levels(dataset$lvm)[4] <- "Luzon"
levels(dataset$lvm)[3] <- "Luzon"
levels(dataset$lvm)[2] <- "Luzon"
levels(dataset$lvm)[1] <- "Luzon"
# Graph the data
dataset %>%
filter(!is.na(lvm)) %>%
ggplot(aes(x = lvm, fill = lvm)) + geom_bar() + geom_text(aes(label = ..count..), stat = "

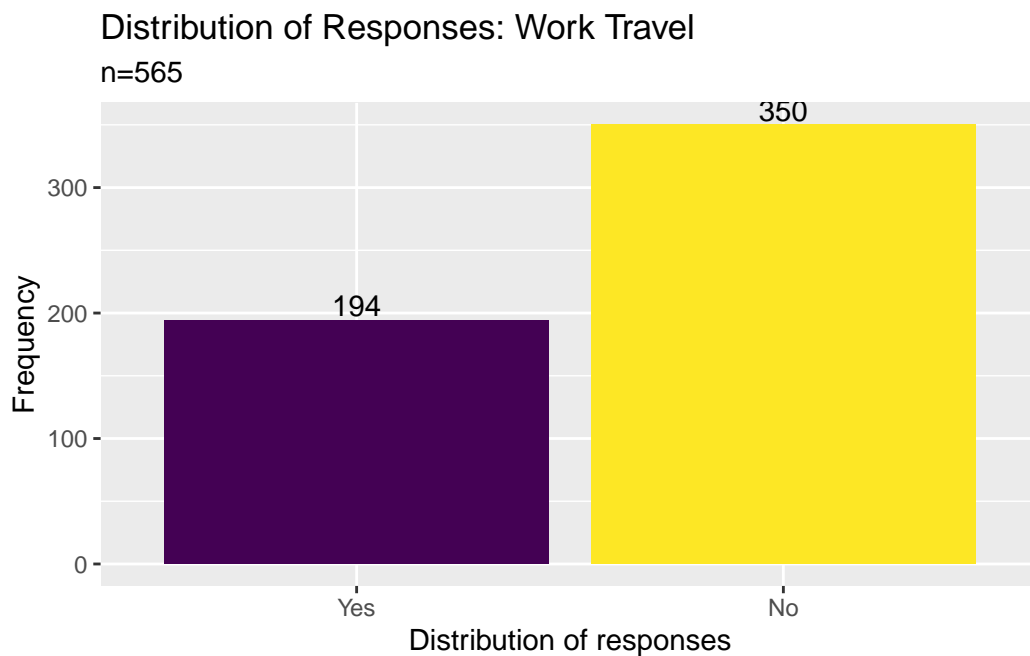
```



Question: Have you moved out of your original / permanent residence in order to work on your current work place?

- Yes
- No

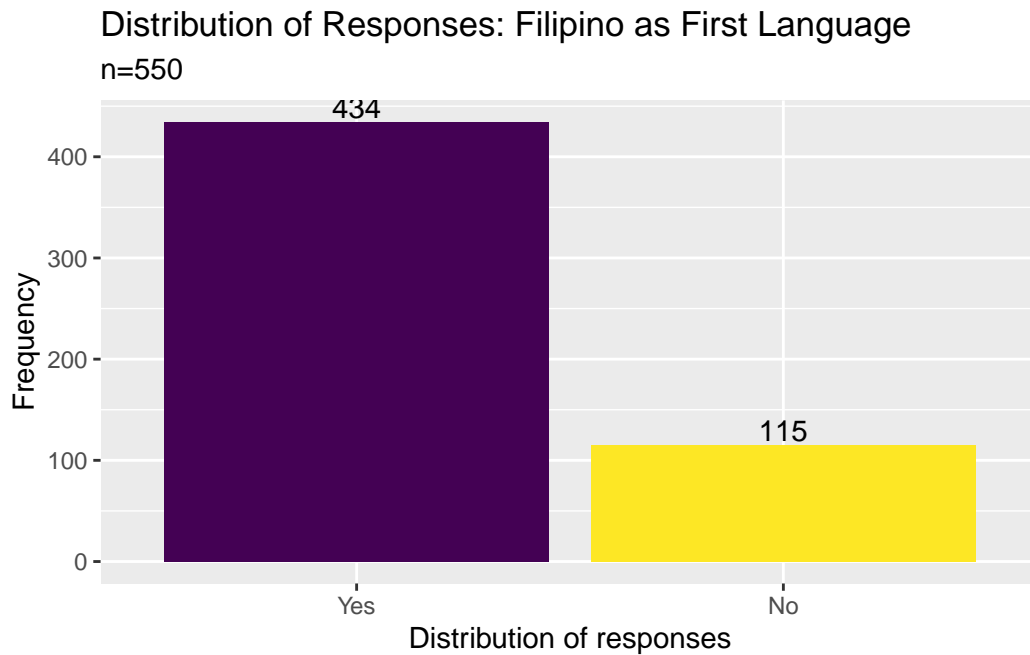
```
# Change the factor type
dataset$worktravel <- as.factor(dataset$worktravel)
# Rearrange the responses
dataset$worktravel <- factor(dataset$worktravel, levels = c("Yes", "No"))
# Graph the data
dataset %>%
  filter(!is.na(worktravel)) %>%
  ggplot(aes(x = worktravel, fill = worktravel)) + geom_bar() + geom_text(aes(label = ..count..))
```



Is Filipino / Tagalog your native language?

- Yes
- No

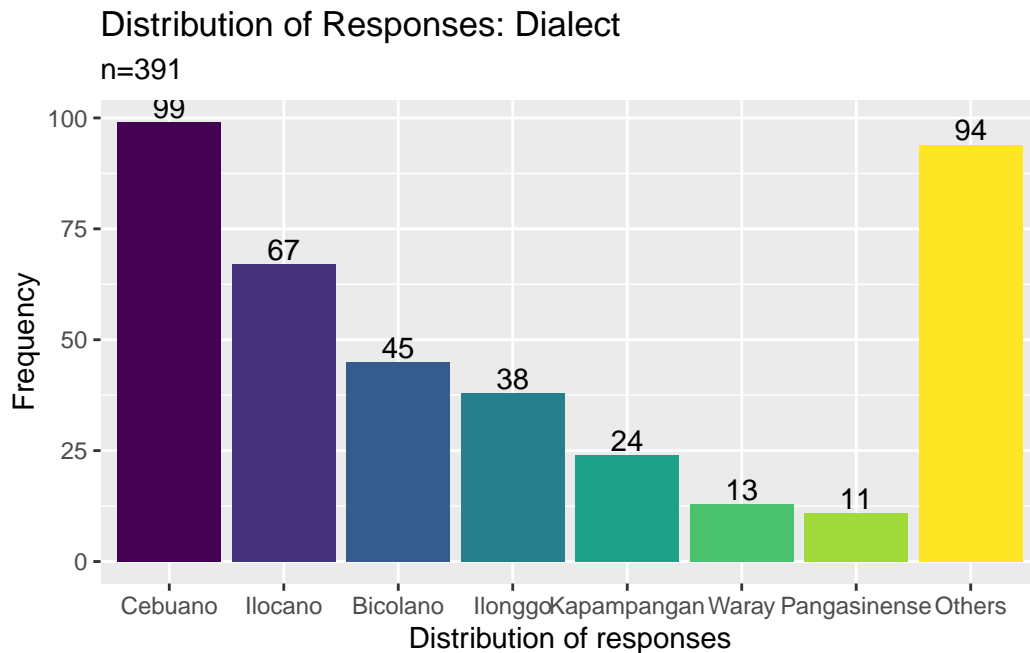

```
# Change the factor type
dataset$filipino <- as.factor(dataset$filipino)
# Rearrange the responses
dataset$filipino <- factor(dataset$filipino, levels = c("Yes", "No"))
# Graph the data
dataset %>%
  filter(!is.na(filipino)) %>%
  ggplot(aes(x = filipino, fill = filipino)) + geom_bar() + geom_text(aes(label = ..count..))
```



Kindly check the local languages or dialects you can speak on a conversational level other than Filipino / Tagalog

- Ilocano
- Cebuano
- Ilonggo
- Kapampangan
- Pangasinense
- Waray
- Bicolano
- Others

```
# Change the factor type
dataset$dialect <- as.factor(dataset$dialect)
# Rearrange the responses
dataset$dialect <- factor(dataset$dialect, levels = c('Cebuano', 'Ilocano', 'Bicolano', 'I
# Graph the data
dataset %>%
filter(!is.na(dialect)) %>%
ggplot(aes(x = dialect, fill = dialect)) + geom_bar() + geom_text(aes(label = ..count..),
```



Kindly indicate any foreign languages or dialects you can speak on a conversational level other than English (if any)

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "none", "can", "kankanaey"))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}
```

```

}
# Create a text corpus
topics.corpus <- dataset$foreignlang %>%
  VectorSource() %>%
  VCorpus()
# Clean the created corpus
topics.corpus <- clean_corpus(topics.corpus)
# Create a wordcount
term.frequency <- topics.corpus %>%
  TermDocumentMatrix() %>%
  as.matrix() %>%
  rowSums() %>%
  sort( , decreasing=TRUE)
term.frequency[1:10]

```

spanish	english	tagalog	french	ilonggo	korean	chinese	filipino
24	15	11	9	8	8	6	6
ilocano	bahasa						
6	4						

Kindly indicate the type of library you are currently working in

```

# Change the factor type
dataset$type <- as.factor(dataset$type)
# Recode responses
dataset$type <- fct_recode(dataset$type,
  "NA" = '#NAME?',
  "NA" = '...',
  "Academic" = 'academic',
  "Academic" = 'Academic',
  "Academic" = 'Academic-Special',
  "Academic" = 'Academic - Seminary',
  "Academic" = 'Academic (College)',
  "Academic" = 'Academic (Government)',
  "Academic" = 'Academic (Private-Catholic School)',
  "Academic" = 'Academic (University)',
  "Academic" = 'academic / private',
  "Academic" = 'Academic and Research Library (SUC-University)',
  "Academic" = 'academic and school',
  "Academic" = 'Academic and school libraries',

```

"Academic" = 'academic and school library',
 "Academic" = 'academic college library',
 "Academic" = 'academic library',
 "Academic" = 'Academic library',
 "Academic" = 'Academic Library',
 "Academic" = 'ACADEMIC LIBRARY',
 "Academic" = 'Academic Library (Medical/Allied Health)',
 "Academic" = 'academic library (university)',
 "Academic" = 'Academic library / State university library',
 "Academic" = 'Academic library in a Government Institution',
 "Academic" = 'Academic library under LGU',
 "Academic" = 'Academic Medical Library',
 "Academic" = 'Academic State University',
 "Academic" = 'academic SUCs',
 "Academic" = 'Academic SUCs HEIs',
 "Academic" = 'Academic.',
 "Academic" = 'Academic/ Research Library',
 "Academic" = 'Academic/Government Institutionalize',
 "Special" = 'Academic/institutional archives',
 "Academic" = 'Academic/Public Library',
 "Academic" = 'Academic/Secondary library',
 "Academic" = 'Academic/Special Library',
 "Academic" = 'Alliance Graduate School Library',
 "Academic" = 'An academic library with a 10, 000 student population. The library has 3 sub',
 "Special" = 'Archives',
 "Special" = 'Archives or Records Management',
 "Special" = 'Archives/Records Management',
 "School" = 'Chinese private school',
 "Academic" = 'college',
 "Academic" = 'college library',
 "Academic" = 'College library',
 "Academic" = 'College Library',
 "Academic" = 'College Library funded by LGU',
 "NA" = 'Confidential',
 "School" = 'DepEd (School) Library',
 "School" = 'DepEd division Library',
 "School" = 'Deped library hub',
 "School" = 'Division library',
 "NA" = 'Freelance',
 "Public" = 'Government',
 "Special" = 'Government Archives',

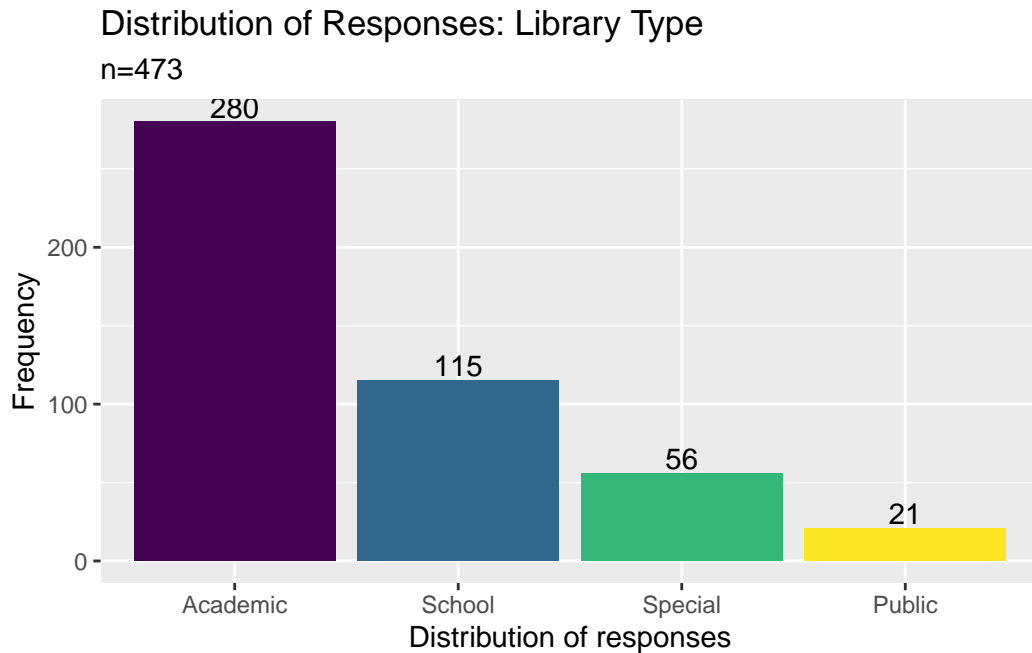
"Public" = 'Government library',
 "Public" = 'Government Public Library',
 "Special" = 'government special library',
 "Academic" = 'Government/ Academic',
 "Academic" = 'Graduate school',
 "School" = 'High school library',
 "Academic" = 'I am currently the audio-visual librarian of our university',
 "School" = 'IBO K-12',
 "School" = 'Junior High School Library',
 "Special" = 'Law library',
 "Special" = 'Law Library',
 "Special" = 'Law Library (Special)',
 "Special" = 'Law Library/Special Library',
 "Public" = 'LGU library',
 "Public" = 'Library Hub',
 "School" = 'Library school',
 "Academic" = 'LIS school',
 "Special" = 'Media Library',
 "Academic" = 'Medical and\r\nAllied Health Library',
 "Academic" = 'Medical/Academic library',
 "Special" = 'Military Library (Special academic library',
 "Public" = 'municipal library',
 "Public" = 'Municipal Library',
 "Special" = 'Museum',
 "NA" = 'n/a',
 "NA" = 'N/A',
 "NA" = 'NA. publishing company',
 "NA" = 'NA. Working in a BPO',
 "Public" = 'National',
 "Special" = 'NGO Digital Library',
 "NA" = 'None',
 "NA" = 'Not applicable',
 "Special" = 'Not in a library but in an archives',
 "NA" = 'Not library. GOCC.',
 "NA" = 'Not working in a library',
 "NA" = 'Not working in a library setting',
 "NA" = 'Not working in a library setting since two years ago',
 "Special" = 'Not working in a library. Working in a Data Office',
 "NA" = 'Private', "Academic" = 'Private academic/school (hybrid) library',
 "Academic" = 'private HEI',
 "Academic" = 'Privately owned libraries (academic)',

"Public" = 'Provincial Library.',
 "Public" = 'Public high school library',
 "Public" = 'public library',
 "Public" = 'Public library',
 "Public" = 'Public Library',
 "Academic" = 'Public Library/ Academic',
 "School" = 'Public school library',
 "School" = 'Public School Library',
 "School" = 'Public School Library',
 "School" = 'Public School Library (Deped Division Office)',
 "Special" = 'records center',
 "School" = 'school',
 "School" = 'School',
 "School" = 'School (primary and secondary)',
 "Academic" = 'School and Academic library',
 "Academic" = 'School and Academic Library',
 "School" = 'school library',
 "School" = 'School library',
 "School" = 'School Library',
 "School" = 'School library (Private)',
 "School" = 'School Library (Elementary)',
 "School" = 'School Library (Senior High School Library)',
 "School" = 'School library (under both DepED and TESDA)',
 "Academic" = 'School Library and academic Library',
 "School" = 'School library and prefer to be with the academic library soon',
 "School" = 'School library and reading center',
 "School" = 'School Library specifically Library Hub, a project of the Department of Education',
 "School" = 'School library, catering exclusively junior high school students and staff',
 "School" = 'School library.',
 "Academic" = 'School/Academic Library',
 "Academic" = 'School/academic library - Technical Section',
 "School" = 'School/Public',
 "School" = 'Secondary school library',
 "Academic" = 'Slis',
 "NA" = 'Spanish, bahasa',
 "Special" = 'special',
 "Special" = 'Special',
 "Special" = 'Special-Public',
 "Special" = 'Special (government law) library',
 "Special" = 'special academic library',
 "Special" = 'special library',

```

"Special" = 'Special library',
"Special" = 'Special Library',
"Special" = 'SPECIAL LIBRARY',
"Special" = 'Special Library in an Auditing Firm',
"Special" = 'Special library/Archives\r\nDigital records',
"Special" = 'Special library/Government',
"Special" = 'Special Public Library',
"Special" = 'Special/public/research library',
"Special" = 'Special/Research Library',
"Special" = 'specual libray',
"Academic" = 'State u',
"Academic" = 'State University',
"Academic" = 'State university library',
"Academic" = 'State/ Academic',
"Academic" = 'SUC',
"Academic" = 'SUC (Academic Library)',
"Academic" = 'SUC library',
"Academic" = 'SUC Library',
"Special" = 'University Archives',
"Academic" = 'University library',
"Academic" = 'University Library',
"Academic" = 'University Library - at the regional campus specializing in Law')
# Rearrange the responses
dataset$type <- factor(dataset$type, levels = c("Academic", "School", "Special", "Public"))
# Graph the data
dataset %>%
filter(!is.na(type)) %>%
ggplot(aes(x = type, fill = type)) + geom_bar() + geom_text(aes(label = ..count..), stat =

```



Other than being a librarian, what other roles do you serve in your institution/organization?

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "library", "librarianship", "
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}
# Create a text corpus
roles.corpus <- dataset$roles %>%
  VectorSource() %>%
  VCorpus()
# Clean the created corpus
roles.corpus <- clean_corpus(roles.corpus)
# Create a dendrogram
roles.corpus %>%
  TermDocumentMatrix() %>%
```

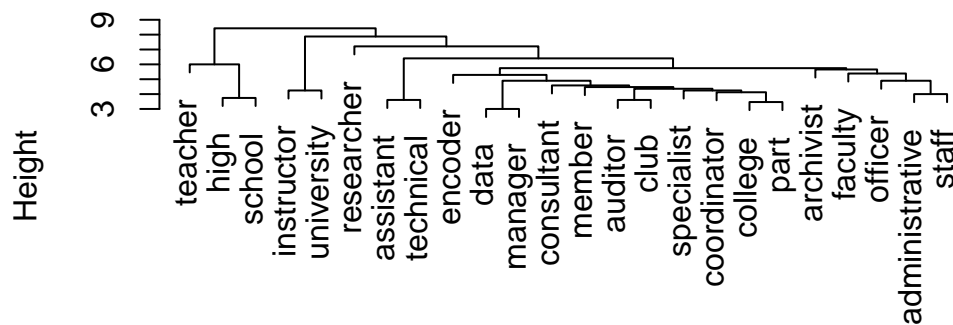


```

removeSparseTerms(sparse = 0.99) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()

```

Cluster Dendrogram



```
hclust (*, "complete")
```

Kindly indicate your current work position

```

# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "iii", "library", "librariansh
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}
# Create a text corpus
institutionposition.corpus <- dataset$institutionposition %>%
VectorSource() %>%

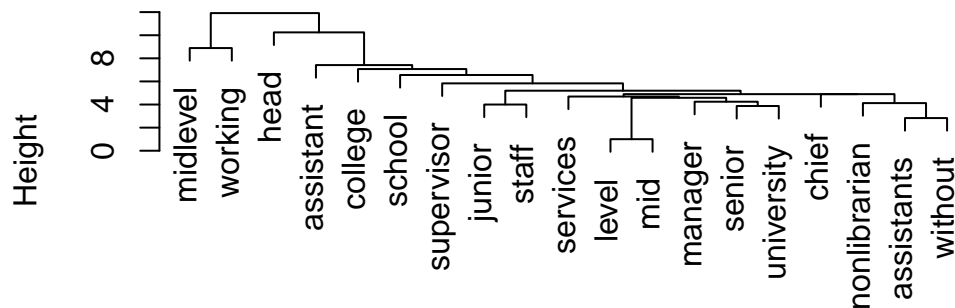
```

```

VCorpus()
# Clean the created corpus
institutionposition.corpus <- clean_corpus(institutionposition.corpus)
# Create a dendrogram
institutionposition.corpus %>%
TermDocumentMatrix() %>%
removeSparseTerms(sparse = 0.99) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()

```

Cluster Dendrogram



hclust (*, "complete")

Kindly indicate your primary job / designation inside the library

```

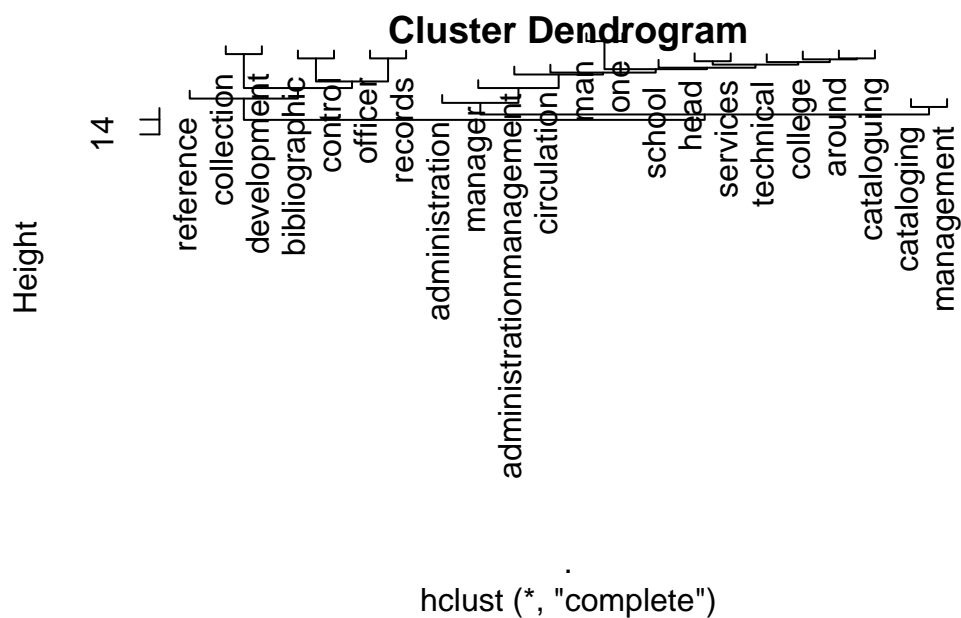
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "iii", "library", "librariansh

```

```

    corpus <- tm_map(corpus, stripWhitespace)
    return(corpus)
}
# Create a text corpus
primaryjob.corpus <- dataset$primaryjob %>%
VectorSource() %>%
VCorpus()
# Clean the created corpus
primaryjob.corpus <- clean_corpus(primaryjob.corpus)
# Create a dendrogram
primaryjob.corpus %>%
TermDocumentMatrix() %>%
removeSparseTerms(sparse = 0.99) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()

```



Kindly indicate your current / existing professional affiliations

```
# Change the factor type
dataset$affiliation <- as.factor(dataset$affiliation)
# Recode factors
levels(dataset$affiliation)[143] <- "PLAI, PAARL, RIzalenos Librarians Association, USTLIS"
levels(dataset$affiliation)[144] <- "PLAI, PAARL, RLA Rizalenos Librarians Association"
dataset$affiliation <- fct_recode(dataset$affiliation,
"ALPS" = 'ALPS',
"ASLP" = 'ASLP' ,
"PASLI" = 'before member of PASLI and PLAI-CeLRLC' ,
"NA" = 'Bplai' ,
"NA" = 'CLAI' ,
"NA" = 'DepEd' ,
"NA" = 'Dont renew yet my membership in PLAI' ,
"NA" = 'Haven\'t join any organization' ,
"NA" = 'I have not yet member seen i have a license' ,
"NA" = 'Inactive member of PLAI' ,
"LAQUEP" = 'Laquep, Plai stlrc' ,
"MAHLAP" = 'MAHLAP' ,
"MAHLAP PATLS" = 'MAHLAP /PATLS /PLAI' ,
"MAHLAP PAARL" = 'MAHLAP, PLAI, PAARL, ABCLI, PUPLSAAI, PAMJE, MLA' ,
"NA" = 'n/a' ,
"NA" = 'N/A' ,
"NA" = 'Na' ,
"NA" = 'NELAI (Nueva Ecija Librarian\'s Association, Inc.)' ,
"PAARL" = 'NELAI, PLAI, PAARL' ,
"NA" = 'none' ,
"NA" = 'None' ,
"NA" = 'NONE' ,
"NA" = 'None this year.' ,
"NA" = 'None yet' ,
"PAARL" = 'Paarl' ,
"PAARL" = 'PAARL' ,
"PAARL PASLI" = 'PAARL, PASLI' ,
"PAARL" = 'paarl, plai' ,
"PAARL" = 'PAARL, PLAI' ,
"PAARL PATLS" = 'PAARL, PLAI, PATLS, CODHLIS' ,
"PASLI" = 'PADLI' ,
"PASLI" = 'Pasli' ,
"PASLI" = 'PASLI' ,
```

"PASLI" = 'PASLI, PLAI' ,
 "PASLI" = 'PASLI, PLAI, MUNPARLAS' ,
 "PASLI" = 'PASLI, PNULISAA' ,
 "PATLS" = 'PATLS' ,
 "NA" = 'PLA' ,
 "NA" = 'plai' ,
 "NA" = 'Plai' ,
 "NA" = 'PLAI' ,
 "NA" = 'PLAI-CARAGA' ,
 "NA" = 'PLAI-Caraga, CLLA, MINLA, MISORLA' ,
 "NA" = 'PLAI-Davao' ,
 "NA" = 'PLAI-IRLC' ,
 "PAARL PATLS" = 'PLAI \r\nPAARL\r\nPATLS' ,
 "MAHLAP" = 'PLAI & MAHLAP\r\nbut not renewed membership with PLAI and MAHLAP for the past
 "PASLI" = 'PLAI & PASLI' ,
 "NA" = 'PLAI (PLAI-BRLC)' ,
 "NA" = 'PLAI , CoLiNet' ,
 "PGLL ASLP SFA" = 'PLAI / PGLL / ASLP / SFA' ,
 "NA" = 'PLAI and Bicol Region Librarians Council' ,
 "NA" = 'PLAI and CLAI' ,
 "MAHLAP" = 'PLAI and MAHLAP' ,
 "MAHLAP" = 'PLAI AND MAHLAP' ,
 "NA" = 'PLAI and MASTLINET' ,
 "PAARL" = 'PLAI and Paarl' ,
 "PAARL" = 'PLAI and PAARL' ,
 "PASLI" = 'PLAI and PASLI' ,
 "SFA" = 'PLAI and SFA' ,
 "ASLP" = 'plai aslp' ,
 "PAARL" = 'PLAI PAARL CAMANAVA' ,
 "PAARL" = 'PLAI PAARL PNULISAA Asso of Laguna Librarians' ,
 "PGLL SFA" = 'PLAI PGLL SFA' ,
 "NA" = 'PLAI STRLC, ALLPI\r\nPAARL (but have not yet renewed)' ,
 "NA" = 'PLAI WVRLC' ,
 "ASLP PAARL" = 'PLAI\r\nASLP\r\nPAARL' ,
 "PATLS" = 'Plai\r\nCodhlis\r\nPatls' ,
 "NA" = 'PLAI\r\nFORATL' ,
 "MAHLAP" = 'PLAI\r\nMAHLAP' ,
 "PAARL" = 'plai\r\nnpaarl' ,
 "PAARL" = 'Plai\r\nPaarl' ,
 "PAARL" = 'PLAI\r\nPAARL' ,
 "PAARL" = 'PLAI\r\nPAARL\r\nAASDI' ,

"PAARL ASLP" = 'PLAI\r\nPAARL\r\nASLP' ,
 "PAARL" = 'PLAI\r\nPAARL\r\nLAQueP-LInc' ,
 "PAARL PATLS" = 'PLAI\r\nPAARL\r\nPATLS' ,
 "PAARL" = 'Plai\r\nPaarl\r\nQCLC' ,
 "PAARL SFA" = 'plai\r\nPaarl\r\nSfa\r\nAasdi\r\nKbp' ,
 "PASLI" = 'PLAI\r\nPASLI' ,
 "PASLI" = 'PLAI\r\nPASLI\r\nLLAI\r\nPLAI -ARMM' ,
 "NA" = 'PLAI\r\nPNULISAA' ,
 "NA" = 'PLAI\r\nPUP LSAA' ,
 "NA" = 'PLAI\r\nSOCOLA' ,
 "NA" = 'PLAI,' ,
 "NA" = 'PLAI, NIRLC' ,
 "PAARL" = 'PLAI, PAARL' ,
 "NA" = 'PLAI, AASDI' ,
 "ALAP" = 'PLAI, ALAP' ,
 "ASFA" = 'PLAI, ALINET, ASFA, IAMSLIC' ,
 "NA" = 'PLAI, ALINET, MASTLINET, LLAI' ,
 "NA" = 'PLAI, ALINET, PLAI CRLC' ,
 "ALPS" = 'PLAI, ALPS' ,
 "ALPS" = 'PLAI, ALPS, LAQueP-LInc' ,
 "PASLI" = 'PLAI, and PASLI' ,
 "ALPS" = 'PLAI, APLIN, ALPS' ,
 "NA" = 'PLAI, APPLI, NEFLI' ,
 "ASLP" = 'PLAI, ASLP' ,
 "ASLP SFA" = 'PLAI, ASLP, ICA, SFA' ,
 "ASLP SLA" = 'PLAI, ASLP, SLA' ,
 "ASLP" = 'PLAI, ASLP, UELISAAI' ,
 "NA" = 'PLAI, ASSDI' ,
 "NA" = 'PLAI, BPLAI' ,
 "NA" = 'PLAI, BRLC' ,
 "NA" = 'PLAI, Cagayan Valley Regional Librarians Association, Inc.; Nueva Vizcaya Librarians Association, Inc.' ,
 "NA" = 'PLAI, Cavite Librarians Inc.,' ,
 "NA" = 'PLAI, CLAI' ,
 "MAHLAP" = 'Plai, clai, mahlap' ,
 "PAARL" = 'PLAI, CLAI, PAARL' ,
 "NA" = 'PLAI, CLLA, Inc.' ,
 "NA" = 'PLAI, DACUN' ,
 "NA" = 'PLAI, DACUN, MASTLINET' ,
 "NA" = 'PLAI, ICA, UPLSAA' ,
 "NA" = 'PLAI, LAQUEPLINC,' ,
 "MAHLAP" = 'Plai, mahlap' ,

```

"MAHLAP" = 'PLAI, MAHLAP' ,
"NA" = 'PLAI, MUNPARLAS' ,
"ASLP" = 'PLAI, MUNPARLAS, ASLP' ,
"NA" = 'PLAI, OMLA' ,
"PAARL" = 'plai, paarl' ,
"PAARL" = 'Plai, paarl' ,
"PAARL" = 'PLAi, PAARL' ,
"PAARL" = 'PLAI, pAARL' ,
"PAARL" = 'PLAI, PAARL' ,
"PAARL" = 'PLAI, PAARL and QCLC' ,
"PAARL" = 'PLAI, PAARL,' ,
"PAARL" = 'PLAI, PAARL, AASDI' ,
"PAARL" = 'PLAI, PAARL, ABCLI, QCLC' ,
"PAARL" = 'PLAI, PAARL, ALA' ,
"PAARL ALAP" = 'PLAI, PAARL, ALAP' ,
"PAARL ALAP" = 'PLAI, PAARL, ALAP, AASDI' ,
"PAARL ALAP" = 'PLAI, PAARL, ALAP, ALLPI' ,
"PAARL ASLP" = 'PLAI, PAARL, ASLP' ,
"PAARL ASLP MAHLAP SLA" = 'PLAI, PAARL, ASLP, MAHLAP, SLA' ,
"PAARL" = 'PLAI, PAARL, BLAI' ,
"PAARL MAHLAP" = 'PLAI, PAARL, MAHLAP' ,
"PAARL MAHLAP" = 'PLAI, PAARL, MAHLAP (Local), MLA (international)' ,
"PAARL PGLL" = 'PLAI, PAARL, PGLL, PUPLSAA' ,
"PAARL" = 'PLAI, PAARL, PLAI-STRLC' ,
"PAARL" = 'PLAI, PAARL, PNULISAA' ,
"PAARL" = 'PLAI, PAARL, PUPLSAA' ,
"PAARL" = 'PLAI, PAARL, QCLC (Quezon City Library Consortium)' ,
"PAARL" = 'PLAI, PAARL, RiZalenos Librarians Association, USTLISAA' ,
"PAARL" = 'PLAI, PAARL, RLA Rizalenos Librarians Association' ,
"PAARL" = 'PLAI, PAARL, UPLSAA' ,
"PAARL" = 'PLAI, PAARL, UPLSAA, PUPLSAA' ,
"PAARL" = 'PLAI, PAARL, Vanuatu Librarians\' Association' ,
"PAARL" = 'PLAI, PAARL, Western Visayas Association of Museums' ,
"PAARL PASLI" = 'PLAI, PAARL,PASLI' ,
"PAARL PASLI" = 'PLAI, PADLI' ,
"PAARL PASLI" = 'PLAI, PASLI' ,
"PAARL PASLI" = 'PLAI, PASLI, LAQueP-LInc, MUNPARLAS' ,
"PAARL PASLI" = 'PLAI, PASLI, UPLSAA' ,
"PAARL PATLS" = 'PLAI, PATLS' ,
"PAARL PATLS SFA" = 'Plai, patls, sfa' ,
"PAARL PATLS SFA" = 'PLAI, PATLS, SFA' ,

```

```

"NA" = 'Plai, pearl, clla' ,
"PGLL" = 'PLAI, PGLL' ,
"PGLL" = 'PLAI, PGLL, and LLIWA' ,
"PGLL SFA" = 'PLAI, PGLL, SFA' ,
"NA" = 'Plai, Plai-Central luzon, Nueva ecija librarian association' ,
"NA" = 'PLAI, PLAI-NIRLC' ,
"NA" = 'PLAI, PLAI-RLC' ,
"ALPS" = 'PLAI, PLAI-STRLC, ALPS, LAQueP-LInc' ,
"NA" = 'PLAI, PLAI-STRLC, OMLA' ,
"NA" = 'PLAI, PLAI - IRLC' ,
"NA" = 'PLAI, PSPL, BOLA' ,
"NA" = 'plai, regional council for librarians' ,
"SFA" = 'PLAI, SFA' ,
"NA" = 'PLAI, UPLSAA' ,
"ASLP" = 'PLAI,ASLP' ,
"PASLI" = 'PLAI,PASLI' ,
"PATLS" = 'PLAI,PATLS, CODHLIS' ,
"NA" = 'PLAI.' ,
"PAARL" = 'PLAI; PAARL' ,
"PAARL" = 'PLAI< PAARL' ,
"PASLI" = 'Plain and Philippine Association of DepEd Librarians,Inc. (PADLI)' ,
"NA" = 'PLAL' ,
"NA" = 'PNULISAA' ,
"NA" = 'QCLC' ,
"NA" = 'RLA' ,
"NA" = 'UPLSAA')

```

Warning: Unknown levels in `f`: PLAI, PAARL, QCLC (Quezon City Library Consortium), PLAI, PAARL, UPLSAA

```

# Create new columns: PAARL, PASLI, ASLP, MAHLAP, PATLS, ALAP, ALPS, PGLL, SFA
dataset <- transform(dataset,
  PAARL = grepl("PAARL",affiliation),
  PASLI = grepl("PASLI", affiliation),
  ASLP = grepl("ASLP", affiliation),
  MAHLAP = grepl("MAHLAP", affiliation),
  PATLS = grepl("PATLS", affiliation),
  ALAP = grepl("ALAP", affiliation),
  ALPS = grepl("ALPS", affiliation),
  PGLL = grepl("PGLL", affiliation),
  SFA = grepl("SFA", affiliation))

```


Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

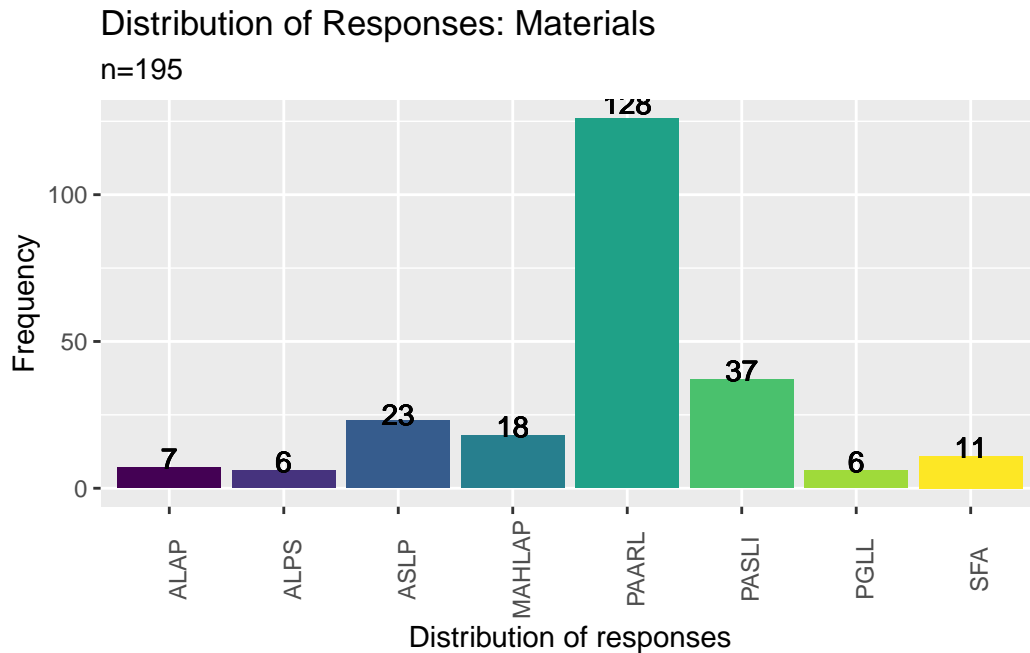
Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RIzale<f1>os Librarians Association,
USTLISAA' to a wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 1 is invalid

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
unable to translate 'PLAI, PAARL, RLA Rizale<f1>os Librarians Association' to a
wide string

Warning in grepl(pattern, c(levels(x), NA_character_), ignore.case, perl, :
input string 30 is invalid

```
# Summarize by counting the TRUE values per materials
dataset %>%
  summarise("PAARL" = sum(PAARL == TRUE),
            "PASLI" = sum(PASLI == TRUE),
            "ASLP" = sum(ASLP == TRUE),
            "MAHLAP" = sum(MAHLAP == TRUE),
            "ALAP" = sum(ALAP == TRUE),
            "ALPS" = sum(ALPS == TRUE),
            "PGLL" = sum(PGLL == TRUE),
            "SFA" = sum(SFA == TRUE)) %>%
  gather() %>%
  ggplot(aes(x = key, y = value, fill = key)) + geom_col(position = "dodge") + xlab("Distribu
  geom_text(x = 1, y = 10, label = "7", colour = "black", size = 4) +
  geom_text(x = 2, y = 9, label = "6", colour = "black", size = 4) +
  geom_text(x = 3, y = 25, label = "23", colour = "black", size = 4) +
  geom_text(x = 4, y = 21, label = "18", colour = "black", size = 4) +
  geom_text(x = 5, y = 131, label = "128", colour = "black", size = 4) +
  geom_text(x = 6, y = 40, label = "37", colour = "black", size = 4) +
  geom_text(x = 7, y = 9, label = "6", colour = "black", size = 4) +
  geom_text(x = 8, y = 14, label = "11", colour = "black", size = 4)
```



Kindly indicate what factors influence your decision when it comes to joining a professional organization?

```
# Create a clean_corpus function
clean_corpus <- function(corpus) {
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, removeWords, c(stopwords("en")))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}

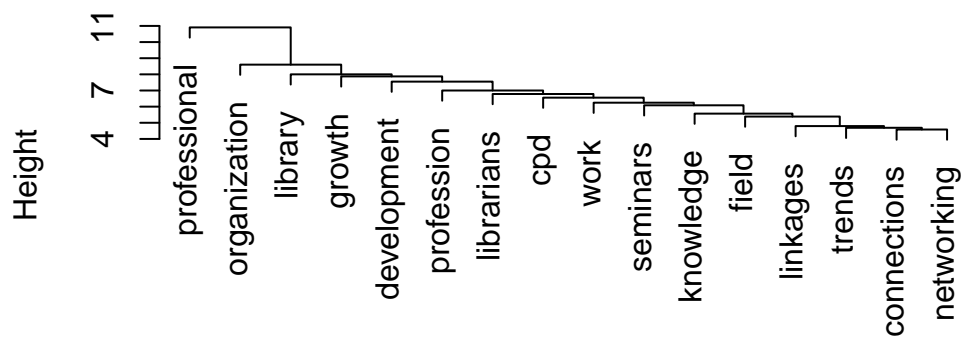
# Create a text corpus
factor.corpus <- dataset$factor %>%
  VectorSource() %>%
  VCorpus()

# Clean the created corpus
factor.corpus <- clean_corpus(factor.corpus)

# Create a dendrogram
factor.corpus %>%
  TermDocumentMatrix() %>%
```

```
removeSparseTerms(sparse = 0.985) %>%
as.matrix() %>%
dist() %>%
hclust() %>%
plot()
```

Cluster Dendrogram



hclust (*, "complete")

```
glimpse(dataset)
```

Rows: 684

Columns: 107

Warning in grepl(",", levels(x), fixed = TRUE): input string 1 is invalid UTF-8

Warning in grepl(",", levels(x), fixed = TRUE): input string 30 is invalid UTF-8

```
$ voter      <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15~
$ licensed   <fct> Licensed librarian, Licensed librarian, Licensed ~
$ philippines <fct> NA, Residing outside of the Philippines, Residing~
$ gender     <fct> NA, Prefer not to say, Female, Male, Male, Male, ~
```

\$ age	<dbl> 59, 100, 40, 46, 32, 46, 48, 44, 25, 24, 34, 39, ~
\$ working	<fct> NA, Retired, Employed, Employed, Employed, Employ~
\$ industry	<fct> NA, ICT, Academe, Academe, Academe, Academe, Acad~
\$ years_service	<dbl> NA, 60, 14, 23, 10, 24, 23, 12, 5, 1, 12, 18, 14,~
\$ institution	<fct> NA, Private, Government, Government, Government, ~
\$ tenure	<fct> NA, Self-employed, Permanent, Permanent, Contract~
\$ gross_salary	<dbl> NA, 100, 51, 53, 40, 56, 36, 38, 19, 11, 95, 58, ~
\$ net_salary	<dbl> NA, 100, 30, 38, 30, 42, 29, 26, NA, 10, 95, 36, ~
\$ benefits1	<fct> NA, night differential pay, paid vacation and/or ~
\$ benefits2	<chr> NA, NA, "monetary allowances", "monetary allowanc~
\$ benefits3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
\$ benefits4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ benefits5	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ otherbenefits1	<chr> NA, "123", "Rice allowance\r\nClothing allowance\~
\$ otherbenefits2	<chr> NA, "123", NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ otherbenefits3	<chr> NA, "Hello Philippines and hello world", NA, NA, ~
\$ otherbenefits4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ materials1	<chr> NA, "law materials and resources", NA, "archives ~
\$ materials2	<chr> NA, NA, NA, "hardcopy theses and dissertations", ~
\$ materials3	<chr> NA, NA, NA, "hardcopy books / pamphlets / etc", N~
\$ materials4	<chr> NA, NA, NA, "institutional files and records", NA~
\$ materials5	<chr> NA, NA, NA, "data sets", NA, "archives (physical)~
\$ materials6	<chr> NA, NA, NA, "electronic material (digital titles,~
\$ materials7	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ othermaterials1	<chr> NA, "Huehs", "Grading sheets\r\nExam papers", "Bo~
\$ othermaterials2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ othermaterials3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ othermaterials4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ position	<fct> NA, Supervisory, Supervisory, Senior level, Senio~
\$ fiveyears	<fct> NA, No, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Undeci~
\$ pursue	<chr> NA, "Hfeudnr", NA, NA, NA, NA, NA, NA, NA, NA, NA~
\$ education	<fct> NA, Ph.D, Master degree, MLIS, Master degree, MLI~
\$ enrolled	<fct> NA, Yes, Yes, Yes, No, Yes, Yes, No, Yes, No, No,~
\$ completing	<fct> NA, Post Graduate non-LIS, Ph.D., Ph.D., Ph.D., P~
\$ study_again	<fct> NA, NA, No, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Ma~
\$ after5years	<fct> NA, NA, NA, Short Course, Ph.D., Short Course, Ph~
\$ cpd1	<chr> NA, NA, "trainings and symposiums facilitated by ~
\$ cpd2	<chr> NA, NA, "government facilitated trainings and sym~
\$ cpd3	<chr> NA, NA, NA, NA, "within the work place (in-house ~
\$ cpd4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ cpdsupport1	<chr> NA, NA, "from the government (scholarships, free ~
\$ cpdsupport2	<chr> NA, NA, NA, "from the government (scholarships, f~
\$ cpdsupport3	<chr> NA, NA, NA, "from the work place (private sponsor~

\$ cpdsatisfaction	<fct> NA, NA, "Yes", "Yes", "Maybe", "Yes", "Yes", "Yes~
\$ cpdbarriers1	<chr> NA, NA, "CPD program sources that I really like",~
\$ cpdbarriers2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ cpdbarriers3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ topics1	<chr> NA, NA, "Ethics\r\nCommunication skills\r\nCookin~
\$ topics2	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ topics3	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ topics4	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ topics5	<chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
\$ location	<fct> NA, NA, NCR, NCR, NCR, NCR, NCR, NCR, Region IV, ~
\$ worktravel	<fct> NA, NA, Yes, Yes, Yes, Yes, No, No, No, No, Yes, ~
\$ filipino	<fct> NA, NA, No, Yes, Yes, Yes, No, Yes, Yes, Yes, Yes~
\$ dialect	<fct> NA, NA, Ilocano, NA, Others, NA, Kapampangan, NA,~
\$ foreignlang	<chr> NA, NA, "None", NA, "Chinese", NA, NA, "English",~
\$ type	<fct> NA, NA, Academic, School, NA, School, NA, NA, Sch~
\$ roles	<chr> NA, NA, "Cleaner\r\nCook", NA, "Faculty", "Asst p~
\$ institutionposition	<chr> NA, NA, "Head", "Asst prof 5", "Asst. Prof.", "As~
\$ primaryjob	<chr> NA, NA, NA, "Education", "Research, Extention, Te~
\$ affiliation	<fct> NA, NA, PATLS, PAARL PATLS SFA, PATLS, PAARL PATL~
\$ factor	<chr> NA, NA, "Friends", "Similar thrusts", "colleagues~
\$ age_group	<fct> 51-60, NA, 41-50, 41-50, 31-40, 41-50, 41-50, 41-~
\$ years_service_group	<fct> NA, NA, 10-14, 20-24, 10-14, 20-24, 20-24, 10-14,~
\$ gross_salary_group	<fct> NA, Greater than 50k, Greater than 50k, Greater t~
\$ net_salary_group	<fct> NA, Greater than 50k, 21-30, 31-40, 21-30, 41-50,~
\$ benefits	<chr> "NA NA NA NA NA", "night differential pay NA NA N~
\$ shuttle	<lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
\$ night_differential	<lgl> FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, F~
\$ paid_leaves	<lgl> FALSE, FALSE, TRUE, TRUE, FALSE, TRUE, TRUE, FALS~
\$ hmo	<lgl> FALSE, FALSE, FALSE, FALSE, TRUE, FALSE, FALSE, F~
\$ allowances	<lgl> FALSE, FALSE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE,~
\$ no_benefits	<lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
\$ benefits6	<fct> NA, With Benefits, With Benefits, With Benefits, ~
\$ materials	<chr> "NA NA NA NA NA NA NA", "law materials and resour~
\$ theses	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, TRU~
\$ books	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, TRU~
\$ electronic_materials	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, TRU~
\$ law_materials	<lgl> FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, TRUE, FA~
\$ datasets	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, FALSE, TR~
\$ archives	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, TRU~
\$ institutional_files	<lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, FAL~
\$ cpd	<chr> "NA NA NA NA", "NA NA NA NA", "trainings and symp~
\$ inhouse	<lgl> FALSE, FALSE, FALSE, FALSE, TRUE, TRUE, FALSE, TR~
\$ government	<lgl> FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, FALSE, TRU~


```

$ private      <lgl> FALSE, FALSE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ~
$ xattended    <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ cpdsupport   <chr> "NA NA NA", "NA NA NA", "from the government (sch~
$ support_self <lgl> FALSE, FALSE, FALSE, TRUE, TRUE, TRUE, FALSE, TRU~
$ support_government <lgl> FALSE, FALSE, TRUE, TRUE, FALSE, TRUE, TRUE, TRUE~
$ support_private <lgl> FALSE, FALSE, FALSE, TRUE, TRUE, TRUE, TRUE, FALS~
$ support_xattended <lgl> FALSE, FALSE, FALSE, TRUE, TRUE, TRUE, FALSE, TRU~
$ lvm          <fct> NA, NA, NCR, NCR, NCR, NCR, NCR, NCR, Luzon, Luzo~
$ PAARL        <lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, FALSE, FA~
$ PASLI        <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ ASLP         <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ MAHLAP       <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ PATLS        <lgl> FALSE, FALSE, TRUE, TRUE, TRUE, TRUE, FALSE, TRUE~
$ ALAP         <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ ALPS         <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ PGLL         <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, ~
$ SFA          <lgl> FALSE, FALSE, FALSE, TRUE, FALSE, TRUE, FALSE, FA~

```

```

# Save the updated dataset
write_rds(dataset, "LibrarianCensus.rds")

```