

STAT 411 Final Project

Rebecca Li, Karen Hong

2025-02-20

```
library(ggplot2)
library(dplyr)

## 
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union

library(tidyr)
library(corrplot)

## Warning: package 'corrplot' was built under R version 4.3.3

## corrplot 0.95 loaded

library(FactoMineR)

## Warning: package 'FactoMineR' was built under R version 4.3.3

library(factoextra)

## Warning: package 'factoextra' was built under R version 4.3.3

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa

library(psych)

## Warning: package 'psych' was built under R version 4.3.3

## 
## Attaching package: 'psych'
```

```

## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha

library(corr)

## Warning: package 'corr' was built under R version 4.3.3

library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## vforcats    1.0.0      vreadr     2.1.5
## v lubridate 1.9.3      v stringr   1.5.1
## v purrr     1.0.2      vtibble    3.2.1

## -- Conflicts ----- tidyverse_conflicts() --
## x psych::%+()%     masks ggplot2::%+]()
## x psych::alpha()   masks ggplot2::alpha()
## 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(candisc)

## Warning: package 'candisc' was built under R version 4.3.3

## Loading required package: heplots

## Warning: package 'heplots' was built under R version 4.3.3

## Loading required package: broom
##
## Attaching package: 'candisc'
##
## The following object is masked from 'package:stats':
##
##     cancor

wildfire_data = read.csv("C:/Users/lenovo/Desktop/STAT 411/Final Project/California Wildfire Damage.csv")

wildfire_data <- wildfire_data %>%
  mutate(across(where(is.numeric), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))

#PCA

wildfire_numeric <- wildfire_data %>%
  select(where(is.numeric))

wildfire_scaled <- scale(wildfire_numeric)

```

```
pca_result <- prcomp(wildfire_scaled, center = TRUE, scale. = TRUE)
```

```
# Summary of PCA
```

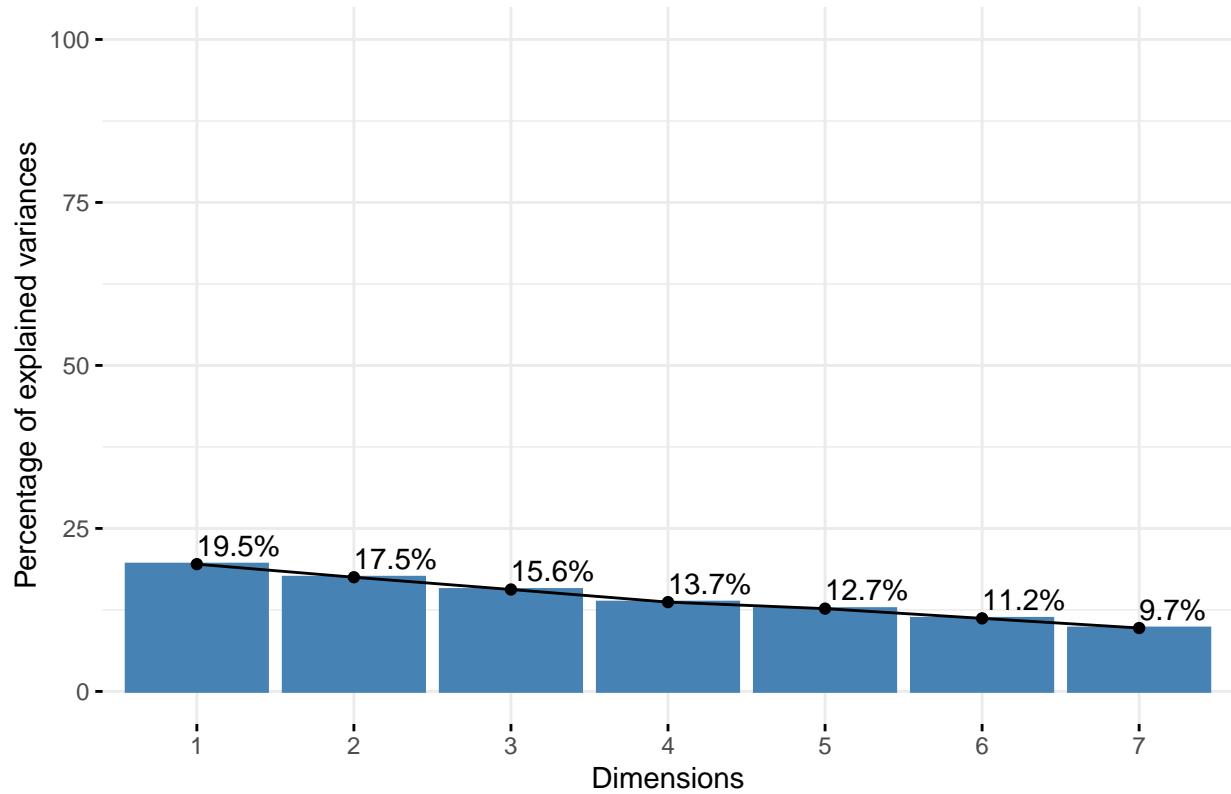
```
summary(pca_result)
```

```
## Importance of components:
```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
## Standard deviation	1.1690	1.1072	1.0461	0.9792	0.9429	0.8857	0.82520
## Proportion of Variance	0.1952	0.1751	0.1563	0.1370	0.1270	0.1121	0.09728
## Cumulative Proportion	0.1952	0.3703	0.5267	0.6637	0.7906	0.9027	1.00000

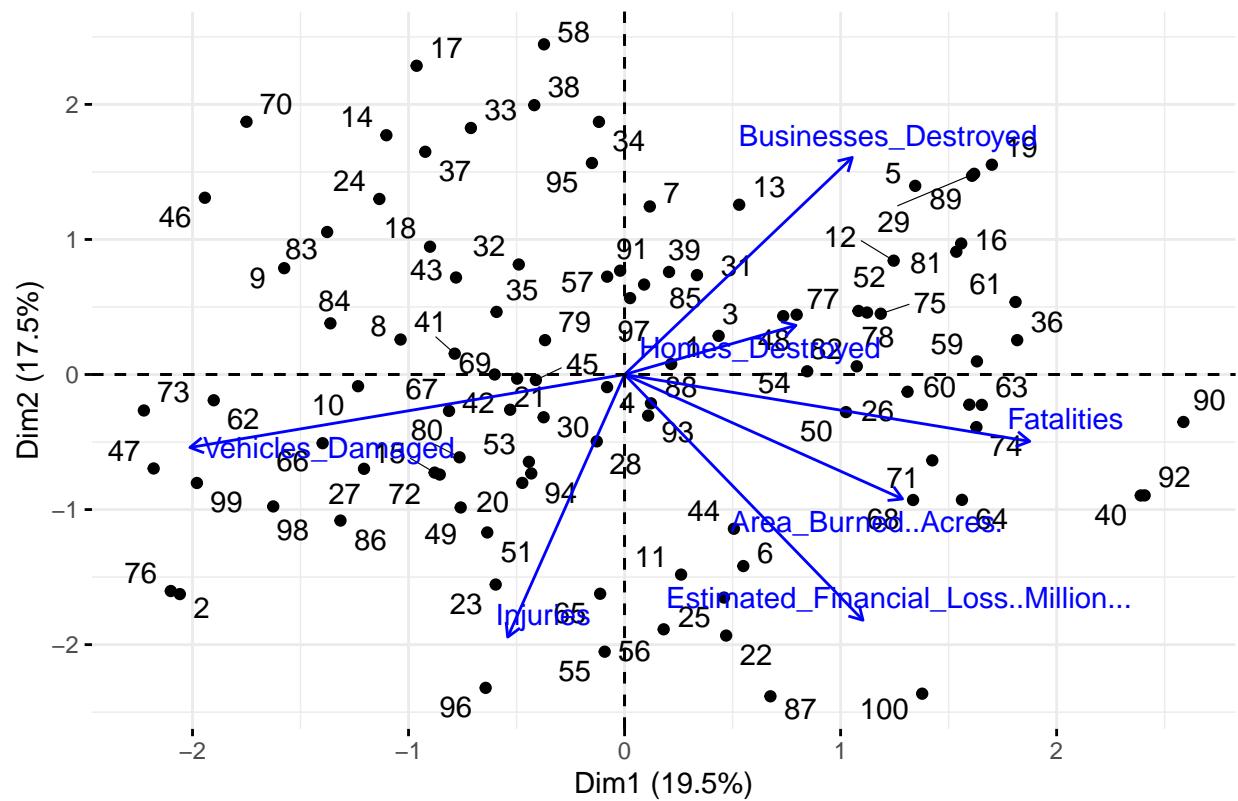
```
fviz_eig(pca_result, addlabels = TRUE, ylim = c(0, 100))
```

Scree plot



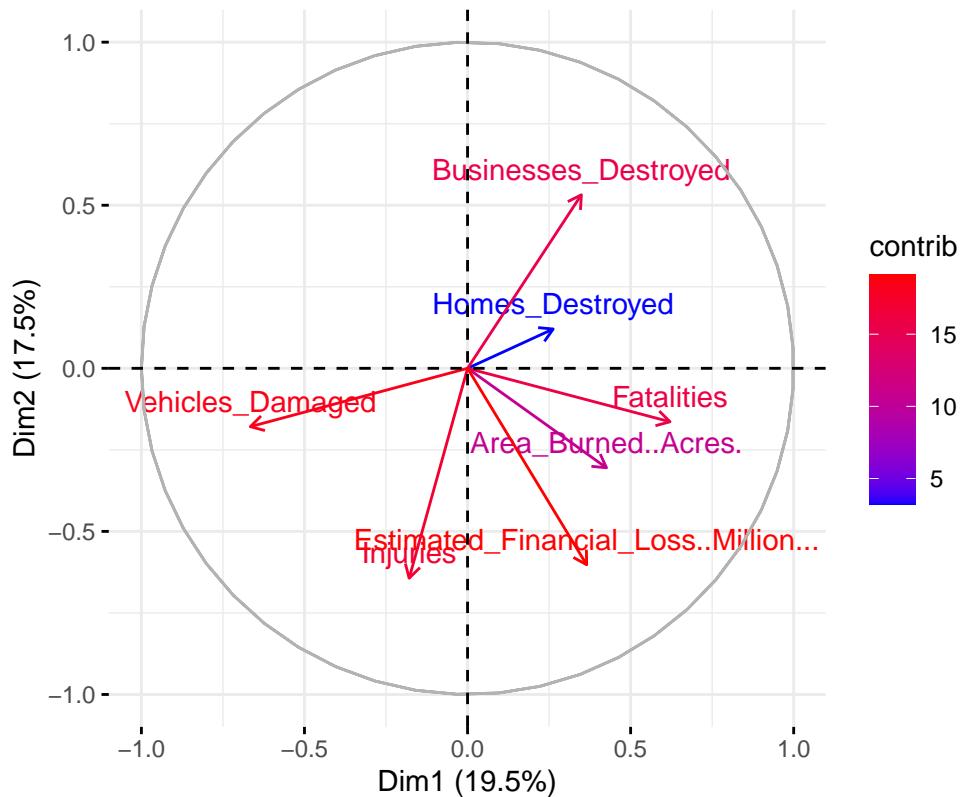
```
fviz_pca_biplot(pca_result, repel = TRUE, col.var = "blue")
```

PCA – Biplot



```
fviz_pca_var(pca_result, col.var = "contrib", gradient.cols = c("blue", "red"))
```

Variables – PCA



```
pca_scores <- as.data.frame(pca_result$x)
```

```
#Factor Analysis
```

```
data_numeric <- wildfire_data[, c("Area_Burned..Acres.", "Homes_Destroyed",
                                 "Businesses_Destroyed", "Vehicles_Damaged",
                                 "Injuries", "Fatalities", "Estimated_Financial_Loss..Million...")]
data_scaled <- scale(data_numeric)
cor_matrix <- cor(data_scaled)
cor_matrix
```

	Area_Burned..Acres.	Homes_Destroyed
## Area_Burned..Acres.	1.00000000	0.05191504
## Homes_Destroyed	0.05191504	1.00000000
## Businesses_Destroyed	0.02819495	0.11349268
## Vehicles_Damaged	-0.13643198	-0.07311470
## Injuries	0.09484300	0.01526998
## Fatalities	0.05039442	-0.04586270
## Estimated_Financial_Loss..Million...	0.07518726	0.04664454
	Businesses_Destroyed	Vehicles_Damaged
## Area_Burned..Acres.	0.02819495	-0.13643198
## Homes_Destroyed	0.11349268	-0.07311470
## Businesses_Destroyed	1.00000000	-0.07556630
## Vehicles_Damaged	-0.07556630	1.00000000
## Injuries	-0.10360711	0.11933077

```

## Fatalities                      0.07356399  -0.17731355
## Estimated_Financial_Loss..Million... -0.07798952  -0.02445021
##                                         Injuries  Fatalities
## Area_Burned..Acres.            0.09484300  0.05039442
## Homes_Destroyed                0.01526998 -0.04586270
## Businesses_Destroyed          -0.10360711  0.07356399
## Vehicles_Damaged              0.11933077 -0.17731355
## Injuries                        1.00000000 -0.03790814
## Fatalities                      -0.03790814  1.00000000
## Estimated_Financial_Loss..Million...  0.07973711  0.18491923
##                                         Estimated_Financial_Loss..Million...
## Area_Burned..Acres.             0.07518726
## Homes_Destroyed                 0.04664454
## Businesses_Destroyed           -0.07798952
## Vehicles_Damaged               -0.02445021
## Injuries                         0.07973711
## Fatalities                       0.18491923
## Estimated_Financial_Loss..Million...  1.00000000

```

```
KMO(cor_matrix)
```

```

## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = cor_matrix)
## Overall MSA =  0.5
## MSA for each item =
##                                         Area_Burned..Acres.      Homes_Destroyed
##                                         0.52                  0.44
##                                         Businesses_Destroyed  Vehicles_Damaged
##                                         0.50                  0.52
##                                         Injuries               Fatalities
##                                         0.50                  0.49
## Estimated_Financial_Loss..Million...
##                                         0.48

```

```
cortest.bartlett(cor_matrix)
```

```

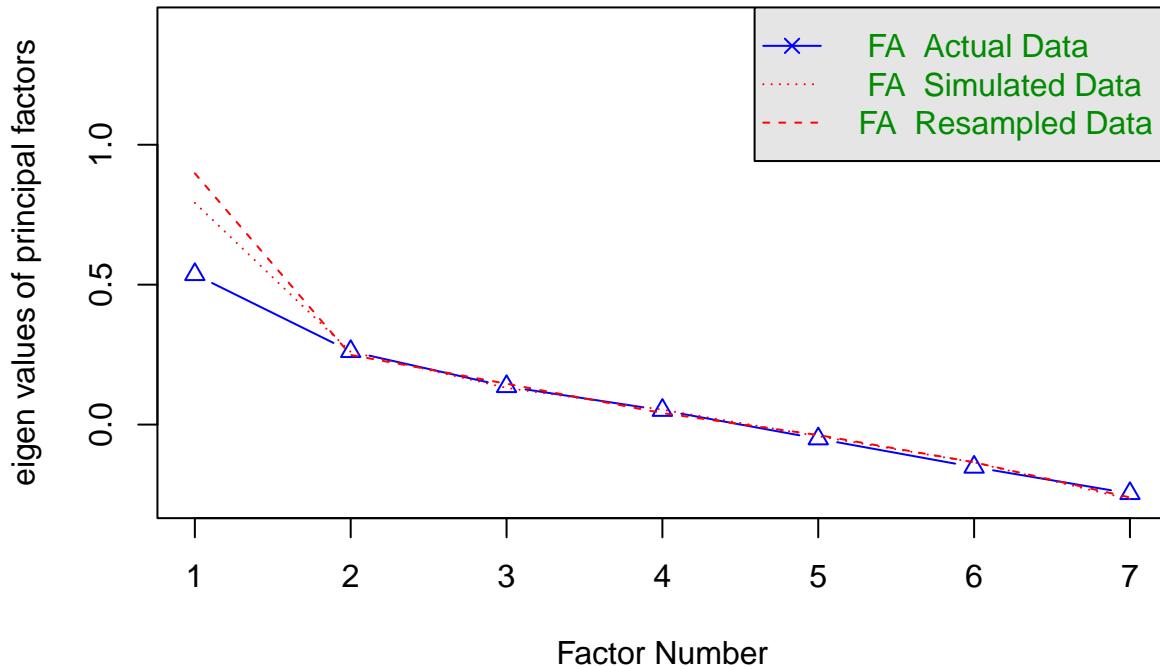
## Warning in cortest.bartlett(cor_matrix): n not specified, 100 used

## $chisq
## [1] 17.30794
##
## $p.value
## [1] 0.6922787
##
## $df
## [1] 21

```

```
# Determine Number of Factors using Scree Plot and Parallel Analysis
fa.parallel(data_scaled, fm="ml", fa="fa")
```

Parallel Analysis Scree Plots



```

## Parallel analysis suggests that the number of factors = 0 and the number of components = NA

fa_result <- fa(data_scaled, nfactors=2, fm="ml", rotate="varimax")

#Factor loading
fa_result$loadings

## 
## Loadings:
##                               ML1     ML2
## Area_Burned..Acres.      -0.177
## Homes_Destroyed          -0.101
## Businesses_Destroyed     -0.162
## Vehicles_Damaged         0.667
## Injuries                  0.169
## Fatalities                0.186 -0.258
## Estimated_Financial_Loss..Million... 0.997
## 
##                               ML1     ML2
## SS loadings    1.051  0.608
## Proportion Var 0.150  0.087
## Cumulative Var 0.150  0.237

```

```

# Compute Factor Scores
factor_scores <- factor.scores(data_scaled, fa_result)
print("Factor Scores:")

## [1] "Factor Scores"

print(factor_scores$scores)

##          ML1         ML2
## [1,] -0.088005569  0.561978119
## [2,] -0.712070670  1.657004597
## [3,]  0.019525900 -1.020083821
## [4,]  1.095666671 -0.647746453
## [5,] -0.415199339 -0.717814427
## [6,]  1.442288265 -0.872403982
## [7,] -1.173733834 -0.940345314
## [8,] -0.277192718  0.058547852
## [9,] -0.970504101  0.820079612
## [10,] 0.148151802  0.504474572
## [11,] 0.095770860 -0.099149158
## [12,] -0.531395326 -1.155435807
## [13,] -1.601772039 -1.459394416
## [14,] -1.118056128  0.852476364
## [15,] -0.635598358  1.442960355
## [16,]  0.827387865 -1.401016118
## [17,] -1.030102177  0.663040774
## [18,] -1.488556233  0.924317965
## [19,] -1.248641934 -1.841350337
## [20,] -1.484894068 -0.720776778
## [21,] -0.493562991  0.406565766
## [22,]  0.247238742 -0.735008059
## [23,]  1.163238506  1.088200324
## [24,] -0.615805587  0.307637200
## [25,]  1.726632514  0.814700611
## [26,]  0.782203919 -0.526659193
## [27,] -0.388967706  0.517015339
## [28,]  0.387736438  0.284621910
## [29,] -0.290759071 -1.486210648
## [30,] -0.454757049  0.419469948
## [31,] -1.222901097 -0.135596128
## [32,] -1.168419446 -0.086918116
## [33,] -1.220002436 -0.491329640
## [34,] -0.707125720 -0.334195515
## [35,] -0.590952256  0.756512329
## [36,]  0.881727996 -0.843949218
## [37,] -0.993057427  0.476600637
## [38,] -0.522391638  0.496026578
## [39,] -1.473002747 -0.204972423
## [40,]  1.260065003 -1.128102596
## [41,] -1.548548656  0.316231056
## [42,]  1.207158963  1.129900009
## [43,] -0.260053454 -0.151447984
## [44,]  1.189603491 -0.825108629

```

```

## [45,]  0.976101837  1.347326265
## [46,] -0.768048335  1.569647580
## [47,]  0.676851401  1.919921257
## [48,]  0.761630662 -0.694797036
## [49,]  0.703546347  1.237197971
## [50,]  0.595103823 -0.847259785
## [51,]  0.680675727  1.473888712
## [52,]  1.570639601 -0.137116011
## [53,] -0.857426983  0.352773731
## [54,]  1.020588655 -0.321280243
## [55,]  1.317372416  0.434019265
## [56,]  1.698728772  1.057012146
## [57,] -0.646822851  0.781316382
## [58,] -0.997666091 -0.514226691
## [59,]  0.131959665 -1.148280630
## [60,]  0.849393757 -1.535400987
## [61,] -0.180290769 -1.356205794
## [62,] -1.616585504  0.898056104
## [63,]  0.709989491 -0.835978901
## [64,]  1.203560022 -1.227568260
## [65,]  1.392707914  0.923406322
## [66,] -1.280315246  1.101657650
## [67,]  0.665208498  1.168395464
## [68,]  0.003153045 -0.790835195
## [69,]  0.696356015  1.279506847
## [70,] -1.486487954  0.045584568
## [71,]  1.075624129 -0.942295584
## [72,]  0.912132543  1.214139549
## [73,] -0.413628587  1.489303699
## [74,]  0.692476062 -1.063523076
## [75,]  1.175363288 -0.959979300
## [76,]  0.485220353  1.920634268
## [77,]  0.416076314 -1.394315068
## [78,]  0.774414060 -1.069171499
## [79,] -1.637702476 -0.095936189
## [80,] -0.756436677  0.646574109
## [81,]  0.903796235 -1.510445805
## [82,] -1.566541243 -1.468241959
## [83,] -0.584736798  1.052954018
## [84,] -0.853808877  0.721711097
## [85,] -0.795043216 -0.476027915
## [86,]  0.953986175  1.291308607
## [87,]  1.158745826 -0.518868745
## [88,]  0.539844845  0.594906525
## [89,] -0.905316463 -1.559688454
## [90,]  1.167217021 -1.409936164
## [91,] -1.381685874 -0.946479401
## [92,]  1.467597419 -1.576445011
## [93,]  0.317128384 -0.080374895
## [94,]  1.475395019  0.439055197
## [95,] -1.210365181 -0.549841210
## [96,]  1.126201647  1.520074632
## [97,] -1.410617290  0.004997727
## [98,] -0.643906019 -0.163309919

```

```
## [99,] 0.632743589 2.041196488
## [100,] 1.317534712 -0.006083609
```