Fall Data Challenge Project

Help Solve Homeless

Chen Wang, Xinyue Tang

October 209

```
library(tidyverse)
library(readr)
set.seed(8888)
Homeless_Overall <- read_csv("Overall_Homeless.csv")
Homeless_Individual <- read_csv("Homeless_Individual.csv")
Homeless_Families <- read_csv("Homeless_in_Families.csv")
Homeless_Sheltered <- read_csv("Sheltered_Homeless.csv")
```

Overall Homeless

Principal Component Analysis

Body:

```
glimpse(Homeless_Overall)
```

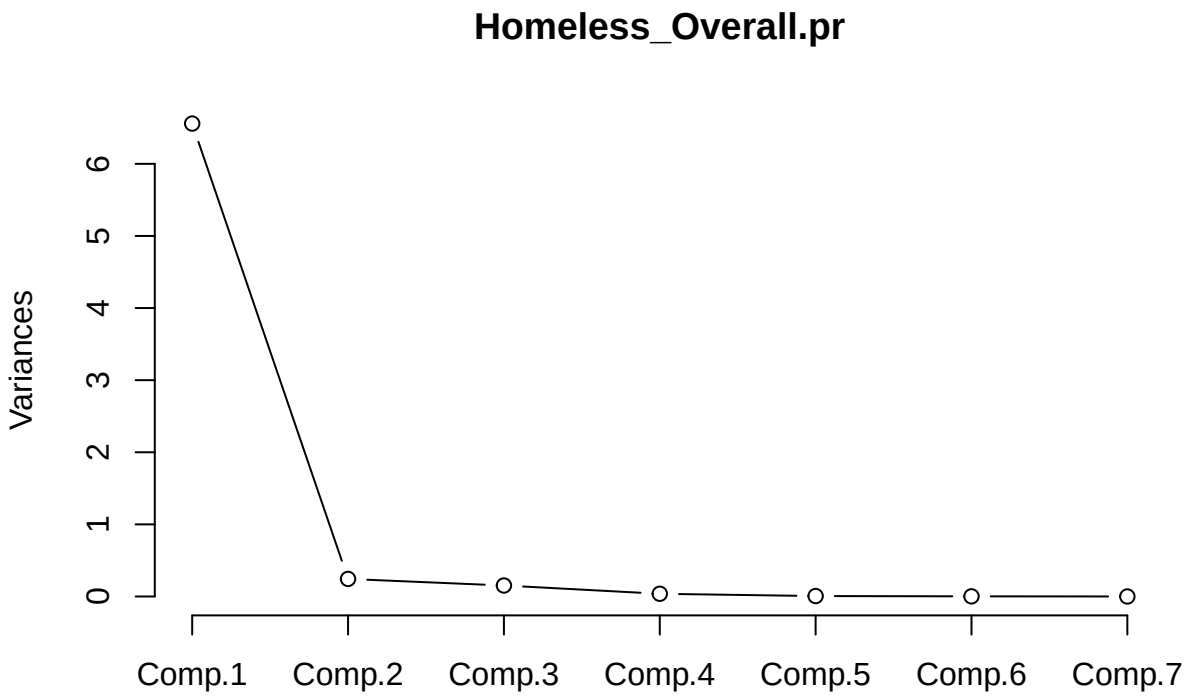
```
## Observations: 9
## Variables: 7
## $ Year                              <dbl> 2010, 2011, 2012, 2013, 2014, 201...
## $ `Overall Homeless`                <dbl> 0.1173554, 0.1769930, 0.0000000, ...
## $ CPI                               <dbl> 0.0000000, 0.1579940, 0.2665358, ...
## $ `Minimum Wages`                   <dbl> 0.0000000, 0.0000000, 0.0000000, ...
## $ `Unemployment Population`         <dbl> 0.00000000, 0.05263158, 0.2105263...
## $ `House Rental Price`              <dbl> 0.01904762, 0.02095238, 0.0000000...
## $ `Drug addict ED visit per 10,000` <dbl> 0.0000000, 0.1315583, 0.1812843, ...
```

```
Homeless_Overall.pr<-princomp(Homeless_Overall,cor=TRUE)
summary(Homeless_Overall.pr,loadings=TRUE)
```

```
## Importance of components:
##                            Comp.1     Comp.2     Comp.3      Comp.4       Comp.5
## Standard deviation     2.5609822 0.49341917 0.39014149 0.194009927 0.0778167323
## Proportion of Variance 0.9369471 0.03478035 0.02174434 0.005377122 0.0008650634
## Cumulative Proportion  0.9369471 0.97172750 0.99347184 0.998848961 0.9997140249
##                              Comp.6       Comp.7
## Standard deviation     0.0439355955 8.455133e-03
## Proportion of Variance 0.0002757624 1.021275e-05
## Cumulative Proportion  0.9999897872 1.000000e+00
##
## Loadings:
##                                Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Year                            0.386  0.238  0.181  0.215  0.235
## Overall Homeless                0.364 -0.557 -0.543  0.503
## CPI                             0.379         0.587  0.224 -0.234  0.559
## Minimum Wages                   0.371 -0.560  0.202 -0.661  0.243
```

```
## Unemployment Population          0.381  0.423 -0.106         0.646 -0.134
## House Rental Price               0.375  0.366 -0.502 -0.446 -0.403  0.331
## Drug addict ED visit per 10,000  0.389         0.152  0.113 -0.502 -0.736
##                                 Comp.7
## Year                             0.812
## Overall Homeless
## CPI                             -0.305
## Minimum Wages
## Unemployment Population         -0.478
## House Rental Price
## Drug addict ED visit per 10,000 -0.114
```

```
screeplot(Homeless_Overall.pr,type="lines")
```




```
p<-predict(Homeless_Overall.pr)
```

Principal Component Analysis - Remove Year

```
home_noyear <-
  Homeless_Overall %>%
  select(-Year)
```

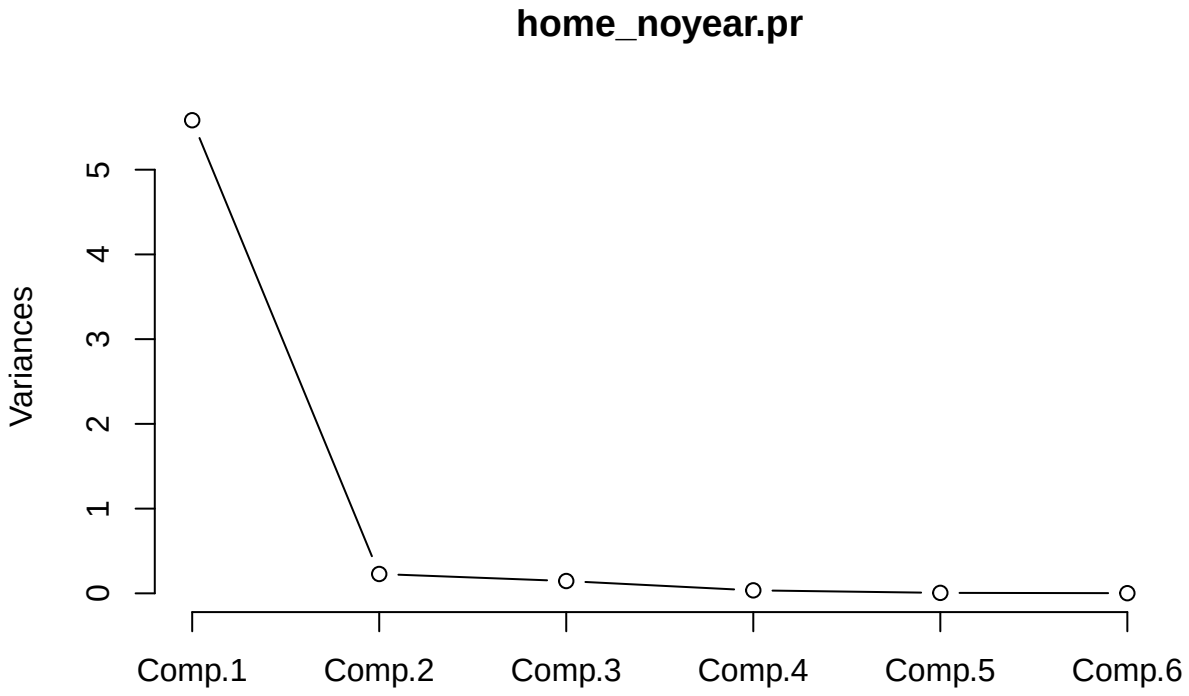
```
home_noyear.pr<-princomp(home_noyear,cor=TRUE)
summary(home_noyear.pr,loadings=TRUE)
```

```
## Importance of components:
##                            Comp.1     Comp.2     Comp.3      Comp.4       Comp.5
## Standard deviation     2.3629095 0.47800174 0.38122350 0.187949659 0.0747441802
## Proportion of Variance 0.9305569 0.03808094 0.02422189 0.005887512 0.0009311154
## Cumulative Proportion  0.9305569 0.96863786 0.99285976 0.998747270 0.9996783854
##                              Comp.6
## Standard deviation     0.0439282085
```

```
## Proportion of Variance 0.0003216146
## Cumulative Proportion  1.0000000000
##
## Loadings:
##                                 Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Overall Homeless                 0.398  0.486  0.608  0.476
## CPI                              0.410        -0.626  0.316 -0.139  0.564
## Minimum Wages                    0.404  0.547 -0.170 -0.675  0.211
## Unemployment Population          0.411 -0.482         0.108  0.756 -0.123
## House Rental Price               0.406 -0.468  0.408 -0.416 -0.410  0.329
## Drug addict ED visit per 10,000  0.421 -0.108 -0.204  0.185 -0.441 -0.735
```

```
# To try to analyze the first 2 components in this output
```

```
screeplot(home_noyear.pr,type="lines")
```




```
p<-predict(home_noyear.pr)
```

Linear Regression

Homeless_Individual

Principal Component Analysis

Body:

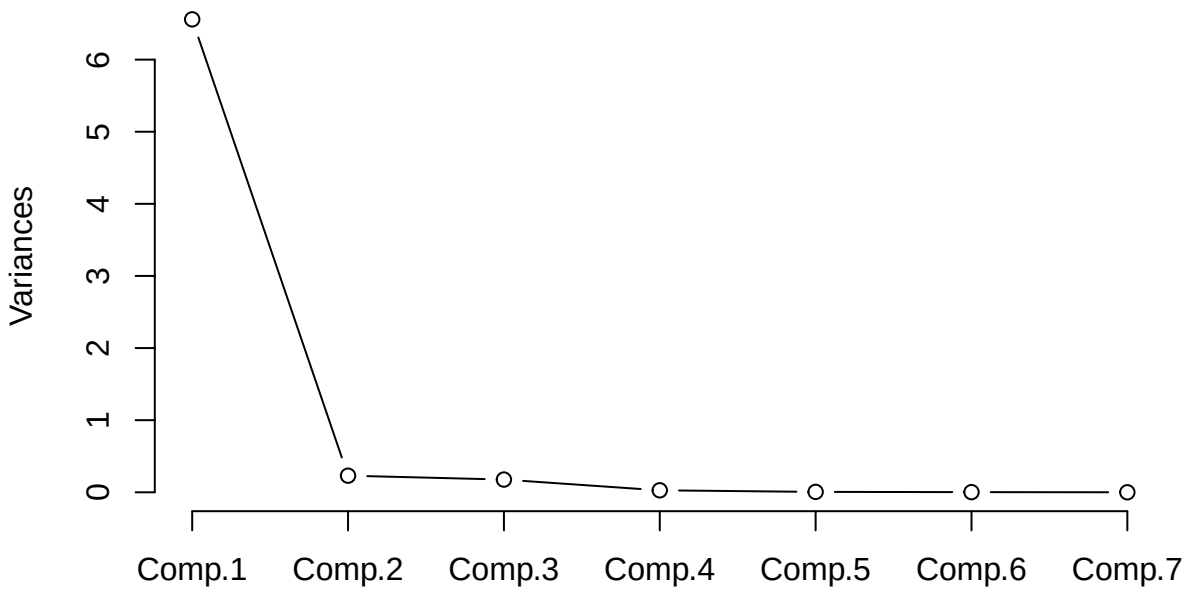
```
glimpse(Homeless_Individual)
```

```
## Observations: 9
## Variables: 7
## $ Year                              <dbl> 2010, 2011, 2012, 2013, 2014, 201...
## $ `Homeless Individual`             <dbl> 0.22899454, 0.05513542, 0.0000000...
## $ CPI                               <dbl> 0.0000000, 0.1579940, 0.2665358, ...
## $ `Minimum Wages`                   <dbl> 0.0000000, 0.0000000, 0.0000000, ...
## $ `Unemployment Population`         <dbl> 0.00000000, 0.05263158, 0.2105263...
## $ `House Rental Price`              <dbl> 0.01904762, 0.02095238, 0.0000000...
## $ `Drug addict ED visit per 10,000` <dbl> 0.0000000, 0.1315583, 0.1812843, ...
```

```
Homeless_Individual.pr<-princomp(Homeless_Individual,cor=TRUE)
summary(Homeless_Individual.pr,loadings=TRUE)
```

```
## Importance of components:
##                            Comp.1     Comp.2     Comp.3     Comp.4       Comp.5
## Standard deviation     2.5608913 0.47998313 0.41981641 0.16532325 0.0745864710
## Proportion of Variance 0.9368806 0.03291197 0.02517797 0.00390454 0.0007947345
## Cumulative Proportion  0.9368806 0.96979257 0.99497054 0.99887508 0.9996698145
##                              Comp.6       Comp.7
## Standard deviation     0.046812453 1.094955e-02
## Proportion of Variance 0.000313058 1.712752e-05
## Cumulative Proportion  0.999982872 1.000000e+00
##
## Loadings:
##                                    Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Year                                0.386  0.262  0.148  0.239  0.172
## Homeless Individual                 0.364 -0.578 -0.514  0.505
## CPI                                 0.379         0.570  0.218 -0.140 -0.623
## Minimum Wages                       0.370 -0.579  0.279 -0.565  0.326  0.151
## Unemployment Population             0.382  0.398 -0.165  0.139  0.615  0.225
## House Rental Price                  0.376  0.298 -0.507 -0.550 -0.249 -0.380
## Drug addict ED visit per 10,000     0.388  0.111  0.166        -0.630  0.618
##                                    Comp.7
## Year                                0.817
## Homeless Individual
## CPI                                -0.275
## Minimum Wages
## Unemployment Population            -0.469
## House Rental Price
## Drug addict ED visit per 10,000    -0.174
```

```
screeplot(Homeless_Individual.pr,type="lines")
```




```
p<-predict(Homeless_Individual.pr)
```

Homeless_Families

Principal Component Analysis

Body:

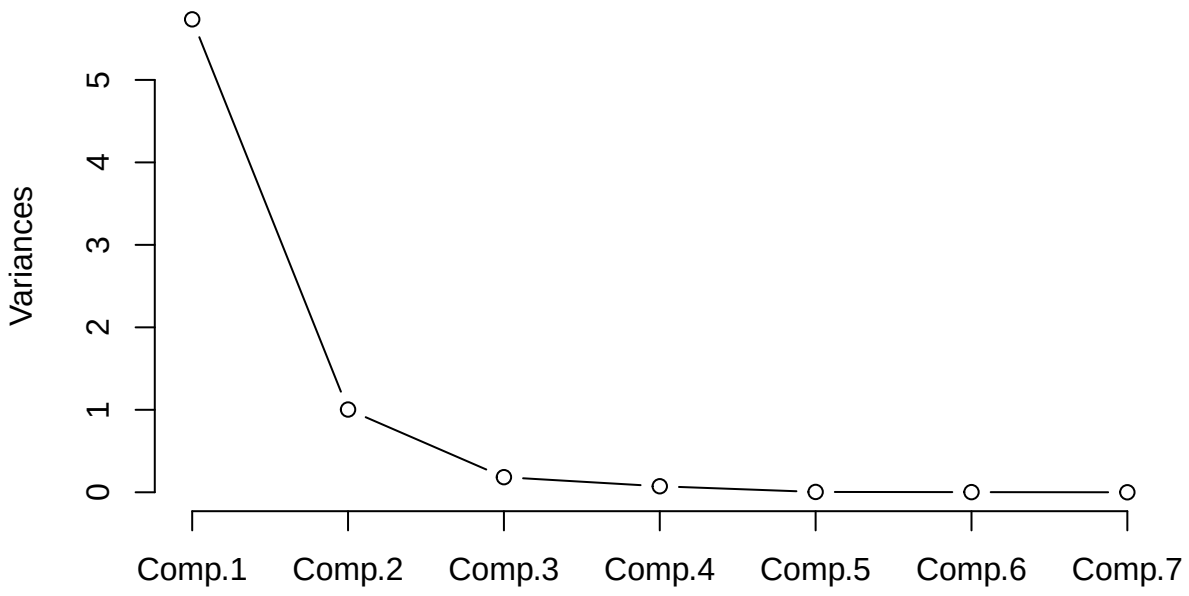
```
glimpse(Homeless_Families)
```

```
## Observations: 9
## Variables: 7
## $ Year                              <dbl> 2010, 2011, 2012, 2013, 2014, 201...
## $ `Homeless in Family`              <dbl> 0.0000000, 1.0000000, 0.4568040, ...
## $ CPI                               <dbl> 0.0000000, 0.1579940, 0.2665358, ...
## $ `Minimum Wages`                   <dbl> 0.0000000, 0.0000000, 0.0000000, ...
## $ `Unemployment Population`         <dbl> 0.00000000, 0.05263158, 0.2105263...
## $ `House Rental Price`              <dbl> 0.01904762, 0.02095238, 0.0000000...
## $ `Drug addict ED visit per 10,000` <dbl> 0.0000000, 0.1315583, 0.1812843, ...
```

```
Homeless_Families.pr<-princomp(Homeless_Families,cor=TRUE)
summary(Homeless_Families.pr,loadings=TRUE)
```

```
## Importance of components:
##                            Comp.1    Comp.2     Comp.3     Comp.4       Comp.5
## Standard deviation     2.3946496 1.0016528 0.42821093 0.27042743 0.0634600679
## Proportion of Variance 0.8191924 0.1433298 0.02619494 0.01044729 0.0005753115
## Cumulative Proportion  0.8191924 0.9625221 0.98871709 0.99916437 0.9997396859
##                              Comp.6       Comp.7
## Standard deviation     0.0425590993 3.304810e-03
## Proportion of Variance 0.0002587538 1.560253e-06
## Cumulative Proportion  0.9999984397 1.000000e+00
##
## Loadings:
##                                  Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Year                              0.415         0.126  0.305  0.280  0.152
## Homeless in Family                      -0.981  0.144
## CPI                               0.409        -0.276  0.556 -0.510  0.333
## Minimum Wages                     0.393        -0.733 -0.485  0.252
## Unemployment Population           0.409  0.115  0.374  0.102  0.591  0.119
## House Rental Price                0.400  0.141  0.458 -0.580 -0.475  0.209
## Drug addict ED visit per 10,000   0.417                      -0.147 -0.893
##                                  Comp.7
## Year                              0.785
## Homeless in Family
## CPI                              -0.269
## Minimum Wages
## Unemployment Population          -0.553
## House Rental Price
## Drug addict ED visit per 10,000
```

```
screeplot(Homeless_Families.pr,type="lines")
```




```
p<-predict(Homeless_Families.pr)
```

Homeless_Sheltered

Principal Component Analysis

Body:

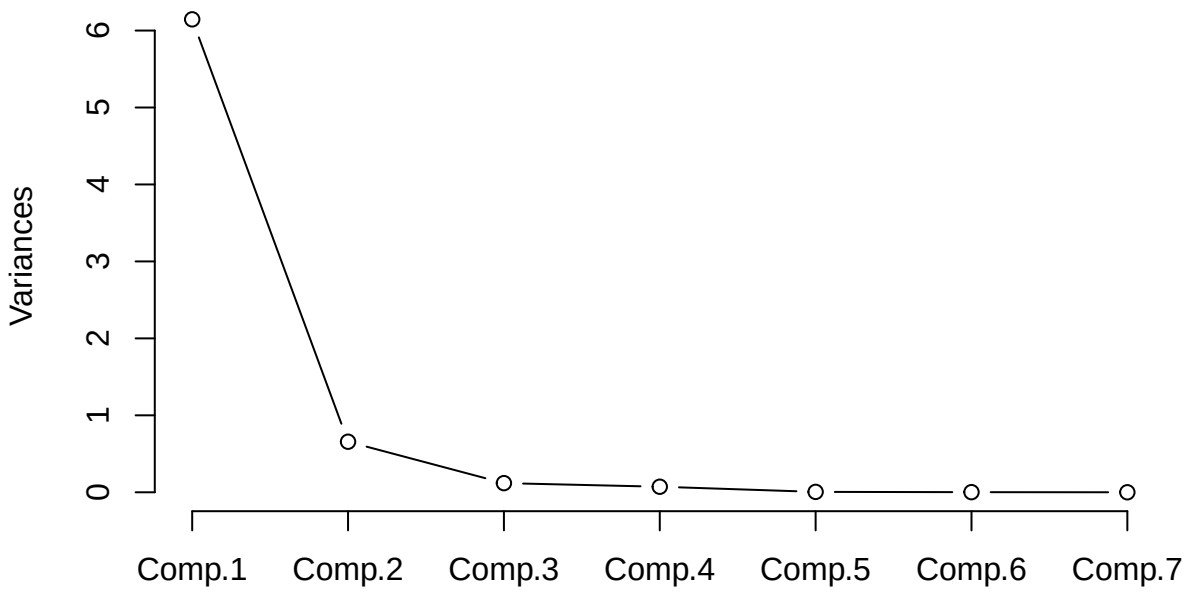
```
glimpse(Homeless_Sheltered)
```

```
## Observations: 9
## Variables: 7
## $ Year                              <dbl> 2010, 2011, 2012, 2013, 2014, 201...
## $ `Sheltered Homeless`              <dbl> 0.5635586, 1.0000000, 0.4894764, ...
## $ CPI                               <dbl> 0.0000000, 0.1579940, 0.2665358, ...
## $ `Minimum Wages`                   <dbl> 0.0000000, 0.0000000, 0.0000000, ...
## $ `Unemployment Population`         <dbl> 0.00000000, 0.05263158, 0.2105263...
## $ `House Rental Price`              <dbl> 0.01904762, 0.02095238, 0.0000000...
## $ `Drug addict ED visit per 10,000` <dbl> 0.0000000, 0.1315583, 0.1812843, ...
```

```
Homeless_Sheltered.pr<-princomp(Homeless_Sheltered,cor=TRUE)
summary(Homeless_Sheltered.pr,loadings=TRUE)
```

```
## Importance of components:
##                            Comp.1     Comp.2     Comp.3     Comp.4       Comp.5
## Standard deviation     2.4790191 0.81042459 0.34476877 0.26850204 0.0717059258
## Proportion of Variance 0.8779337 0.09382686 0.01698079 0.01029905 0.0007345343
## Cumulative Proportion  0.8779337 0.97176056 0.98874134 0.99904039 0.9997749264
##                              Comp.6       Comp.7
## Standard deviation     0.0396901336 4.563008e-04
## Proportion of Variance 0.0002250438 2.974434e-08
## Cumulative Proportion  0.9999999703 1.000000e+00
##
## Loadings:
##                                  Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Year                              0.401                0.341  0.287  0.159
## Sheltered Homeless               -0.280  0.874  0.340  0.188
## CPI                               0.390  0.198 -0.459  0.392 -0.379  0.441
## Minimum Wages                     0.370  0.402 -0.349 -0.728  0.221
## Unemployment Population           0.399         0.305  0.190  0.653
## House Rental Price                0.390         0.677 -0.305 -0.451  0.296
## Drug addict ED visit per 10,000   0.400  0.145         0.189 -0.305 -0.830
##                                  Comp.7
## Year                              0.783
## Sheltered Homeless
## CPI                              -0.325
## Minimum Wages
## Unemployment Population          -0.528
## House Rental Price
## Drug addict ED visit per 10,000
```

```
screeplot(Homeless_Sheltered.pr,type="lines")
```




```
p<-predict(Homeless_Sheltered.pr)
```