

# PCA

2023-09-08

## Load data

```
df = read_csv("../PCA_data_01.csv")

## Rows: 414 Columns: 9
## -- Column specification -----
## Delimiter: ","
## dbl (9): PolPop, PolTerr, CapPop, levels, government, infrastr, writing, tex...
##
## 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.
summary(df)
```

	PolPop	PolTerr	CapPop	levels
## Min.	:1.069	Min. : -0.7632	Min. : 0.9861	Min. : 0.2243
## 1st Qu.:	:4.226	1st Qu.: 3.7052	1st Qu.: 3.4988	1st Qu.: 2.4151
## Median :	:6.006	Median : 5.2282	Median : 4.4746	Median : 3.5706
## Mean :	:5.559	Mean : 4.8039	Mean : 4.2784	Mean : 3.4651
## 3rd Qu.:	:6.832	3rd Qu.: 5.9542	3rd Qu.: 5.1640	3rd Qu.: 4.4078
## Max.	:8.534	Max. : 7.4018	Max. : 6.1995	Max. : 7.2363

	government	infrastr	writing	texts
## Min.	:0.0000	Min. : 0.0000	Min. : 0.0000	Min. : 0.0000
## 1st Qu.:	:0.2333	1st Qu.: 0.3500	1st Qu.: 0.2375	1st Qu.: 0.1000
## Median :	:0.6455	Median : 0.7646	Median : 0.8167	Median : 0.9310
## Mean :	:0.5576	Mean : 0.6464	Mean : 0.6430	Mean : 0.6419
## 3rd Qu.:	:0.8612	3rd Qu.: 0.9000	3rd Qu.: 0.8571	3rd Qu.: 0.9825
## Max.	:1.0000	Max. : 1.0000	Max. : 1.0000	Max. : 1.0312

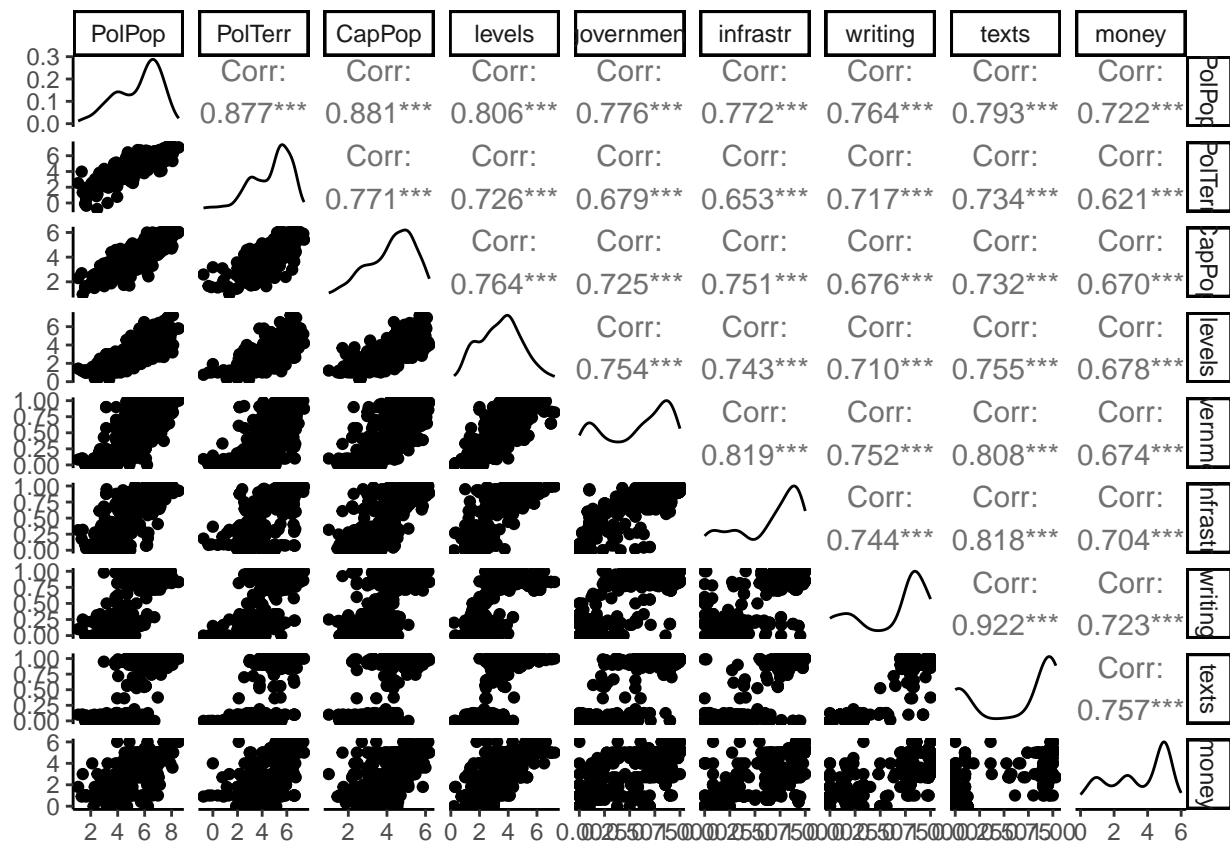
  

	money
## Min.	:0.000
## 1st Qu.:	:1.800
## Median :	:4.000
## Mean :	:3.397
## 3rd Qu.:	:5.000
## Max.	:6.000

## Correlation

```
library(GGally)

## Registered S3 method overwritten by 'GGally':
## method from
## +.gg ggplot2
ggpairs(df) + theme_classic()
```



## PCA

```
library(ggfortify)

pca_res <- prcomp(df, scale. = TRUE)

autoplot(pca_res) + theme_classic()
```

