

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ ДНР  
**ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ**

**ФИЗИКО-ТЕХНИЧЕСКИЙ ФАКУЛЬТЕТ**

**ЛАБОРАТОРНАЯ РАБОТА №4**

По теме: «Анализ данных с помощью метода главных компонент средствами языка R.»

По предмету: «Машинное обучение»

Выполнил:

студент 2 курса

группы ИВТ-3

Дворников Г.Н.

Проверил:

Савенков И.Н.

Донецк  
2021

**Лабораторная работа № 4**

**Тема**: Анализ данных с помощью метода главных компонент средствами языка R.

**Цель:** изучить возможности анализа данных, предоставляемые PCA-моделированием, средствами языка R.

**Ход** **работы**

Переменные:

* mstat: Математика и статистика
* econ: Экономика
* eng: Английский язык
* polth: История политических учений
* mstat2: Математика и статистика (часть 2)
* phist: Политическая история
* law: Право
* phil: Философия
* polsoc: Политическая социология
* ptheo: Политическая теория
* compp: Сравнительная политология
* game: Теория игр
* wpol: Мировая политика и международные отношения
* male: пол (1 - мужской, 0 - женский)

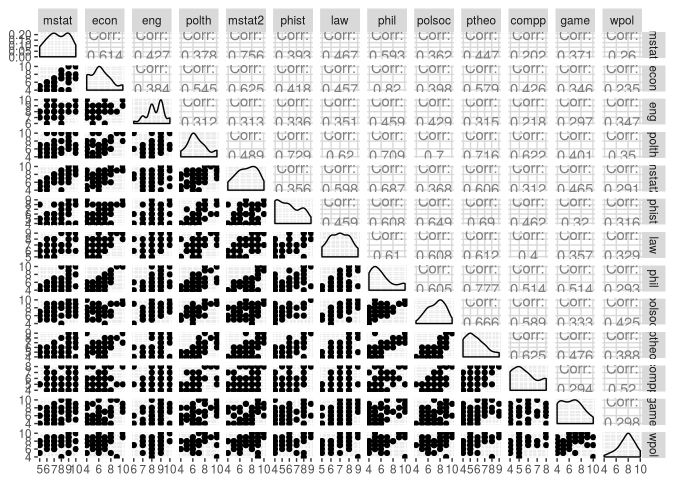


Рисунок 1 – Матрица диаграмм рассеяния

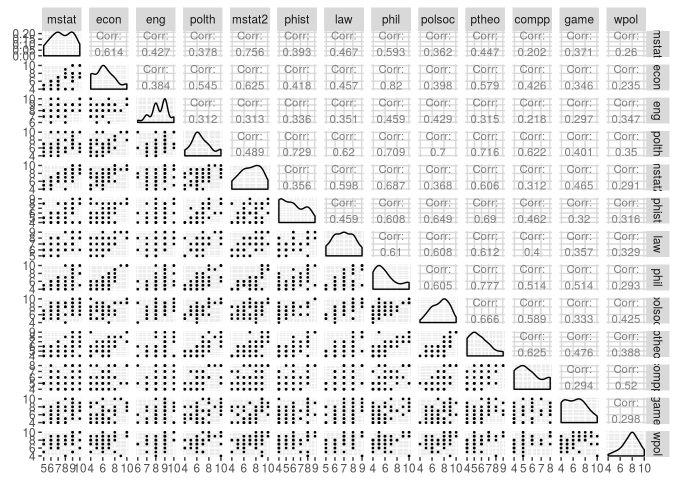


Рисунок 2 – Матрица рассеяния 2

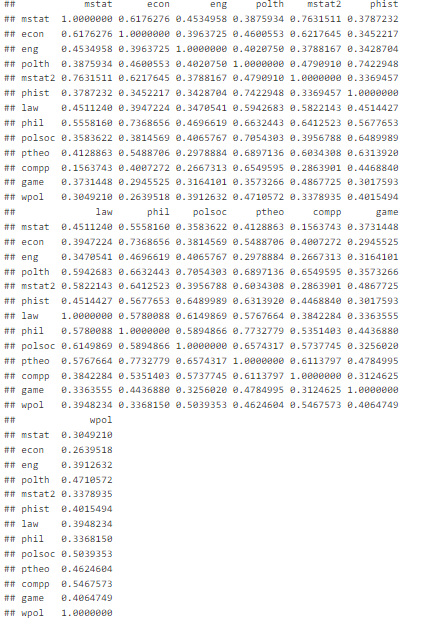


Рисунок 3 – Коэффициенты корреляции Спирмена

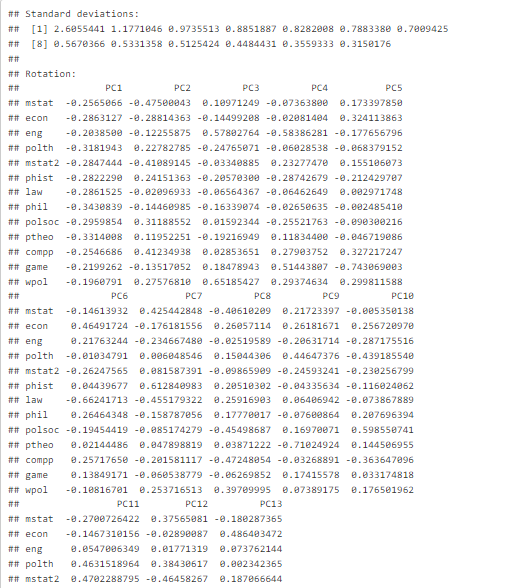


Рисунок 4 – Стандартные отклонения данных методом главных компонент



Рисунок 5 – Отдельные отклонения

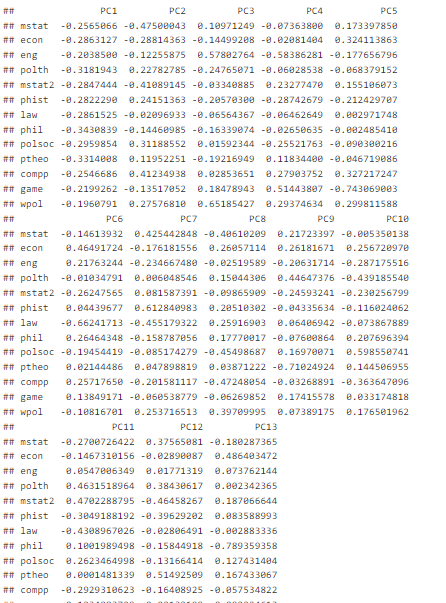


Рисунок 6 – Матрица перехода

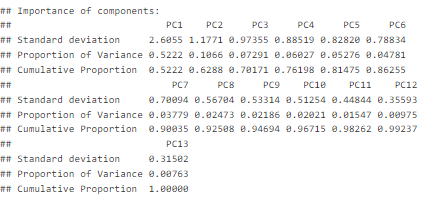


Рисунок 7 – Информативность главных компонент

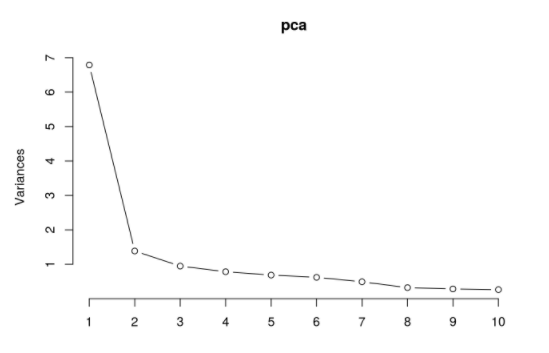


Рисунок 8 – График матрицы

Листинг программы:

df <- read.csv("scores3.csv")

head(df)

df <- na.omit(df)

d <- df[1:ncol(df)-1]

ggpairs(d)

ggpairs(d,

upper = list(continuous = wrap("cor", size = 3)),

lower = list(continuous = wrap("points", size = 0.3)))

cor(d, method = "spearman")

pca <- prcomp(d, center = TRUE, scale = TRUE)

pca

pca$sdev

pca$rotation

summary(pca)

plot(pca, type = "l")