Лабораторная работа №1

Кудлай Я.

24 Декабрь 2020

# Раздел 1.

В этом разделе необходимо собрать данные по социально-экономическим показателям в регионах РФ и изучить их взаимосвязи.

# **Необходимо построить:**

* Гистограммы распределения с наложенными теоретическими кривыми плотности нормального закона, разместить графики на одном полотне.
* Тест Шапиро-Уилка на нормальность: таблица со статистикой теста, P-значением и выводом по каждой переменной.
* Точечные графики взаимного разброса, сделать маркеры полупрозрачными.
* Взаимосвязь показателей: графическое представление корреляционной матрицы. Незначимые коэффициенты корреляции скрыть, остальные показать числами.

## Обозначение показателей:

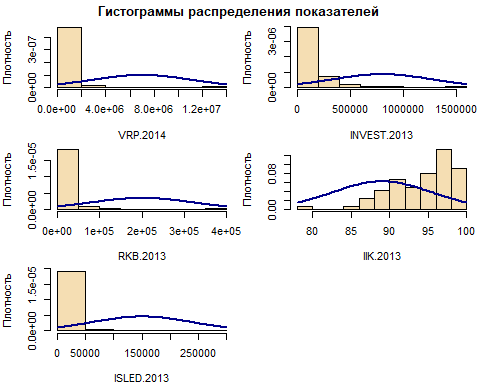
* *VRP.2014* - Валовый региональный продукт
* *INVEST.2013* - Инвестиции в основной капитал
* *RKB.2013* - Расходы консолидированных бюджетов субъектов Российской Федерации: на национальную экономику
* *IIK.2013* - Использование информационных и коммуникационных технологий в организациях: персональные компьютеры
* *ISLED.2013* - Внутренние затраты на научные исследования и разработки

## Описательная статистика

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VRP.2014 | INVEST.2013 | RKB.2013 | IIK.2013 | ISLED.2013 |
| Среднее | 715713.1 | 157593.5 | 20975.8 | 94.4 | 9143.4 |
| Стандартное отклонение | 1474868.8 | 212058.3 | 42846.5 | 4.0 | 32363.7 |
| Коэффициент вариации, % | 206.1 | 134.6 | 204.3 | 4.2 | 354.0 |

## Анализ распределния данных

Построим гистограммы с теоретической нормальной плотностью распределения.



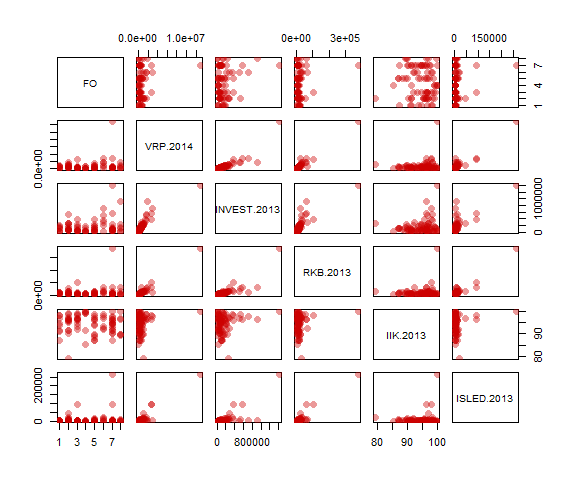
## Тест Шапиро-Уилка на нормальность

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VRP.2014 | INVEST.2013 | RKB.2013 | IIK.2013 | ISLED.2013 |
| Статистика | 0.35276 | 0.61157 | 0.32271 | 0.93184 | 0.26884 |
| P-значение | 0.00000 | 0.00000 | 0.00000 | 0.00030 | 0.00000 |

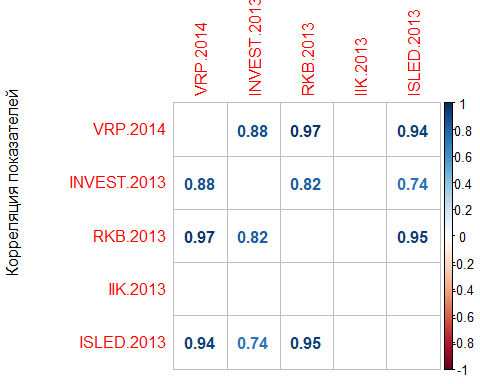
**Вывод:** По итогам теста Шапиро-Уилка на нормальность распредления логарифмированных переменных, при уровне значимости 0,05, отвергаются все переменные *‘VRP.2014’*, *‘INVEST.2013’*, *‘RKB.2013’*,*‘IIK.2013’* и *‘ISLED.2013’*, т.к. P-значения < 0.05.

## Анализ взаимосвязей показателей

### Графики разброса



### Корреляционная матрица



**Вывод:** Наблюдается сильная корреляционная связь между показателями *‘VRP.2014’* и *‘INVEST.2013’*, *‘VRP.2014’* и *‘RKB.2013’*, *‘VRP.2014’* и *‘ISLED.2013’*, *‘INVEST.2013’* и *‘RKB.2013’*,*‘INVEST.2013’* и *‘ISLED.2013’*,*‘RKB.2013’* и *‘ISLED.2013’*.

# Раздел 2.

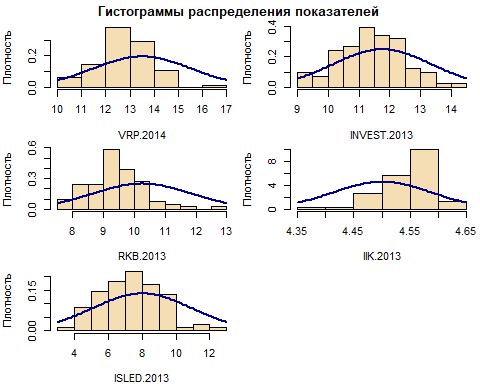
В этом разделе необходимо произвести теже самые действия, с логарифмированными исходными данными.

## Описательная статистика логарифмированных данных:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VRP.2014 | INVEST.2013 | RKB.2013 | IIK.2013 | ISLED.2013 |
| Среднее | 12.8 | 11.4 | 9.4 | 4.5 | 7.3 |
| Стандартное отклонение | 1.1 | 1.0 | 0.9 | 0.0 | 1.8 |
| Коэффициент вариации, % | 8.6 | 8.8 | 9.6 | 0.0 | 24.7 |

## Анализ распределния логарифмированных данных

Построим гистограммы логарифмированных данных с их теоретической нормальной плотностью распределения.



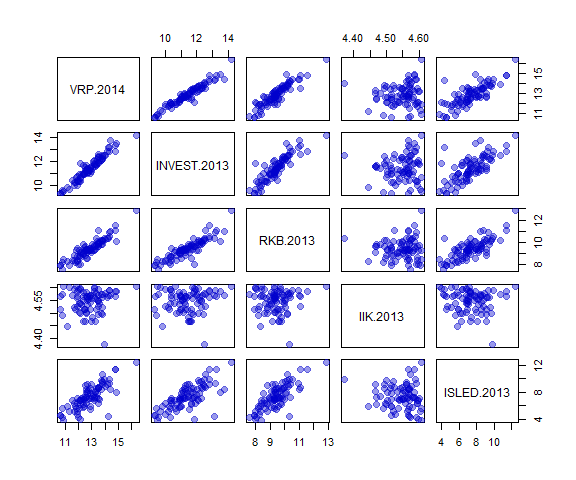
## Тест Шапиро-Уилка на нормальность для логарифмированных данных

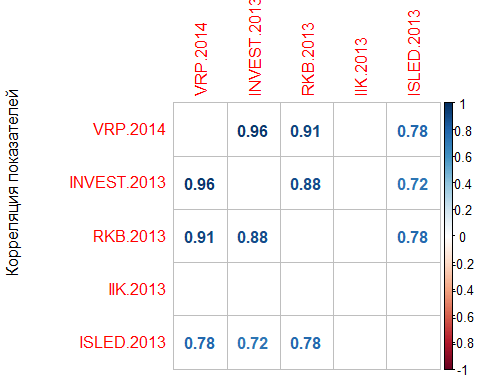
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | VRP.2014 | INVEST.2013 | RKB.2013 | IIK.2013 | ISLED.2013 |
| Статистика | 0.98504 | 0.99068 | 0.96586 | 0.91856 | 0.98598 |
| P-значение | 0.45914 | 0.82509 | 0.02786 | 0.00007 | 0.51566 |

**Вывод:** По итогам теста Шапиро-Уилка на нормальность распредления логарифмированных переменных, при уровне значимости 0,05, отвергаются все переменные, кроме *‘IIK.2013’* и *‘RKB.2013’*, т.к. P-значения < 0.05. А переменные *‘VRP.2014’*, *‘INVEST.2013’* и *‘ISLED.2013’* не отвергаются.

## Анализ взаимосвязей логарифмированных показателей

### Графики разброса логарифмированных величин





**Вывод:** Наблюдается сильная корреляционная связь между показателями *‘VRP.2014’* и *‘INVEST.2013’*, *‘VRP.2014’* и *‘RKB.2013’*, *‘VRP.2014’* и *‘ISLED.2013’*, *‘INVEST.2013’* и *‘RKB.2013’*,*‘INVEST.2013’* и *‘ISLED.2013’*,*‘RKB.2013’* и *‘ISLED.2013’*.