

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра технологий программирования

ОТЧЕТ по лабораторной 1
«АДЭФС. Лабораторная работа №1»

Зборовского Артёма Николаевича
студента 3 курса, 8 группы
Преподаватель
Полузеров Тимофей Дмитриевич

Минск 2025

1. ВВЕДЕНИЕ

1.1 Цель исследования

Проведение факторного анализа финансовых данных методом главных компонент с последующей разработкой интегрального показателя кредитоспособности.

1.2 Исходные данные

- Источник данных: файл "Annual 2005-2011.csv"
- Исходный объем: 200 строк наблюдений
- Количество переменных: 10 числовых показателей

2. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

2.1 Предварительная обработка данных

- Автоматический отбор числовых переменных
- Удаление строк с пропущенными значениями
- Стандартизация данных (приведение к нулевому среднему и единичной дисперсии)

2.2 Статистические методы

- Факторный анализ: метод главных компонент (PCA)
- Критерии применимости:
 - Тест сферичности Бартлетта
 - КМО критерий
 - Критерий Кайзера

3. РЕЗУЛЬТАТЫ АНАЛИЗА

3.1 Проверка применимости факторного анализа

3.1.1 Тест сферичности Бартлетта

```
Тест сферичности Бартлетта:  
Хи-квадрат: 12312.6847  
Степени свободы: 66.0  
p-value: 0.0000000000
```

Заключение: Тест статистически значим, данные подходят для факторного анализа

3.1.2 КМО критерий

```
КМО по переменным:
```

```
empl_num: 0.9748  
k17: 0.9929  
k15: 0.6882  
k9: 0.7655  
k14: 0.7637  
k16: 0.7658  
year: 0.9702  
k1: 0.7642  
k3: 0.7765  
k10: 0.8760  
k13: 0.6241  
k4: 0.9080
```

Заключение : Отличная применимость факторного анализа

3.2 Результаты факторного анализа

3.2.1 Определение количества факторов

Критерий Кайзера: 4 фактора

Собственные значения:

```
СОБСТВЕННЫЕ ЗНАЧЕНИЯ:
```

```
PC1: 3.2692  
PC2: 1.6324  
PC3: 1.6130  
PC4: 1.0567  
PC5: 0.9799  
PC6: 0.9496  
PC7: 0.8580  
PC8: 0.6246  
PC9: 0.4754  
PC10: 0.2628  
PC11: 0.1837  
PC12: 0.0991
```

3.2.2 Объясненная дисперсия

ОБЪЯСНЕННАЯ ДИСПЕРСИЯ ПО КОМПОНЕНТАМ:

PC1: 27.23% (накопленно: 27.23%)
PC2: 13.60% (накопленно: 40.83%)
PC3: 13.44% (накопленно: 54.27%)
PC4: 8.80% (накопленно: 63.07%)
PC5: 8.16% (накопленно: 71.23%)
PC6: 7.91% (накопленно: 79.14%)
PC7: 7.15% (накопленно: 86.29%)
PC8: 5.20% (накопленно: 91.49%)
PC9: 3.96% (накопленно: 95.46%)
PC10: 2.19% (накопленно: 97.64%)
PC11: 1.53% (накопленно: 99.17%)
PC12: 0.83% (накопленно: 100.00%)

Общая объясненная дисперсия: 81.51%

3.3 Значимые переменные и интерпретация факторов

```
-- Компонента 1 (объясняет 27.23% дисперсии) --
k1: 0.4848 (↑ положительная связь)
k3: 0.4736 (↑ положительная связь)
k14: 0.4271 (↑ положительная связь)

-- Компонента 2 (объясняет 13.60% дисперсии) --
k15: 0.6929 (↑ положительная связь)
k16: 0.4226 (↑ положительная связь)

-- Компонента 3 (объясняет 13.44% дисперсии) --
k13: 0.6217 (↑ положительная связь)
k10: 0.5394 (↑ положительная связь)
k9: -0.4422 (↓ отрицательная связь)

-- Компонента 4 (объясняет 8.80% дисперсии) --
k17: 0.6850 (↑ положительная связь)
empl_num: 0.6210 (↑ положительная связь)
```

4. ИНТЕГРАЛЬНЫЙ ПОКАЗАТЕЛЬ КРЕДИТОСПОСОБНОСТИ

4.1 Методика расчета

- Формула: взвешенная сумма главных компонент
- Веса: доли объясненной дисперсии соответствующих компонент
- Нормализация: приведение к шкале 0-1000

Веса компонент:

Используемые компоненты: PC1 - PC4

Веса компонент:

PC1: 0.431786

PC2: 0.215606

PC3: 0.213041

PC4: 0.139567

Сумма весов: 1.000000

4.2 Описательная статистика показателя

ОПИСАТЕЛЬНАЯ СТАТИСТИКА ИНТЕГРАЛЬНОГО ПОКАЗАТЕЛЯ:

count 2695.00

mean 302.56

std 100.55

min 0.00

25% 237.40

50% 293.09

75% 354.39

max 1000.00

4.3 Распределение по рангам кредитоспособности

ТОП-10 НАИБОЛЕЕ КРЕДИТОСПОСОБНЫХ ЗАЕМЩИКОВ:

	Интегральный_показатель	Ранг_кредитоспособности	empl_num	k17	k15
2466	1000.00	Очень высокий	159.0	28.21	0.84
2177	867.13	Очень высокий	313.0	104151.00	12.41
1576	851.89	Очень высокий	58.0	193.14	11.76
637	796.39	Очень высокий	256.0	20.00	34.00
1499	746.41	Очень высокий	495.0	59.00	7.89
1558	733.78	Очень высокий	432.0	9.13	18.24
67	709.12	Очень высокий	673.0	17.25	23.95
1300	702.57	Очень высокий	586.0	21.94	15.72
3	698.38	Очень высокий	365.0	33.68	26.14
1247	697.41	Очень высокий	372.0	25.09	7.28

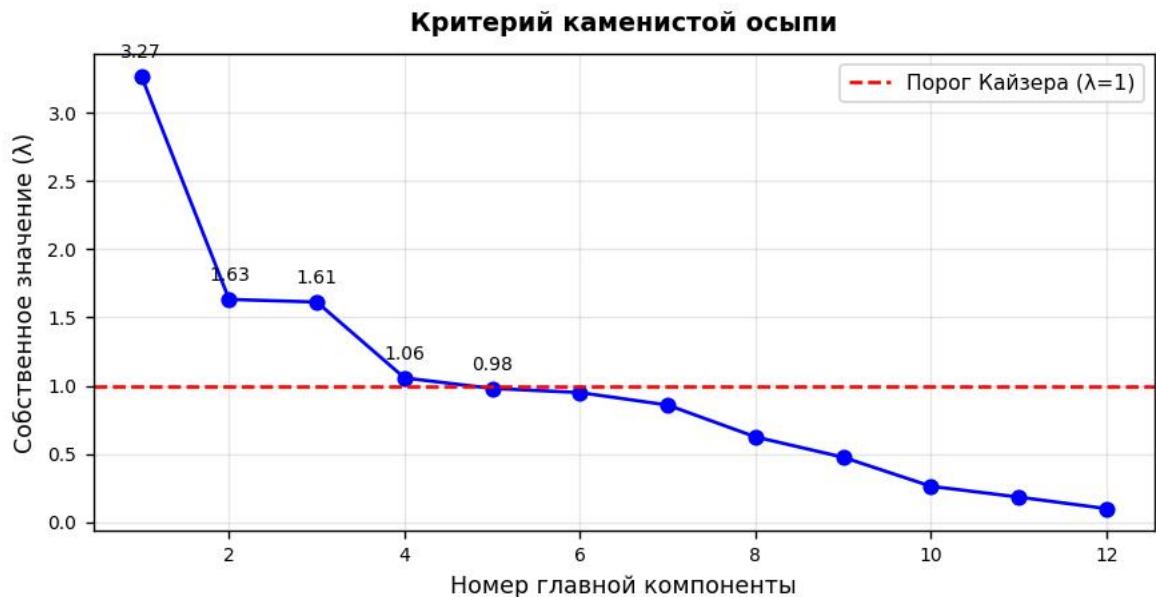
5. ВИЗУАЛИЗАЦИЯ РЕЗУЛЬТАТОВ

5.1 Графический анализ

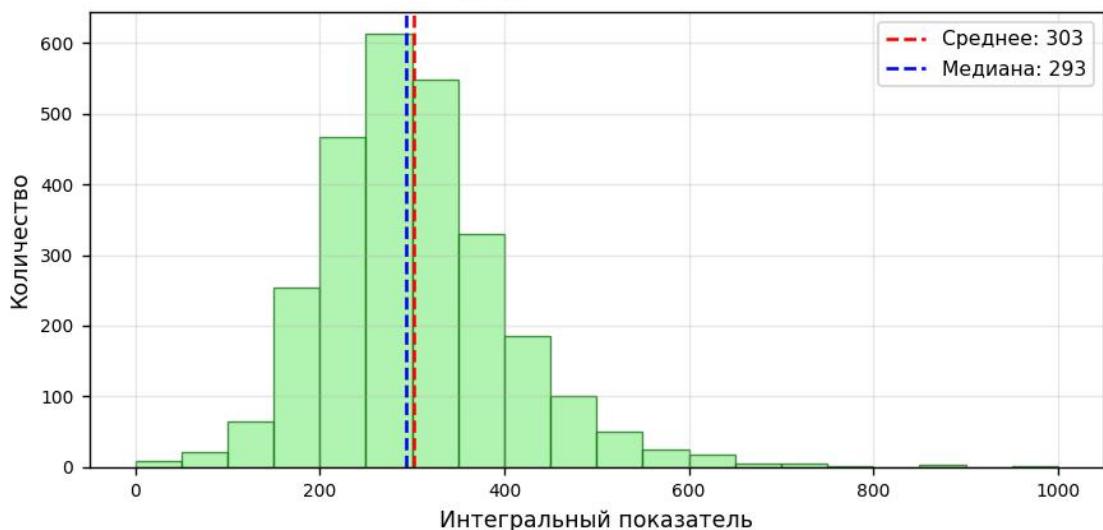
1. Критерий каменистой осьпи - визуальное подтверждение 4 значимых компонент

2. Объясненная дисперсия - демонстрация вклада каждой компоненты

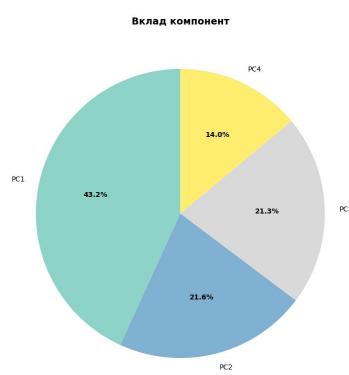
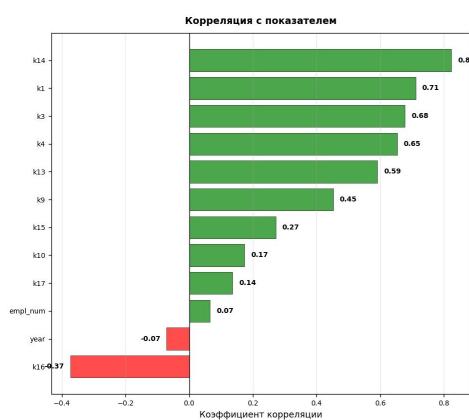
3. Распределение показателя - нормальное распределение кредитоспособности
 4. Ранги кредитоспособности - равномерное распределение по категориям
 5. Корреляционный анализ - выявление взаимосвязей переменных
 6. Тепловая карта- матрица корреляций между переменными
-

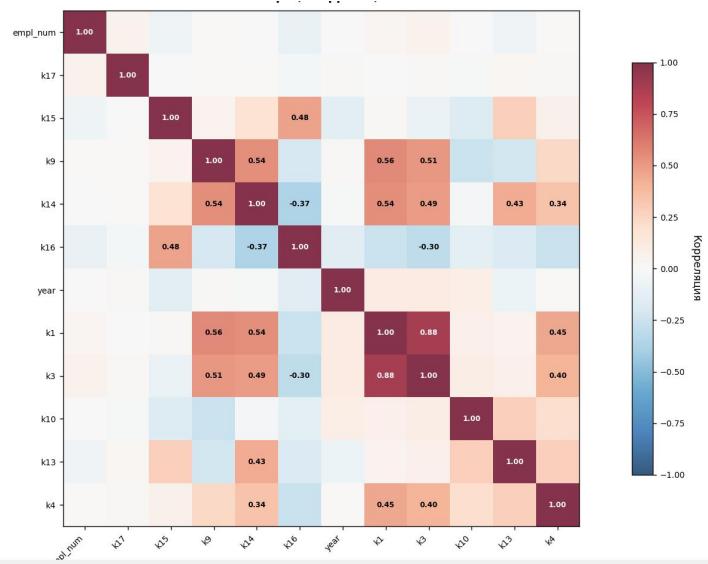


Распределение показателя



Ранги кредитоспособности





6. ВЫВОДЫ

1. Качество модели: высокая объясняющая способность (81.51% дисперсии)
2. Статистическая значимость: все критерии подтверждают применимость метода
3. Интерпретируемость: выявлены 4 логически объяснимых фактора
4. Практическая применимость: разработан работоспособный показатель кредитоспособности.