



Selling Pandas

1. Иванов Герман
2. Иванова Полина
3. Михайлов Андрей
4. Кириченко Дмитрий
5. Спирихин Андрей
6. Кром Илья

Ты еще не совсем зеленый, сынок



Структура данных и очистка от выбросов

Информация

Опрос компаний

48 вопросов

1323 компании

2019 год

Очистка

Вопросы: VMGc23a – VMGc23j

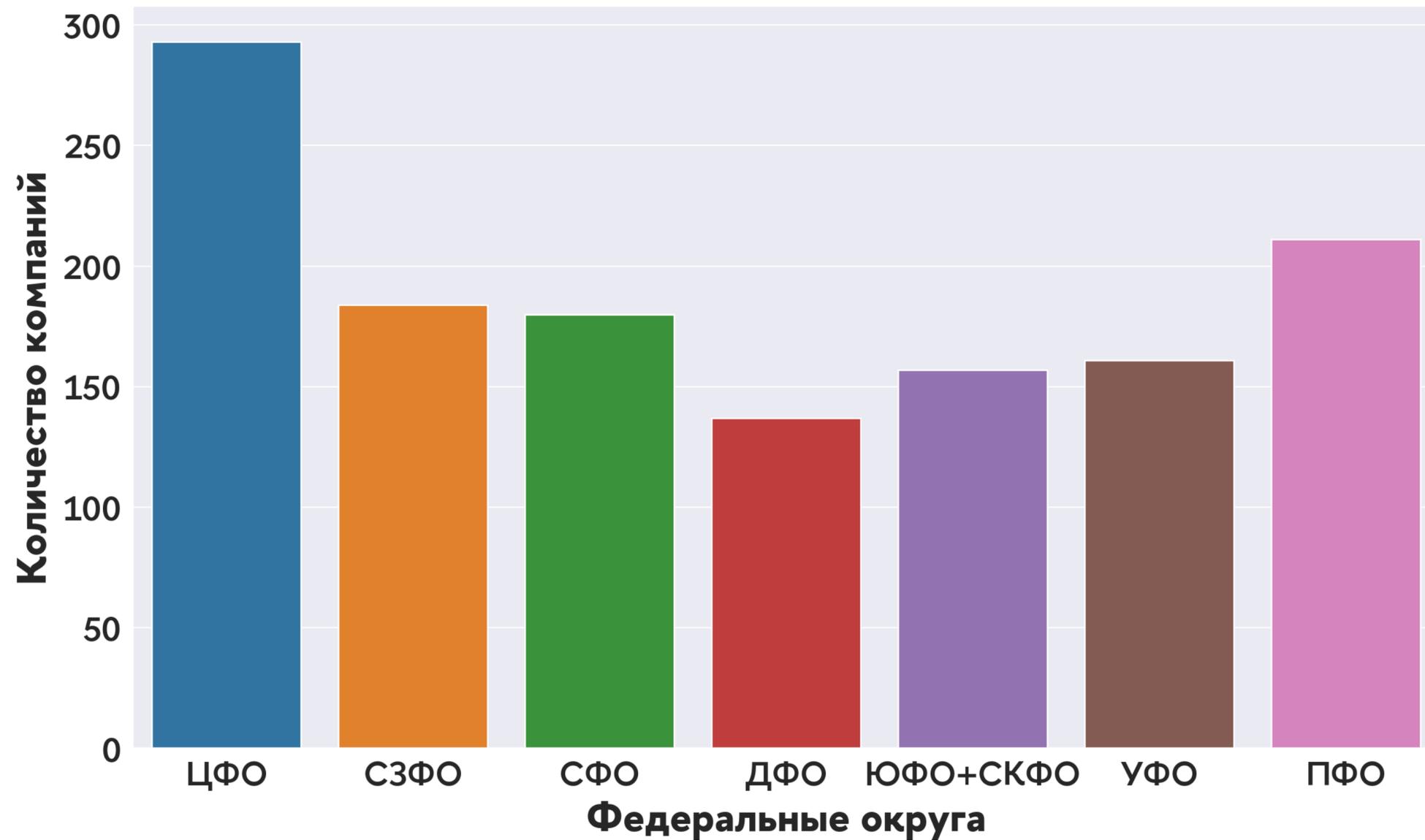
≥ 1 "Не применимо"

1323 строки → 1228 строк



Предварительный анализ

Распределение по федеральным округам

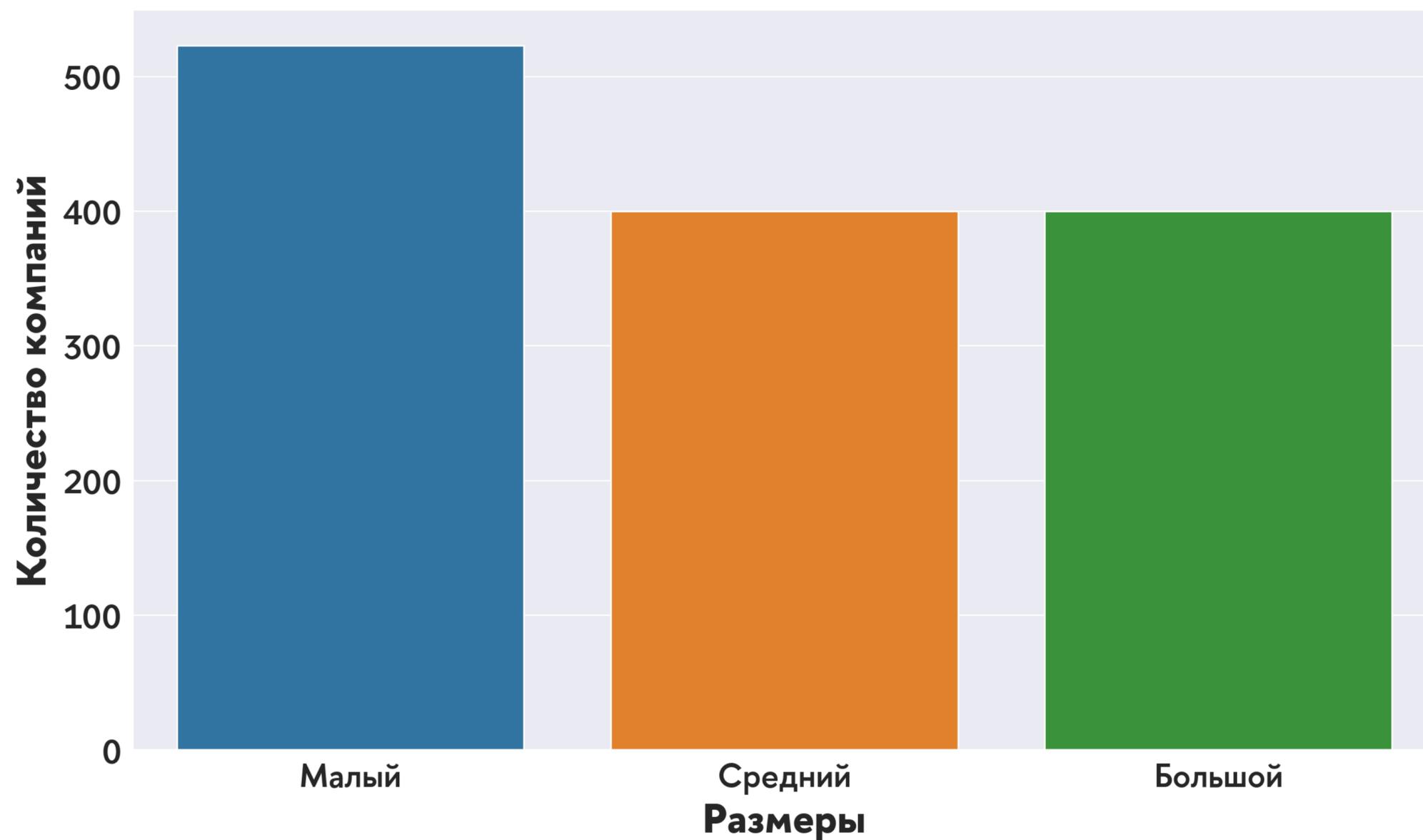


Мода: Центральный
Федеральный
Округ



Предварительный анализ

Распределение по размеру



Малый: 474

Средний: 372

Большой: 382



Индекс экологичности компании

Набор мер, принятых с целью озеленения производства

Индекс обоснован с использованием материалов ООН¹

Алгоритм

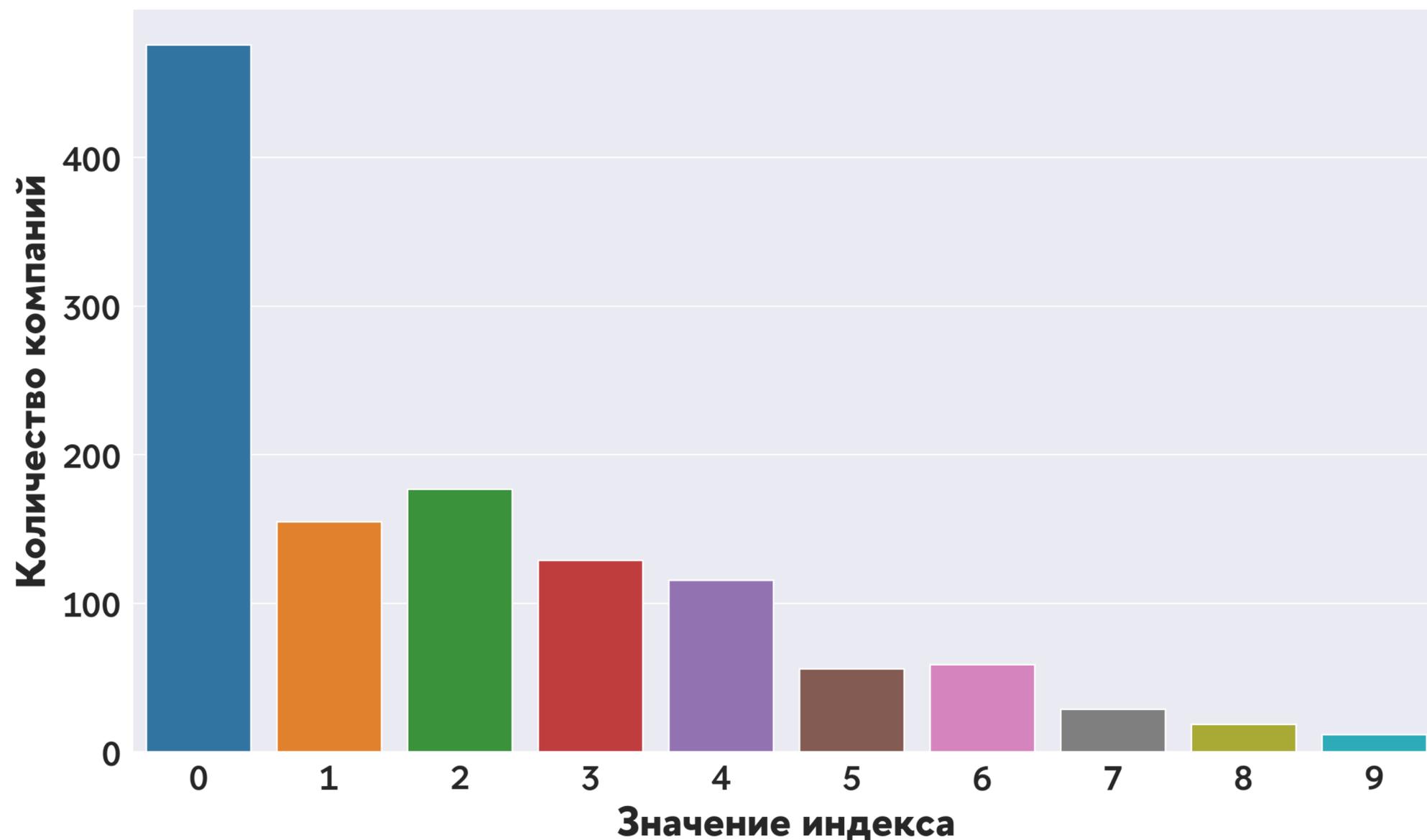
Да	+1
Нет	0
Не знаю	0
Не применимо	Drop

$$0 \leq \text{score} \leq 9$$



Индекс экологичности

Распределение индекса вовлечённости в "озеленение" производства



Среднее: 2.23

Медиана: 1

Стандартное отклонение: 2.22



Матрица корреляций

Размер фирмы	1	-0.29	-0.2	0.12	0.32
Год создания фирмы	-0.29	1	0.095	0.097	-0.23
Женщина ли глава?	-0.2	0.095	1	-0.065	-0.15
Насколько экологичность препятствие?	0.12	0.097	-0.065	1	0.038
Индекс вовлечён- ности	0.32	-0.23	-0.15	0.038	1
	Размер фирмы	Год создания фирмы	Женщина ли глава?	Насколько экологичность препятствие?	Индекс вовлечён- ности



Исследовательский вопрос и гипотеза

Исследовательский вопрос

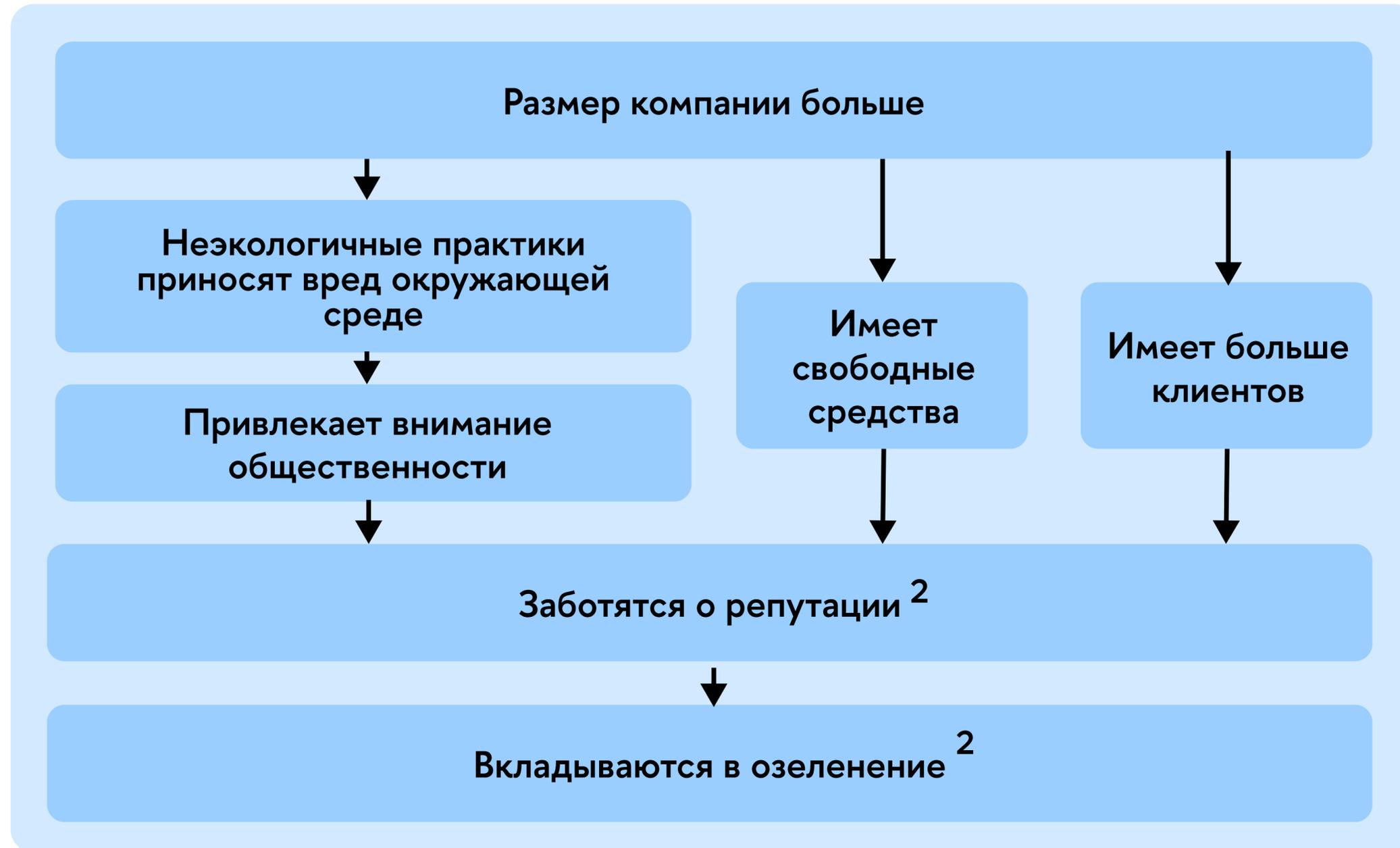
Как взаимосвязаны
характеристики фирмы с
вовлечённостью данной фирмы в
процесс "озеленения"
производства?

Гипотеза

Компании **большого размера**
более вовлечены в "озеленение"
производства



Механизм





Обработка данных

Замена заголовков

аба



company_size

Кодирование порядковых категориальных переменных

Крупная

3

Средняя



2

Малая

1

Создание «Индекса экологичности компании»



Математическая модель

Итоговый объем выборки

1228

Уровень стат. значимости

1%

H0

Связи между исследуемыми переменными нет или она отрицательна

H1

Связь между исследуемыми переменными есть и она положительна



Линейная регрессия

Уравнение регрессионной прямой

$$involved_index_i \sim \beta_0 + \beta_1 \cdot company_size_i + \epsilon_i$$

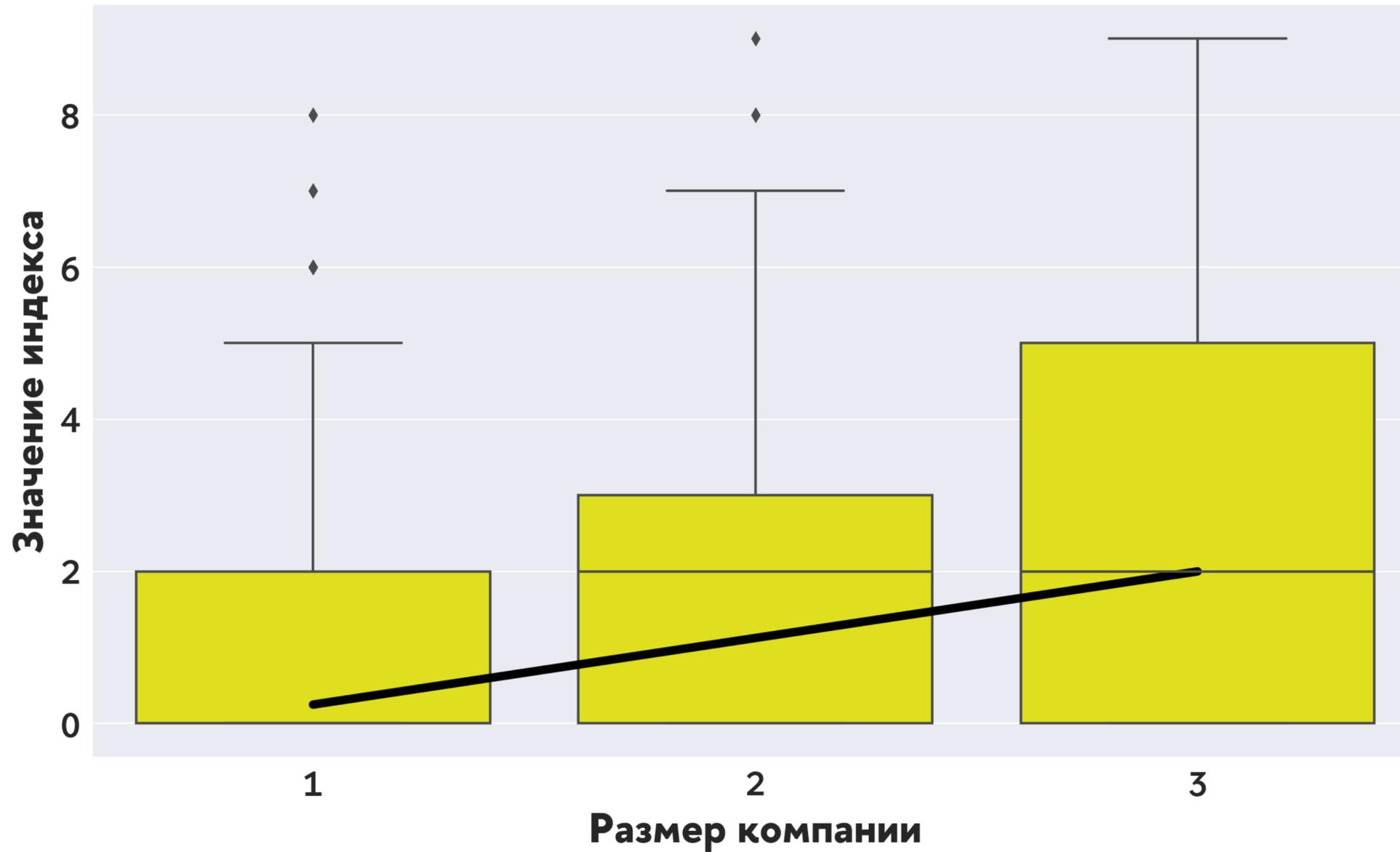
Результат

Показатель	Переменная	β_i	p-value	Значимость
Размер компании	company_size	0.866	<0.01	✓
Свободный член	-	0.334	-	-



Линейная регрессия

Зависимость индекса вовлечённости от размера компании





Контрольные переменные

Переменная	Объясняемая характеристика компании
year	Год основания
flag_woman	Является ли главой женщина?
eco_trouble	Насколько экологичность является препятствием?



Множественная линейная регрессия

Регрессионная прямая

$$\begin{aligned} \text{involved_index}_i &\sim \beta_0 + \beta_1 \cdot \text{company_size}_i + \\ &+ \beta_2 \cdot \text{year}_i + \beta_3 \cdot \text{flag_woman}_i + \beta_4 \cdot \text{eco_trouble}_i + \epsilon_i \end{aligned}$$



Множественная линейная регрессия

Результат

Переменная	β_i	p-value	Значимость
Свободный член	38.3	-	-
company_size	0.786	<0.01	✓
year	-0.019	<0.01	✓
flag_woman	-0.33	0.13	✗
eco_trouble	-0.262	0.08	✗



Проверка модели на устойчивость

Алгоритм

Разделение на группы по федеральным округам



Построение линейной регрессии для каждой группы



Проверка результата



Разделение на группы по федеральным округам

Результат			
Группа	Объем	p-value	Значимость
Центральный федеральный округ	269	<0.01	✓
Приволжский федеральный округ	194	<0.01	✓
Северо-Западный федеральный округ	163	<0.01	✓
Сибирский федеральный округ	163	<0.01	✓
Уральский федеральный округ	159	<0.01	✓
Южный федеральный округ и Северо-Кавказский федеральный округ	149	<0.01	✓
Дальневосточный федеральный округ	132	<0.01	✓



Интерпретация



Гипотеза подтвердилась

НО на таргет влияет и год создания
компании

Произведена проверка устойчивости



Разделение по
федеральным округам



Множественная линейная регрессия



Ограничения и перспективы

Ограничения

Представлено мало компаний

Только 2019 год

Мало количественных переменных

Перспективы

Использование данных о репутации компаний для проверки механизма

Сбор большего числа финансовых данных компаний

Проведение нового опроса



Policy implication

Снижение или отмена налога на прибыль за переработку отходов

Снижение налога на прибыль для более экологичных компаний

Обязывать компании устанавливать СК потребления энергии

Привлекать внимание общ. к проблеме загрязнения окружающей среды



Спасибо за внимание!

Остались ли вопросы?

Ссылка на папку с материалами:

