



# Возможности и ограничения кластерного анализа в региональной системе качества

Александр Пронин

Ирина Богданова

ГБУ ДПО «Санкт-Петербургский центр оценки качества образования  
и информационных технологий»

# Запрос на кластерный анализ со стороны региональной системы образования

## При оценке региональной системы оценки качества образования

- Проведение анализа результатов мониторинга (отчет по результатам одной процедуры -1 балл, по результатам нескольких процедур - 2 балла, комплексный анализ по нескольким процедурам - 3 балла, **использование кластерного анализа** - дополнительно 1 балл)

## При запросе Комитета по образованию сведений о результатах работы районных систем оценки качества образования

- Необходимо включить **кластеры** образовательных организаций района Санкт-Петербурга по результатам комплексного анализа результатов оценочных процедур;
- Отчеты должны включать сравнительный анализ результатов 2018 и 2019 годов, объяснение изменений (при наличии) и сравнительный анализ результатов различных процедур в одних и тех же образовательных организациях. **По возможности следует использовать методы кластерного анализа**

## При организации городских экспериментальных площадок

- "Вариативная модель оценки качества образования в системе образования района на основе интеграции результатов оценочных процедур"



# Кластеризация и кластерный анализ в образовании

Кластеризация – выделение различных групп объектов с общими признаками.

В образовании это знакомая практика:

- «Статусные» школы;
- школы со «сложным» контингентом обучающихся;
- школы, показывающие результаты ниже средних по городу и пр.

Кластерный анализ в образовании чаще всего сводится к группировке объектов (например, школ) по заранее известному признаку: статус школы, тип района...

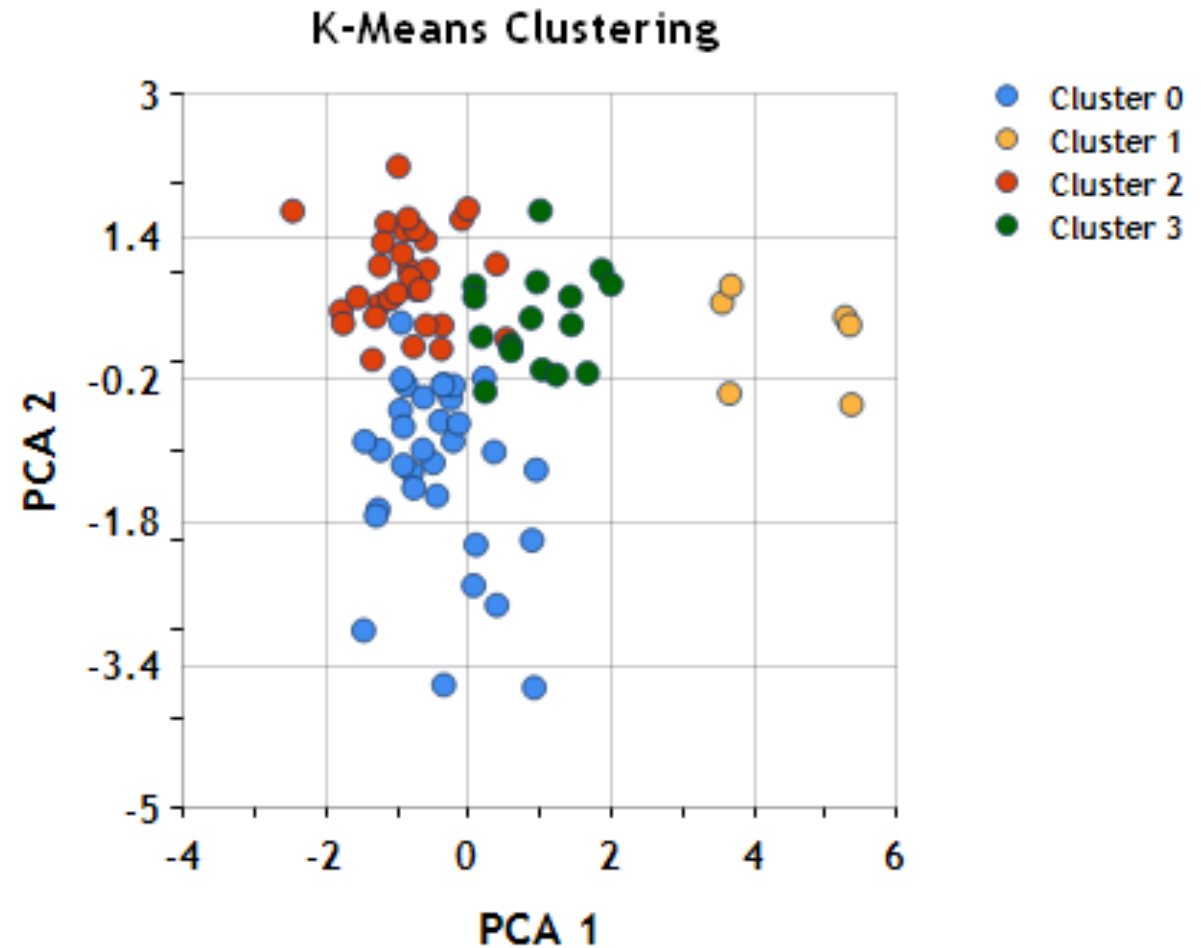


# Попробуем перевернуть задачу

Нам известны показатели многих разных объектов, но нет информации, на какие группы они в реальности делятся.

- Кластерный анализ - статистическая процедура многомерного анализа, позволяет группировать наблюдения в классы (кластеры), исходя из их сходства по измеренным признакам (переменным)

В качестве результата, получаем скрытые группы объектов по заданным переменным.



<https://www.centerspace.net/clustering-analysis-part-ii-k-means-clustering>

# Возможности работы с кластерным анализом

Использование многих переменных, на основании которых можно проводить группировку объектов

- Например, достижения педагогов, результаты учеников, социальный портрет школы в совокупности

Получение реальных групп, на которые распределяются объекты

- Мы можем узнать, какие школы похожи друг на друга вне зависимости от их территориального расположения или статуса

# Возможности работы с кластерным анализом

Сравнение тех групп, на которые мы привыкли делить объекты, с их реальными группировками

- Например, мы можем узнать, все ли школы, относящиеся к определенному виду, образуют по своим результатам единый кластер

Адресная помощь школам, демонстрирующим схожие результаты

- С 2020 года переход от фронтальных предметных региональных работ к адресным

Сравнение результатов школ, работающих в одинаковых условиях

- Анализ социального портрета и выделение, например, резильентных школ

# Этапы проведения кластерного анализа

Выбор переменных-критериев для проведения кластерного анализа

Выбор способа измерения расстояния между объектами, или кластерами

Выбор метода формирования кластеров

Интерпретация результатов





# Использования элементов кластерного анализа в региональной системе оценки качества образования: вид образовательного учреждения

Данные по ОГЭ и ЕГЭ для всех школ региона по русскому языку и математике

Будет ли структура кластеров соответствовать группам школ по их виду?

# Результаты кластерного анализа и ВИД ШКОЛ



		Кластеры				Всего
		1	2	3	4	
Вид ОУ	Гимназия	65,3%		34,7%		100%
	Лицей	71,7%	4,3%	23,9%		100%
	Средняя общеобразовательная школа	14,9%	20,8%	59,2%	5,1%	100%
	Средняя общеобразовательная школа с углубленным изучением отдельных	49,2%	2,4%	46,8%	1,6%	100%
	Центр образования		33,3%	22,2%	44,4%	100%
Всего		32,2%	13,4%	50,4%	3,9%	100%

		ЕГЭ русский язык			ОГЭ русский язык			ЕГЭ математика			ОГЭ математика		
		Среднее	Медиана	Стандартное отклонение	Среднее	Медиана	Стандартное отклонение	Среднее	Медиана	Стандартное отклонение	Среднее	Медиана	Стандартное отклонение
Кластеры	1	79,5	79,1	4,3	33,3	33,3	1,5	68,0	66,9	5,7	19,2	18,8	2,5
	2	64,8	65,4	3,4	28,8	28,7	1,6	43,9	43,8	4,8	14,0	14,0	1,5
	3	72,5	72,4	4,4	31,2	31,3	1,6	55,8	56,2	5,1	16,3	16,2	1,8
	4	58,5	59,1	2,6	26,5	26,3	3,0	35,5	35,4	4,1	12,4	12,7	1,9

# Использования элементов кластерного анализа в региональной системе оценки качества образования: профили школ конкретного района города

Четыре профиля образовательных учреждений

Насколько соотносятся результаты школ с реализуемыми профилями?

Каким образом будут кластеризоваться школы внутри каждого профиля?

Используются данные по ГИА и ВСОШ разного уровня



# Результаты кластерного анализа и направления профилей школ

Профиль	Количество ОУ	
	Кластер 1	Кластер 2
Гуманиарный профиль	10	17
Технологический профиль	23	5
Естественно-научный профиль	25	1
Социально-экономический профиль	11	9

# Ограничения кластерного анализа



Невозможно использовать для районов с небольшим количеством образовательных учреждений

Сложность в отказе от привычных категорий для группировки объектов

Необходимость тщательной первичной обработки данных

Требование в специальном программном ограничении

# Кластерный анализ может применяться на разных уровнях

## Макро уровень

- Группировка регионов, стран
- **Пример №1:** группировка субъектов федерации по данным рейтинга «Уровень развития дошкольного образования» (*Ковалева, Кузнецова, Нечаева, 2008*)
- **Пример №2:** группировка стран по уровню профессионального стресса и удовлетворенности работой сотрудников ВУЗов (*Давыдова, Козьмина, 2014*)

## Мезо уровень

- Группировка городов, районов, школ
- **Пример №1:** Группировка мест по расстоянию от них до дома – радиусов городской и районной мобильности детей, по ответам 819 школьников (*Сивак, Глазков, 2017*)
- **Пример №2:** Типология школ по показателям результата, входа и процесса (*Константиновский, Вахштайн, Куракин, 2013*)

## Микро уровень

- Группировка учителей, учеников, школ, параллелей внутри 1 школы
- Группировка заданий, задач, контрольных
- **Пример:** 83 текстовых задания по математике были сгруппированы в 3 типа, отличающиеся друг от друга по ситуационной значимости, математическому моделированию и новизне формулировки (*Ларина, 2016*)

# Примеры группировок на микро-уровне

## Учителя

**1623 сотрудника 25 российских ВУЗов** (*Давыдова, Козьмина, 2014*) и сотрудники ВУЗов из **20 стран мира**

**4 кластера стран** в зависимости от уровня профессионального стресса и удовлетворенности работой сотрудников

## Ученики

**7778 первоклассников**, поступивших в 2015 году в школы 4 российских городов: Москвы, Набережных Челнов, Севастополя и Тамбова (*Карданова, Иванова, Сергоманов, Канонир, Антипкина, Кайки, 2018*)

**4 кластера** первоклассников с типичными паттернами развития на момент поступления в школу

**Учащиеся 9-10 классов** средних школ г. Пскова (*Ковалевская, 2014*)

**3 кластера** старшеклассников, обладающих различными личностными особенностями

## Выпускники, студенты

**Студенты 2 ведущих факультетов** НГТУ 2013 года набора (*Тимофеев, Борисова, Аврунев, 2014*)

**6 кластеров** студентов с разной готовностью состояться в профессии (мотивация, успеваемость и т.д.)

**Студенты 1 курса ВУЗов Москвы и Санкт-Петербурга, 1922 человека** (*Попов, Тюменева, Кузьмина, 2013*)

**7 кластеров** по уверенности в будущей карьере и другим характеристикам

## Семьи/домохозяйства

**9000 российских семей**, общенациональная территориальная случайная выборка, данные 2006 года (*Петренко, Галицкая, 2007*)

**5 кластеров** домохозяйств – семей с детьми в возрасте от 4 до 22 лет, сходных по социальному капиталу и установкам на те или иные образовательные стратегии детей

# Группировать МОЖНО ПО разным данным

---

ЕГЭ, средний балл в аттестате, средний балл в 1 семестре обучения (*Тимофеев, Борисова, Аврунев, 2014*)

---

Содержание и характеристика задач по математике (*Ларина, 2016*)

---

Балл ЕГЭ, академическая активность (*Попов, Тюменева, Кузьмина, 2013*)

---

**iPIPS** – данные о когнитивных, социальных и эмоциональных навыках ребёнка (*Карданова, Иванова, Сергоманов, Канонир, Антипкина, Кайкы, 2018*)

---

"**Мониторинг экономики образования**", опрос родителей по индикаторам размера экономического, социального и культурного капитала (*Петренко, Галицкая, 2007*)

---

Интегральный рейтинг "**Состояние дошкольного образования**" 2004-2007 (*Ковалева, Кузнецова, Нечаева, 2008*)

---

**CAP - Changing academic profession** - данные о профессиональном стрессе и удовлетворенности работой сотрудников ВУЗов (*Давыдова, Козьмина, 2014*)

---

Образовательные ресурсы по **PIRLS** (*Попов, Тюменева, Кузьмина, 2013*)





Александр Пронин  
socrpron@gmail.com

Ирина Богданова  
irina.al.bogdanova@gmail.com

ГБУ ДПО «Санкт-Петербургский центр оценки качества образования и информационных технологий»

<https://rcokoit.ru>

<https://monitoring.spbcokoit.ru>