



Санкт-Петербургский
политехнический университет
Петра Великого

Сергей Нестеров, Олеся Хворова

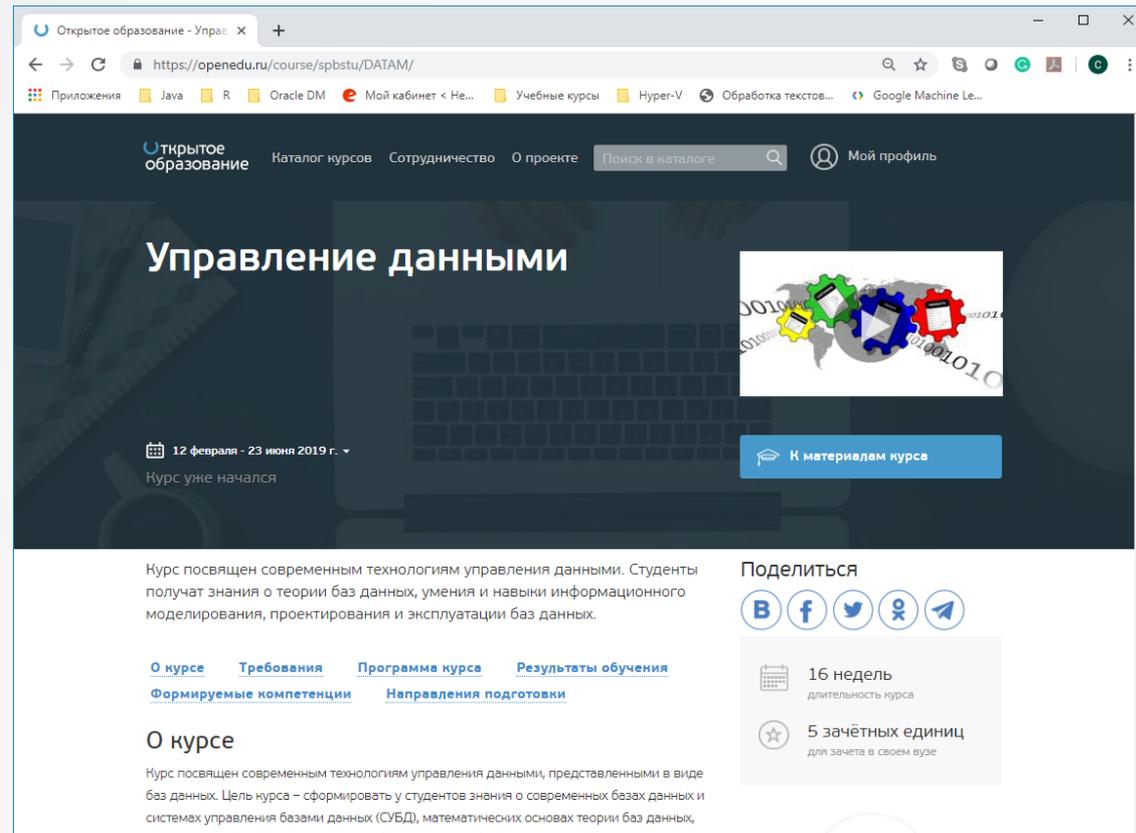
Кластерный анализ результатов изучения массового открытого онлайн-курса

<https://openedu.ru/course/spbstu/DATAM/>

Авторы курса – преподаватели ИКНТ СПбПУ
Нестеров С.А. и Андреева Н.В.

Курс был впервые запущен осенью 2016 г. и три года проводился в семестровом формате с началом в сентябре и феврале.

В сентябре 2019 г. был запуск на весь учебный год.

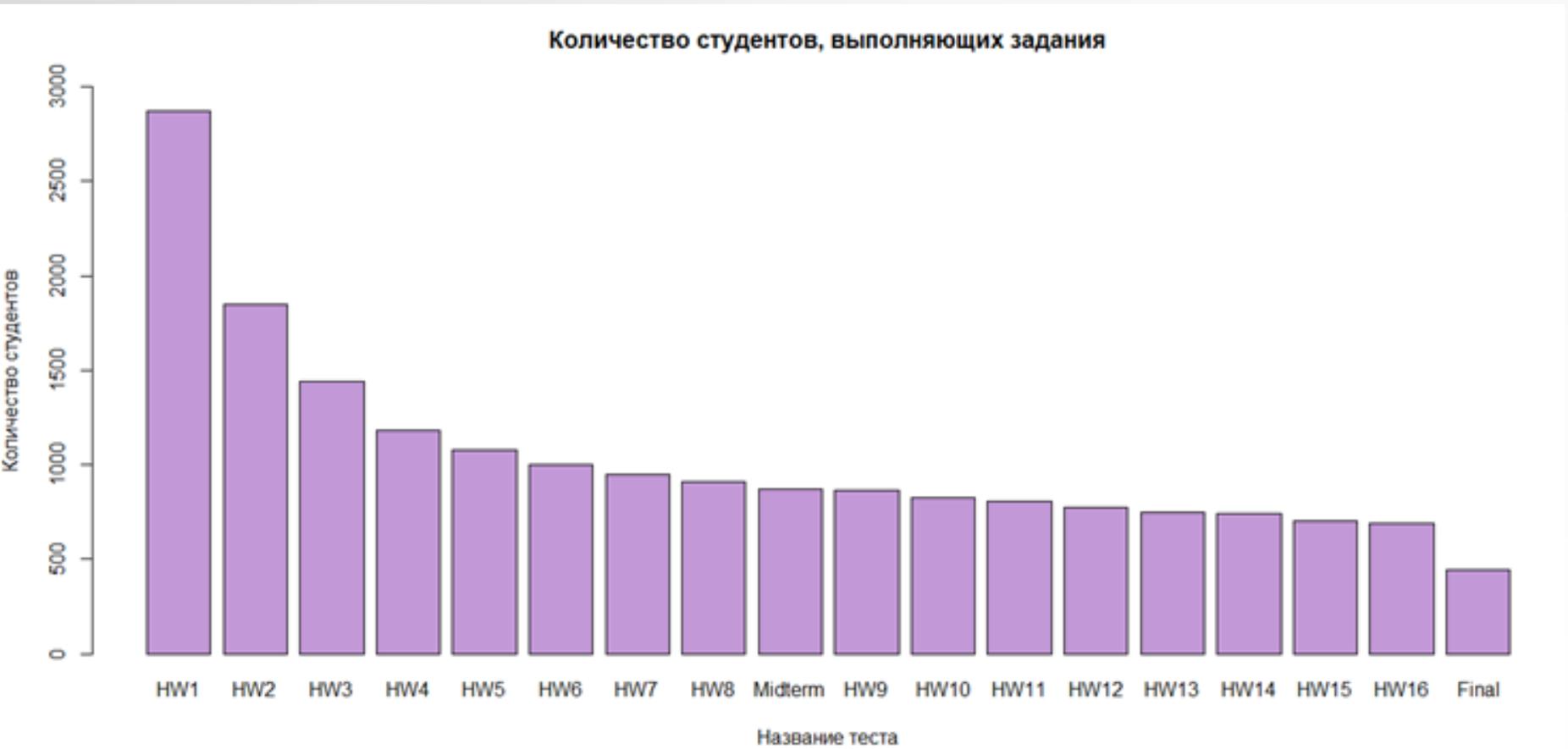


The screenshot shows a web browser window displaying the course page for "Управление данными" (Data Management) on the Open Education Platform (openedu.ru). The page features a dark header with navigation links, a search bar, and a user profile icon. The main content area includes the course title, a date range from February 12 to June 23, 2019, and a blue button labeled "К материалам курса". Below this, there is a paragraph describing the course's focus on modern data management technologies, followed by several links for course details, requirements, and competencies. On the right side, there is a "Поделиться" (Share) section with social media icons and a box indicating a 16-week duration and 5 credit units.

1. Введение. Обобщенная архитектура систем баз данных
2. Этапы проектирования баз данных, понятие модели данных, обзор основных моделей данных
3. Реляционная модель данных: допустимые структуры и ограничения
4. Реляционная алгебра: операции и примеры
5. Нормализация. Нормальные формы 1-3
6. НФБК и старшие нормальные формы
7. Модель сущность-связь, ER-диаграммы Чена, Мартина и Баркера, IDEF1x
8. IDEF1x (продолжение), IE, использование CASE-средств, переход от логической модели к физической
9. История SQL, подязыки (DDL, DML ...), типы данных, функции, создание домена/пользовательского типа, создание таблицы, определение ограничений
10. SQL: добавление, обновление, удаление, записей (INSERT, DELETE, UPDATE, MERGE)
11. SELECT: составление простых запросов, выборка данных из нескольких таблиц
12. Подзапросы. Реализация теоретико-множественных операций реляционной алгебры средствами SQL
13. Представления
14. Транзакции
15. Переменные, операторы, временные таблицы
16. Хранимые процедуры, функции, курсоры, триггеры

Семестр	Количество записавшихся на курс	Количество слушателей, выполняющих задания	Количество слушателей, не приступивших ни к одному заданию	Количество получивших сертификат
Осень 2016	2 690	900	1 790	97
Весна 2017	1 618	539	1 079	84
Осень 2017	1 823	440	1 383	79
Весна 2018	1 504	290	1 214	18
Осень 2018	2 346	540	1 806	63
Весна 2019	1 572	270	1 302	23

Суммарные данные о числе студентов, выполнявших задания по неделям





В 2019-2020 учебном году мы изменили формат запуска курса, сделав его доступны весь учебный год. Записаться на курс можно было с сентября по середину апреля, все сроки сдачи тестов были установлены на вторую половину июня.

Количество записавшихся на курс	Количество слушателей, выполняющих задания	Количество слушателей, не приступивших ни к одному заданию	Количество получивших сертификат
5 937	1 220	4 717	85

- На курс записывается большое количество слушателей, но заканчивает курс лишь небольшой процент записавшихся
- С теми кто записался и не начал ничего сдавать, будем считать, что понятно – случайные посетители
- Обратной связи от слушателей явно не хватает, чтобы понять почему бросают в середине курса

- Платформа orepedu.ru представляет преподавателю отчеты об оценках за тесты (у нас 16 еженедельных, промежуточный и финальный экзамен)
- Отчеты выгружались с orepedu.ru, из них удалялись персональные данные, дальше мы их анализировали, используя язык R
- Проводили кластеризацию, используя алгоритм k-средних. Число кластеров определяли используя эвристические методы (метод локтя).
- Для каждого кластера строились столбчатые диаграммы, аналогичные показанной выше

Для семестровых сессий можно говорить о наличии следующих характерных групп слушателей:

- Слушатели, в большинстве своем, не приступившие к выполнению заданий;
- Слушатели, приступившие к выполнению, но бросившие изучение курса практически сразу (1-2 недели);
- Слушатели, забросившие курс примерно во второй его половине;
- Слушатели, стабильно выполнявшие все задания.

Для сессии на весь учебный год можно говорить о наличии следующих характерных групп слушателей:

- слушатели, которые записались на курс и, в большинстве своем, не приступили к выполнению заданий;
- слушатели, которые бросили курс после первых недель;
- слушатели, стабильно выполнявшие все задания на курсе.

Более продолжительный период записи на сессия 2019-2020 уч. года привел к увеличению числа слушателей (более чем в 2 раза, по сравнению с лучшими из семестровых сессий).

Можно предположить, что отсутствие четвертой группы для сессии 2019-2020 (бросившие учиться в середине курса), связано с тем, что более длинная сессия и сдвинутые на ее конец дедлайны дали возможность закончить курс тем, в принципе кто хотел доучиться.



Спасибо за внимание!

Сергей Нестеров
к.т.н., доцент высшей школы
Киберфизических систем и управления
СПбПУ