

Департамент образования и науки города Москвы Государственное  
автономное образовательное учреждение высшего образования города  
Москвы «Московский городской педагогический университет» Институт  
цифрового образования Департамент информатики, управления и технологий

ДИСЦИПЛИНА:

Инструменты для хранения и обработки больших данных

Практическая работа №1

Тема:

Визуализация данных из CSV-файла

Выполнила: Шепелева Е. В., группа: АДЭУ-201

Преподаватель: Т. М. Босенко

Москва

2022

## Вариант №3

Homework (<https://www.kaggle.com/datasets/yonkotoshiro/homework>)

### Этап №1. Создание подключения

Users / liza080302 / Homework ☆ ...

Создать датасет Сохранить изменения

File

+ Загрузить файлы ?

Homework\_eng.csv ...

field1	month	year	russian language 5 grade	english language 5 grade	mathematics 5 grade	computer science 5 grade	history 5 grade	geography 5 grade	biology 5 grade
0	2007-01-01	2007	0	0	0	0	0	0	0
1	2007-02-01	2007	0	0	0	0	0	0	0
2	2007-03-01	2007	0	0	0	0	0	0	0
3	2007-04-01	2007	0	0	0	0	0	0	0
4	2007-05-01	2007	0	0	4	0	0	0	0
5	2007-06-01	2007	0	0	0	0	0	0	0
6	2007-07-01	2007	0	0	0	0	0	0	0
7	2007-08-01	2007	0	0	0	0	0	0	0
8	2007-09-01	2007	0	0	0	0	0	0	0
9	2007-10-01	2007	0	0	0	0	0	0	0
10	2007-11-01	2007	0	0	0	0	0	0	0
11	2007-12-01	2007	0	0	0	0	0	0	0
12	2008-01-01	2008	0	0	0	0	0	0	0
13	2008-02-01	2008	3	0	0	0	0	0	0
14	2008-03-01	2008	0	0	0	0	0	0	0
15	2008-04-01	2008	0	0	0	0	0	0	0

### Этап №2. Создание датасета

Создание поля «месяц»

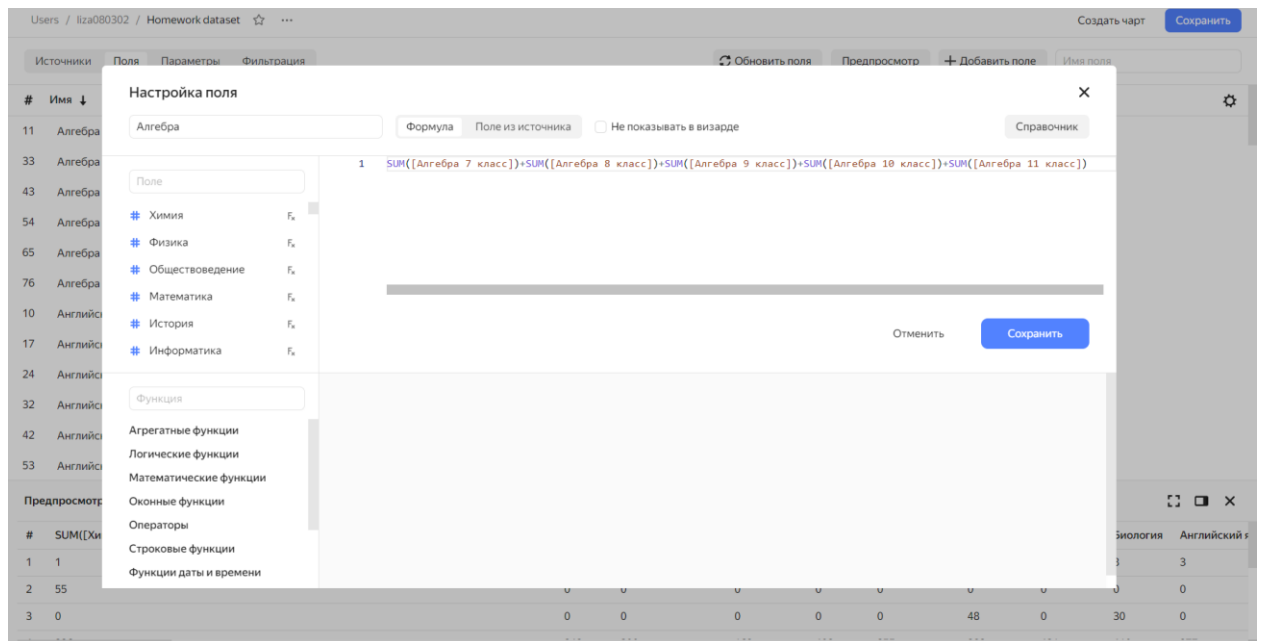
1 MONTH([month])

#	Имя ↓	Источник поля ↓	Тип ↓	Агрегация ↓	Описание ↓
1	month	csv.month	Дата	Нет	
2	Месяц	$F_x$	Целое число	Нет	
3	Год	csv.year	Целое число	Нет	

Предпросмотр    Количество строк: 10    не больше 1 000

#	month	Месяц	Год	Русский язык 5 класс	Английский язык 5 класс	Математика 5 класс	Информатика 5 класс	История 5 класс	География 5 класс	Биология 5 к
1	2007-01-01	1	2007	0	0	0	0	0	0	0
2	2007-02-01	2	2007	0	0	0	0	0	0	0
3	2007-03-01	3	2007	0	0	0	0	0	0	0
4	2007-04-01	4	2007	0	0	0	0	0	0	0
5	2007-05-01	5	2007	0	0	4	0	0	0	0

Для создания чартов необходимо создать поля с общим кол-вом запросов по каждому предмету и классу:



Users / liza080302 / Homework dataset

Создать чарт Сохранить

Источники Поля Параметры Фильтрация Обновить поля Предпросмотр + Добавить поле Имя поля

#	Имя	Источник поля	Тип	Агрегация	Описание
1	11 класс	F <sub>x</sub>	Целое число	Авто	
2	10 класс	F <sub>x</sub>	Целое число	Авто	
3	9 класс	F <sub>x</sub>	Целое число	Авто	
4	8 класс	F <sub>x</sub>	Целое число	Авто	
5	7 класс	F <sub>x</sub>	Целое число	Авто	
6	6 класс	F <sub>x</sub>	Целое число	Авто	
7	5 класс	F <sub>x</sub>	Целое число	Авто	
8	Химия	F <sub>x</sub>	Целое число	Авто	
9	Физика	F <sub>x</sub>	Целое число	Авто	
10	Обществоведение	F <sub>x</sub>	Целое число	Авто	

Users / liza080302 / Новый датасет

Создать чарт Сохранить

Источники Поля Параметры Фильтрация Обновить поля Предпросмотр + Добавить поле Имя поля

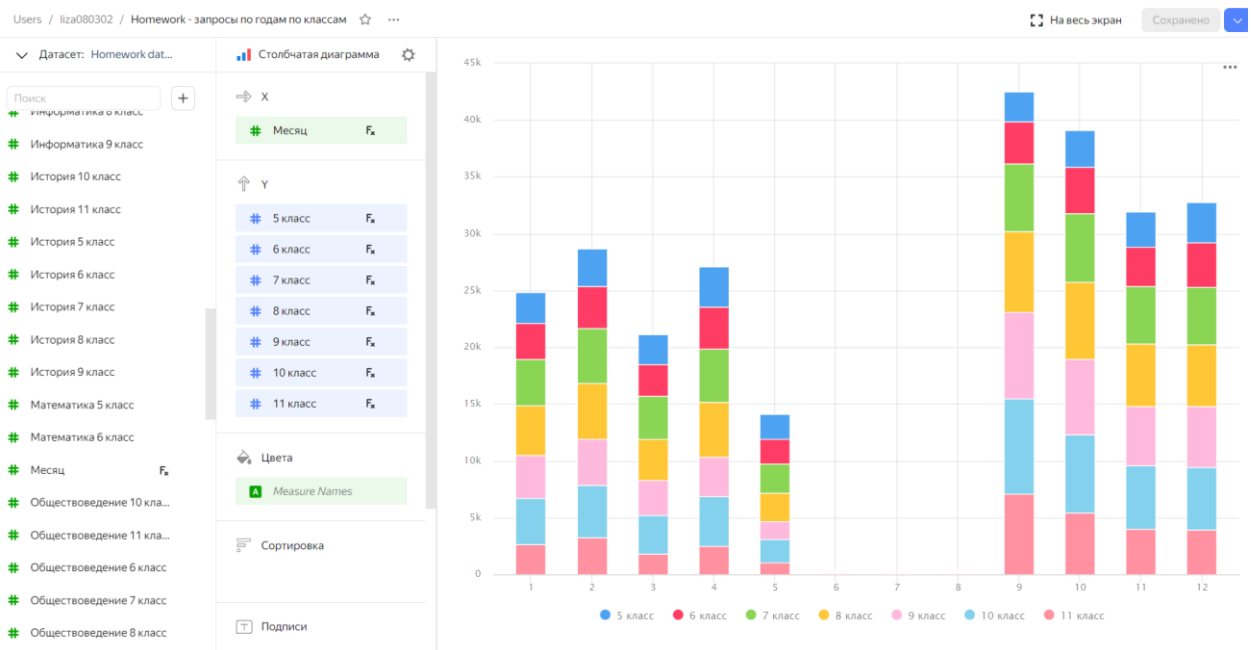
#	Имя	Источник поля	Тип	Агрегация	Описание
1	Месяц	csv.month	Дата	Нет	
2	Год	csv.year	Целое число	Нет	
3	Русский язык 5 класс	csv.russian_language_5_grade	Целое число	Нет	
4	Английский язык 5 класс	csv.english_language_5_grade	Целое число	Нет	
5	Математика 5 класс	csv.mathematics_5_grade	Целое число	Нет	
6	Информатика 5 класс	csv.computer_science_5_grade	Целое число	Нет	
7	История 5 класс	csv.history_5_grade	Целое число	Нет	
8	География 5 класс	csv.geography_5_grade	Целое число	Нет	
9	Биология 5 класс	csv.biology_5_grade	Целое число	Нет	
10	Русский язык 6 класс	csv.russian_language_6_grade	Целое число	Нет	

Предпросмотр Количество строк: 15 не больше 1 000

#	month	year	Русский язык 5 класс	english language 5 класс	mathematics 5 класс	computer science 5 класс	history 5 класс	geography 5 grade	biology 5 grade	russian language 6 grade	english language 6 grade
1	2007-01-01	2007	0	0	0	0	0	0	0	0	0
2	2007-02-01	2007	0	0	0	0	0	0	0	0	0
3	2007-03-01	2007	0	0	0	0	0	0	0	5	0
4	2007-04-01	2007	0	0	0	0	0	0	0	0	0
5	2007-05-01	2007	0	0	4	0	0	0	0	0	0

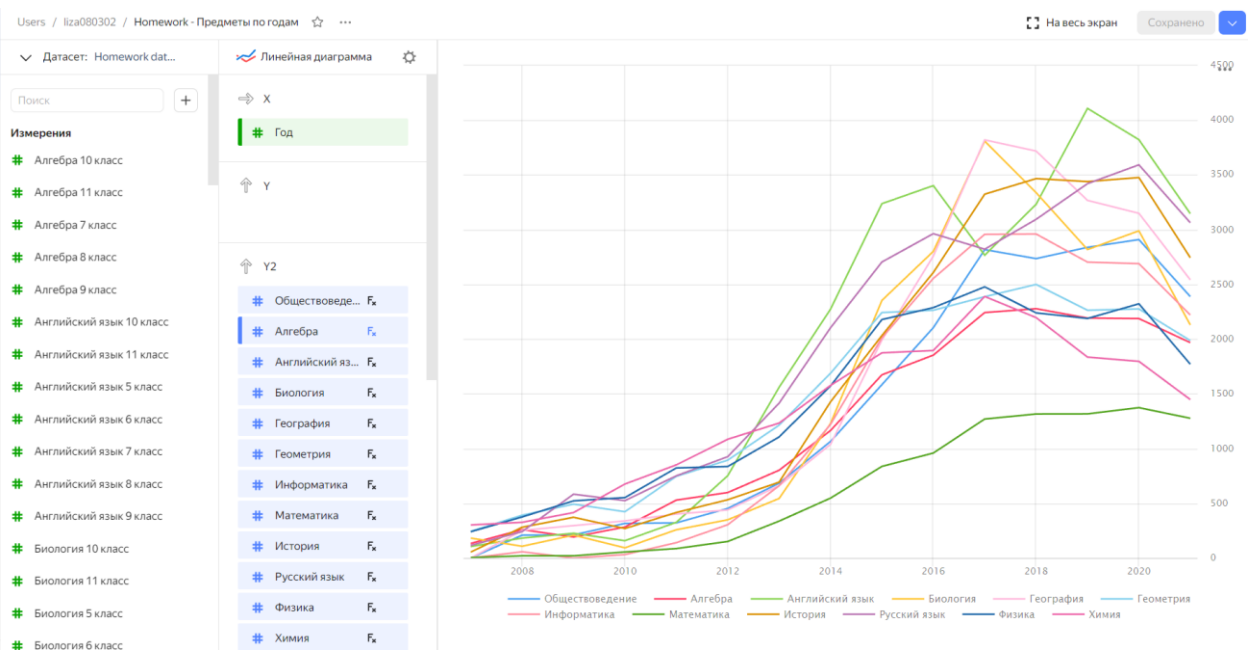
## Этап №3. Создание первого чарта

Пусть первый чарт отображает кол-во запросов по классам, разделенным по месяцам.



## Этап №4. Создание второго чарта

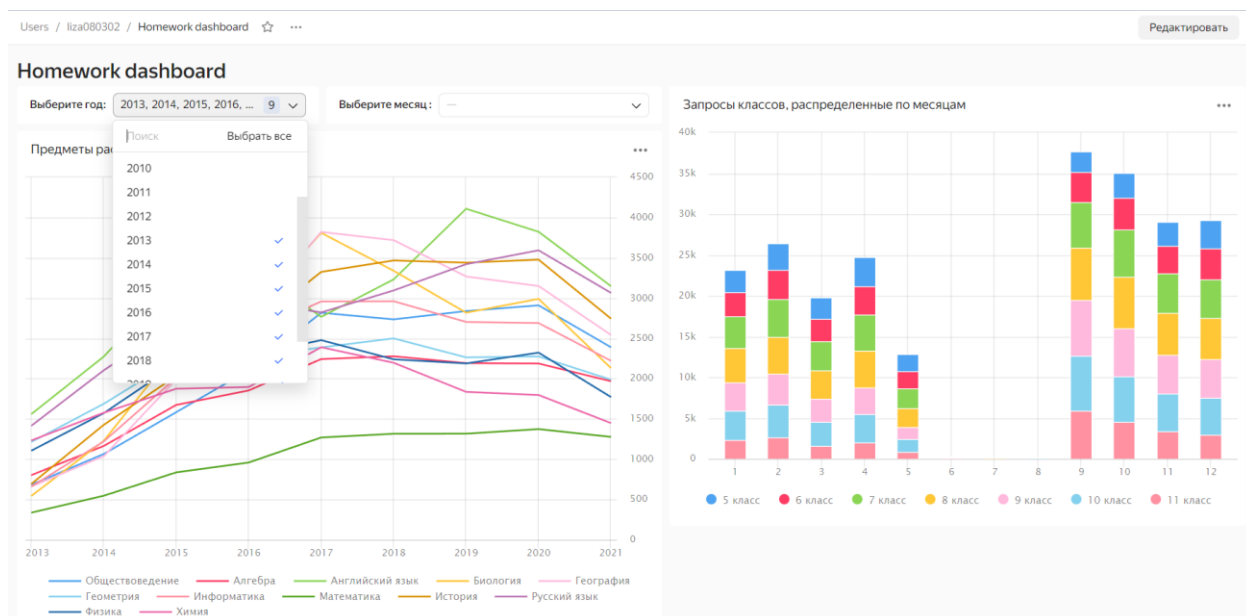
Второй чарт будет отображать кол-во запросов по каждому предмету, разделенных по годам



## Этап №5. Создание дашборда

Создам дашборд под названием «Homework dashboard», расположу на нем два чарта, созданных на этапах выше и несколько селекторов.

Первый селектор будет позволять делать выборку по годам, второй по месяцам:

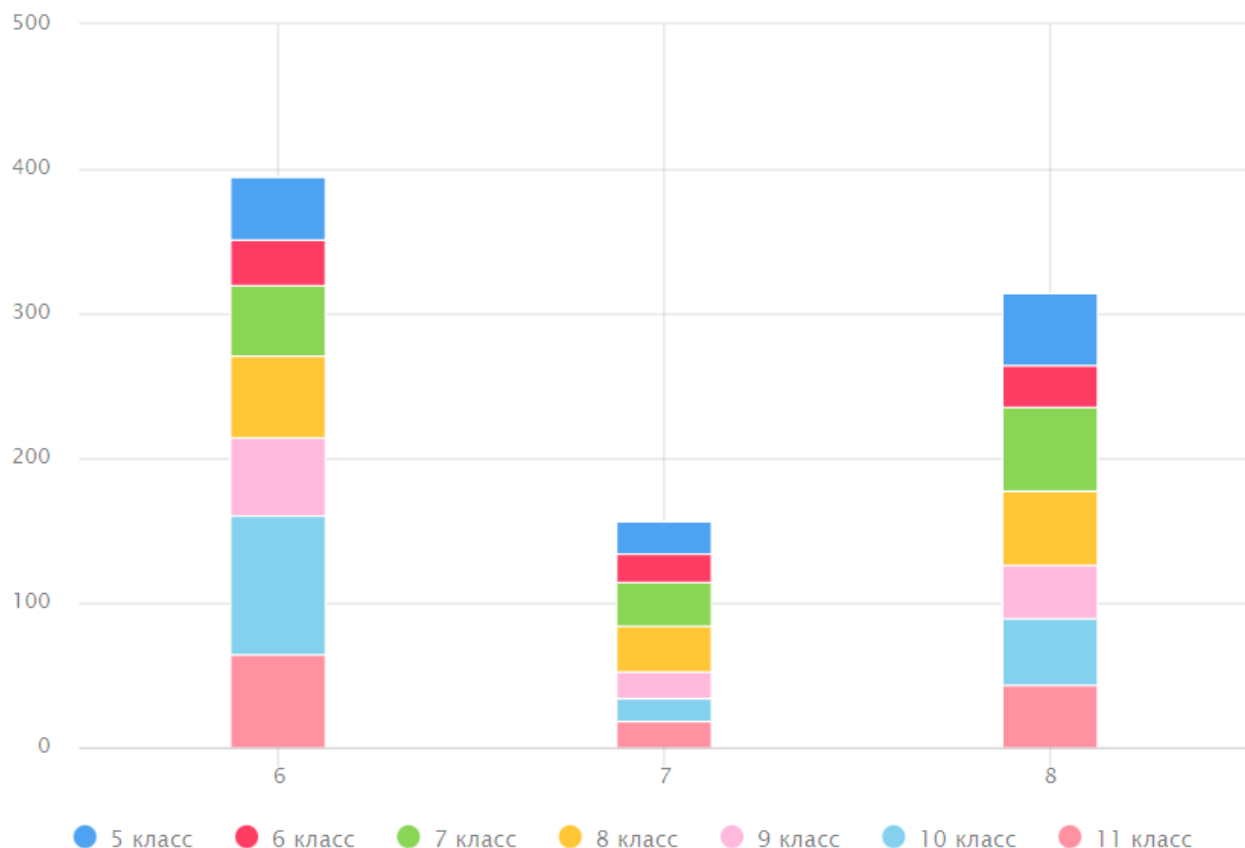


## Этап №6. Выводы

Подведу несколько выводов, которые можно сделать по диаграммам:

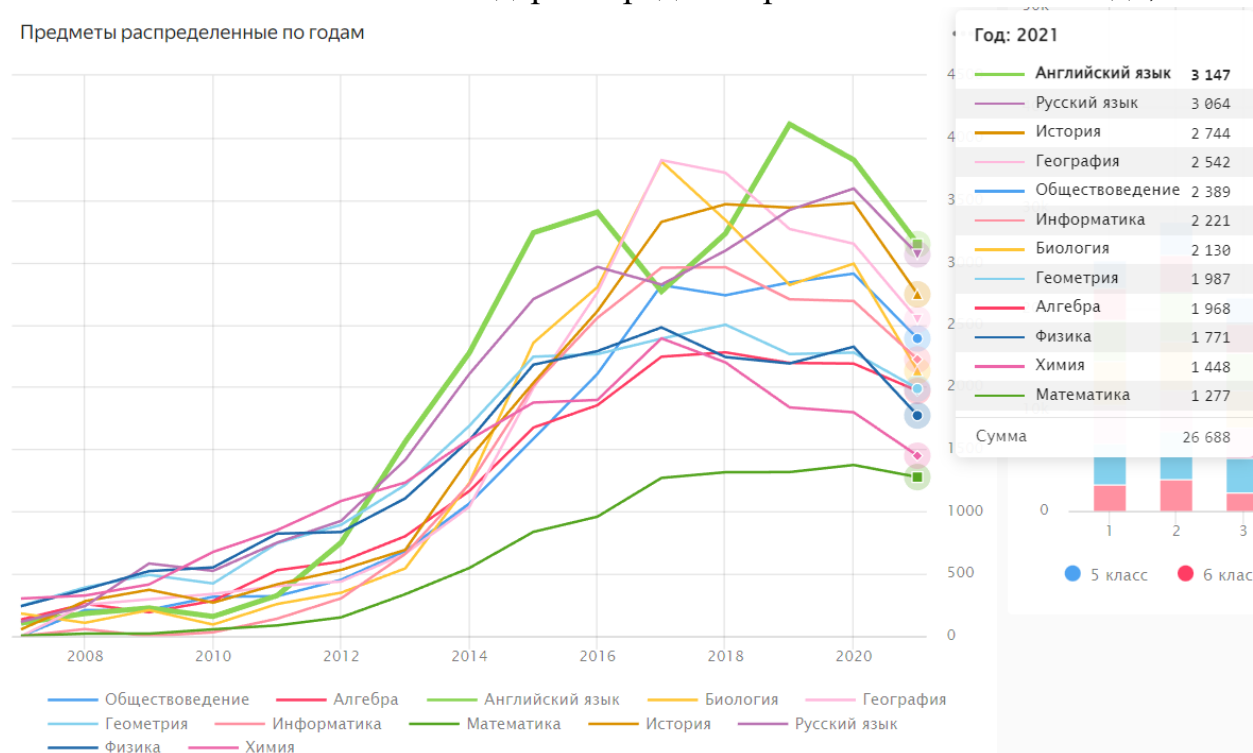
- Самое наименьшее количество запросов поступает в следующие месяцы: июнь, июль и август;

Запросы классов, распределенные по месяцам



- Английский язык является лидером среди запросов с 2019-2021 года;

Предметы распределенные по годам



- Наибольшее кол-во запросов с 2007-2021 год поступило в 2019 году;

- Три лидирующих класса по кол-ву запросов (от наибольшего к наименьшему): 10, 8 и 7.

### **Ссылка на дашборд**

Ссылка на дашборд: <https://datalens.yandex/ff0fmfrorki5>