

# Основы программирования в R

**Преподаватель:** Алла Тамбовцева (allatambov@gmail.com)

3 ноября 2017

# Формат занятий

**Лекции:** демонстрация решения различных задач в R + теоретические аспекты работы с данными

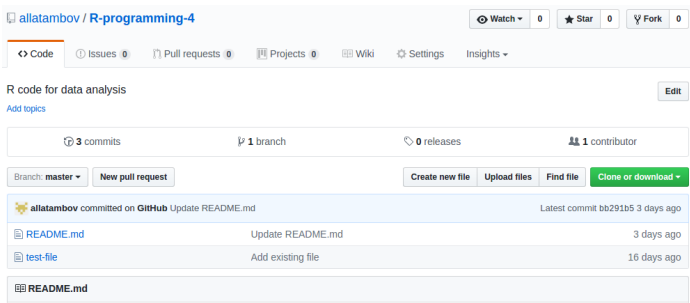
**Семинары:** практикум по программированию в R

Материалы занятий выкладываются на GitHub:  
<https://github.com/allatambov/R-programming-3>

# Формат занятий

## Как скачать файлы с GitHub?

- 1 Зайти на страницу репозитория (см.предыдущий слайд).
- 2 Нажать зеленую кнопку Clone or download.



- 3 Выбрать Download ZIP, скачать zip-архив и распаковать.

# Оценивание

## Формула оценки

$$O_{\text{результатирующая}} = 0.6 \cdot O_{\text{накопленная}} + 0.4 \cdot O_{\text{экзамен}}$$

$$O_{\text{накопленная}} = 0.4 \cdot O_{\text{ДЗ}} + 0.2 \cdot O_{\text{аудиторная}} + 0.4 \cdot O_{\text{см.р.}}$$

# Оценивание

Что есть что?

- **ДЗ:** небольшое исследование, которое предполагает несколько этапов: сбор данных, описание полученной базы данных, первичная обработка данных, визуализация и анализ данных.
- **Аудиторная работа:** практические задания (в R) и проверочные работы (без R)
- **Самостоятельная работа:** регулярные домашние задания
- **Экзамен:** набор задач по пройденным темам, которые выполняются на компьютере, в R (RStudio).

# Оценивание

## Правила игры

- Задачи, для решения которых предоставлен неработающий код (код, который невозможно запустить из-за наличия грубых ошибок/опечаток), не засчитываются, даже если при этом зафиксирован верный результат.
- Если при проверке работ установлен факт нарушения академической этики, студент получает оценку «0» за данную работу. Работа студента, предоставившего свою работу для списывания, также аннулируется.
- Домашние задания, сданные после срока, оцениваются так: опоздание в пределах **часа** – штраф **10%** от полученной оценки, в пределах **суток** – штраф **20%**, в пределах **недели** – штраф **50%**. Домашние задания, сданные через неделю после указанного срока и позже, не принимаются.

# Оценивание

## Правила игры

Домашние задания состоят из двух блоков: базовый и продвинутый. Студенту на выбор предлагается решить задачи одного из блоков. Если студент хочет решить задачи продвинутого блока, но не уверен, что сделает их полностью верно или в требуемом объеме, возможен следующий вариант.

Студент решает 50% задач из базового блока, решает задачи продвинутого блока, и тогда его оценка за домашнюю работу считается как округленное среднее арифметическое оценок, полученных за базовый блок и продвинутый блок.

# Что нас ждет?

## Знакомство с элементами reproducible research

Rmarkdown и RPubs – как показать свой код миру и сделать это красиво.

**RPubs** brought to you by RStudio

### Занятие 1. Введение в R: переменные

Алла Тамбовцева

- Переменные
  - Присваивание значений
  - Типы переменных
  - Преобразование типов
- Логические выражения
- Калькулятор
- Числовые переменные
- Текстовые переменные
- Установка пакетов

#### Переменные

Названия переменных в R могут содержать буквы, цифры, точки и знаки подчеркивания, при этом название переменной не может начинаться с цифры. Название переменной не должно совпадать со служебными словами в R (операторами): `if`, `else`, `for`, `while` и другим.

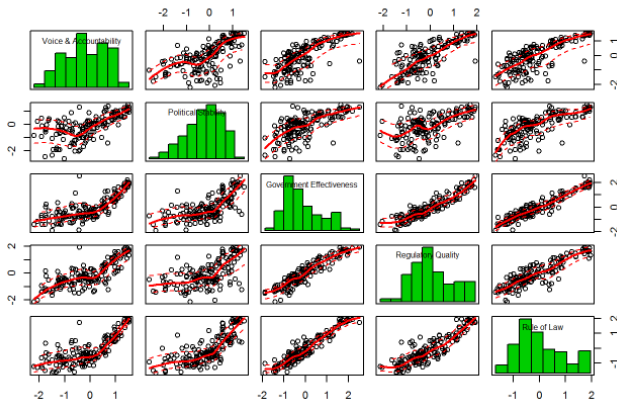
Вообще R позволяет создавать переменные, названные кириллицей, но лучше этого никогда не делать и использовать латинские буквы.



Что нас ждет?

Много красивых графиков

Correlations of Worldwide Governance Indicators



# Что нас ждет?

## Web scraping

Было:

```
</table>
</td><td width="90%"><div style="width:100%; bgcolor:white;
style="width:100%;overflow:scroll" vspace="0" align="left">
<tr bgcolor="#FFFFFF" valign="top">
<td style="color:black" align="center"><nobr><a style="text
action=show&vtd=100100067795912&vrn=100100067795849
233">Алтайская</a></nobr></td><td style="color:black" align=
action=show&vtd=100100067795912&vrn=100100067795849
233">Барнаул, Железнодорожная</a></nobr></td><td style="col
href="http://www.vybory.izbirkom.ru/region/region/izbirkom?
action=show&vtd=100100067795912&vrn=100100067795849
233">Барнаул, Центральная</a></nobr></td><td style="color:b
href="http://www.vybory.izbirkom.ru/region/region/izbirkom?
action=show&vtd=100100067795912&vrn=100100067795849
233">Белокурихинская городская</a></nobr></td><td style="co
href="http://www.vybory.izbirkom.ru/region/region/izbirkom?
action=show&vtd=100100067795912&vrn=100100067795849
233">Бийская</a></nobr></td><td style="color:black" align="
action=show&vtd=100100067795912&vrn=100100067795849
233">Быстроистокская</a></nobr></td><td style="color:black"
href="http://www.vybory.izbirkom.ru/region/region/izbirkom?
```

# Что нас ждет?

Web scraping

Стало:

	1	2	3	4	5	6	7
<b>Алтайская</b>	20240	19600	0	8009	917	10674	917
<b>Барнаул, Железнодорожная</b>	79031	59300	0	30347	1009	27944	1009
<b>Барнаул, Центральная</b>	103384	101898	174	39431	2236	60057	2409
<b>Белокурихинская городская</b>	12690	10550	0	6329	89	4132	89
<b>Бийская</b>	28130	28161	0	9969	522	17670	522
<b>Быстроистокская</b>	7982	8200	0	3066	402	4732	402
<b>Змеиногорская</b>	17076	13630	1	6506	327	6796	328
<b>Калманская</b>	11750	11720	0	3697	258	7765	258

# Что нас ждет?

## Работа с API

### Информация

Название:

Описание:

Категория:

Прочее



Тип турнирной таблицы:

Не поддерживается



Группа приложения:

Групп не найдено



Иконка 16×16:



Выбрать файл

Сохранить изменения

# Что нас ждёт?

## Работа с API

	id	first_name	last_name	sex	home_town	num_friends	country
1	12374247	Александра	Елизарова	0	Нижний Новгород	89	Россия
2	54281447	Злата	Сергеева	0		578	Россия
3	68892630	Иван	Родин	1	Саратов	543	Россия
4	50467050	Kirill	Chmel	1	Самара	407	Россия
5	4161834	Роман	Киселёв-Август	1	Третий Рим	905	Россия
6	171544496	Наталья	Василёнок	0	Мурманск	285	Россия
7	8761370	Дмитрий	Турчанинов	1	Калининград	2016	Россия
8	3990163	Андрей	Савельев	1	Москва	247	Россия
9	16260739	Дмитрий	Ефимов	1	Курган	797	Россия
10	20473269	Алла	Тамбовцева	0	Москва	201	Россия

# Что нам потребуется

- R
- RStudio

# В поисках помощи

Где искать (ну, кроме Google)

- <https://stackoverflow.com/>
- <http://www.statmethods.net/>
- <https://www.r-bloggers.com/>
- учебники из программы курса