



ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ  
УПРАВЛЕНИЯ

Основан в 1919 году

# Обработка данных в среде офисных приложений: введение в R

Светлана Андреевна Суязова (Аксюк)

[sa\\_aksyuk@guu.ru](mailto:sa_aksyuk@guu.ru)

4 сентября 2021

# Лабораторная 1

- "Загрузка данных из различных источников"
- Методические указания: код + текст, pdf
- Основная часть: установка R и RStudio, установка пакетов (библиотек) в R, импорт данных из .csv, загрузка данных с помощью API из баз данных международной статистики (ООН, Всемирный банк)
- Дополнительная часть (опционально): импорт данных из xml, html, веб-скраппинг сайтов средней сложности на примере imdb.com

# Установка R и RStudio

- R (свободная лицензия): <https://cran.r-project.org/bin/windows/base/>
- RStudio Desktop (свободная лицензия): <https://www.rstudio.com/products/rstudio/>

Важно: пути к директориям установки R и Rstudio под Windows не должны содержать русских букв, и не должны быть очень длинными.

# Рабочая директория

- `getwd()` – узнать, какая директория является рабочей в данный момент;
- `setwd('G:\\R-Work\\masters_lab-01')` – сделать рабочей директорию `G:/R-Work/masters_lab-01`;
- Меню RStudio: Session >> Set Working Directory >> To Source File Location – установить рабочей ту директорию, в которой лежит скрипт.

# Установка и подключение пакетов

```
# установить пакет (один раз)  
install.packages('WDI')  
  
# подключить пакет к текущей сессии R  
library('WDI')
```

После успешного запуска `library()` можно пользоваться функциями и объектами пакета.

# Создание директорий

- `dir.create(имя-папки)` - создать директорию внутри рабочей;
- `file.exists(имя-папки-или-файла)` - проверить, есть ли внутри рабочей директории папка или файл (возвращает `TRUE` или `FALSE`)

```
# проверим, нет ли внутри рабочей директории папки data
if (!file.exists('data')) {
  # если нет, создать такую папку
  dir.create('data')
}
```

# Важные горячие клавиши

`Alt + -` - оператор присваивания `<-`

`Ctrl + Enter` - запустить код: строку, в которой стоит курсор, либо выделенный участок кода

`Tab` - посмотреть подсказку с именем функции / объекта рабочего пространства / пути к файлу

# Полезные функции для просмотра объектов

- `is.data.frame()` - является ли объект фреймом данных (таблицей)
- `str()` - посмотреть структуру объекта
- `dim()` - посмотреть количество измерений объекта
- `length()` - посмотреть длину, т.е. количество элементов объекта (первого измерения, если их несколько)