

Факультативный курс БИВНИ-1 с использованием языка Python

Илья Кочергин

кафедра экономической информатики ЭФ МГУ

19 ноября 2019 г.

Что за БИВНИ?

Б

Базовые

И

Инструменты

В

Воспроизводимых

Н

Научных

И

Исследований

R

Reproducible

R

Research



Зачем нужна воспроизводимость научных исследований

- ▶ Методология **RR (Reproducible Research)** нацелена на максимальное упрощение воспроизведения содержащихся в публикации методик
 - Публикация содержит блоки кода на языках программирования
 - Шаги получения данных, импорта, очистки, преобразования данных в другое представление, обработка математическими методами происходят автоматически
- ▶ Возможность выявить ошибки, воспроизводя этапы работы с данными
 - Результаты исследований ложатся в основу экономических решений и законов (порядок цены $\$10^{10}$)
 - Данные и многоэтапные вычисления, на которые опирается исследование могут содержать ошибки
- ▶ Распространение знаний
 - возможность изучить методику автора научной публикации
 - можно увидеть формулы и алгоритмы работы с данными

Процесс подготовки научной публикации

- ▶ получение данных
- ▶ очистка данных
- ▶ разведочный анализ
- ▶ создание гипотез о закономерностях
- ▶ оценка параметров модели
- ▶ Получение объясняющих графиков, таблиц и т.д.
- ▶ "Вплетение" результатов вычислений и алгоритмов в язык верстки научных публикаций (\LaTeX или Markdown)

Почему язык Python?

- ▶ Универсальный язык программирования
- ▶ Открытый проект — это нужно для воспроизводимости
- ▶ Самые современные выразительные средства
- ▶ Является общепринятым
- ▶ Получает возможности языка R (своего конкурента) при подключении научных библиотек (pandas, numpy, matplotlib, scipy)

Цели курса

- ▶ Дать возможность научиться программировать
- ▶ Курс должен стать базой для курсов
 - ▶ БИВНИ-2-python (2й курс)
 - ▶ По машинному обучению
 - ▶ По анализу данных
 - ▶ Эконометрики
- ▶ Позволит применить полученные навыки для автоматизии ежедневных рутинных задач

Программа курса

Подробнее примерную программу курса можете прочитать [в заметке Evernote](#)

презентация сверстана на языке разметки для научных публикаций L^AT_EX