

Python для анализа данных

МИРЭК | 4 модуль

Лектор

Татьяна Рогович

rogovich@gmail.com

Координация

- Рассылки в ЛМС
- [Wiki страничка курса](#)
- Канал курса + чаты курса
https://t.me/joinchat/WJU_iizMJ7q7UrBi
- Карточка курса:
<https://www.hse.ru/ba/we/courses/339556904.html>

Цели курса

- Уверенно овладеть синтаксисом языка Python и основных библиотек для работы с данными (pandas, matplotlib, numpy, etc.)
- Собирать данные с помощью Python (скрейпинг)
- Собирать и визуализировать данные с помощью инструментов Python
- Уметь с помощью средств Python представлять отчет об исследовании от разведывательного анализа данных и гипотезы до интерпретации выводов.

Темы курса

1. Программирование в Python
2. Сбор данных с Python
3. Работа с данными в Python
(Numpy, pandas)
4. Визуализация данных
(matplotlib, plotly)
5. Разведывательный анализ
данных
6. Регрессии и статистический
анализ в Python

Формула

$$\begin{aligned} & 0.4 * \text{среднее(ДЗ)} \\ & + 0.2 * \text{Контрольная работа} \\ & + 0.1 * \text{среднее(ОПРОСЫ)} \\ & + 0.1 * \text{Оценка за онлайн-курс} \\ & + 0.2 * \text{Письменный экзамен} \end{aligned}$$

1. **Контрольная работа (Питон)**
Контрольная может проводиться онлайн в письменной форме с использованием прокторинга (Zoom)
2. **Мини-опросы на лекциях** — лекции 2-5 (среднее всех)
3. **Домашнее задание** — 3 ДЗ (среднее всех)
4. **Оценка за онлайн-курс**
5. **Письменный экзамен** (анализ данных)
Экзамен проводится на платформе Онлайн-образование в НИУ ВШЭ (<https://online.hse.ru/>), прокторинг на платформе Экзамус (<https://hse.student.examus.net>).

Почему Python?



Язык программирования Python



- Высокоуровневый язык программирования, созданный в 1980-х годах.
- Наиболее актуальные версии - 3.xx (более старые версии могут отличаться в синтаксисе).
- Популярность стремительно растет.
- Гид по стилю: <https://www.python.org/dev/peps/pep-0008/>

Почему Python?



python machine le...

Search term

R machine learning

Search term

Java machine lear...

Search term

c machine learning

Search term

javascript machine...

Search term

Worldwide ▼

Past 5 years ▼

All categories ▼

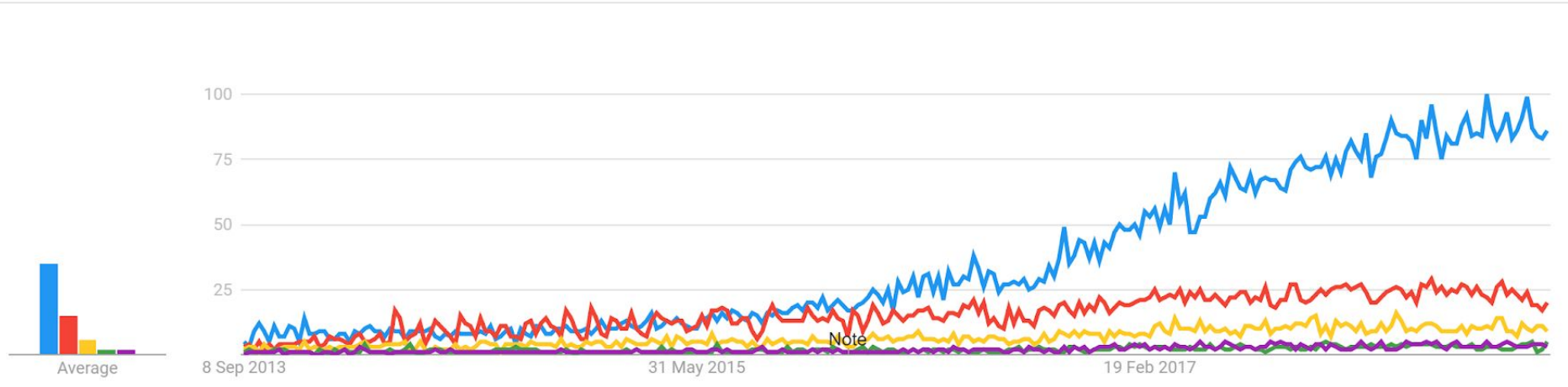
Web Search ▼

Interest over time ?

↓

<>

↻



How many Python?

python data analysis

Search term

r data analysis

Search term

+ Add comparison

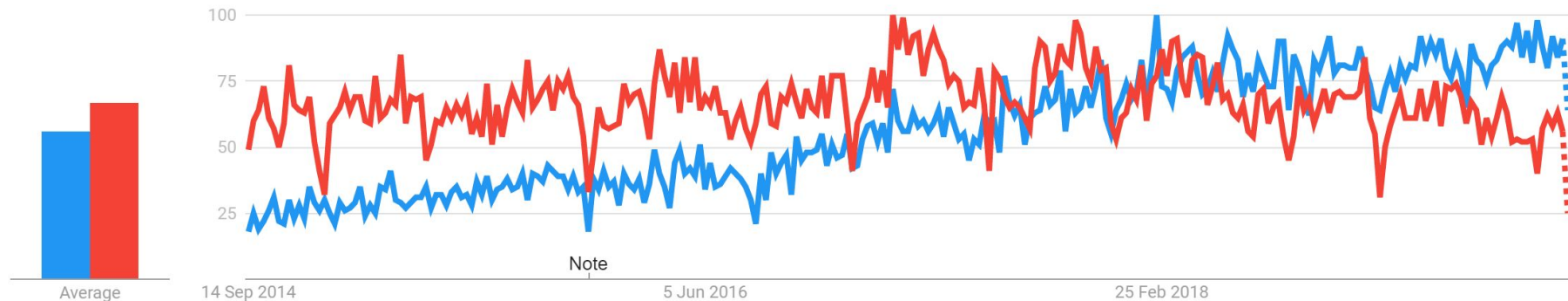
Worldwide ▼

Past 5 years ▼

All categories ▼

Web Search ▼

Interest over time ?



Как Python может пригодиться вам?

- Сбор данных (scraping и API)
- Работа с табличным представлением данных (pandas), возможность выполнения SQL запросов
- Создание визуализаций
- Анализ текста
- Статистический анализ
- Использование нейронных сетей и машинного обучения для анализа больших данных или автоматизации анализа (замена кодеров)
- Оформление готовых отчетов из Jupyter notebook (с публикацией кода или нет)