

# Основы программирования и анализа данных в Python

Лекция 1. Введение.

Максим Карпов  
[mekarpov@hse.ru](mailto:mekarpov@hse.ru)

НИУ ВШЭ, 5 апреля 2022



# Почему Python?



# Почему Python?



## C++

```
#include <iostream>

int main()
{
    std::cout << "Hello, World!";
    return 0;
}
```

CPP

## Java

```
class Greeting {
    public static void main(String[] args) {
        System.out.println("Hello World!");
    }
}
```

JAVA

## Python

```
print('Hello world!')
```

PYTHON

# Google Trends

● Python

Язык программиров...

● C++

Язык программиров...

● Java

Язык программиров...

● Go

Язык программиров...

● C#

Поисковый запрос

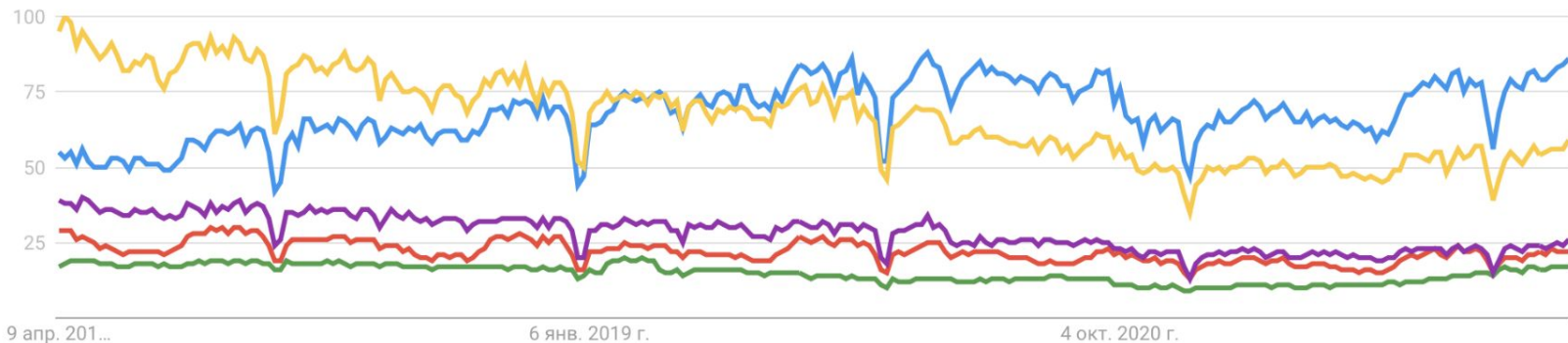
По всему миру ▼

Последние 5 лет ▼

Все категории ▼

Веб-поиск ▼

Динамика популярности ?



В среднем

# 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/>

# 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)

# Востребованный навык работы с данными

- К 2030 году грамотность в работе с данными станет самым востребованным навыком
- Спрос на навыки работы с данными во многом связан с развитием **искусственного интеллекта (AI)**.
- Профессии аналитика данных и дата-сайентиста уже сейчас являются наиболее востребованными
- Дефицит кадров будет только расти

<https://rb.ru/news/data-literacy/>

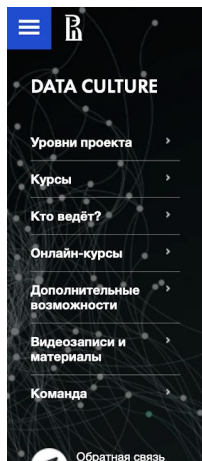




- Привет, Малыш!  
- Докурился, вентилятор разговаривает!

# Общевышкинский проект Data Culture

- Наш курс реализуется в рамках *общеуниверситетского проекта* **Data Culture**
- Этот проект стартовал в 2017/2018 уч. г.
- Охватывает **все бакалаврские программы** Вышки, часть магистерских и аспирантских программ
- Обучение **культуре работы с данными**, основам программирования, ключевым цифровым компетенциям современности



[Национальный исследовательский университет «Высшая школа экономики»](#) → Data Culture

RU EN   

## Data Culture

### Что такое Data Culture?

Data Culture — это общий термин для обозначения навыков и культуры работы с данными. Современный специалист в любой области постоянно сталкивается с задачами по обработке данных: юрист изучает сотни дел, историк — тысячи документов на различных языках, экономист — разрабатывает прогнозную модель на основе большого объема данных.

### Новости 3

2 июля 2020

**От основ информационной безопасности до методов машинного обучения**

Какие цифровые компетенции будут развивать студенты Вышки

29 июня 2020

**Вышкинский курс по цифровой грамотности открывается для внешних слушателей**

# Команда курса

## Семинары 201, 202 группы

**Аброскин Илья Дмитриевич**

Приглашенный преподаватель: [Факультет компьютерных наук](#) /  
[Департамент больших данных и информационного поиска](#)

## Семинары 203, 204 группы

**Быков Кирилл Валерьевич**

Приглашенный преподаватель: [Факультет компьютерных наук](#) /  
[Департамент больших данных и информационного поиска](#)

# Команда курса

## Лекции. Семинары 205, 206 группы

### Карпов Максим Евгеньевич

Аспирант, Младший научный сотрудник: [Факультет компьютерных наук](#) / [Научно-учебная лаборатория методов анализа больших данных](#)

Старший преподаватель: [Факультет компьютерных наук](#) / [Департамент больших данных и информационного поиска](#)

*Начал работать в НИУ ВШЭ в 2017 году.*

#### Образование

- 2020 Магистратура: Национальный исследовательский университет «Высшая школа экономики», факультет: компьютерных наук, специальность «Прикладная математика и информатика», квалификация «магистр»
- 2014 Специалитет: Московский государственный университет им. М.В. Ломоносова, факультет: глобальных процессов, специальность «Международные отношения», квалификация «специалист в области международных отношений»

# Команда курса

## Учебные ассистенты

201-202	Ярик Скобликов @Braye4 Илья Рябиков @ilyar11
203-204	Настя Егорова @wwhatisitt Лиза Бердина @berdina_elis
205-206	Коля Васильев @nnvasilevkk Миша Никитюк @iamchris0

# Логистика курса

Главный информационный ресурс – [wiki-страница](#) курса. На ней находятся все полезные ссылки, а именно:

- ПУД
- Форма обратной связи
- Плейлист с лекциями
- Контакты в телеграм
- Репозиторий с материалами на гитхабе

# Логистика курса



ФАКУЛЬТЕТ КОМПЬЮТЕРНЫХ НАУК

НИУ ВШЭ

[Заглавная страница](#)

▼ Инструменты

[Ссылки сюда](#)

[Связанные правки](#)

[Загрузить файл](#)

[Спецстраницы](#)

[Версия для печати](#)

[Постоянная ссылка](#)

[Сведения о  
странице](#)

Bunta

Статья [Обсуждение](#)

[Читать](#) [Пр](#)

## Основы программирования в Python (Мирэк) 2022

**Содержание** [\[убрать\]](#)

- 1 [О курсе](#)
- 2 [Преподаватели и ассистенты](#)
- 3 [Формула оценки](#)
- 4 [Материалы занятий](#)

### О курсе [\[править\]](#)

"Основы программирования в Python" читается на 2 курсе, в 3 и 4 модуле. 3 модуль - самостоятельное прохождение студентами онлайн-курса.

[Карточка курса и программа](#)

Канал и чат в Telegram (для всего потока), чаты групп: см. на [online.hse.ru](https://online.hse.ru)

[Плейлист с лекциями](#).

Ведомость с оценками – в разделе "Оценки" на курсе на [online.hse.ru](https://online.hse.ru)

[Форма для обратной связи](#).

# Цели курса

- Закрепление базовых навыков программирования на языке Python
- Формирование навыков работы со специализированными библиотеками для сбора, обработки, визуализации и анализа данных
- Развитие навыков чтения, интерпретации и оценки качества анализа и визуализации количественных данных.



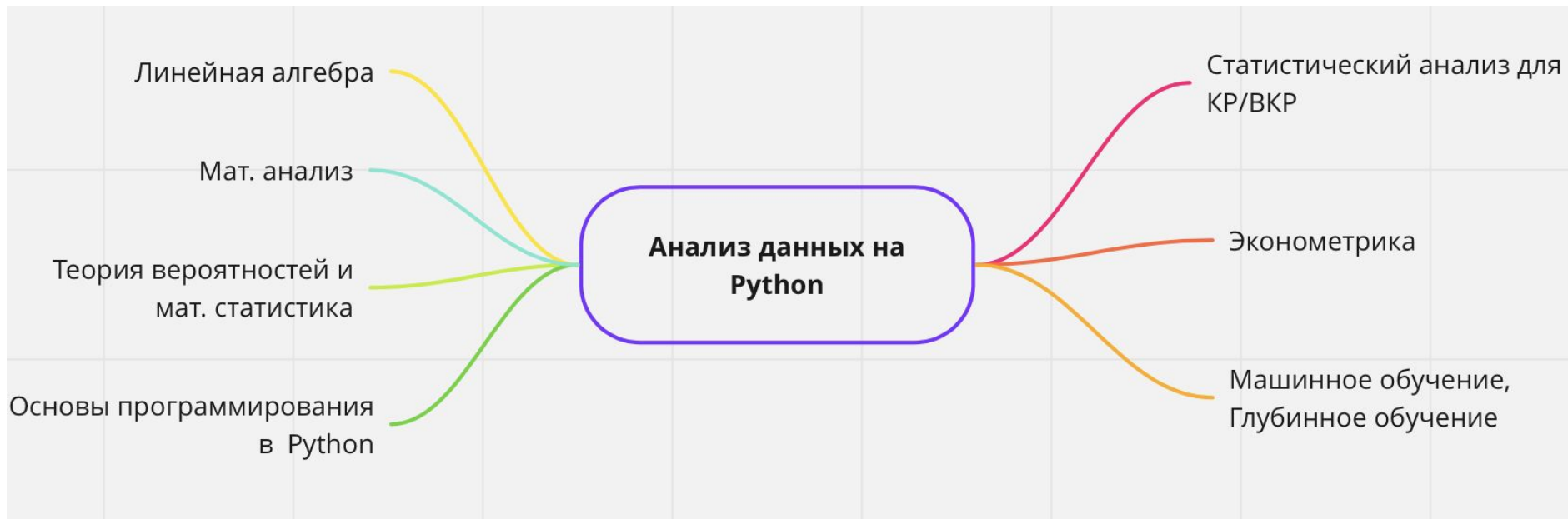
# Темы курса



## **Содержание учебной дисциплины**

- Основы программирования в Python
- Основы статистики
- Введение в анализ данных на Python
- Визуализация данных
- Введение в машинное обучение

# Место дисциплины в учебном плане ОП “Мировая экономика”



# Формула оценки



## Промежуточная аттестация

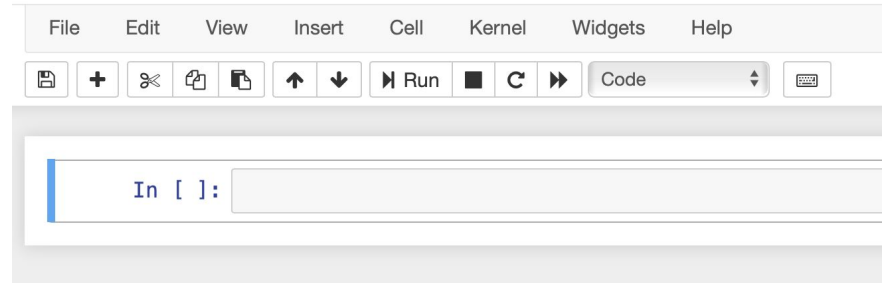
— 2021/2022 учебный год 4 модуль

*Итог = Минимум (0.15 \* Онлайн курс + 0.001 \* Квизы по читингу + 0.1 \* Домашние задания (среднее всех) + 0.2 \* Контрольная работа (Питон) + 0.25 \* Проект + 0.3 \* Экзамен (анализ данных) + 0.1 \* Бонусные баллы, 10)*

# Полезные горячие клавиши в джупитере

Находясь во внешней области ячейки (загорается **голубая** вертикальная линия)

 **jupyter notebook** Last Checkpoint: a few seconds ago (unsaved changes)



shift + enter    #запуск ячейки

x                    #удаление ячейки

a (above)        #добавление ячейки сверху

b (below)        #добавление ячейки снизу