



Введение в Python

ДПО Компьютерная лингвистика 2024.
Установочная встреча



Анна Сенина,
ст. преп. кафедры
гуманитарных дисциплин
НИУ ВШЭ - Пермь, историк,
преподаватель курсов по
Python и анализу данных



Аркадий Шалдов,
ассистент курса

Как устроен курс

Python с нуля

1. Знакомство с Python, git, PyCharm, Colab, Jupyter Notebook, VS Code
2. Операторы ввода/вывода (print, display, input), списки, **циклы**
3. Кортежи, множества, словари
4. Работа с **функциями**, методы
5. Работа с файлами, модуль os

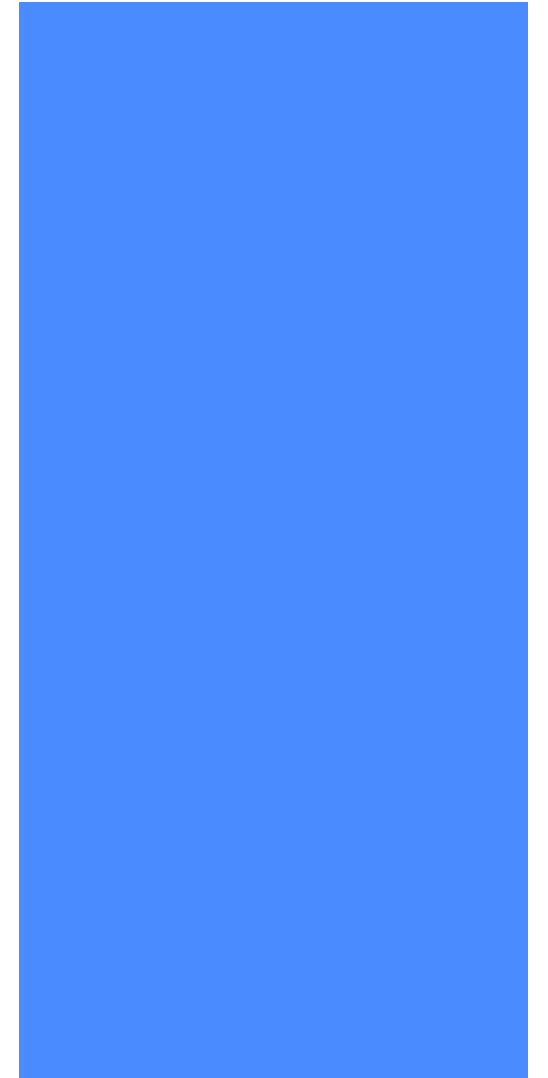


Python для CL

1. **Регулярные выражения**, препроцессинг
2. Частотные списки и n-граммы
3. rymorphy, mystem и др
4. pandas, работа с датасетами
5. Анализ данных и визуализация
6. собираем корпус: **парсеры...**



1. Classes and objects
2. Пользовательские интерфейсы
3. numru, pandas к машинному обучению



Оценка за курс

Домашние работы

4 домашних задания,
первые 3 задания
дают по 1 баллу
каждое, парсинг - 2
балла

Проект

Вес 0,5 в итоговой
оценке

Зачет по курсу

4 балла из 10 = зачет

Нет итогового теста (!)

Гвидо ван Россум



«Давняя гегемония Java и C закончилась. Простота обучения, огромное количество библиотек и широкое использование во всех областях сделали Python самым популярным языком программирования на сегодняшний день», – Пол Янсен, генеральный директор Tiobe.



Хабр



Как мы узнаём, какой язык программирования самый популярный

Менеджмент

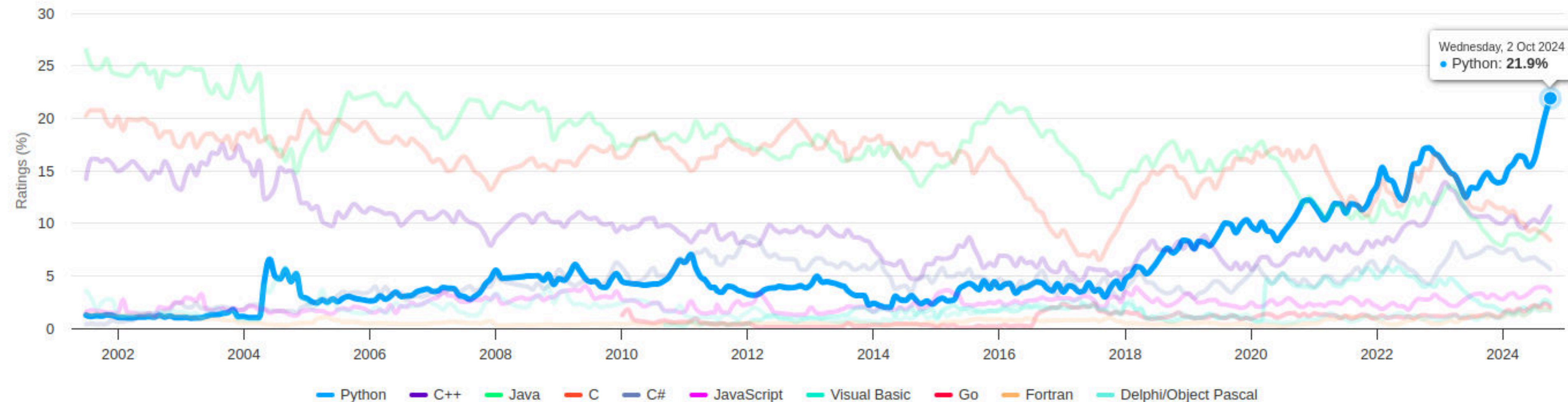
Как мы узнаём, какой язык программирования самый популярный

Выясним, как вычисляются рейтинги языков программирования, какие языки программирования наиболее популярны, и что нам с этой информацией делать. В декабре 2021 года Python остается одним из самых...

Хабр / Dec 15, 2021

TIOBE Programming Community Index

Source: www.tiobe.com



***Единственный способ
научиться
программированию -
это программировать***



1. Свобода манипуляции объектами цифрового мира
2. Свобода создания инструментов для ваших нужд
3. Свобода трудоустройства и передвижения
(Д. Скоринкин)

План занятия

1. Установим Python*
2. Выберем среду программирования
3. Зарегистрируемся на GitHub
4. Начнем знакомство с типами данных



Установка Python

Самая новая версия: Python 3.13 (любая версия, начиная с 3.7 и выше, подойдет, если вы ранее скачивали Python)

Важно! Python при установке спросит, **требуется ли установить pip** - **проверьте, что поставили галочку!**

для Windows <https://www.python.org/downloads/windows/>
MacOS <https://www.python.org/downloads/macos/>
Linux <https://www.python.org/downloads/source/>

Ждем всех и проверяем

- Windows: открываем Windows PowerShell
- MacOS / Linux: открываем терминал

Пишем `python` (иногда `python3` / `py`)

Программы-помощники: среды программирования

- Цветовые подсказки
- Элементы синтаксиса (автоматически закрываются кавычки, скобки, расставляются правильные отступы и т.д.)
- Ошибки в коде могут подчеркиваться как в MS Word
- Некоторые программы могут подсказывать вам команду (похоже на набор текста с телефона)

Важно!

Сейчас будет много незнакомых слов...
Мы смотрим 3 программы-помощника и
выбираем для себя ***ОДНУ*** из них



Разные типы программ-помощников

пишем код

смотрим
результат

PyCharm и др.

код

результат

код

результат

код

результат

**Jupyter Notebook /
Google Colab и др.**

!!! удобнее на курсе

Сравнение программ-помощников (программа с окошками)

PyCharm

Преимущества: debugger - помогает с поиском ошибок; программисты предпочитают этот вариант, поскольку можно связывать библиотеки с конкретными проектами

Недостатки: нужно запускать программу целиком, доведя ее до вывода данных (сложнее начинающим)

Сравнение программ-помощников (программа-тетрадка)

Jupyter Notebook (базовый вариант для нас)

Преимущества: тетрадки с кодом (пошаговый запуск отдельных команд)

- Установить через командную строку (сложнее настроить)

ИЛИ

- Установить программу Anaconda (это сборка программ, в нее входит Jupyter Notebook, а еще вы сможете в нее добавить программу для работы с R позднее)

Сравнение программ-помощников (программа-тетрадка)

Visual Studio Code

- **Преимущества:** тетрадки с кодом, отладчик, большой набор расширений

Сравнение программ-помощников (программа-тетрадка)

Google Colab (без установки)

Преимущества: тетрадки с кодом; **не требует установки** (нужен аккаунт в Google)

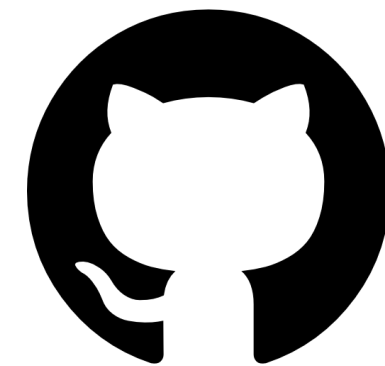
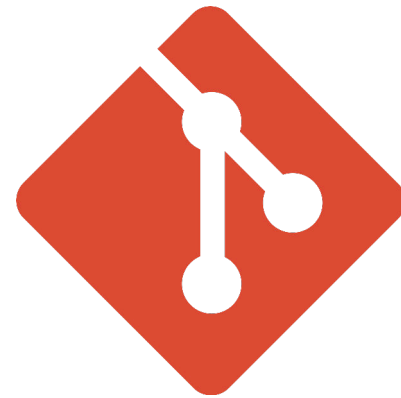
Недостатки: есть неудобства с хранением данных, а также с длительным запуском программ (парсинг)

Последовательно
смотрим все
программы,
**выбираем для
работы ОДНУ**

Можно пока выбрать Google
Colab, не требующий установки, а
к концу курса перейти в похожий
Jupyter



Git и GitHub



1. **Git** - система контроля версий
2. **GitHub** - сайт/соцсеть/коллаборация



диплом.docx



диплом2.docx



диплом3.docx



диплом_финал.docx



диплом_финал3финал_
ver100500aaaaaa...docx

Концепция git

этапа становления и основы российской общественной медицины говорит и тот факт, что исторический опыт земств, опыт земских врачей служит одним из истоков решения современных проблем развития здравоохранения в регионах России.

Земская медицина традиционно рассматривается как важный этап организации системы общественного здравоохранения в России.

Земские собрания и управы в области медицины и здравоохранения занимались организационно-хозяйственными, управленческими, кадровыми и иными вопросами, в том числе: строительством и ремонтом больниц, приемом на работу врачей и фельдшеров, организацией курсов для повышения квалификации врачей, постановкой профилактической и санитарной службы устройством санитарной части. Во многом, земская медицина создала базу для развития научной медицины в России и способствовала повышению квалификации местных врачей и фельдшеров.

Сентябрь

▶ 1 сентября, 19:57

Текущая версия

● Анна Сенина

● Все анонимные пользо

▶ 1 сентября, 18:50

● Все анонимные пользо

● Анна Сенина

1 сентября, 17:43

● Все анонимные пользо

▶ 1 сентября, 16:35

● Анна Сенина

▶ 1 сентября, 14:10

● Анна Сенина

Концепция git: контроль версий / изменения

Регистрация на GitHub

1. Заходим на github.com
2. Регистрируем аккаунт через электронную почту (не .ru, только .com)
3. Создаем репозиторий
4. Работаем через веб-интерфейс