



ПРЕЗИДЕНТСКАЯ
АКАДЕМИЯ

Работа с API Московской биржи

От установки ПО до первых результатов

Москва

2023

РАНХиГС

Python и R

Python – высокоуровневый язык программирования

- Минималистичный синтаксис
- Подключаемые библиотеки
- Огромное сообщество

Python используется повсеместно

- Анализ данных
- Машинное обучение
- Веб-разработка

R – применяется в основном для обработки данных, работы с графикой. Широко используется для статистического анализа данных.



Среда разработки

Среда разработки (IDE, **I**ntegrated **D**evelopment **E**nvironment) – программная оболочка используемая для написания кода.

Наиболее популярные IDE для Python:

- PyCharm
- Jupyter Notebook
- Atom

IDE для R:

- R-Studio

Все перечисленное – либо бесплатно полностью, либо имеет бесплатную версию.



Jupyter Notebook

Jupyter Notebook позволяет писать код в браузере.

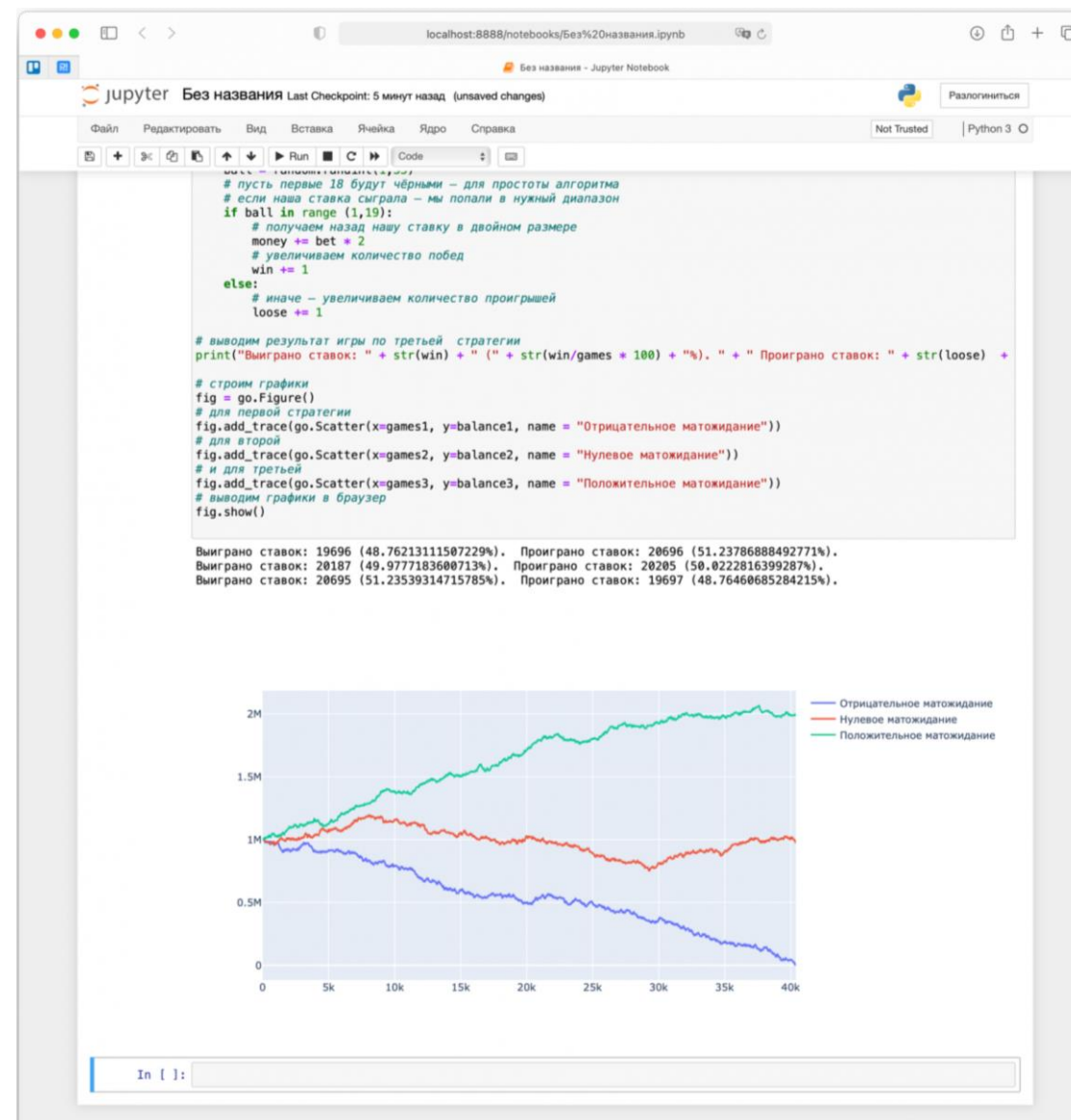
Написанный код можно выполнять блоками в любом порядке. Кроме того можно оставлять примечания используя Markdown.

Используется для написания небольших скриптов. Для больших проектов – PyCharm.

Поддерживает Python, R и некоторые другие языки.

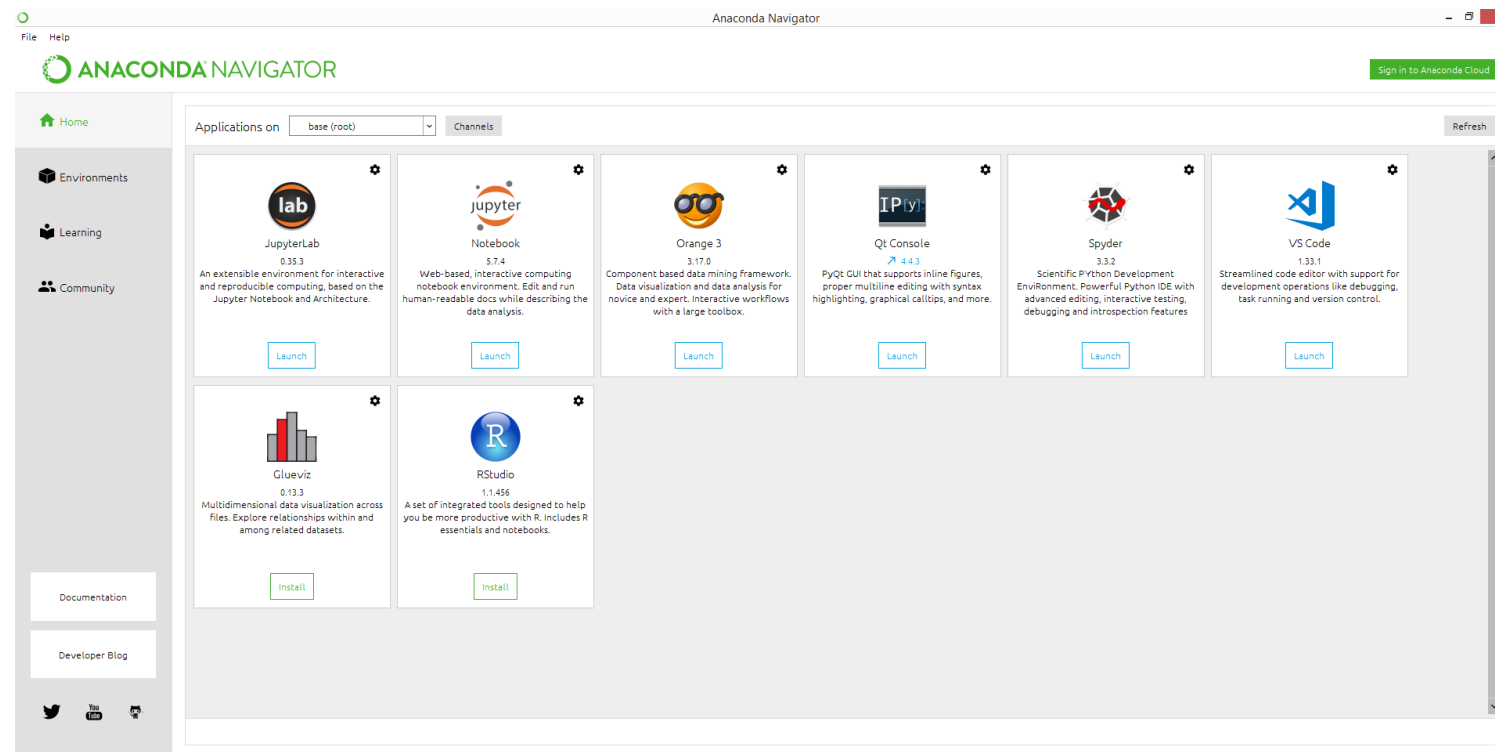
При использовании дополнительных библиотек может дополнять и форматировать код, делать цветовую подсветку синтаксиса.

Облачная версия от Google – Google Colab – позволяет работать даже на слабых устройствах, необходимо лишь наличия интернета (все вычисления выполняются облачно).



Anaconda

Anaconda – программный комплекс, позволяющий без сложностей установить как и сам Python так среду разработки PyCharm, Jupyter Notebook, R-Studio и дополнительные библиотеки.



Запуск Jupyter Notebook

Ярлык для запуска Jupiter Notebook после установки Anaconda можно найти либо запустив Anaconda Navigator или в списке всех программ Anaconda -> Jupyter Notebook.

При запуске появляется черный экран (командная строка) и через некоторое время открывается новая страница в браузере.

Как правило открывается папка пользователя (C:\Users\UserName).

Для удобства можно настроить любой другой каталог.

В этом окне можно создавать новые папки и файлы (Кнопка «New», вверху справа)



Навигация в Jupyter Notebook

Основные кнопки на панели – запуск и остановка выполнения кода (выделено красным). Запускать ячейку кода можно и с клавиатуры – CTRL + ENTER

Стрелки вниз/вверх позволяют перемещать ячейки кода, + – добавляет ячейку после

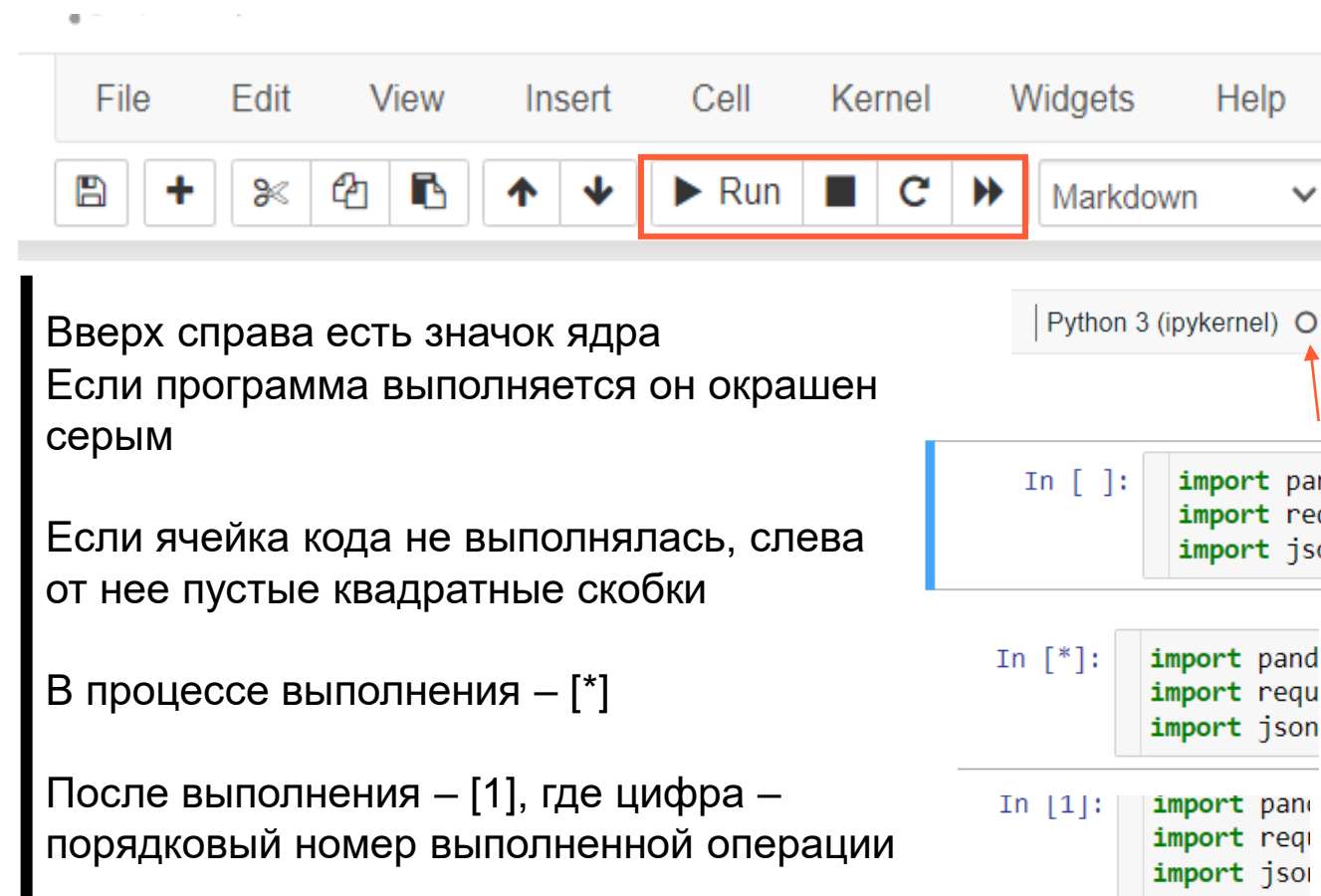
Так же понадобятся пункты меню Cell и Kernel:

Cell – позволяет запустить выбранные ячейки либо весь код целиком

Kernel – пригодится после установки модулей для перезагрузки ядра либо для отчистки результатов вывода

Выпадающее меню (Markdown) позволяет выбрать тип ячейки – для написания текста (Markdown) или для программного кода (Code).

Ячейки Markdown при выполнении кода программы игнорируются



Вверх справа есть значок ядра

Если программа выполняется он окрашен серым

Если ячейка кода не выполнялась, слева от нее пустые квадратные скобки

В процессе выполнения – [*]

После выполнения – [1], где цифра – порядковый номер выполненной операции

Парсинг данных

Парсинг – автоматизированный сбор структурированной информации с сайтов.

В ряде случаев парсинг довольно сложный процесс. Многие сайты имеют защиту от парсинга, которую в ряде случаев сложно обойти.

Библиотеки Python для парсинга:

- Requests – простые запросы к сайтам, чаще для работы с API
- Selenium – средство авто тестирования, эмулирует работу браузера
- BeautifulSoup – разбор HTML кода полученного от Requests или Selenium

Аналогичные библиотеки существуют так же на R

Перечисленные библиотеки не входят в стандартную установку, их необходимо устанавливать отдельно.

```
<!DOCTYPE html>
<html lang="ru" style="margin-top: 0 !important;">

<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0, maximum-scale=1.0, user-scalable=no">
  <meta name="theme-color" content="#ffffff">
  <link rel="preload" href="https://thecode.media/wp-content/themes/thecode_refresh/assets/fonts/Steinbeck.woff2" as="font" type="font/woff2" crossorigin">
  <link rel="preload" href="https://thecode.media/wp-content/themes/thecode_refresh/assets/fonts/Yandex-Sans-Text-Regular.woff2" as="font" type="font/woff2">
  <meta name="robots" content="index, follow, max-image-preview:large, max-snippet:-1, max-video-preview:-1" />

  <title>Что такое jupyter-ноутбук и для чего он нужен – Журнал «Код»</title>
  <meta name="description" content="Если вы хотите писать на Python или заниматься дата-сайенсом, вот классная штука для вас. Это уже, считай, стандарт" />
  <link rel="canonical" href="https://thecode.media/jupyter/" />
  <meta property="og:locale" content="ru_RU" />
  <meta property="og:type" content="article" />
  <meta property="og:title" content="Что такое jupyter-ноутбук и для чего он нужен – Журнал «Код»" />
  <meta property="og:description" content="Если вы хотите писать на Python или заниматься дата-сайенсом, вот классная штука для вас. Это уже, считай, стандарт" />
  <meta property="og:url" content="https://thecode.media/jupyter/" />
  <meta property="og:site_name" content="Журнал «Код» программирование без снобизма" />
  <meta property="article:publisher" content="https://www.facebook.com/thecode.media/" />
  <meta property="article:published_time" content="2021-05-28T09:38:14+00:00" />
  <meta property="article:modified_time" content="2023-05-05T09:58:34+00:00" />
  <meta property="og:image" content="https://thecode.media/wp-content/uploads/2021/05/photo_2021-05-27-16-13-58.jpg" />
  <meta property="og:image:width" content="1080" />
  <meta property="og:image:height" content="718" />
  <meta property="og:image:type" content="image/jpeg" />
  <meta name="author" content="Никита Кучеров" />
  <meta name="twitter:card" content="summary_large_image" />
  <meta name="twitter:creator" content="@thecode_media" />
  <meta name="twitter:site" content="@thecode_media" />
  <script type="application/ld+json" class="yoast-schema-graph">{
    "@context": "https://schema.org",
    "@graph": [
      {
        "@type": "Article",
        "@id": "https://thecode.media/jupyter/",
        "headline": "Что такое jupyter-ноутбук и для чего он нужен – Журнал «Код»",
        "description": "Если вы хотите писать на Python или заниматься дата-сайенсом, вот классная штука для вас. Это уже, считай, стандарт",
        "image": "https://thecode.media/wp-content/uploads/2021/05/photo_2021-05-27-16-13-58.jpg",
        "author": {
          "@type": "Person",
          "name": "Никита Кучеров"
        },
        "publisher": {
          "@type": "Organization",
          "name": "Журнал «Код»",
          "logo": "https://thecode.media/wp-content/uploads/2021/05/logo.png"
        },
        "publishedTime": "2021-05-28T09:38:14+00:00",
        "modifiedTime": "2023-05-05T09:58:34+00:00"
      }
    ]
  }</script>

  <link rel="stylesheet" id="wp-block-library-css" href="https://thecode.media/wp-includes/css/dist/block-library/style.min.css?ver=6.3.1" type="text/css">
  <link rel="stylesheet" id="prismatic-blocks-css" href="https://thecode.media/wp-content/plugins/prismatic/css/styles-blocks.css?ver=6.3.1" type="text/css">
  <style id="classic-theme-styles-inline-css" type="text/css">
    /* This file is auto-generated */
  </style>
```


API

API сокращение от **A**pplication **P**rogramming **I**nterface – интерфейс программирования приложения, то есть описание способов взаимодействия одной программы с другими.

Многие мобильные приложения, например Telegram, работают по средствам API, отправляя и получая сообщения в виде специально сформированной строки запроса (JSON)

JSON напоминает собой структуру вложенных словарей, с доступом к информации по ключу словаря.

Для работы с JSON в Python есть библиотека с аналогичным названием, которая входит в стандартную установку.

```
{  
  "history": {  
    "metadata": {  
      "BOARDID": {"type": "string", "bytes": 12, "max_size": 0},  
      "SECID": {"type": "string", "bytes": 36, "max_size": 0},  
      "TRADEDATE": {"type": "date", "bytes": 10, "max_size": 0},  
      "SHORTNAME": {"type": "string", "bytes": 189, "max_size": 0},  
      "NAME": {"type": "string", "bytes": 765, "max_size": 0},  
      "CLOSE": {"type": "double"},  
      "OPEN": {"type": "double"},  
      "HIGH": {"type": "double"},  
      "LOW": {"type": "double"},  
      "VALUE": {"type": "double"},  
      "DURATION": {"type": "int32"},  
      "YIELD": {"type": "double"},  
      "DECIMALS": {"type": "int32"},  
      "CAPITALIZATION": {"type": "double"},  
      "CURRENCYID": {"type": "string", "bytes": 9, "max_size": 0},  
      "DIVISOR": {"type": "double"},  
      "TRADINGSESSION": {"type": "string", "bytes": 3, "max_size": 0},  
      "VOLUME": {"type": "double"}  
    },  
    ...  
  },  
  ...  
}
```

ИСС Московской биржи

ИСС – Информационно-статистический сервер Московской биржи.

ИСС как раз и есть API, созданный для работы с базой данных Московской биржи.

Может быть использован в том числе и для разработки приложений связанных с фондовым рынком.

Имеется платный вариант (подписка, чаще всего оплачивается определенное количество запросов), в этом случае данные предоставляются в реальном времени (например для создания торговых приложений, отображающих текущие котировки акций). При бесплатном доступе данные имеют небольшую задержку.

В рамках интерфейса доступны следующие типы информации:

- статические данные о рынках (режимы торгов и их группы, финансовые инструменты и их описание)
- данные для построения графиков ("свечей")
- сделки (анонимно)
- котировки
- исторические данные
- различные метаданные.

Доступные форматы ИСС

Через API Мосбиржи, в зависимости от запроса, можно получить данные в виде JSON, XML, HTML или сформированный csv файл. Для обработки в Python наиболее привычен JSON или чуть менее популярный XML.

history

BOARDID	SECID	TRADEDATE	SHORTNAME	NAME	CLOSE	OPEN	HIGH	LOW
(string:12)	(string:36)	(date:10)	(string:189)	(string:765)	(double)	(double)	(double)	(double)
SNDX	MOEXOG	2004-12-30	Индекс нефти и газа	МосБиржи нефти и газа	1000	1000	1000	1000
SNDX	MOEXOG	2005-01-11	Индекс нефти и газа	МосБиржи нефти и газа	1007.46	1000.7	1016.19	998.14
SNDX	MOEXOG	2005-01-12	Индекс нефти и газа	МосБиржи нефти и газа	1009.62	1007.46	1012.27	1001.41

HTML

history						
BOARDID	SECID	TRADEDATE	SHORTNAME	NAME	CLOSE	OPEN
SNDX	MOEXOG	#####	Индекс не	Индекс М	1000	1000
SNDX	MOEXOG	#####	Индекс не	Индекс М	1007.46	1000.7
SNDX	MOEXOG	#####	Индекс не	Индекс М	1009.62	1007.46
SNDX	MOEXOG	#####	Индекс не	Индекс М	1008.79	1009.46
SNDX	MOEXOG	#####	Индекс не	Индекс М	1022.41	1008.77
SNDX	MOEXOG	#####	Индекс не	Индекс М	1016.77	1022.42

CSV

JSON

```
{
  "history": {
    "metadata": {
      "BOARDID": {"type": "string", "bytes": 12, "max_size": 0},
      "SECID": {"type": "string", "bytes": 36, "max_size": 0},
      "TRADEDATE": {"type": "date", "bytes": 10, "max_size": 0},
      "SHORTNAME": {"type": "string", "bytes": 189, "max_size": 0},
      "NAME": {"type": "string", "bytes": 765, "max_size": 0},
      "CLOSE": {"type": "double"},
      "OPEN": {"type": "double"},
      "HIGH": {"type": "double"},
      "LOW": {"type": "double"},
      "VALUE": {"type": "double"},
      "DURATION": {"type": "int32"},
      "YIELD": {"type": "double"},
      "DECIMALS": {"type": "int32"},
      "CAPITALIZATION": {"type": "double"},
      "CURRENCYID": {"type": "string", "bytes": 9, "max_size": 0},
      "DIVISOR": {"type": "double"},
      "TRADINGSESSION": {"type": "string", "bytes": 3, "max_size": 0},
      "VOLUME": {"type": "double"}
    },
    "columns": ["BOARDID", "SECID", "TRADEDATE", "SHORTNAME", "NAME", "CLOSE", "OPEN", "HIGH", "LOW", "VALUE", "DURATION", "YIELD", "DECIMALS", "CAPITALIZATION", "CURRENCYID", "DIVISOR", "TRADINGSESSION", "VOLUME"],
    "data": [
      ["SNDX", "MOEXOG", "2004-12-30", "Индекс нефти и газа", "МосБиржи нефти и газа", 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000],
      ["SNDX", "MOEXOG", "2005-01-11", "Индекс нефти и газа", "МосБиржи нефти и газа", 1007.46, 1000.7, 1016.19, 998.14, 1007.46, 1000.7, 1016.19, 998.14, 1007.46, 1000.7, 1016.19, 998.14, 1007.46, 1000.7],
      ["SNDX", "MOEXOG", "2005-01-12", "Индекс нефти и газа", "МосБиржи нефти и газа", 1009.62, 1007.46, 1012.27, 1001.41, 1009.62, 1007.46, 1012.27, 1001.41, 1009.62, 1007.46, 1012.27, 1001.41, 1009.62, 1007.46, 1012.27, 1001.41]
    ]
  }
}
```

XML

```
<document>
  <data id="history">
    <metadata>
      <columns>
        <column name="BOARDID" type="string" bytes="12" max_size="0"/>
        <column name="SECID" type="string" bytes="36" max_size="0"/>
        <column name="TRADEDATE" type="date" bytes="10" max_size="0"/>
        <column name="SHORTNAME" type="string" bytes="189" max_size="0"/>
        <column name="NAME" type="string" bytes="765" max_size="0"/>
        <column name="CLOSE" type="double"/>
        <column name="OPEN" type="double"/>
        <column name="HIGH" type="double"/>
        <column name="LOW" type="double"/>
        <column name="VALUE" type="double"/>
        <column name="DURATION" type="int32"/>
        <column name="YIELD" type="double"/>
        <column name="DECIMALS" type="int32"/>
        <column name="CAPITALIZATION" type="double"/>
        <column name="CURRENCYID" type="string" bytes="9" max_size="0"/>
        <column name="DIVISOR" type="double"/>
        <column name="TRADINGSESSION" type="string" bytes="3" max_size="0"/>
        <column name="VOLUME" type="double"/>
      </columns>
    </metadata>
    <rows>
      <row BOARDID="SNDX" SECID="MOEXOG" TRADEDATE="2004-12-30" SHORTNAME="Индекс нефти и газа" NAME="МосБиржи нефти и газа" CLOSE="1000" OPEN="1000" HIGH="1000" LOW="1000" VALUE="1000" DURATION="1000" YIELD="1000" DECIMALS="1000" CAPITALIZATION="1000" CURRENCYID="RUB" DIVISOR="1" TRADINGSESSION="1" VOLUME="1000"/>
      <row BOARDID="SNDX" SECID="MOEXOG" TRADEDATE="2005-01-11" SHORTNAME="Индекс нефти и газа" NAME="МосБиржи нефти и газа" CLOSE="1007.46" OPEN="1000.7" HIGH="1016.19" LOW="998.14" VALUE="1007.46" DURATION="1000.7" YIELD="1016.19" DECIMALS="998.14" CAPITALIZATION="1007.46" CURRENCYID="1000.7" DIVISOR="1016.19" TRADINGSESSION="998.14" VOLUME="1007.46"/>
      <row BOARDID="SNDX" SECID="MOEXOG" TRADEDATE="2005-01-12" SHORTNAME="Индекс нефти и газа" NAME="МосБиржи нефти и газа" CLOSE="1009.62" OPEN="1007.46" HIGH="1012.27" LOW="1001.41" VALUE="1009.62" DURATION="1007.46" YIELD="1012.27" DECIMALS="1001.41" CAPITALIZATION="1009.62" CURRENCYID="1007.46" DIVISOR="1012.27" TRADINGSESSION="1001.41" VOLUME="1009.62"/>
    </rows>
  </data>
</document>
```

Простые запросы

Запросы к API Мосбиржи можно выполнять не только через IDE но и просто написав его в браузере.

Простой запрос информации об инструменте:

`http://iss.moex.com/iss/securities/GAZP.формат`

Начинается запрос с протокола HTTP (HTTPS используется для платной подписки, т.к. передаются данные авторизации).

Затем основной адрес запроса - `iss.moex.com/iss/...`

В конце указывается тикер ценной бумаги – GAZP (Газпром).

Тикеры можно найти в любой поисковой системе или на сайте Мосбиржи.

Завершает запрос формат получаемых данных - `.json`, `.xml`, `.csv` (без указания формата будет получен HTML).

Простые запросы есть в документации ИСС на сайте Мосбиржи <https://www.moex.com/a2193>

Там же описание метаданных и методов, пример небольшой программы на Python и Visual Basic.

Информация о ценной бумаге (HTML)

При запросе мы получим 2 блока данных

description – описание ценной бумаги. В этом блоке содержится коды переменных, их описание и значение, а так же формат данных (строка, число).

boards – блок описания режимов торгов для данной бумаги, наличие и периоды исторических данных, валюту торгов и другую вспомогательную информацию, которая может пригодится для построения более сложных запросов

description

name (string:189)	title (string:765)	value (string:6144)	t
SECID	Код ценной бумаги	GAZP	s
NAME	Полное наименование	"Газпром" (ПАО) ао	s
SHORTNAME	Краткое наименование	ГАЗПРОМ ао	s
ISIN	ISIN код	RU0007661625	s
REGNUMBER	Номер государственной регистрации	1-02-00028-A	s
ISSUESIZE	Объем выпуска	23673512900	r
FACEVALUE	Номинальная стоимость	5	r
FACEUNIT	Валюта номинала	SUR	s
ISSUEDATE	Дата начала торгов	2006-01-23	d

boards

secid (string:51)	boardid (string:12)	title (string:381)	board_group_id (int32)	market_id (int32)	market (string:45)	engine_id (int32)	engine (string:45)	is_traded (int32)	decimals (int32)	history_from (date:10)	history_till (date:10)	listed_from (date:10)	listed_till (date:10)	is_primary (int32)	currencycid (string:9)
GAZP	TQBR	T+: Акции и ДР - безадрес.	57	1	shares	1	stock	1	2	2014-06-09	2023-10-06	2014-06-09	2023-10-09	1	RUB
GAZP	TQBS	T+: А2-Акции и пан - безадрес. T+: Акции пан и ЛР	57	1	shares	1	stock	0	2	2013-12-26	2014-06-06	2013-12-26	2014-06-06	0	RUB

В формате JSON информация менее приспособлена для визуального просмотра, но более удобна для обработки на компьютере.

metadata – содержит имена и описание формата колонок (в терминах python – это вложенные словари, *dictionary*)

`columns` – блок названия колонок (значение словаря – список, *list*)

data – основная информация, по структуре похожа на csv файл, где колонки разделены запятой (значение словаря – список списков)

```
{
  "description": {
    "metadata": {
      "name": {"type": "string", "bytes": 189, "n
      "title": {"type": "string", "bytes": 765, "
      "value": {"type": "string", "bytes": 6144,
      "type": {"type": "string", "bytes": 21, "ma
      "sort_order": {"type": "int64"}},
      "is_hidden": {"type": "int64"}},
      "precision": {"type": "int64"}
    },
    "columns": ["name", "title", "value", "type", "sort
    "data": [
      ["SECID", "Код ценной бумаги", "GAZP", "str
      ["NAME", "Полное наименование", "\Газпром
      ["SHORTNAME", "Краткое наименование", "ГАЗП
      ["ISIN", "ISIN код", "RU0007661625", "strir
      ["REGNUMBER", "Номер государственной регистр
      ["ISSUESIZE", "Объем выпуска", "23673512906
      ["FACEVALUE", "Номинальная стоимость", "5",
      ["FACEUNIT", "Валюта номинала", "SUR", "str
      ["ISSUEDATE", "Дата начала торгов", "2006-6
      ["LATNAME", "Английское наименование", "Gaz
      ["LISTLEVEL", "Уровень листинга", "1", "nul
      ["ISQUALIFIEDINVESTORS", "Бумаги для квалиф
      ["MORNINGSSESSION", "Допуск к утренней допол
      ["EVENINGSESSION", "Допуск к вечерней допол
      ["TYPENAME", "Вид\категория ценной бумаги"
      ["GROUP", "Код типа инструмента", "stock_sl
      ["TYPE", "Тип бумаги", "common_share", "str
      ["GROUPNAME", "Типа инструмента", "Акции",
      ["EMITTER_ID", "Код эмитента", "1243", "nul
    ]
  },
  "boards": {
    "metadata": {
      "secid": {"type": "string", "bytes": 51, "n
      "boardid": {"type": "string", "bytes": 12,
```

```
[
    {
        "GROUPNAME": "Типа инструмента", "Акции", "string",
        "EMITTER_ID": "Код эмитента", "1243", "number", 11
    }
],
"boards": {
    "metadata": {
        "secid": {"type": "string", "bytes": 51, "max_size": 51},
        "boardid": {"type": "string", "bytes": 12, "max_size": 12},
        "title": {"type": "string", "bytes": 381, "max_size": 381},
        "board_group_id": {"type": "int32"},
        "market_id": {"type": "int32"},
        "market": {"type": "string", "bytes": 45, "max_size": 45},
        "engine_id": {"type": "int32"},
        "engine": {"type": "string", "bytes": 45, "max_size": 45},
        "is_traded": {"type": "int32"},
        "decimals": {"type": "int32"},
        "history_from": {"type": "date", "bytes": 10, "max_size": 10},
        "history_till": {"type": "date", "bytes": 10, "max_size": 10},
        "listed_from": {"type": "date", "bytes": 10, "max_size": 10},
        "listed_till": {"type": "date", "bytes": 10, "max_size": 10},
        "is_primary": {"type": "int32"},
        "currencyid": {"type": "string", "bytes": 9, "max_size": 9}
    },
    "columns": ["secid", "boardid", "title", "board_group_id"],
    "data": [
        ["GAZP", "TQBR", "Т+: Акции и ДР - безадрес.", 57, 1],
        ["GAZP", "TQBS", "Т+: А2-Акции и паи - безадрес.", 1],
        ["GAZP", "TQNE", "Т+: Акции, паи и ДР внесписочные", 1],
        ["GAZP", "EQNE", "Основной режим: Акции и паи внесписочные", 1],
        ["GAZP", "SPEQ", "Поставка по СК (акции)", 125, 1, 1],
        ["GAZP", "SMAL", "Т+: Неполные лоты (акции) - безадрес.", 1],
        ["GAZP", "TQDP", "Крупные пакеты - Акции - безадрес.", 1],
        ["GAZP", "EQDP", "Крупные пакеты - Акции - безадрес.", 1],
        ["GAZP", "RPMO", "РЕПО-М - адрес.", 201, 3, "repo"]
    ]
}
```

Запрос исторических данных

Рассмотрим запроса исторических данных по акции Сбера:

<http://iss.moex.com/iss/history/engines/stock/markets/shares/boards/TQBR/securities/SBER.json>

Облигациям РЖД:

<http://iss.moex.com/iss/history/engines/stock/markets/bonds/boards/TQCB/securities/RU000A0JSGV0.json>

Индексу МосБиржи:

<http://iss.moex.com/iss/history/engines/stock/markets/index/boards/SNDX/securities/imoex.json>

Метаданные

API Мосбирж позволяет добавлять к запросу дополнительные параметры:

- Указать за какой период нужна информация (за конкретный день или за промежуток между датами)
- Наличие в ответе метаданных
- Выбрать определенные столбцы

Описание методов, а так же допустимые параметры для разных запросов можно найти на страницах документации.

Параметры указываются в запросе либо явно:

Для этого после основного запроса ставится '?', затем идет параметр '=' его значение, между собой параметры разделяются '&':

<http://iss.moex.com/iss/history/engines/stock/markets/shares/boards/TQBR/securities/SBER.json?iss.meta=off&iss.only=history&from=2020-10-10>

Параметры можно задать явно, либо передаются в виде словаря Python (ключ – параметр, значение ключа – значения параметра) как аргумент к функции `requests.get(url, params=dict)`, где `url` – адрес запроса, `dict` – словарь параметров – `{“iss.meta”:“off”, “iss.only”:“history”, “from”:“дата в формате ГГГГ-ММ-ДД”}`

XML

ФНС предоставляет открытые данные в виде файлов XML.

Работать с ними не сложнее, чем с API МосБиржи.

Набор данных о среднесписочной численности работников, если открыть его в браузере:

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```

<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<Файл ИдФайл="VO_OTKRDAN_3_9965_9965_20230925_000c9d82-efe9-4fbc-8069-ccb8abbeb704" ВерсФорм="4.01" ВерсПрог="1.0" ТипИнф="ОТКРДАННБЗ" КолДок="900">
  <ИдОтпр>
    <ФИООтв Фамилия="_" Имя="_" />
  </ИдОтпр>
  <Документ ИдДок="56a34f7d-a260-4047-9898-aa3b60c9b50a" ДатаДок="25.09.2023" ДатаСост="31.12.2022">
    <СведНП НаимОрг="ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "УРАЛ-ПРЕСС СИБИРЬ"" ИННЮЛ="4205136916" />
    <СведССЧР КолРаб="20" />
  </Документ>
  <Документ ИдДок="0d737ea8-151d-4155-95d3-96930e169b3a" ДатаДок="25.09.2023" ДатаСост="31.12.2022">
    <СведНП НаимОрг="ПУБЛИЧНОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "КОЖЕВНИКОВСКОЕ АВТОТРАНСПОРТНОЕ ПРЕДПРИЯТИЕ"" ИННЮЛ="7008000020" />
    <СведССЧР КолРаб="17" />
  </Документ>
  <Документ ИдДок="0ae69004-244e-47bf-a835-d7920554d5f0" ДатаДок="25.09.2023" ДатаСост="31.12.2022">
    <СведНП НаимОрг="ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "АГРОВЕТСНАБ"" ИННЮЛ="7008005780" />
    <СведССЧР КолРаб="1" />
  </Документ>
  <Документ ИдДок="0d1d1630-61a6-41b1-8d3b-0ebb0cfa4a57" ДатаДок="25.09.2023" ДатаСост="31.12.2022">
    <СведНП НаимОрг="ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ГНБ - НЕФТЕСТРОЙ"" ИННЮЛ="7017047592" />
    <СведССЧР КолРаб="0" />
  </Документ>
</Файл>

```

ИСС Московской

Видно, что данные структурированы определенным образом

Python. Парсинг XML.

`xml.etree.ElementTree` – стандартная библиотека Python для работы с XML

При помощи библиотек `xml.etree.ElementTree`, `pandas` и `os` можно быстро получить данные из XML документов ФНС

Полезные ссылки

Репозитории - <https://github.com/kosyrev-aq>

Anaconda - <https://www.anaconda.com/>

PyCharm - <https://www.jetbrains.com/pycharm/download/>

Google Colab - <https://colab.research.google.com/>