Министерство образования Республики Беларусь

Учреждение образования

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет компьютерных систем и сетей

Кафедра информатики

Дисциплина: ТПпУБД

Отчёт по лабораторной работе на тему

**Обработка данных курсов валют**

Выполнил:

Магистрант гр.

956241 Шуба И.А.

Проверил:

Доцент, кандидат технических наук

Стержанов М.В.

Минск 2020

**СОДЕРЖАНИЕ**

1. [Постановка задачи…………………………………………………………3](#_TOC_250004)
2. [Получение данных…………………………………………………………...4](#_TOC_250003)
3. [Работа с API…………………………………………….……………………](#_TOC_250002).4
4. Функции программы………………………………………………………..7
5. Анализ данных……………………………………………………………….9

Вывод…………………………………………………………………………...12

# Постановка задачи

Целью данного проекта является обработка данных курсов валют национального банка НБ РБ. Для достижения данной цели были поставлены следующие задачи:

* Реализация сбора данных;
* Организация хранения и доступа к собранным данным;
* Построение графов курсов валют.

# Получение данных

Официальный курс белорусского рубля по отношению к доллару США, евро и российскому рублю устанавливается по данным биржевых торгов в режиме торгов с установлением фиксинга или как средневзвешенный курс по сделкам, совершенным на биржевых торгах в режиме непрерывного двойного аукциона, проводимых [открытым акционерным обществом "Белорусская валютно-фондовая биржа"](http://www.bcse.by/) по доллару США, евро и российскому рублю. Национальный банк РБ, предоставляет доступ к своему сайту посредством API. Это очень удобно, поскольку нет необходимости скачивать HTML страницы и парсить их. Вместо этого мы можем получить доступ к информации сразу в подходящем формате, а именно в виде JSON файла.

**2. Работа с АPI**

# Национальный банк РБ предоставляет API для получения данных официального курса белорусского рубля по отношению к иностранным валютам, устанавливаемого Национальным банком Республики Беларусь:

### Полный перечень иностранных валют, по отношению к которым Национальным банком устанавливается официальный курс белорусского рубля.

### Официальный курс белорусского рубля по отношению к иностранным валютам, устанавливаемый Национальным банком на конкретную дату.

### Получение динамики официального курса белорусского рубля по отношению к заданной иностранной валюте (не более чем за 365 дней).

API — описание способов, которыми одна компьютерная программа может взаимодействовать с другой программой. Обычно входит в описание какого-либо интернет-протокола, программного каркаса или стандарта вызовов функций операционной системы.

# Ниже расписано API получения официального курса белорусского рубля по отношению к иностранным валютам, устанавливаемого Национальным банком Республики Беларусь.

### **Полный перечень иностранных валют, по отношению к которым Национальным банком устанавливается официальный курс белорусского рубля.**

**Адрес запроса:**: http://www.nbrb.by/API/ExRates/Currencies[/{Cur\_ID}]

**Результат:** Возвращает массив объектов класса Currency. Если указан Cur\_ID, то возвращается один объект Currency.

Класс **Currency** содержит следующие свойства:

* Cur\_ID – внутренний код
* Cur\_ParentID – этот код используется для связи, при изменениях наименования, количества единиц к которому устанавливается курс белорусского рубля, буквенного, цифрового кодов и т.д. фактически одной и той же валюты\*.
* Cur\_Code – цифровой код
* Cur\_Abbreviation – буквенный код
* Cur\_Name – наименование валюты на русском языке
* Cur\_Name\_Bel – наименование на белорусском языке
* Cur\_Name\_Eng – наименование на английском языке
* Cur\_QuotName – наименование валюты на русском языке, содержащее количество единиц
* Cur\_QuotName\_Bel – наименование на белорусском языке, содержащее количество единиц
* Cur\_QuotName\_Eng – наименование на английском языке, содержащее количество единиц
* Cur\_NameMulti – наименование валюты на русском языке во множественном числе
* Cur\_Name\_BelMulti – наименование валюты на белорусском языке во множественном числе\*
* Cur\_Name\_EngMulti – наименование на английском языке во множественном числе\*
* Cur\_Scale – количество единиц иностранной валюты
* Cur\_Periodicity – периодичность установления курса (0 – ежедневно, 1 – ежемесячно)
* Cur\_DateStart – дата включения валюты в перечень валют, к которым устанавливается официальный курс бел. рубля
* Cur\_DateEnd – дата исключения валюты из перечня валют, к которым устанавливается официальный курс бел. рубля

### **Официальный курс белорусского рубля по отношению к иностранным валютам, устанавливаемый Национальным банком на конкретную дату.**

**Адрес запроса:**: http://www.nbrb.by/API/ExRates/Rates[/{Cur\_ID}]

**Параметры (GET):**

onDate\*\* – дата, на которую запрашивается курс (если не задана, то возвращается курс на сегодня)

Periodicity – периодичность установления курса (0 – ежедневно, 1 – ежемесячно)

ParamMode – формат аргумента Cur\_ID: 0 – внутренний код валюты, 1 – трехзначный цифровой  код валюты в соответствии со стандартом ИСО 4217, 2 – трехзначный буквенный код валюты (ИСО 4217). По умолчанию = 0

При использовании буквенного или цифрового кода валюты (ИСО 4217) учитывайте его значение на запрашиваемую дату.

**Результат:** Возвращает массив объектов класса Rate. Если указан Cur\_ID, то возвращается один объект Rate.

Класс **Rate** содержит следующие свойства:

Cur\_ID – внутренний код

Date – дата, на которую запрашивается курс

Cur\_Abbreviation – буквенный код

Cur\_Scale – количество единиц иностранной валюты

Cur\_Name – наименование валюты на русском языке во множественном, либо в единственном числе, в зависимости от количества единиц

Cur\_OfficialRate – курс\*

### **Получение динамики официального курса белорусского рубля по отношению к заданной иностранной валюте (не более чем за 365 дней).**

**Адрес запроса:**: http://www.nbrb.by/API/ExRates/Rates/Dynamics/{Cur\_ID}

**Параметры (GET):**

startDate\*\* – дата начала запрашиваемого периода

endDate\*\* – дата окончания запрашиваемого периода

**Результат:** Возвращает массив объектов класса RateShort.

Класс **RateShort** содержит следующие свойства:

Cur\_ID – внутренний код

Date – дата, на которую запрашивается курс

Cur\_OfficialRate – курс\*

**Примечание.** Если в любой из перечисленных методов передается неверный код валюты, то возвращается 404 ошибка. Если на запрашиваемую дату курс не установлен, возвращается пустой массив.

1. **Функционал программы.**

Сохранение данных проходят обычной записью в файл, для этого используется JSON (чтобы распарсить данные) и библиотека pandas, чтобы преобразовать словарь Python в csv формат.

validate\_args() - проверяет валидность введённых с терминала аргументов. Если они не валидны то выбрасывается исключение ValueError.

1. Если их больше 4 (*main.py* тоже считается).

2.Если первый аргумент не совпадает ни с одним из ключей словаря currencies/currencies.

3. Если второй и третий аргумент не соответствуют шаблону **%d-%m-%Y.**

Далее вызывается get\_rates() из scraper.py и проверяет есть ли уже csv файл с данными, если есть просто выходит. Если нету, то формируется ссылка к API национального банка, даты просто переворачиваются, т.к. нужен формат год-месяц-день, мы выводили по другому формату. safe\_request() из utils.py делает зацикливание, это нужно потому что Национальный банк очень слабо отвечает, иногда присылает ошибки. Чтобы фалы csv загрузились, а потом всё быстро работало. safe\_request() возвращает response с json структурой, поэтому используется json.loads(), он преобразует json в словарь Python. Далее с помощью библиотеки pandas создаётся DataFrame, сделано это только для того чтобы легко создать csv файл. ([date\_format(i["Date"]) for i in data] это генераторы списков <https://younglinux.info/python/feature/generators>). С помощью менеджера контекста (оператор with), создаётся и открывается файл, в который далее записывается dataframe. generate\_file\_path() просто генерирует уникальное имя файла, этот файл нужен для того чтобы в следующий раз не посылать запрос.

В main.py вызывается csv\_parse(), который возвращает два списка, для «Х» и для «У» он проходит по файлу csv и построчно записывает все значения в списки для простоты используется стандартная библиотека csv и csv.reader из неё исключается первая строчка first\_line, т.к. там по умолчанию названия полей, а нам нужны только данные.

label это подпись к графику.

plt.plot(x, y, "-", linewidth=1, label=label) с помощью matplotlib рисуется график, "-" это значит соединить точки линией, linewidth ширина линии.

plt.xticks(rotation=90) повернуть надписи для оси x.

plt.xticks(shorten\_dates(x)) так как надписей даты очень много, может быть до 365 дней, здесь говориться что нужно показывать только те что вернёт shorten\_dates(x). Он проходит по всему списку и оставляет каждый пятый элемент, остальные удаляет, если длина списка всё равно больше 31, то это повторяется.

plt.legend() добавляет надпись

plt.show() выводит на экран график.

1. **Анализ данных**

Для построения графиков был использован инструмент matplotlib.pyplot.plot. На рисунке 1 представлена официальный курс белорусского рубля по отношению к евро, устанавливаемый Национальным банком Республики Беларусь за месяц.

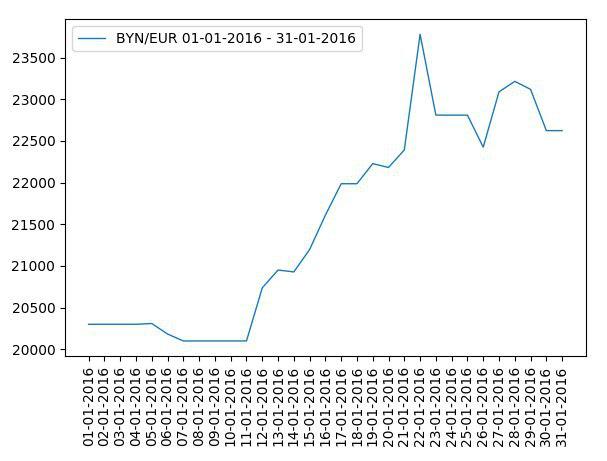


Рисунок 1 - Курс белорусского рубля по отношению к евро за месяц

# На рисунке 2 представлена официальный курс белорусского рубля по отношению к евро, устанавливаемый Национальным банком Республики Беларусь за месяц.

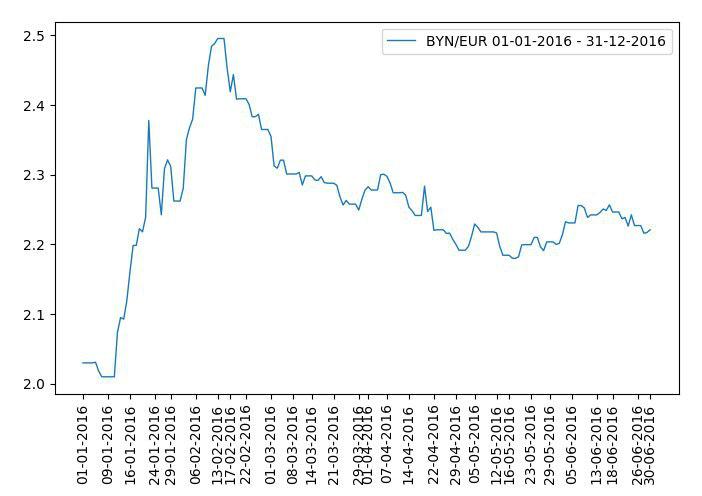


Рисунок 2 - Курс белорусского рубля по отношению к евро за выбранный период

# На рисунке 3 представлена официальный курс белорусского рубля по отношению к евро, устанавливаемый Национальным банком Республики Беларусь за месяц.

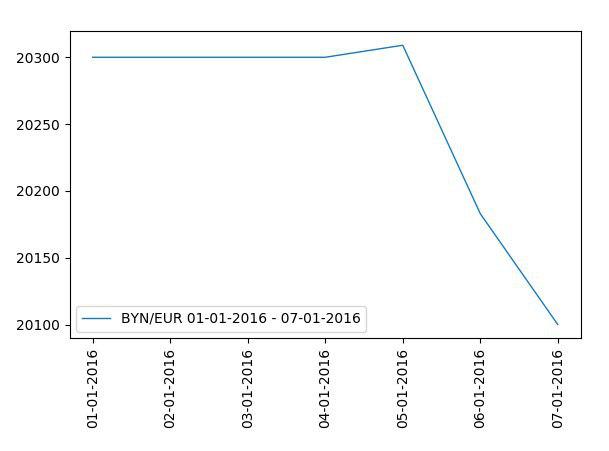


Рисунок 3 - Курс белорусского рубля по отношению к евро за пол года

# На рисунке 4 представлена официальный курс белорусского рубля по отношению к евро, устанавливаемый Национальным банком Республики Беларусь за месяц.

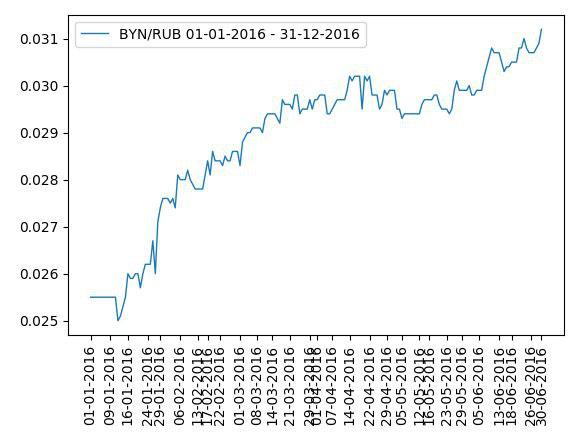


Рисунок 4 - Курс белорусского рубля по отношению к российскому

рублю за год

Рассмотрим первичный анализ данных на примере данных курса доллара к белорусскому рублю в период с 1 января 2019 по 31 декабря 2019.

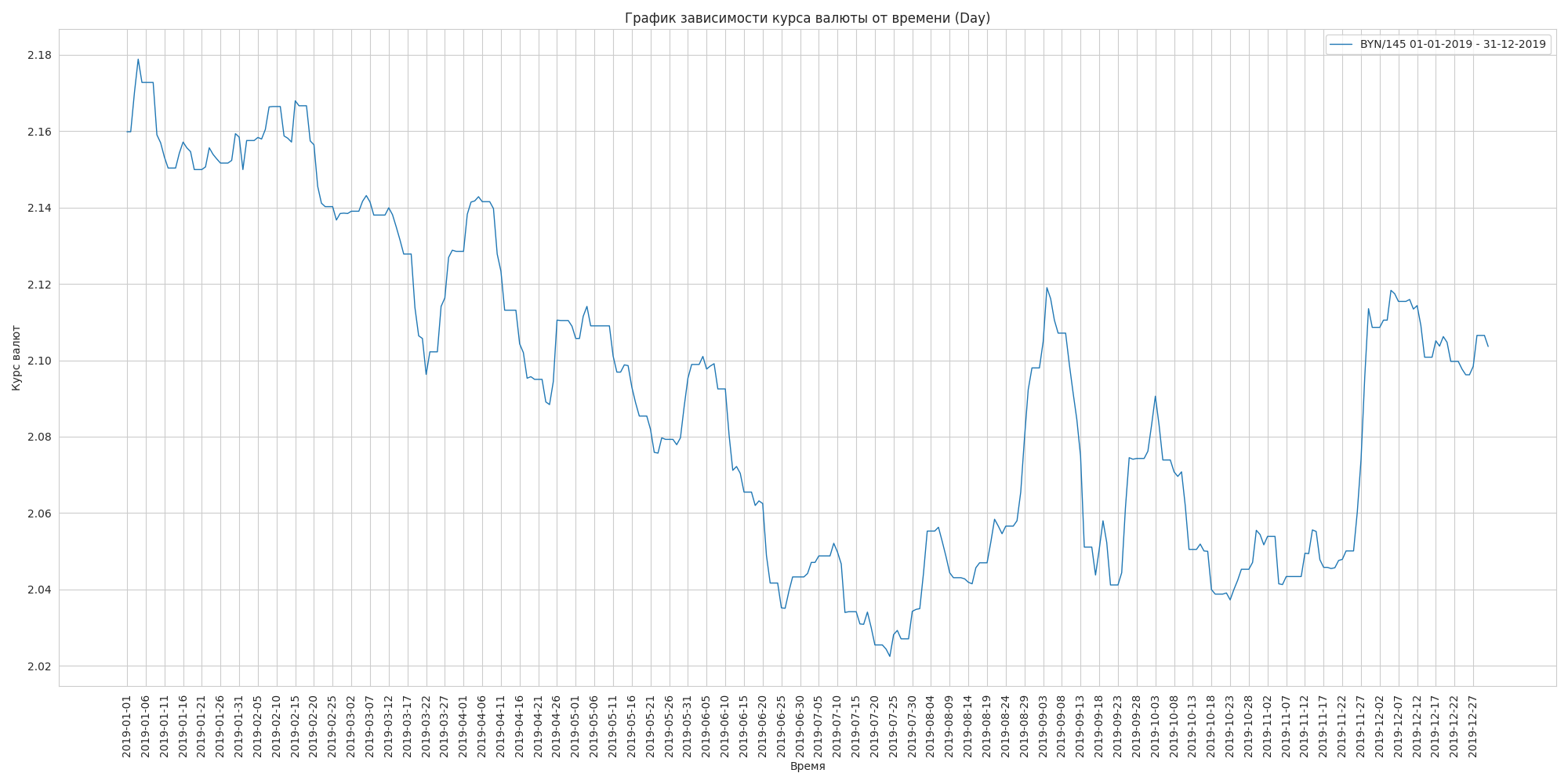


Рисунок 5 – График зависимости курса

валюты от времени за один день

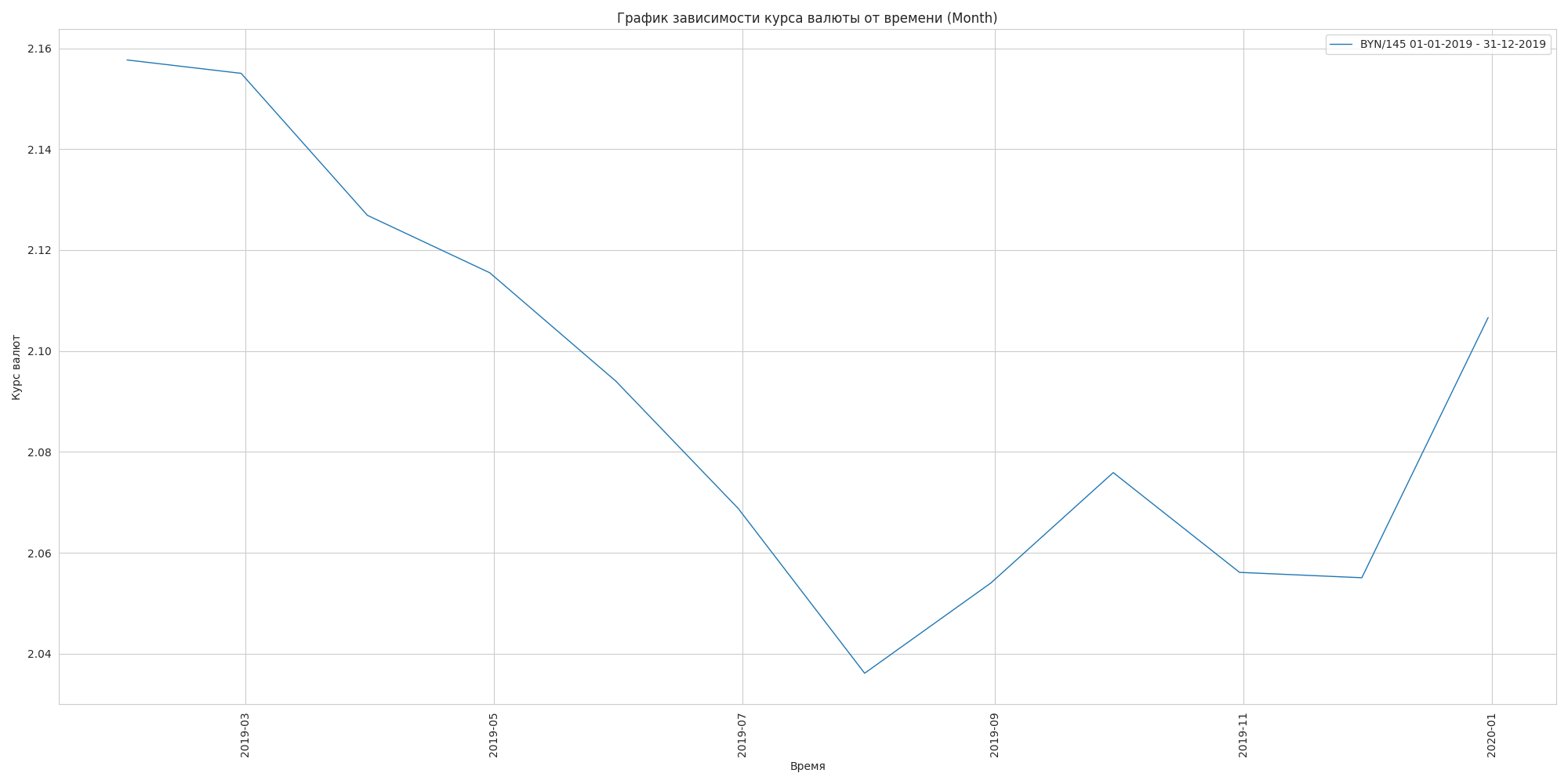


Рисунок 6 – График зависимости курса

валюты от времени помесячно

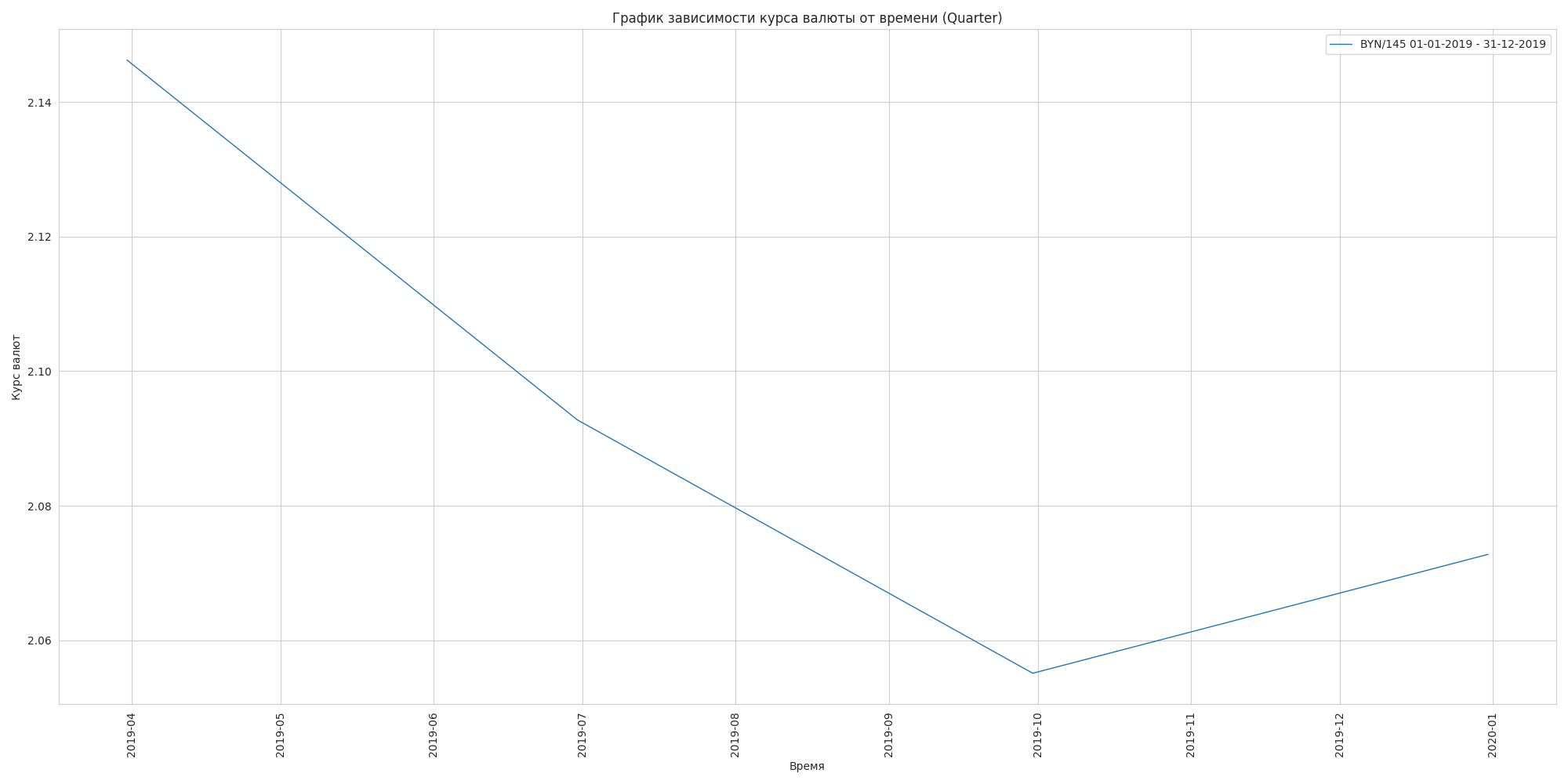


Рисунок 7 – График зависимости курса

валюты от времени поквартально

Декомпозируем временной ряд на составляющие: тренд, сезонность и остаток.

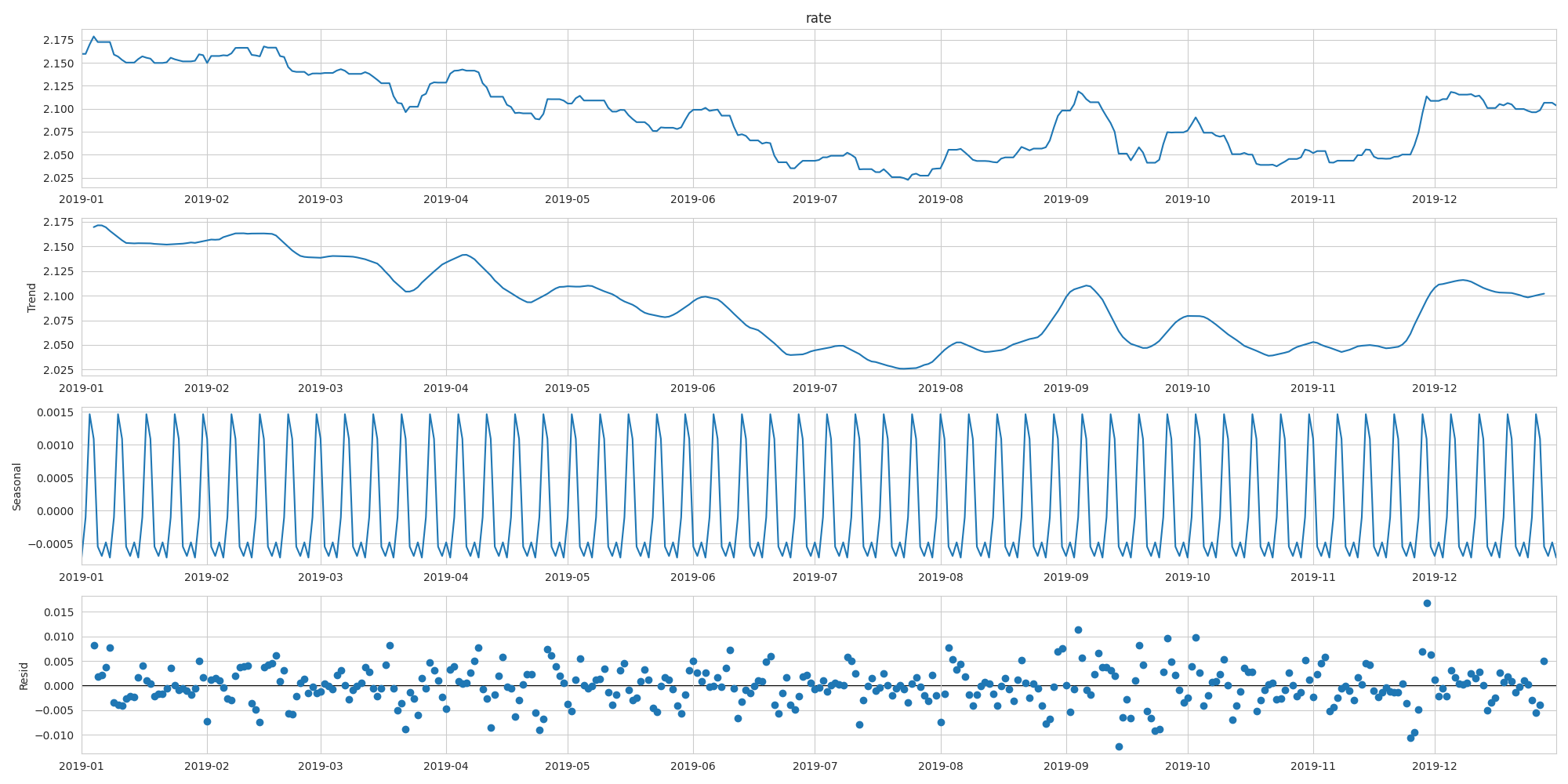


Рисунок 8 – Декомпозиция временного ряда

Найдем изменение курса валюты за предыдущий день.

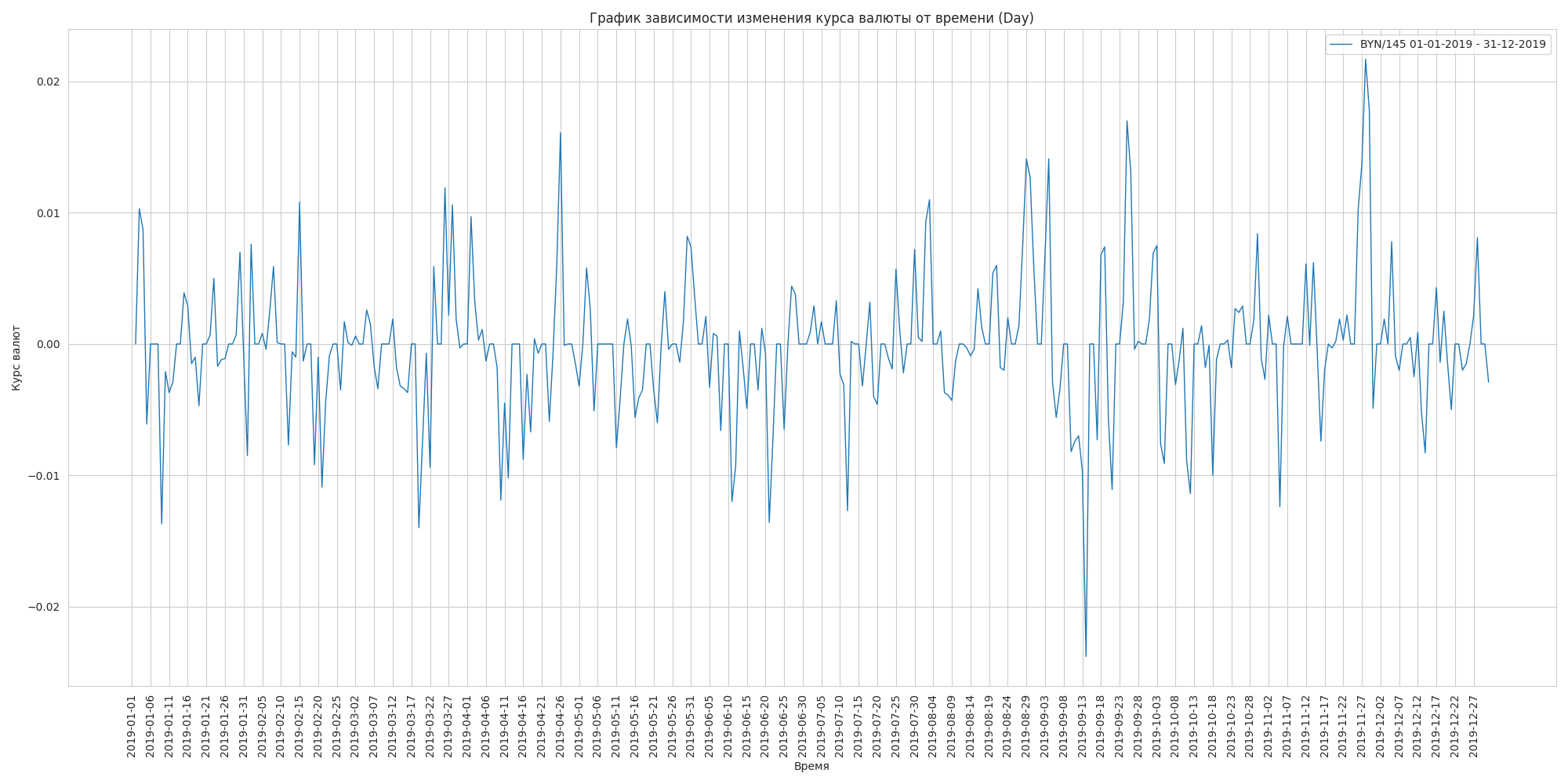


Рисунок 9 – Изменение курса валюты за предыдущий день

Найдем среднеквадратическое отклонение курса валюты за каждые 5 дней.

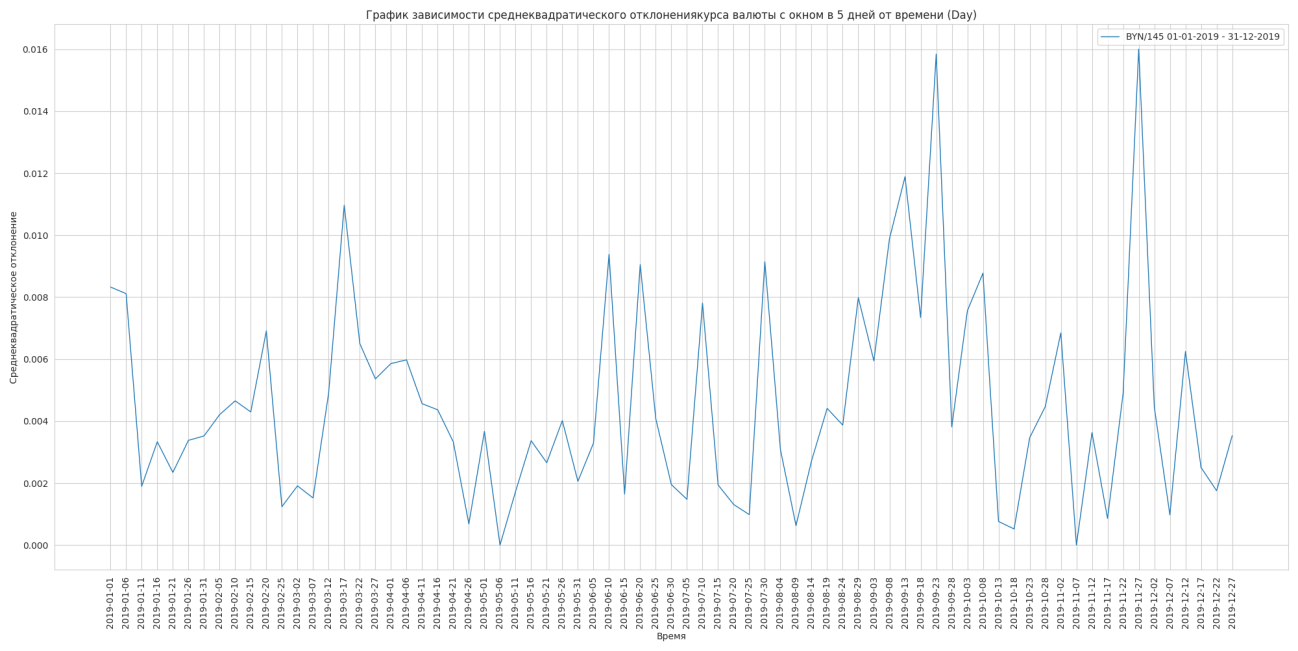


Рисунок 10 – среднеквадратическое отклонение курса

валюты за каждые 5 дней

По приведенным выше графикам видно, что в период с 9 по 28 сентября, а также с 22 по 28 ноября были наибольшие выбросы. По рисунку 9 определяем, что на первом интервале было наибольшие снижение курса, а на втором – наибольшие возрастание.

Построим гистограмму исходного временного ряда.

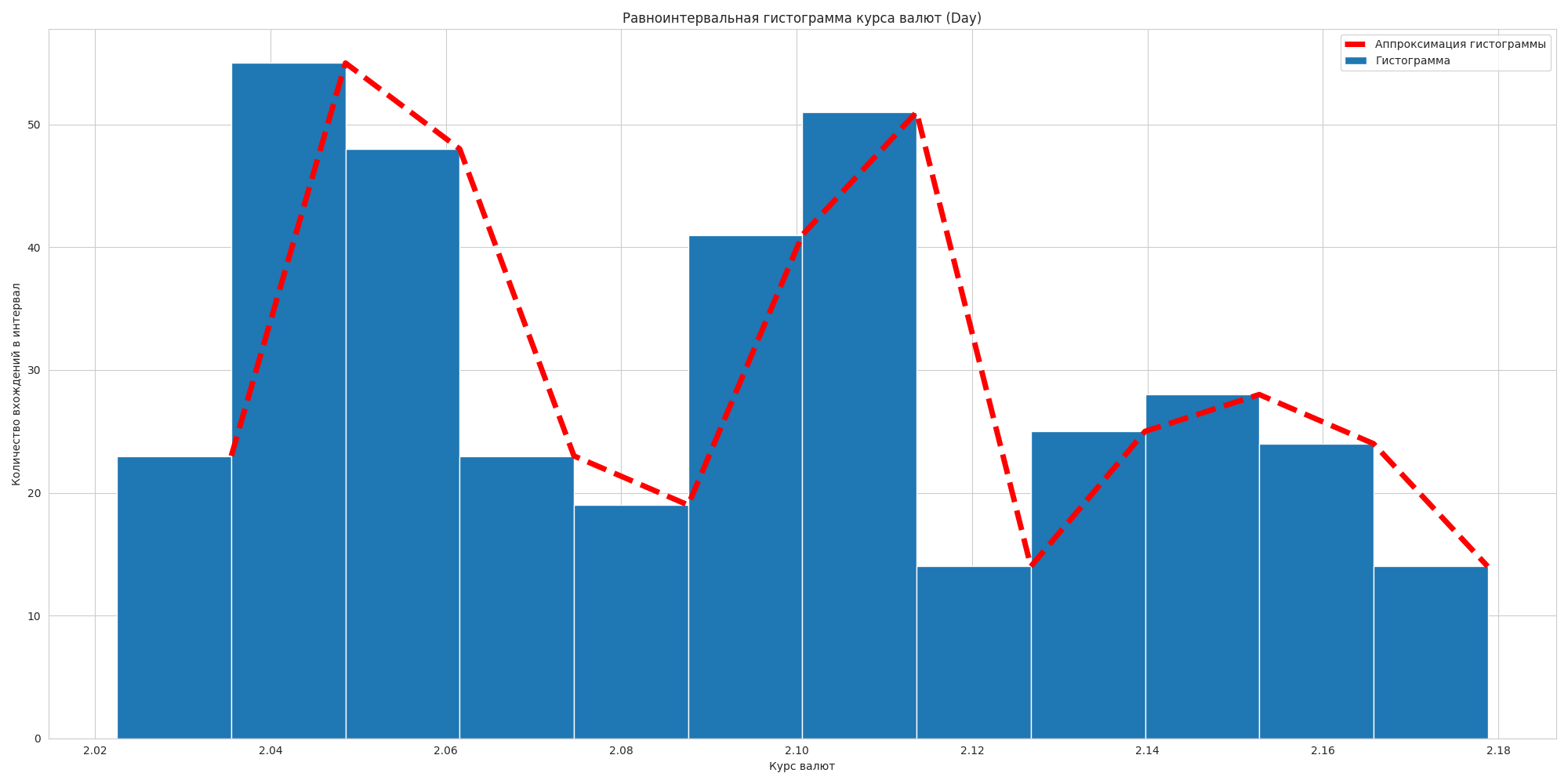


Рисунок 11 – Равноинтервальная гистограмма курса валюты

Из рисунка 11 видно, что существуют 3 более-менее различимые гаусианы, это значит, что на протяжении рассматриваемого периода это были наиболее характерные зоны курса валюты.

**Вывод**

В ходе данной лабораторной работы была разработана программа для обработки данных курсов валют Национального банка Республики Беларусь. Для этого использовался язык программирования Python. Официальный курс брался с сайта <http://www.nbrb.by/>. Сайт отдавал нам API.

Были проанализированы и построены графики по следующим пунктам:

* Курс белорусского рубля по отношению к иностранной валюте за год;
* Курс белорусского рубля по отношению к иностранной валюте за месяц;
* Курс белорусского рубля по отношению к иностранной валюте за заданный период;
* Использовалась следующая валюта: доллары, евро, российский рубль.