

Документация к блоку анализа и предобработки данных

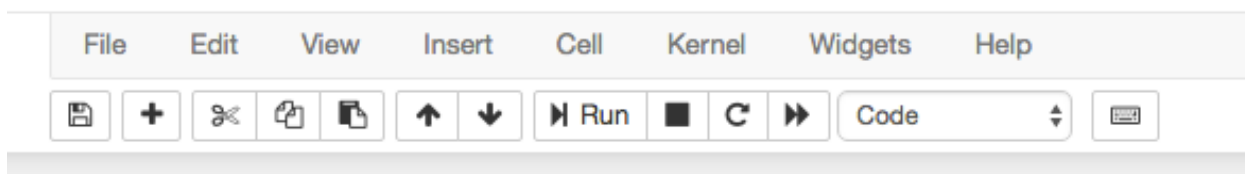
1. Для работы необходимо установить БД PostgreSQL и Jupyter Notebook

<https://jupyter.org>

<https://www.postgresql.org>

Следуйте инструкциям на сайте.

Запуск блока в Jupyter Notebook осуществляется путем клика на блок и нажатием на кнопку RUN.



2. Для загрузки исходных excel таблиц в БД, необходимо сохранить их в формате XLSX или CSV и загрузить их в БД одним из трех способов:

1) Загрузить excel данные в SQL с помощью питона:

```
!pip install psycopg2 #установка библиотеки

import psycopg2 #подключение библиотеки

from sqlalchemy import create_engine

engine = create_engine('postgresql://username:password@localhost:5432/
dmitry') #создание соединения

#создание соединения

conn = psycopg2.connect("dbname = 'dmitry' user = 'dmitry' ")

bond_description = pd.read_excel('bond_description.xlsx', sheet_name =
'bond_discription') #загрузка данных в питон

#импорт данных в таблицу в БД

bond_description.to_sql("bond_description", engine)
```

2) Загрузить данные в БД (PostgreSQL) из интерфейса БД с помощью следующего запроса:

```
COPY table_name FROM '/path_to_csv_file.csv' DELIMITERS ',' CSV;
```

3) Заранее подготовить таблицу и загрузить из интерфейса БД вручную пред обработанный CSV файл нажав кнопку IMPORT.

3. Блок программы «Описательная статистика» выводит информацию по загруженной базе данных облигаций, такую как период БД, количество эмитентов и т.д. Для запуска блока вмешательство юзера не требуется, только нажать RUN.

Пример исполнения:

ОПИСАТЕЛЬНАЯ СТАТИСТИКА БАЗЫ ДАННЫХ:

| Начало_периода_размещений | Конец_периода_размещений | Длина_периода_лет | Количество_заемщиков | Количество_выпусков |
|---------------------------|--------------------------|-------------------|----------------------|---------------------|
| 1998-12-23 | 2015-11-05 | 16 | 952 | 2941 |

ОПИСАТЕЛЬНАЯ СТАТИСТИКА БАЗЫ ДАННЫХ:

| Максимальный_объем_эмиссии_млрд | Минимальный_объем_эмиссии_млн | Средний_объем_эмиссии_млрд | Краткосрочных_выпусков |
|---------------------------------|-------------------------------|----------------------------|------------------------|
| 90.0 | 2.24 | 5.86115 | 1262 |

| Среднесрочных_выпусков | Долгосрочных_выпусков | Средний_срок_обращения_дней |
|------------------------|-----------------------|-----------------------------|
| 1023 | 633 | 2304.56141 |

4. Следующий блок выводит список всех эмитентов и количество их выпусков для выбора и дальнейшего анализа, для запуска блока вмешательство юзера не требуется, только нажать RUN.

Пример вывода:

```
№ 158 ) ГАЗ-финанс [ Выпусков: 1 ]
№ 159 ) Пульсар-1 ИА [ Выпусков: 1 ]
№ 160 ) РОСНАНО [ Выпусков: 9 ]
№ 161 ) ИА Уралсиб 03 [ Выпусков: 2 ]
№ 162 ) Медведь-Финанс [ Выпусков: 3 ]
№ 163 ) ИА ВТБ24 — 1 [ Выпусков: 2 ]
№ 164 ) Роснефть НК [ Выпусков: 33 ]
№ 165 ) Расини [ Выпусков: 1 ]
№ 166 ) РТК [ Выпусков: 4 ]
№ 167 ) АИЖК 2014-3 [ Выпусков: 2 ]
№ 168 ) МОЭК [ Выпусков: 1 ]
№ 169 ) ВКМ-Финанс [ Выпусков: 1 ]
№ 170 ) СМАРТС [ Выпусков: 1 ]
```

5. Блок «Описательная статистика по заданному эмитенту и моменту времени» показывает какие выпуски находятся в обращении и их характеристики на момент времени и по названию эмитента которые указал юзер. Для запуска блока нужно на него кликнуть и нажать RUN. Сначала вводится дата, на которую юзер хочет получить информацию, например:

'2019-01-01'

ВВЕДИТЕ ДАТУ В ОДИНАРНЫХ КАВЫЧКАХ НА МОМЕНТ КОТОРОЙ БУДЕТ ОСУЩЕСТВЛЕН ПОИСК

После нажатия Enter дата будет принята и появится поле для ввода названия эмитента:

'2019-01-01'

СКОПИРУЙТЕ НАЗВАНИЕ ЭМИТЕНТА И ВСТАВЬТЕ В ОДИНАРНЫХ КАВЫЧКАХ (нажмите ENTER) :

'РОСБАНК'

После нажатия на Enter программа выдаст описательную статистику выпусков по входным данным:

Описательная статистика по облигациям в обращении эмитента 'РОСБАНК' на дату '2019-01-01'

| | ИСИН | Статус | Тип облигации | Вид облигации | Тип купона | Наличие обеспечения | Валюта выпуска | Дней до погашения | Наличие |
|---|--------------|-------------|---------------|---------------|---------------|---------------------|----------------|-------------------|---------|
| 0 | RU000A0JUGH5 | В обращении | Корп | Биржевые | Прочий | Необеспеченные | Нет данных | 1882 | |
| 1 | RU000A0JUL67 | В обращении | Корп | Биржевые | Фиксированный | Необеспеченные | Нет данных | 1946 | |
| 2 | RU000A0JUL91 | В обращении | Корп | Биржевые | Фиксированный | Необеспеченные | Нет данных | 1946 | |
| 3 | RU000A0JULA0 | В обращении | Корп | Биржевые | Фиксированный | Необеспеченные | Нет данных | 1946 | |
| 4 | RU000A0JULC6 | В обращении | Корп | Биржевые | Фиксированный | Необеспеченные | Нет данных | 1946 | |

| Наличие оферты | Амортизация | Конвертируемость | Наличие дефолта |
|----------------|-------------|------------------|-----------------|
| Есть | Нет | Нет | Нет |
| Нет | Нет | Нет | Нет |
| Нет | Нет | Нет | Нет |
| Нет | Нет | Нет | Нет |
| Нет | Нет | Нет | Нет |

Если пользователь укажет название эмитента, у которого в анализируемой БД нет выпусков в обращении программа покажет предупреждение:

'2000-11-02'

СКОПИРУЙТЕ НАЗВАНИЕ ЭМИТЕНТА И ВСТАВЬТЕ В ОДИНАРНЫХ КАВЫЧКАХ (нажмите ENTER):

'ВсероссБРР'

ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!
ОШИБКА! У ЭМИТЕНТА НЕТ ВЫПУСКОВ В ОБРАЩЕНИИ! ВЫБЕРИТЕ ДРУГОЙ ЭМИТЕНТ!

Если пользователь укажет дату, на которую все выпуски эмитента уже погашены, программа покажет другое предупреждение:

'2035-11-02'

СКОПИРУЙТЕ НАЗВАНИЕ ЭМИТЕНТА И ВСТАВЬТЕ В ОДИНАРНЫХ КАВЫЧКАХ (нажмите ENTER):

'Банк ВТБ'

ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!
ОШИБКА! ВСЕ ВЫПУСКИ ДО ЭТОГО ГОДА УЖЕ ПОГАШЕННЫ! ВВЕДИТЕ ДАТУ РАНЕЕ!

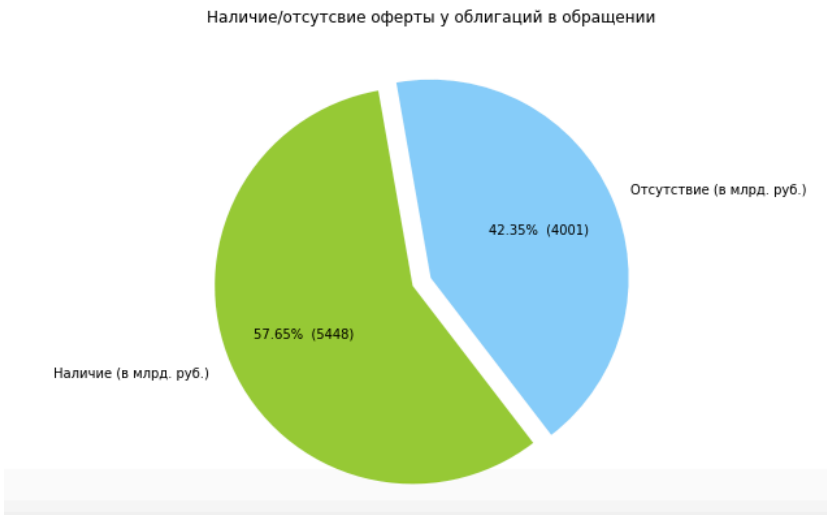
5.5) Перед блоком рисования PIE CHART есть несколько вспомогательных блоков с кодом, которые не взаимодействуют с юзером, но должны быть запущены нажатием на кнопку RUN.

6. Блок PIE CHART рисует две круговые диаграммы (*диаграммы рисуются по всей БД, так как рисовать их по нескольким выпускам одного эмитента не целесообразно*) по заданному показателю облигации (по количеству выпусков в обращении и по объему в обращении) – всего предложено 9 показателей согласно заданию.

Для запуска блока юзеру необходимо выделить его и нажать RUN, затем ввести соответствующую цифру в открывшееся поле и нажать Enter:

Для получения PIE CHART облигаций в обращении по наличию oferty – введите 1, амортизации – введите 2, конвертируемости – введите 3, де фолтам – введите 4, по видам купонов – введите 5, по типам облигаций – введите 6, по видам облигаций – введите 7, по видам гарантии у облигаций – введите 8, по виду срочности облигаций – введите 9:

Пример графиков при вводе «1»:



Пример графиков при вводе «5»:





При вводе значения отличного от 1 до 9, программа выдаст ошибку:

```

10
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!
ВВЕДЕН НЕВЕРНЫЙ НОМЕР! ПОВТОРИТЕ ВВОД!

```

7. Последний блок сохраняет полученные в результаты предыдущих блоков.

Для запуска блока нужно его выделить и нажать RUN.

Описательные статистики сохраняются в формате excel таблицы, рисунки графиков в формате PNG. У юзера есть возможность выбрать какую именно информацию сохранить. При запуске блока появляется сообщение:

Для сохранения описательной статистики по БД введите "1"

Для сохранения описательной статистики по эмитенту и заданному времени введите "2"

Для сохранения графиков PIE CHART введите "3"

В зависимости от введенной цифры, соответствующие файлы будут сохранены на рабочий стол (при условии, что у юзера MacBook и его зовут

dmitry, в противном случае путь до места сохранения нужно поменять на свой).

Например, ввод «2» и Enter:

2

Описательная статистика по эмитенту сохранена на рабочий стол!

Сохраненный файл имеет название и дату аналогично тем, которые вводил юзер:

 stat_emitent_'РОСБАНК"2019-01-01'_.xlsx

Вид сохраненной таблицы:

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|---|--------------------------|--------|---------------|---------------|-------------|------------------|---------------|--------------|------------------|------------|--------------|--------------|---|
| 1 | | ИСИН | Статус | Тип облигации | Вид облигации | Тип купона | Иные обеспечения | Сумма выпуска | До погашения | Иные особенности | Мортизация | Верифицируем | Иные дефолты | |
| 2 | 0 | RU000A0JUGH! В обращении | Корп | Биржевые | Прочий | Необеспечен | Нет данных | 1882 | Есть | Нет | Нет | Нет | Нет | |
| 3 | 1 | RU000A0JUL67 В обращении | Корп | Биржевые | Фиксирован | Необеспечен | Нет данных | 1946 | Нет | Нет | Нет | Нет | Нет | |
| 4 | 2 | RU000A0JUL91 В обращении | Корп | Биржевые | Фиксирован | Необеспечен | Нет данных | 1946 | Нет | Нет | Нет | Нет | Нет | |
| 5 | 3 | RU000A0JULA0 В обращении | Корп | Биржевые | Фиксирован | Необеспечен | Нет данных | 1946 | Нет | Нет | Нет | Нет | Нет | |
| 6 | 4 | RU000A0JULC6 В обращении | Корп | Биржевые | Фиксирован | Необеспечен | Нет данных | 1946 | Нет | Нет | Нет | Нет | Нет | |
| 7 | 5 | RU000A0JUV57 В обращении | Корп | Классические | Переменный | Необеспечен | Нет данных | 2106 | Есть | Нет | Нет | Нет | Нет | |
| 8 | 6 | RU000A0JV9A4 В обращении | Корп | Биржевые | Прочий | Необеспечен | RUB | 2248 | Есть | Нет | Нет | Нет | Нет | |
| 9 | 7 | RU000A0JVS60 В обращении | Корп | Классические | Переменный | Необеспечен | RUB | 2463 | Есть | Нет | Нет | Нет | Нет | |
| 10 | | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | | |

Как видно таблице требуется косметическое форматирование, но порядок ячеек полностью соблюден.