

## Определение доходности облигаций по данным ММВБ

Disclaimer: данное задание подготовлено исключительно для тренировки навыков анализа данных, результаты выполнения задания не являются рекомендациями по операциям с облигациями, автор задания не является инвестиционным консультантом.

Имеются данные по облигациям, которые торгуются на ММВБ на текущую дату.

Требуется выполнить следующие действия с данными по облигациям на дату формирования данных:

1. Оценить рынок рублевых облигаций по распределению купонной доходности, предложить разделение на кластеры по купонной доходности.  
Предложить графическое представление полученных групп/кластеров. (*график*)
2. По *низкорисковым* рублевым облигациям с датой промежуточной выплаты по купонам не более 20% купонного периода назад до текущей даты и облигациям выпущенным в аналогичный период, рассчитать доходность последующих купонов по отобранным облигациям с учетом разницы номинала и цены продажи, а также с учетом налоговых отчислений на период год с текущей даты. (*график и/или таблица*).
3. Допускается обоснованное изменение условий как по критериям оценки рынка, так с учетом выбранной стратегии инвестирования.

Датасет с данными расположен по адресу:

<https://iss.moex.com/iss/apps/infogrid/emission/rates.csv?iss.dp=comma&iss.df=%25d.%25m.%25Y&iss.tf=%25H:%25M:%25S&iss.dtf=%25d.%25m.%25Y%25H:%25M:%25S&iss.only=rates&limit=unlimited&lang=ru>

На усмотрение слушателя:

- можно скачать и загрузить в вашем скрипте файл csv (например, `pd.read_csv` с указанием русской кодировки, релевантного знака сепарирования и номера строки заголовков),
- загружать on-line:

```
s=requests.get(bond_url).content
```

```
pd.read_csv(io.StringIO(s.decode('УКАЗАТЬ КОДИРОВКУ')), sep='СЕПАРАТОР', header=№ СТРОКИ ЗАГОЛОВКОВ)
```

Могут быть использованы следующие поля (а также все остальные) в любых комбинациях:

'SECID' - Код ценной бумаги,

'SHORTNAME' - Краткое наименование,

'NAME' Полное наименование,

'FACEVALUE' - Номинальная стоимость,

FACEUNIT' - валюта,

'DURATION' - дюрация,

'IS\_QUALIFIED\_INVESTORS' - признак продукта для квалифицированных инвесторов,

'HIGH\_RISK' - признак уровня риска,

'COUPONFREQUENCY' – число купонов в год,

'COUPONDATE' - Дата выплаты купона,

'COUPONPERCENT' - Ставка купона, %,

'COUPONVALUE' - Сумма купона, в валюте номинала,

'OFFERDATE' - Дата oferty,  
'EMITENTNAME' – наименование эмитента,  
'INN' – ИНН эмитента,  
'PRICE' – текущий % от FACEVALUE,  
'PRICE\_RUB' - котировка,

Пояснения:

1. Для получения текущей даты можно использовать метод `datetime.date.today()`;
2. Период расчета доходности рекомендуется выбирать 365 дней с текущей даты;
3. Нужно учитывать накопленный купонный доход (НКД) и разницу котировки и номинала ;
4. В общем случае процентный доход при ставке НДФЛ 13% будет  
$$\text{Profit} = (\text{Цена\_Номинальная} - \text{Цена\_покупки} - \text{НКД} + 0.87 * \text{Купонный\_доход\_руб} * X\_лет) * 100\% / (\text{Цена\_покупки} + \text{НКД})$$
  
(комиссией брокера пренебрегаем)
5. Период расчета НКД: `df['COUPONDAYS PASSED']/df['COUPONLENGTH']` , рекомендуется выбирать меньше или равный 0.2.  
Если удовлетворяет условию, полученная величина умножается на `COUPONPERCENT/COUPONFREQUENCY`.
6. Датасет содержит данные, которые нужно преобразовывать в форматы python, а также пустые и искаженные данные, с которыми нужно что-то решить.