

## ЗАДАЧА 11

Условие:

Имеется 3 строки с данными.

rating\_agencies - содержит данные о рейтинговых агентствах. Состав данных:

id рейтингового агентства

наименование рейтингового агентства

rating\_scales - содержит данные о рейтинговых шкалах. Состав данных:

id рейтинговой шкалы

id рейтингового агентства

наименование рейтинговой шкалы

ratings\_data - содержит данные о присвоенных рейтингах ценным бумагам и датах их присвоения. Состав данных:

дата присвоения рейтинга

наименование выпуска ценных бумаг (эмиссии)

isin code выпуска (эмиссии) ценных бумаг (идентификатор)

id рейтингового агентства

id рейтинговой шкалы

буквенное обозначение рейтинга

Подробно ознакомиться с значениями обозначений можно в интернете (например, сайт АКРА <https://www.acra-ratings.ru/about-ratings/scales/>, в нашем примере не используется приставка "(RU)").

Всего имеется 113 уникальных выпусков и их кодов ISIN.

ПРИМЕР: Для 'Выпуск 100' в дату '2021-06-08' агентством 'Рейтинговое агентство 3' присвоен рейтинг на уровне 'BBB+'. А уже в дату '2021-07-13' для этого же выпуска и этим же агентством рейтинг был пересмотрен и установлен на уровне 'A-'. Таким образом, мы знаем, что в период с '2021-06-08' до '2021-07-13' действовал рейтинг 'BBB+' от агентства 3 и по рейтинговой шкале 3, а уже '2021-07-13' и предположительно до сих пор действует рейтинг 'A-'.

От Вас требуется:

подготовить данные для записи в СУБД и представить в структуре следующего вида:

- Идентификатор (уникальный) рейтинга
- Идентификатор выпуска
- Название выпуска
- Код ISIN выпуска
- Название рейтингового агентства
- Название рейтинговой шкалы
- Значение рейтинга, буквенное обозначение
- Дата присвоения рейтинга - дата начала срока жизни записи
- Последняя дата актуальности этого рейтинга - дата окончания срока жизни записи

проанализировать данные и дать ответы на следующие вопросы:

- проанализировать выборку на численное распределение по уровням рейтингов
- проанализировать динамику активности изменений рейтингов во времени
- предложить свои варианты анализа данной выборки

- дать несколько самостоятельных вариантов оценки и анализа выборки
- самостоятельно визуализировать результаты
- предложить варианты нормализации данных

Выбор методов за Вами. Удачи.

Решение.

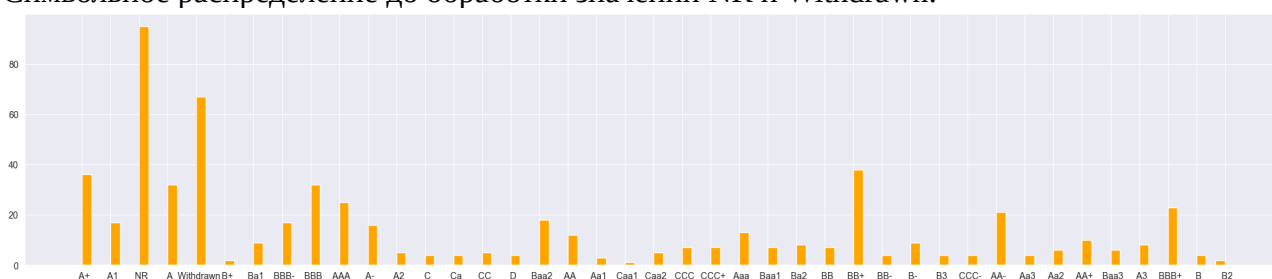
1. Данные сохранены «как есть» в формат csv, далее загружены во фреймы и очищены от «мусора».

Для определения цифрового значения рейтинга принимались данные с сайта АКРА, кроме того было определено:

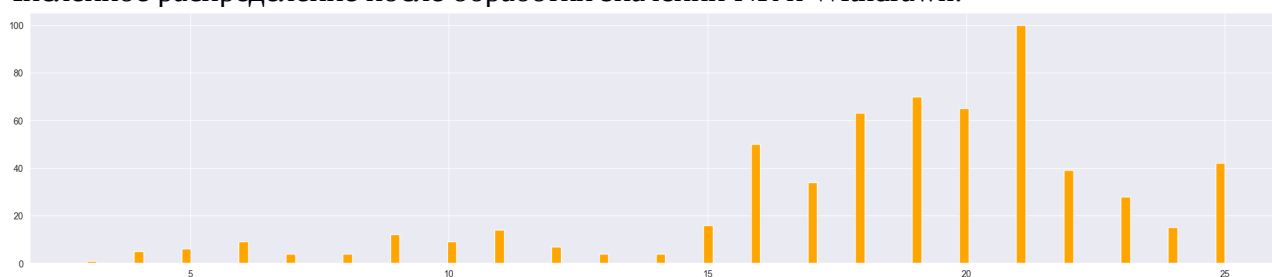
NR – так как рейтинг не присвоен, ринималось предыдущее значение рейтинга;

Withdrawn — понижение предыдущего значения рейтинга на 1 балл.

Символьное распределение до обработки значений NR и Withdrawn:

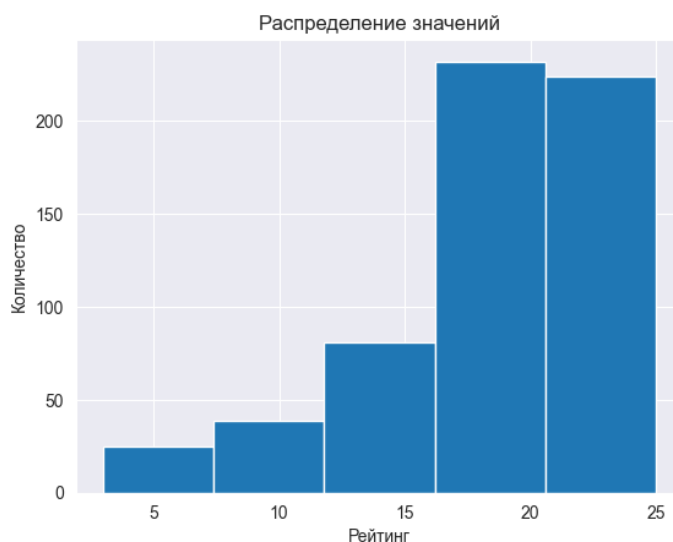


Численное распределение после обработки значений NR и Withdrawn:

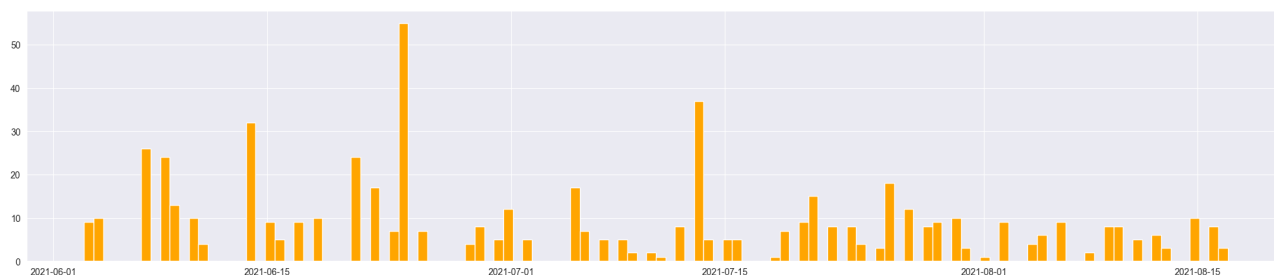


Наиболее часто присваиваемый рейтинг — A или A+(RU) (в числовой интерпретации – 21).

Почти все значения (506 из 601) — от умеренного уровня кредитоспособности (BBB-) и выше (в числовой интерпретации – от 16 и выше):



## 2. Динамика активности изменений рейтингов во времени:



Почти половина изменений рейтингов произошла в июне (288).

В июле произошло 232 изменения.

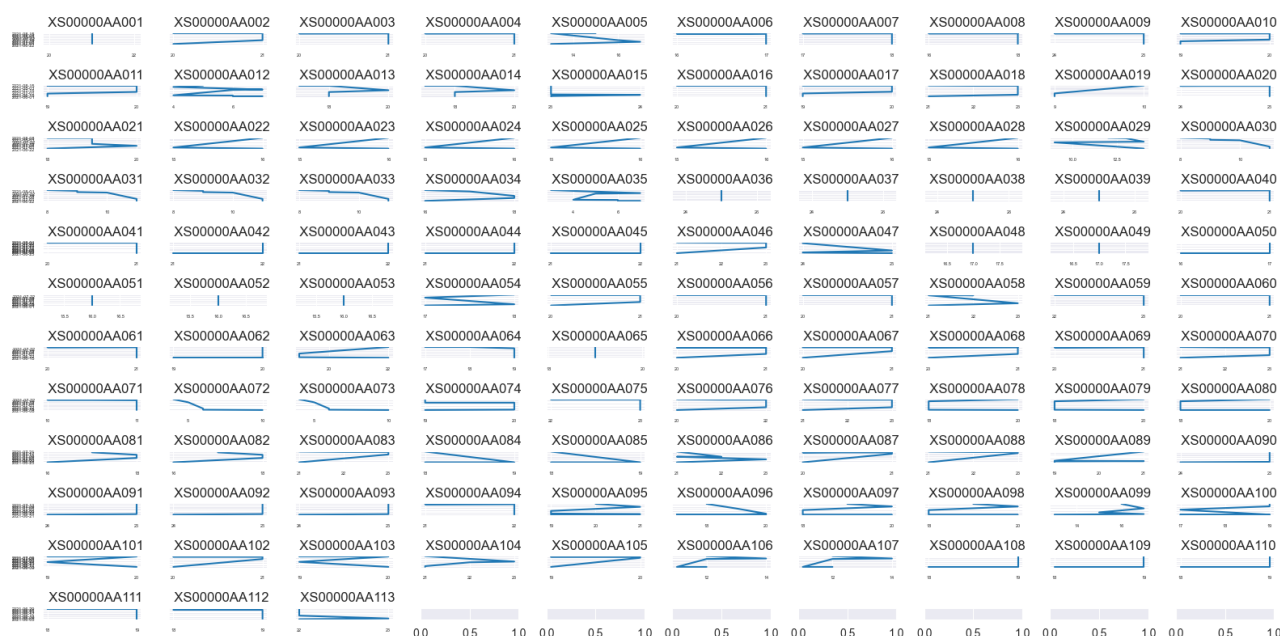
В августе - всего 89 изменений.

Пиковый по активностям день — 24 июня (54 записи).

## 3. Дополнительный анализ:

Для начала отметим, что отсутствуют данные о выпуске №8.

### а) Графическое отображение устойчивости выпусков:



### б) Наиболее устойчивые выпуски:

Выпуск	Движение
1	0
8	0
36	0
37	0
38	0
39	0
48	0
49	0

Выпуск	Движение
51	0
52	0
53	0
65	0
113	0

Т.е. рейтинг выпусков №№ 1, 8, 36, 37, 38, 39, 48, 49, 51, 52, 53, 65, 113 ни разу не изменялся.

в) Самые неустойчивые выпуски:

Выпуск	Движение
5	10
29	10
34	6
58	6
72	10
73	10
86	6
99	6
104	6

Т.е. рейтинг выпусков 34, 58, 86, 99, 104 менялся на 6 баллов, выпусков 5, 29, 72, 73 — на 10 баллов.

г) Также отмечено, что в целом разность первоначальных оценок Международной и Суверенной шкал со временем уменьшается (например, выпуски 14, 18, 19, 21, 30, 70, 76, 77, 81, 82, 86, 87, 95, (99, 106 — исключение из правила), 101, 102, 104, 105).

д) Выпуски с наибольшим изменением рейтинга:

	name	max_rg	max_rg_date	min_rg	min_rg_date
<b>4</b>	XS00000AA005	17	2021-07-05	13	2021-06-24
<b>28</b>	XS00000AA029	14	2021-06-17	9	2021-07-16
<b>71</b>	XS00000AA072	10	2021-06-17	4	2021-07-23
<b>72</b>	XS00000AA073	10	2021-06-17	4	2021-07-23
<b>98</b>	XS00000AA099	17	2021-06-25	13	2021-06-24

Таким образом, выпуски 5 и 99 показали наибольший рост, выпуски 29, 72 и 73 — наибольшее падение.

е) Выпуски с лучшими и худшими рейтингами:

	name	max_rg	max_rg_date	min_rg	min_rg_date
35	XS000000AA036	25	2021-06-21	25	2021-06-21
36	XS000000AA037	25	2021-06-14	25	2021-06-14
37	XS000000AA038	25	2021-06-04	25	2021-06-04
38	XS000000AA039	25	2021-06-04	25	2021-06-04

Т.е. выпуски 36-39 обладают наивысшим рейтингом.

	name	max_rg	max_rg_date	min_rg	min_rg_date
11	XS000000AA012	7	2021-06-18	5	2021-06-23
18	XS000000AA019	10	2021-08-09	9	2021-06-03
34	XS000000AA035	7	2021-06-18	5	2021-06-23
71	XS000000AA072	10	2021-06-17	4	2021-07-23
72	XS000000AA073	10	2021-06-17	4	2021-07-23

Т.е. выпуски 12, 19, 35, 72 и 73 — обладают наихудшими рейтингами.

4) Нормализация данных.

а) Для приведения к первой нормальной форме нужно удалить столбец «Код ISIN выпуска» - т. к. он полностью дублирует значения столбца «Название выпуска» и ввести дополнительное отношение:

Название выпуска	Код ISIN выпуска
Выпуск 1	XS000000AA001
Выпуск 2	XS000000AA002
...	...

Для экономии дискового пространства (и увеличения скорости обработки данных) следует:

а) заменить столбец «Значение рейтинга, буквенное обозначение» - численными соответствиями;

б) заменить столбец «Название рейтингового агентства» - численными соответствиями и ввести дополнительное отношение:

Название рейтингового агентства	Код рейтингового агентства
Рейтинговое агентство 1	1
Рейтинговое агентство 1	2
Рейтинговое агентство 1	3

в) аналогично — для столбца «Название рейтинговой шкалы», заменить значения численными соответствиями и ввести дополнительное отношение.