ЗАДАЧА 11

Условие:

Имеется 3 строки с данными.

rating_agencies - содержит данные о рейтинговых агенствах. 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 AKPA 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 'ВВВ+'. А уже в дату '2021-07-13' для этого же выпуска и этим же агентством рейтинг был пересмотрен и установлен на уровне 'А-'. Таким образом, мы знаем, что в период с '2021-06-08' до '2021-07-13' действовал рейтинг 'ВВВ+' от агентства 3 и по рейтинговой шкале 3, а уже '2021-07-13' и предположительно до сих пор действует рейтинг 'А-'.

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 окнончания 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.

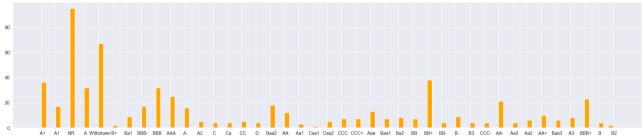
Решение.

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:

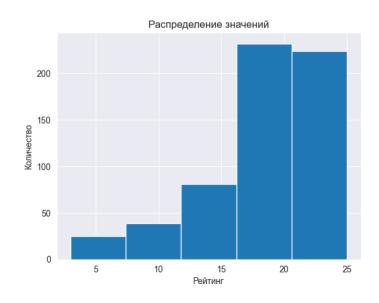


Численное распределение после обработки значений NR и Withdrawn:

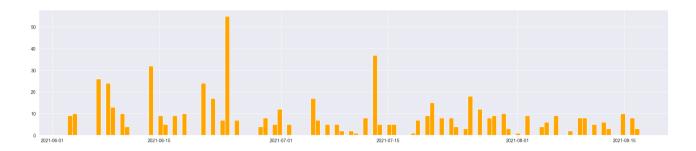


Наиболее часто присваиваемый рейтинг — А или A+(RU) (в числовой интерпретации – 21).

Почти все значения (506 из 601) — от умеренного уровня кредитоспособности (ВВВ-) 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 усточивости выпусков:

翻着着	XS00000AA001	XS00000AA002	XS00000AA003	XS00000AA004	XS00000AA005	XS00000AA006	XS00000AA007	XS00000AA008	XS00000AA009	XS00000AA010
	XS00000AA011	XS00000AA012	XS00000AA013	XS00000AA014	XS00000AA015	XS00000AA016	XS00000AA017	XS00000AA018	XS00000AA019	XS00000AA020
	XS00000AA021	XS00000AA022	XS00000AA023	XS00000AA024	XS00000AA025	XS00000AA026	XS00000AA027	XS00000AA028	XS00000AA029	XS00000AA030
	XS00000AA031	XS00000AA032	XS00000AA033	XS00000AA034	XS00000AA035	XS00000AA036	XS00000AA037	XS00000AA038	XS00000AA039	XS00000AA040
1111	XS00000AA041	XS00000AA042	XS00000AA043	XS00000AA044	XS00000AA045	XS00000AA046	XS00000AA047	XS00000AA048	XS00000AA049	XS00000AA050
MAA	XS00000AA051	XS00000AA052	XS00000AA053	XS00000AA054	XS00000AA055	XS00000AA056	XS00000AA057	XS00000AA058	XS00000AA059	XS00000AA060
翻鎖	XS00000AA061	XS00000AA062	XS00000AA063	XS00000AA064	XS00000AA065	XS00000AA066	XS00000AA067	XS00000AA068	XS00000AA069	XS00000AA070
MAG	XS00000AA071	XS00000AA072	XS00000AA073	XS00000AA074	XS00000AA075	XS00000AA076	XS00000AA077	XS00000AA078	XS00000AA079	XS00000AA080
	XS00000AA081	XS00000AA082	XS00000AA083	XS00000AA084	XS00000AA085	XS00000AA086	XS00000AA087	XS00000AA088	XS00000AA089	XS00000AA090
###	XS00000AA091	XS00000AA092	XS00000AA093	XS00000AA094	XS00000AA095	XS00000AA096	XS00000AA097	XS00000AA098	XS00000AA099	XS00000AA100
	XS00000AA101	XS00000AA102	XS00000AA103	XS00000AA104	XS00000AA105	XS00000AA106	XS00000AA107	XS00000AA108	XS00000AA109	XS00000AA110
##S	XS00000AA111	XS00000AA112	XS00000AA113	0.0 0.5 1.0	0.0 0.5 1.0	0.0 0.5 1.0	0.0 0.5 1.0	0.0 0.5 1.0	0.0 0.5 1.0	0.0 0.5 1.0

б) Наиболее устойчивые выпуски:

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

Выпуск	Движение
-	дынение
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

T.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
4	XS00000AA005	17	2021-07-05	13	2021-06-24
28	XS00000AA029	14	2021-06-17	9	2021-07-16
71	XS00000AA072	10	2021-06-17	4	2021-07-23
72	XS00000AA073	10	2021-06-17	4	2021-07-23
98	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	XS00000AA036	25	2021-06-21	25	2021-06-21
36	XS00000AA037	25	2021-06-14	25	2021-06-14
37	XS00000AA038	25	2021-06-04	25	2021-06-04
38	XS00000AA039	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	XS00000AA012	7	2021-06-18	5	2021-06-23
18	XS00000AA019	10	2021-08-09	9	2021-06-03
34	XS00000AA035	7	2021-06-18	5	2021-06-23
71	XS00000AA072	10	2021-06-17	4	2021-07-23
72	XS00000AA073	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	XS00000AA001		
Выпуск 2	XS00000AA002		

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.