**Задача №1**

Разработка и обоснование оценки эффективности работы коллекторов

**Задание**

На основе имеющихся данных по выработке и результатам сборов коллекторов, разработать нормативы и показатели эффективности работы, предложить схему мотивации и премирования сотрудников. Обосновать свое предложение.

**Входные данные**

1. Учет времени работы коллекторов. Файл - Учет времени коллекторов.xlsx
2. Сборы коллекторов. Файл - Сборы коллекторов.xlsx

Было замечено, что данные не совсем структурированы, присутствуют аномальные значения, буквенные и цифровые символы, пропуски в данных.

1. Во-первых, скорее всего, если данные заносятся непосредственно самими сотрудниками, необходимо отладить этот процесс, потому что без чистки выбросов будут неверные результаты. Возможно, введение некой системы штрафов за ненадлежащее внесение данных с последующей лишении премии конкретным сотрудникам.
2. По часовым и рабочим нормам, видно, что график похож на типичный 7,5-8 часовой рабочий день и, в целом, медианные и средние значения по данному параметру не сильно отличаются. Поэтому, можно отслеживать медианное значение рабочего дня каждого сотрудники и, по окончании месяца, смотреть, меньше ли оно 7:30 минут и если меньше, то разбираться почему.  
   Но рабочее время является косвенным признаком, так как эффективность у каждого сотрудника разная. Кто-то может работать без выходных, но при этом иметь меньшую эффективность.
3. По имеющимся данным можно сказать, что стандартное отклонение сборов от месяца к месяцу достаточно высоко, поэтому премирование имеет смысл рассчитывать в каждый месяц отдельно. Очевидно, требуется премировать самых эффективных и лучших сотрудников, поэтому, можно сравнивать сборы за месяц каждого сотрудника с квантильным значением. Я для примера взял третий квартиль.

**Задача №2**

Вывод риска NPL1 по месяцам в разрезе первичных и повторных клиентов.

**Задание**

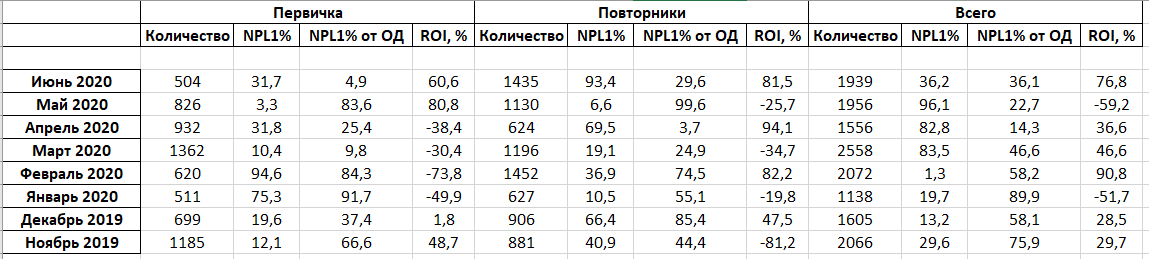
Необходимо по данным о выданных займам вывести долю просроченных займов, доходность клиентов и для наглядности вывести графики. Агрегировать информацию по месяцу выдачи. Вывести данные отдельно по всем займам, первичным клиентам и повторным клиентам.

Для вывода отчета использовать язык программирования Python (R). Написать скрипт, на вход которого подается исходный файл, на выходе отчет в Excel в форме предоставленном в примере.

**Входные данные**

1. Данные о выданных займах. Файл - loan\_date.csv
2. Пример итогового отчета.

**Пример отчета**



***Пояснение***

NPL1 – флаг, клиент ушёл в просрочку на 1 день.

ROI – доходность, (оплаченная сумма – сумма займа) / сумма займа.

NPL1% от ОД – процент суммы займов в просрочке 1 день.

**Задача №3**

**Задача**

Автоматизировать процесс проверки корректности предоставленных лимитов. При подаче заявки на получения кредита на основе данных о клиенте принимается решение о размере предоставляемого лимита.

Важно не только найти ошибки, но и предложить способ автоматизации процесса поиска.

**Входные данные**

Заявки на получение займа с данными о клиенте - define\_limit.csv

**Правила о размере предоставления лимита**

|  |  |
| --- | --- |
| **Правило** | **Лимит** |
| Просрочка в нашем прошлом займе >50 | 0 |
| Просрочка в банке > 15000 или просрочка в МФО > 7000 | 5000 |
| Просрочка в нашем последнем займе > 15 и просрочка МФО >5000 | Лимит из последнего закрытого займа |
| Просрочка в банке < 1000 и просрочка в МФО < 8000 и просрочка в нашем прошлом займе < 20 | Лимит из последнего займа + 10\*оплаченные проценты в прошлом займе |
| Иначе | 2000 |

**Задача №4**

**Задача**

Построить модель, которая будет предсказывать вероятность неплатежа клиента в следующем месяце.

Необходимо описать все выполняемые этапы построения, например, обработка данных, нормализация, отбор предикторов, оценка результатов построения, сравнение различных моделей и тд.

**Входные данные**

Набор данных содержит информацию о дефолтных платежах, демографических факторах, кредитных данных, истории платежей и выписках по счетам клиентов по кредитным картам на Тайване с апреля 2005 г. по сентябрь 2005 г. Наименование файла с данными - UCI\_Credit\_Card.csv

**Описание переменных**

Всего 25 переменных:

1. ID : ID каждого клиента
2. LIMIT\_BAL : сумма предоставленного кредита в тайваньских долларах (включает индивидуальный и семейный / дополнительный кредит)
3. SEX : Пол (1 = мужской, 2 = женский)
4. EDUCATION: (1 = аспирантура, 2 = университет, 3 = средняя школа, 4 = другие, 5 = неизвестно, 6 = неизвестно)
5. MARRIAGE: Семейное положение (1 = женат, 2 = холост, 3 = другие)
6. AGE: Возраст в годах
7. PAY\_0 : Статус погашения в сентябре 2005 г. (-1 = своевременная оплата, 1 = задержка платежа на один месяц, 2 = задержка платежа на два месяца,… 8 = задержка платежа на восемь месяцев, 9 = задержка платежа на девять месяцев и более)
8. PAY\_2 : Статус погашения в августе 2005 г. (шкала такая же, как указано выше)
9. PAY\_3 : Статус погашения в июле 2005 г. (шкала такая же, как указано выше)
10. PAY\_4 : Статус погашения в июне 2005 г. (шкала такая же, как указано выше)
11. PAY\_5 : Статус погашения в мае 2005 г. (шкала такая же, как указано выше)
12. PAY\_6 : Статус погашения в апреле 2005 г. (шкала такая же, как указано выше)
13. BILL\_AMT1 : Сумма выписки по счету за сентябрь 2005 г. (NT-доллар)
14. BILL\_AMT2 : Сумма выписки по счету в августе 2005 г. (NT-доллар)
15. BILL\_AMT3 : Сумма выписки по счету за июль 2005 г. (NT-доллар)
16. BILL\_AMT4 : Сумма выписки по счету в июне 2005 г. (NT-доллар)
17. BILL\_AMT5 : Сумма выписки по счету в мае 2005 г. (NT-доллар)
18. BILL\_AMT6 : Сумма выписки по счету в апреле 2005 г. (NT-доллар)
19. PAY\_AMT1 : Сумма предыдущего платежа в сентябре 2005 г. (NT-доллар)
20. PAY\_AMT2 : Сумма предыдущего платежа в августе 2005 г. (NT-доллар)
21. PAY\_AMT3 : Сумма предыдущего платежа в июле 2005 г. (NT-доллар)
22. PAY\_AMT4 : Сумма предыдущего платежа в июне 2005 г. (NT-доллар)
23. PAY\_AMT5 : Сумма предыдущего платежа в мае 2005 г. (NT-доллар)
24. PAY\_AMT6 : Сумма предыдущего платежа в апреле 2005 г. (NT-доллар)
25. default.payment.next.month : дефолт платежа в следующем месяце (1 = да, 0 = нет)