

Задание 1. Предварительный анализ Таблицы лицевых счетов (ЛС).

Общий вид таблицы ЛС представлен на рис. 1. Таблица содержит 8 столбцов и 49 543 строк с записями.

ИД дома	ИД подъезда	ИД квартиры	Лицевой счет (ЛС)	Услуга	Ставка тарифа	Статус услуги	Дата начала услуги
0	100	1624	43036	ТО МКД	5.2	Активна	2008-01-23
1	100	1624	43037	ТО МКД	5.4	Активна	2008-01-23
2	100	1624	43038	ТО МКД	5.2	Активна	2008-01-23
3	100	1624	43039	ТО МКД	5.2	Активна	2008-01-23
4	100	1624	43040	ТО МКД	5.2	Активна	2008-01-23
...
49538	5219	1884	49529	ТО МКД	5.4	Активна	2019-03-28
49539	5219	1884	49530	ТО МКД	5.2	Активна	2019-03-28
49540	5219	1884	49531	ТО МКД	5.4	Активна	2019-03-28
49541	5219	1884	49532	ТО МКД	5.4	Активна	2019-03-28
49542	5219	1884	49535	ТО МКД	5.2	Активна	2019-03-28

49543 rows × 8 columns

Рисунок 1. Общий вид таблицы ЛС

Таблица ЛС содержит 227 пропущенных значения. Все пропущенные значения приходятся на столбец «Ставка тарифа» и соответствуют четырем типам услуг. Распределение пропущенных значений по типу услуг показано на рисунке 2.

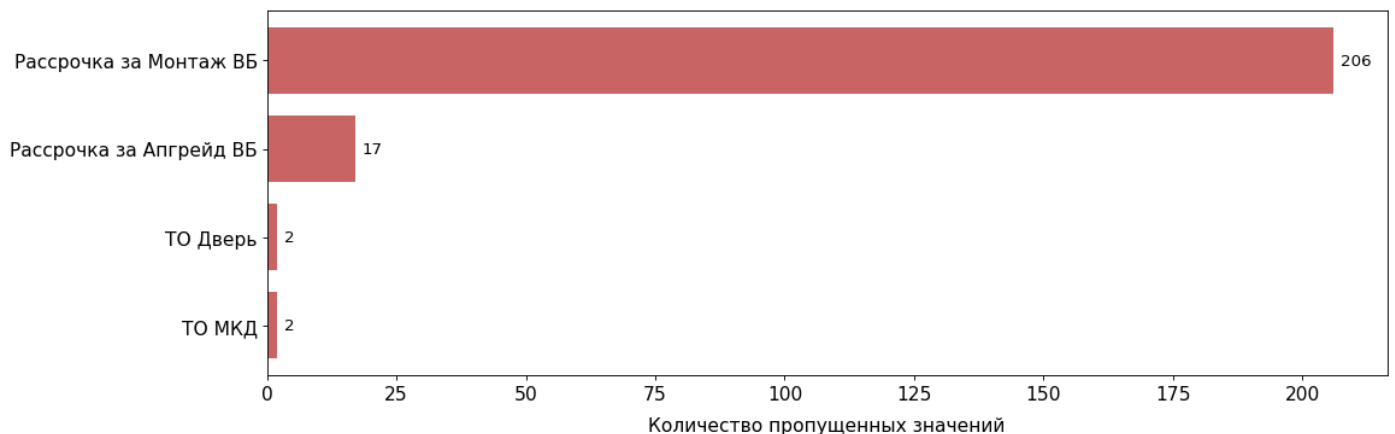


Рисунок 2. Пропущенные значения в таблице ЛС

Как будет показано далее (см. рис 5), количество пропущенных значений по услугам «Рассрочка за монтаж ВБ» и «Рассрочка за апгрейд ВБ» равно общему количеству данных услуг в таблице ЛС. Из этого делаем вывод, что по услугам «Рассрочка за монтаж ВБ» и «Рассрочка за апгрейд ВБ» не предусмотрена ставка тарифа. При дальнейших расчетах, пропущенные значения по этим услугам заменим нулевыми значениями.

ИД дома	ИД подъезда	ИД квартиры	Лицевой счет (ЛС)	Услуга	Ставка тарифа	Статус услуги	Дата начала услуги
3605	632	1242	34945	ТО МКД	NaN	Завершена	2010-01-01
10618	1365	62	2506	ТО Дверь	NaN	Будет активна	2021-11-01
45067	5108	1463	39540	ТО Дверь	NaN	Не предоставляется	2018-11-01
47093	5165	1768	46177	ТО МКД	NaN	Завершена	2010-07-01

Рисунок 3. Пропущенные значения по услугам "ТО МКД" и "ТО Дверь"

Произведем замену пропущенных значений по услугам «ТО МКД» и «ТО Дверь» (рис. 3).

Для лицевых счетов № 48222 и 48399 (услуги «ТО МКД» со статусом «Завершена») в таблице ЛС отсутствуют какие-либо другие услуги. Таким образом, сделать предположение о ставке тарифа для услуг «ТО МКД» не представляется возможным. Для проведения дальнейших расчетов примем значение тарифа равным 5.4, в связи с тем, что это наиболее распространенное значение тарифа для услуги «ТО МКД» (см. далее рис 8).

Для лицевого счета № 39540 (услуга «ТО Дверь» со статусом «Не предоставляется») отсутствуют активные услуги, сделать предположение о ставке тарифа также не представляется возможным. Примем значение тарифа равным 5.0, т.к. это наиболее часто встречающееся значение тарифа для услуги «ТО Дверь».

Для лицевого счета № 2506 (услуга «ТО Дверь» со статусом «Будет активна») в качестве ставки тарифа используем нулевое значение, т.к. отсутствует информация о дате активации тарифа, а также, был ли активен тариф ранее.

На рис. 4 показано количество уникальных элементов в каждом столбце таблицы ЛС. Таблица ЛС содержит 42 531 уникальный лицевой счет, 42 531 квартиру, 6 различных услуг и т.д.

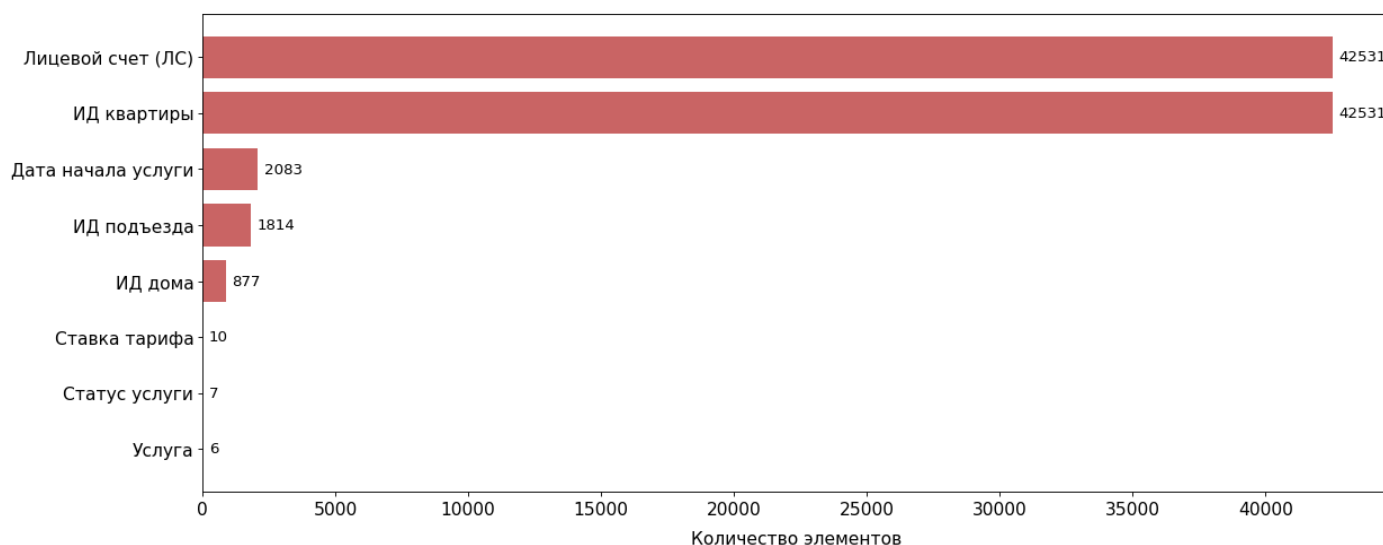


Рисунок 4. Количество уникальных элементов в столбцах таблицы ЛС

Принимая во внимание общее количество записей в таблице, ЛС делаем вывод, что порядка 7 тысяч лицевых счетов имеют более одной услуги. В тоже время, количество ИД квартир совпадает с количеством лицевых счетов, т.е. на одну квартиру приходится один лицевой счет.

На рис. 5 показана количественная оценка услуг, представленных в таблице ЛС. Большая часть услуг приходится на «ТО МКД».

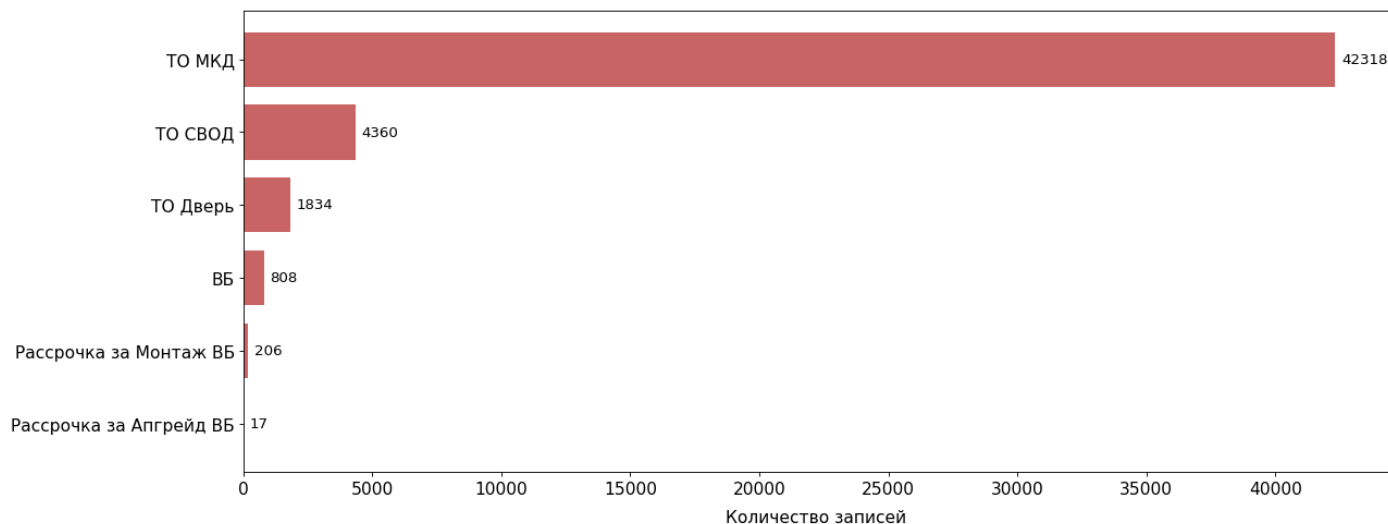


Рисунок 5. Услуги

На рис. 6 показан статус услуг, представленных в таблице ЛС. Большая часть услуг находится в статусе «Активна».

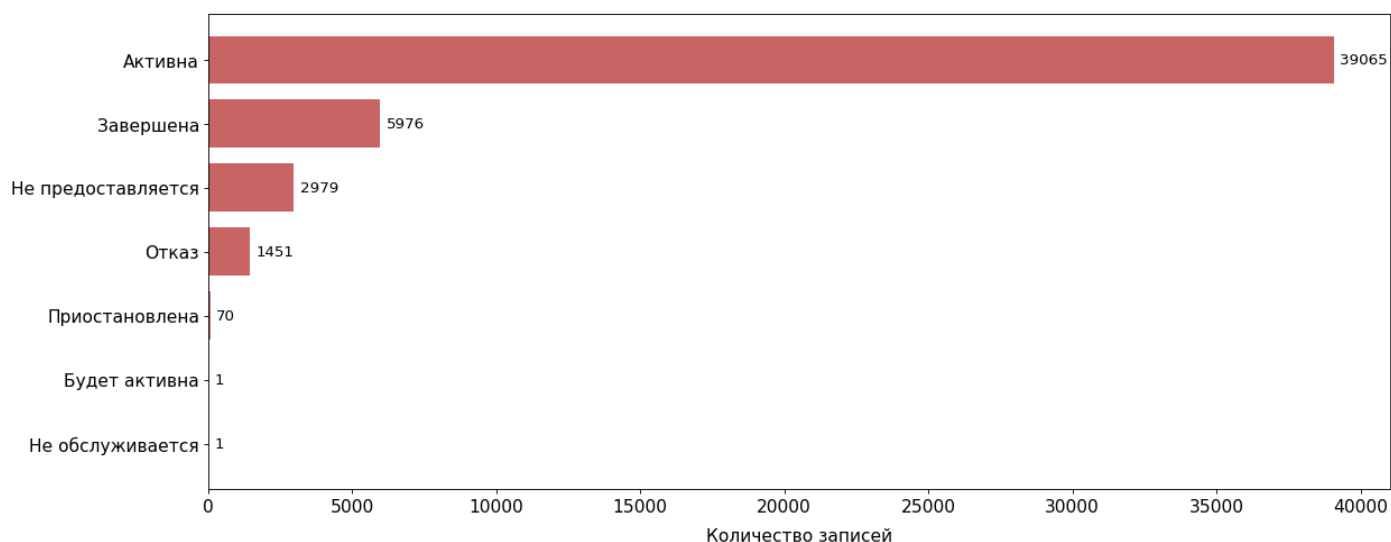


Рисунок 6. Статус услуги

На рис. 7 показано количество предоставленных услуг с распределением по годам. Наибольшее количество предоставленных услуг наблюдалось в 2006, 2007 годах, что обусловлено привлечением новых клиентов (за данный промежуток времени отсутствуют повторяющиеся лицевые счета), после чего последовал спад вплоть до 2015 года. В 2016, 2017, 2018 произошел небольшой всплеск, обусловленный в первую очередь заключением новых услуг по старым лицевым счетам (доля повторяющихся лицевых счетов в сравнении с периодом 2006, 2007 годов составляет 61,4%).

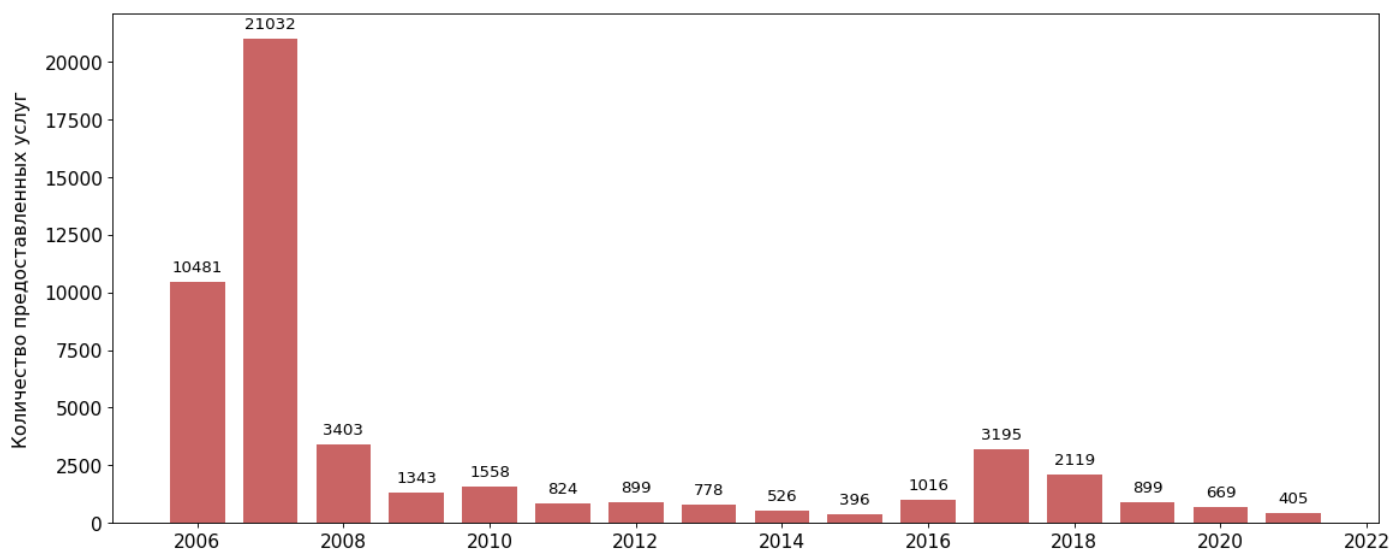


Рисунок 7. Количество предоставленных услуг

На рис. 8 показаны ставки тарифов для услуг «ТО МКД», «ТО СВОД», «ТО Дверь», «ВБ» (первоначальный вариант без учета замен пропущенных значений). Следует обратить внимание на диаграммы «ТО Дверь», «ВБ». Каждая из них содержит по одному значению тарифа (4.3 и 9.7 соответственно), которые выбиваются из общей массы тарифов. Либо в данных случаях имеют место исключительные условия для абонентов, либо допущены ошибки. Услуги «Рассрочка за монтаж ВБ» и «Рассрочка за апгрейд ВБ» не представлены на диаграммах по причине отсутствия для них ставок по тарифам, о чем было сказано ранее.

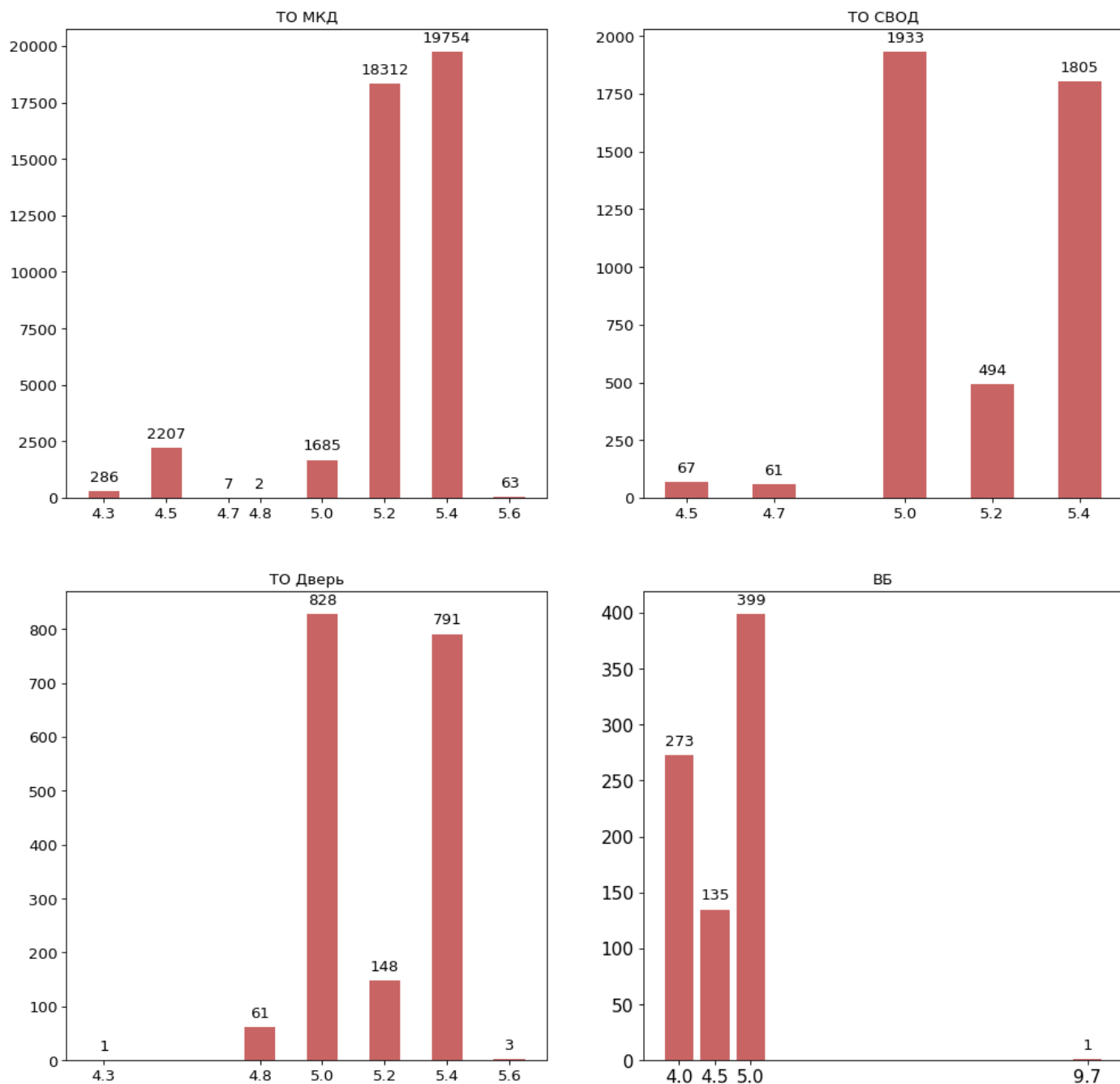


Рисунок 8. Ставки тарифов для разных услуг

Задание 2. Рассчитать для каждого ЛС сумму всех поступивших платежей с учетом возвратов, но исключая списания.

Общий вид таблицы платежей без учета списаний представлен на рис. 9.

	ЛС	Сумма платежа	Дата загрузки	Источник платежа	способ оплаты	Комментарий к платежу
0	1	4.80	2019-04-16	Ручной ввод	Офис	NaN
1	1	14.40	2018-10-17	Ввод из реестра	Сбербанк России	NaN
2	1	14.40	2019-01-21	Ввод из реестра	КС банк	NaN
3	1	32.40	2020-12-11	Ввод из реестра	КС банк	NaN
4	1	56.40	2018-02-09	Ввод из реестра	КС банк	NaN
...
356841	49148	31.20	2021-04-09	Ввод из реестра	Сбербанк России	NaN
356842	49149	31.20	2021-04-27	Ввод из реестра	КС банк	NaN
356843	49150	25.81	2021-06-10	Ручной ввод	Выписки	NaN
356844	49151	20.63	2021-04-16	Ввод из реестра	Сбербанк России	NaN
356845	49151	45.00	2021-07-07	Ручной ввод	Выписки	NaN

352948 rows × 6 columns

Рисунок 9. Таблица платежей

Для таблицы расчетов проведен предварительный анализ, аналогичный представленному в задании № 1 для таблицы лицевых счетов. Следует отметить, что количество лицевых счетов (ЛС) в таблице платежей составляет 41 422, в то время как в таблице ЛС количество счетов составляет 42 531. Кроме того, часть лицевых счетов, имеющих в таблице платежей, отсутствует в таблице ЛС. Общее количество лицевых счетов для обеих таблиц составляет 41 382. Таким образом, в дальнейших расчетах не удастся добиться полного соответствия поступивших платежей планируемой выручке, а расчет сальдо не будет являться в полной мере корректным, в том числе и по причинам, о которых будет сказано ниже.

Для проведения расчетов по заданию № 2 из таблицы исключены платежи из следующих источников: «Авто. списание», «Ручное списание», «Возврат списания». Поступление платежей для каждого ЛС представлено в приложении 1. На рис. 10 показана общая сумма платежей за период с 2018 по 2021 год.

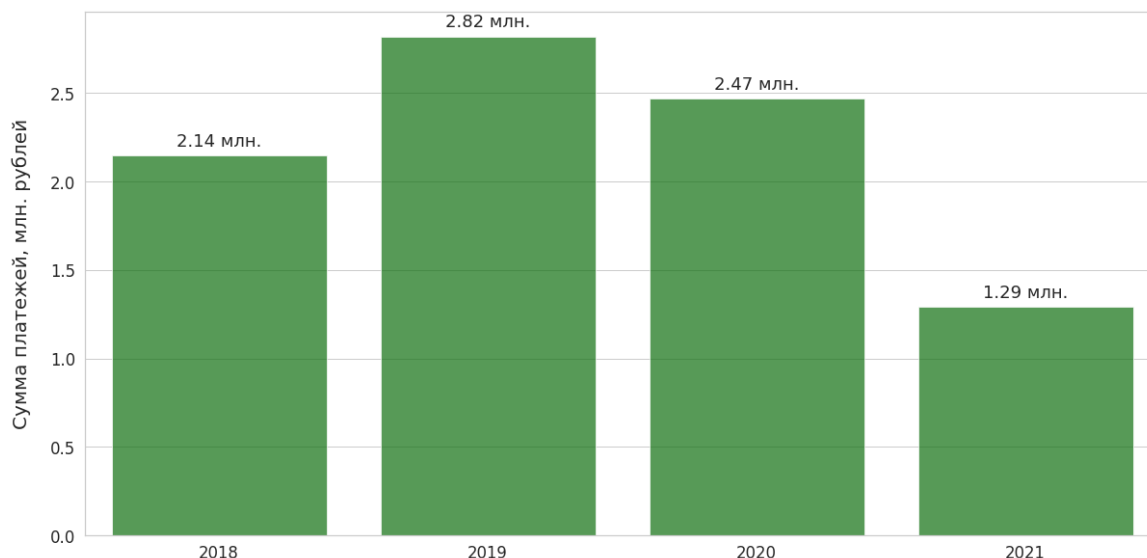


Рисунок 10. Платежи

Задание 3. Рассчитать для каждого ЛС плановую выручку с учетом дат начала услуг на ЛС и ставки тарифа.

Расчет выручки произведен за период с 01.01.2018 по 13.10.2021 (аналогично периоду платежей). Расчет проводился для всех услуг и всех статусов услуг. Следует обратить внимание, что в исходных данных отсутствует информация о дате завершения услуг. То есть не представляется возможным корректно рассчитать значение выручки для лицевых счетов, которые содержат услуги в статусе «Завершена», «Не предоставляется», «Отказ», «Приостановлена», «Не обслуживается». Для лицевых счетов, имеющих активные услуги, можно предположить, что дата окончания завершенных услуг равна дате начала активных услуг, при условии, что услуги одинаковы. Тем не менее, имеется 537 лицевых счетов, имеющих только услуги в статусе «Завершена». Расчет выручки для данных счетов невозможен без предоставления дополнительной информации. Таким образом, в имеющихся условиях, расчет выручки не может являться в полной мере корректным.

Выручка для каждого ЛС представлена в приложении 1. На рис. 11 показана плановая выручка за период с 2018 по 2021 год.

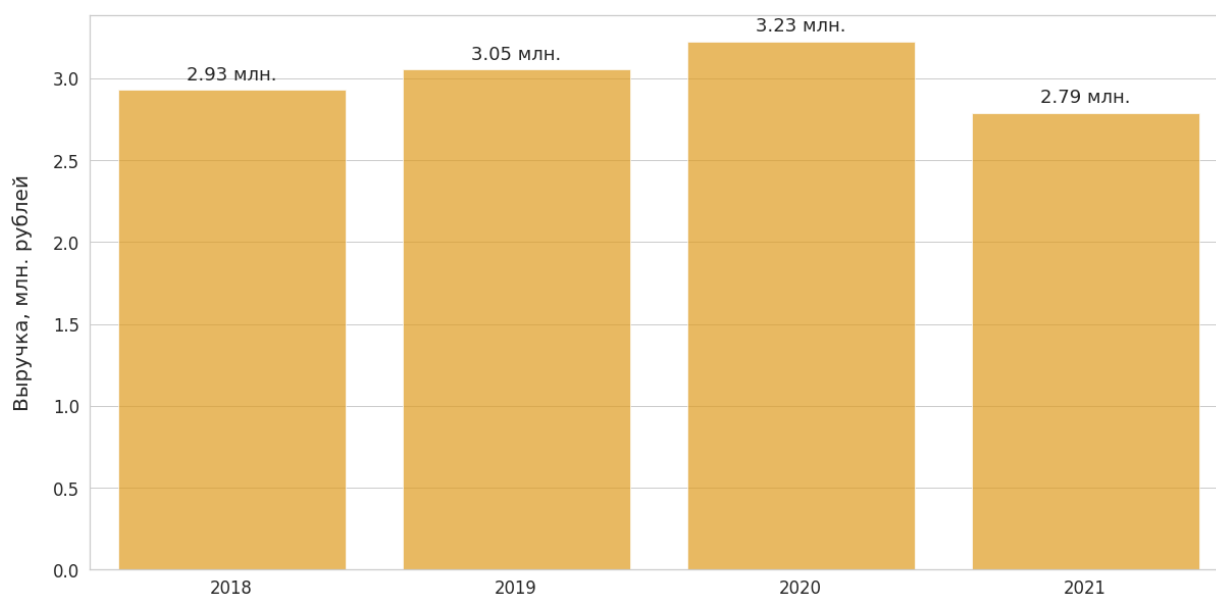


Рисунок 11. Плановая выручка

Задание 4. Рассчитать для каждого ЛС сальдо: плановая выручка (п. 3) минус сумма платежей (из п. 2)

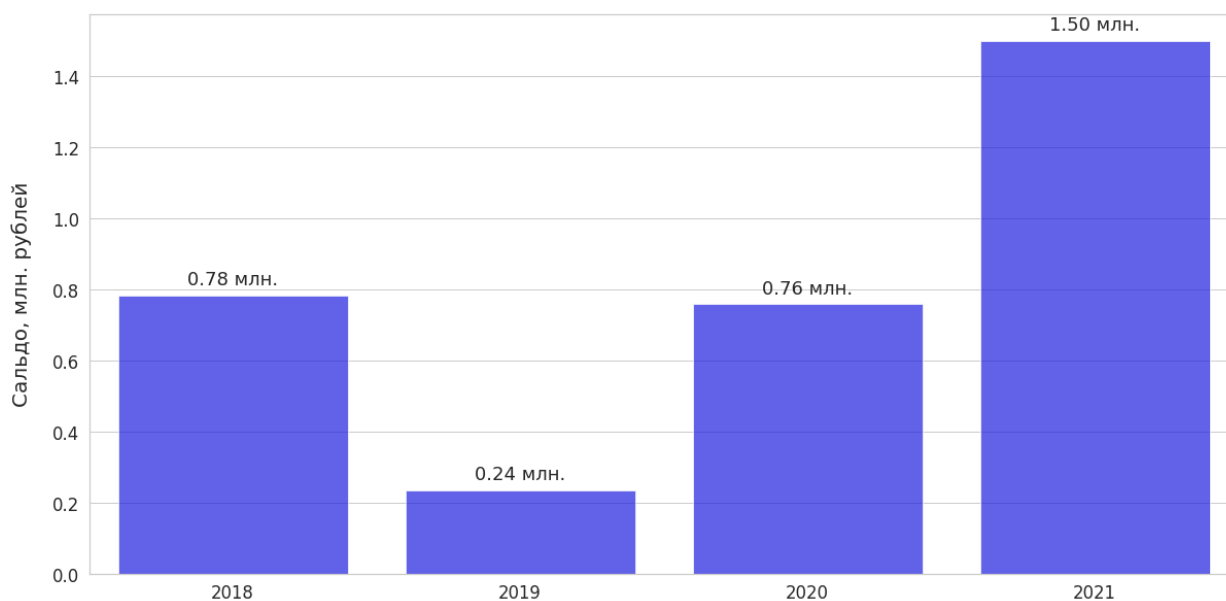


Рисунок 12. Сальдо

Сальдо для каждого ЛС представлена в приложении 1. На рис. 12 показано сальдо за период с 2018 по 2021 год.

Обращаю внимание, что в приложении 1 (application_1.xlsx) имеются незаполненные ячейки. Это связано с тем, что количество лицевых счетов в таблице ЛС и таблице платежей не соответствует друг другу.

Задание 5. Рассчитать для каждой квартиры частоту платежей и количество оплаченных тарифов за календарный год.

Информация представлена в приложении 2 (application_2.xlsx).

Задание 6. Предложите другие метрики для анализа платежной дисциплины абонентов за услуги.

В качестве дополнительных метрик платежной дисциплины можно использовать:

1. Ежемесячный (ежеквартальный) контроль за сроком оплаты. А именно, назначить крайний срок оплаты и фиксировать наличие отклонений.

При наличии достаточного количества данных, возможно создать модель, способную прогнозировать поступление платежей.