

# RFM-сегментация для [the-koleso.ru](http://the-koleso.ru)

## Этапы анализа

Предобработка.

Создание RFM-модели.

## Предобработка

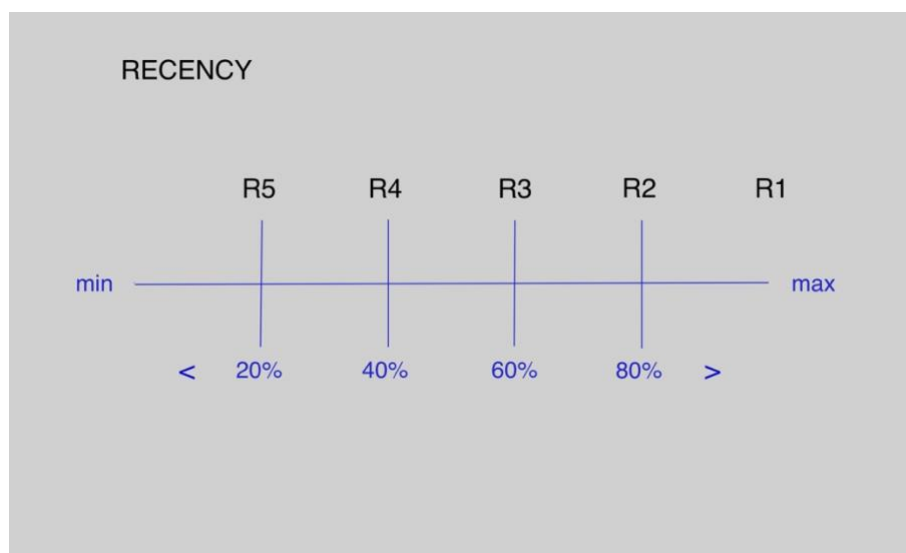
В ходе предобработки были приведены к одному виду номера телефонов (начинаются с +7, длина 12 символов). Номер телефона является уникальным идентификатором пользователя. Строки с номера, которые не прошли стандартизацию были удалены. Также были удалены ненастоящие заказы, дубликаты. Изначальный объем данных был 16477 записей. В ходе предобработки удалено 2753 строки – 16,7%.

## Создание RFM-модели

Временной период рассмотрения данных составил 1790 дней, начало – 20 октября 2018, конец – 14 сентября 2022. Для каждого уникального пользователя (уникальность определялась по номеру телефона) рассчитана последняя дата покупки, частота покупок и сумма, потраченная за все время. Для каждого параметра рассчитаны квантили длиной 20%-40%-60%-80%.

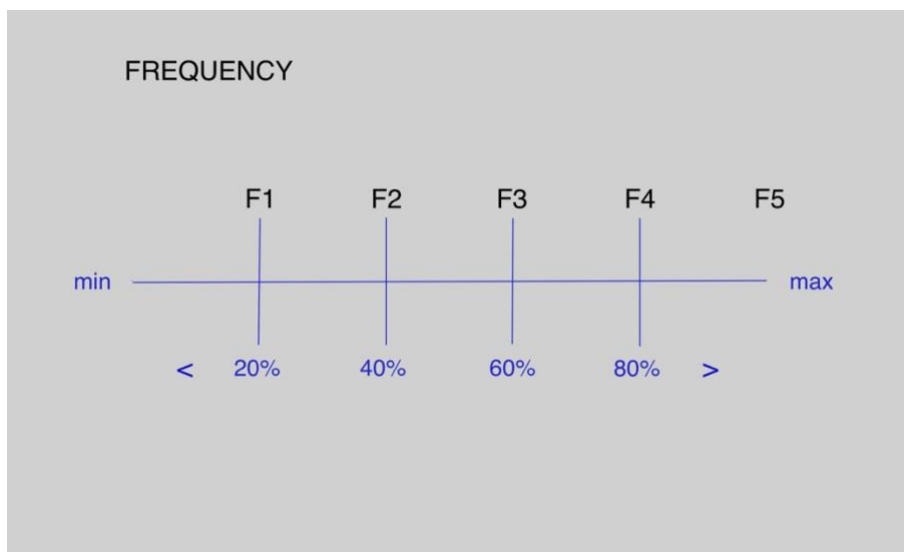
### Recency (давность)

- 20% клиентов в первом квантиле имеют верхнюю границу давности 189 дней, в этом отрезке клиенты, которые покупали совсем недавно по сравнению с остальными - получают балл R5 ,
- 20% клиентов во втором отрезке - давность от 190 до 474 дней - R4,
- 20% клиентов в третьем отрезке - от 475 до 639 дней - R3,
- 20% клиентов в четвертом отрезке - от 640 до 831 дней - R2,
- 20% клиентов в пятом отрезке - от 832 дней и больше - R1.



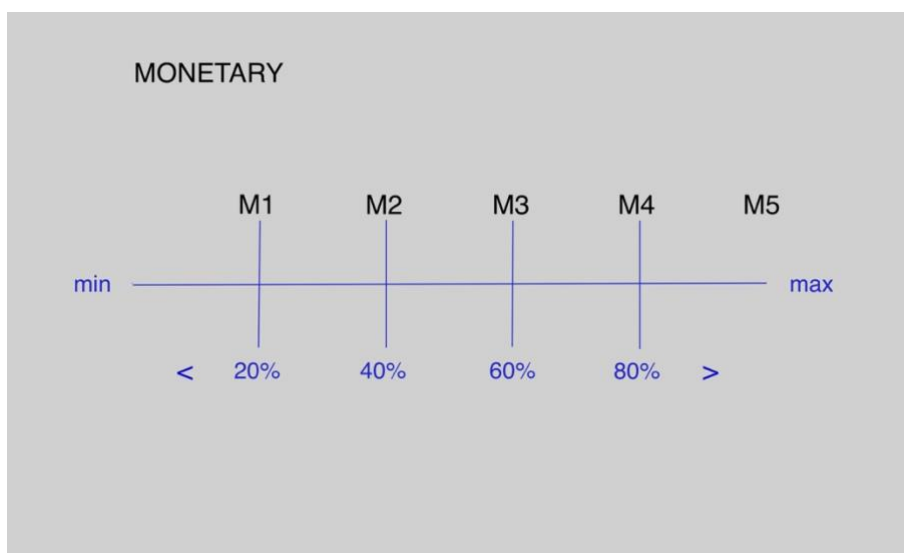
## Frequency (Частота)

- 60% клиентов с первого по третий отрезки - покупали 1 раз - F1,
- 20% в четвертом отрезке - 2 раза - F4,
- 20% в пятом отрезке - больше 2 раз - F5.



## Monetary (Потраченная сумма)

- 20% клиентов в первом отрезке - верхняя граница трат 9780 руб. - M1,
- 20% клиентов во втором отрезке - 18294 - M2,
- 20% клиентов в третьем отрезке - 29503 - M3,
- 20% клиентов в четвертом отрезке - 54400 - M4,
- 20% клиентов в пятом отрезке - более 54400 - M5.



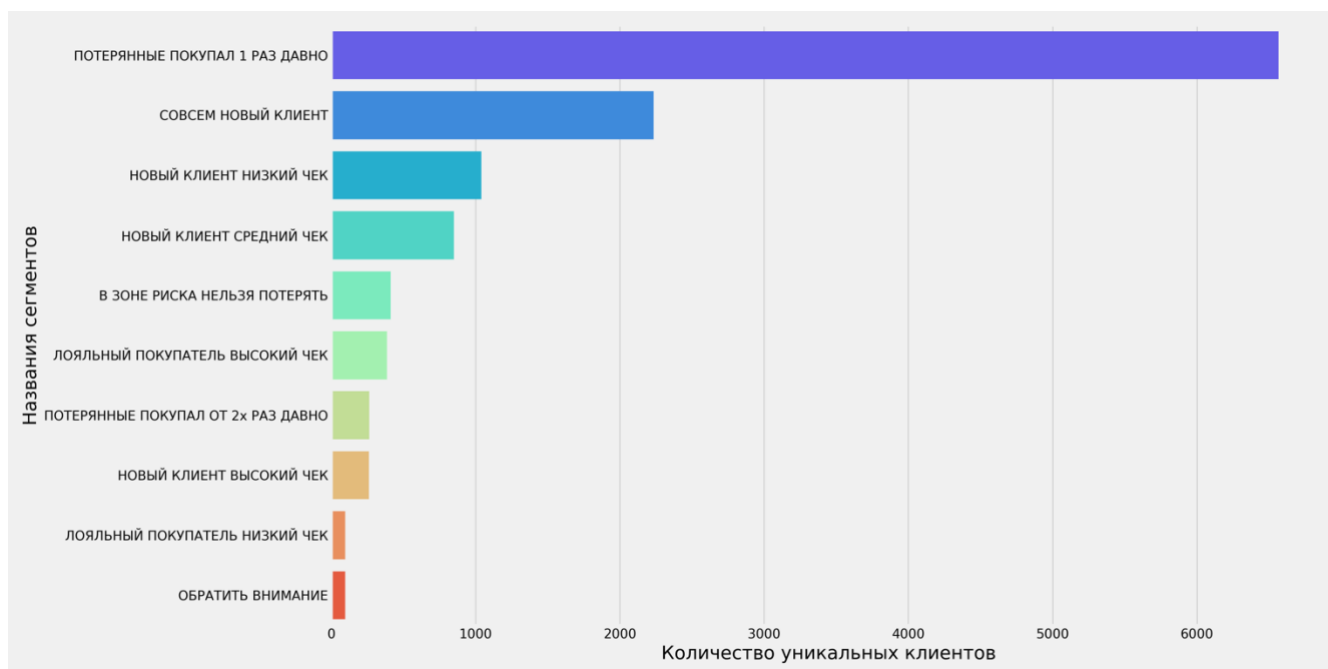
Таким образом за каждым клиентом закреплены баллы RFM, например 415 – это означает, что клиент купил первый раз недавно на крупную сумму, 355 – клиент, который покупал больше 2х раз, на большие суммы, но довольно давно, этот клиент может быть потерян.

По распределению можно поверхностно сделать сказать, что самые недавние клиенты (R) находятся в пределах 474 дней, средняя давность 475-639 дней, давние от 831 дня.

Частота покупок (F) у 60% клиентов - 1 раз, у 20% 2 раза, еще у 20% больше 2 раз.

Потраченная сумма (M) до 9780 руб. имеет низкий ранг, от 18294 до 29503 - средний ранг, больше 54400 - высокий ранг.

## Описание полученных сегментов



Построена столбчатая диаграмма, которая отражает распределение уникальных пользователей по сегментам.

ПОТЕРЯННЫЕ ПОКУПАЛ 1 РАЗ ДАВНО – самый крупный сегмент, 1 раз покупали более 832 двух дней назад. С одной стороны можно сказать, что эти клиенты потеряны навсегда, с другой – возможно какая-то часть этих покупателей до сих пор использует ранее купленные шины.

ПОТЕРЯННЫЕ ПОКУПАЛ ОТ 2х РАЗ ДАВНО – похожая группа, но покупали от 2х раз более 832 двух дней назад.

СОВСЕМ НОВЫЙ КЛИЕНТ – второй по объему сегмент, в котором клиенты совершали покупки совсем недавно – до 190 дней назад.

НОВЫЙ КЛИЕНТ НИЗКИЙ ЧЕК, НОВЫЙ КЛИЕНТ СРЕДНИЙ ЧЕК, НОВЫЙ КЛИЕНТ ВЫСОКИЙ ЧЕК – клиенты совершали покупки относительно недавно – от 190 до 474 дней назад.

В ЗОНЕ РИСКА НЕЛЬЗЯ ПОТЕРЯТЬ – покупали несколько раз и на крупные суммы, но последний раз покупали довольно давно - от 475 до 639 дней назад. Некоторые клиенты совершали

несколько покупок подряд (несколько комплектов шин или дисков), некоторые возвращались в течение нескольких лет – эта информация указана в поле «Интервал покупок».

ОБРАТИТЬ ВНИМАНИЕ – похожая группа, покупали несколько раз на разные суммы, последний раз покупали довольно давно - от 475 до 639 дней назад.

ЛОЯЛЬНЫЙ ПОКУПАТЕЛЬ НИЗКИЙ ЧЕК, ЛОЯЛЬНЫЙ ПОКУПАТЕЛЬ ВЫСОКИЙ ЧЕК – покупали недавно, покупали от 2х раз (часто). Может показаться, что это клиенты, которые действительно лояльны и постоянно совершают покупки в магазине, но в эту группу попало много людей, которые совершили много покупок в течение короткого промежутка времени. Например, клиент совершил 13 заказов в течение 3х месяцев:

4	Дмитрий	-	tartis@yanc	+79013062939	36132	-	2022-05-07 19:34:44	130	545	ЛОЯЛЬНЫЙ	2	35818	Покупал в течение 1 года
5	Дмитрий	-	tartis@yanc	+79013062939	35504	Санкт-Пете	2022-05-19 20:19:15	118	545	ЛОЯЛЬНЫЙ	2	35818	Покупал в течение 1 года
6	Артур Дычи	-	+79013704892	39344	-	-	2022-07-27 17:10:12	49	555	ЛОЯЛЬНЫЙ	3	29530,67	Покупал в течение 1 года
7	Артур Дычи	-	+79013704892	39412	-	-	2022-08-16 17:10:22	29	555	ЛОЯЛЬНЫЙ	3	29530,67	Покупал в течение 1 года
8	Артур Дычи	-	+79013704892	9836	-	-	2022-07-27 17:09:32	49	555	ЛОЯЛЬНЫЙ	3	29530,67	Покупал в течение 1 года
9	Вячеслав	-	+79013897307	17600	-	-	2021-09-22 17:02:20	357	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
10	Вячеслав	Кондрашов	boss694@li	+79013897307	7530	Димитровг	2021-12-08 16:12:21	280	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
11	Вячеслав	Кондрашов	boss694@li	+79013897307	18880	Димитровг	2021-11-25 11:22:36	294	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
12	Вячеслав	Кондрашов	boss694@li	+79013897307	15248	Димитровг	2021-11-22 00:47:38	297	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
13	Вячеслав	Кондрашов	boss694@li	+79013897307	7294	Москва	2021-11-10 05:09:24	309	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
14	Вячеслав	Кондрашов	boss694@li	+79013897307	12744	Москва	2021-11-09 14:30:04	309	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
15	Вячеслав	Кондрашов	boss694@li	+79013897307	11880	Москва	2021-11-09 01:04:07	310	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
16	Вячеслав	-	boss694@li	+79013897307	7268	Москва	2021-09-26 14:21:53	353	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
17	Вячеслав	Кондрашов	boss694@li	+79013897307	39924	Москва	2021-11-04 14:55:00	314	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
18	Вячеслав	-	boss694@li	+79013897307	25932	Москва	2021-10-19 12:04:32	330	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
19	Вячеслав	-	boss694@li	+79013897307	10600	Москва	2021-11-04 05:11:48	315	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
20	Вячеслав	-	boss694@li	+79013897307	15052	Москва	2021-11-04 04:52:17	315	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
21	Вячеслав	-	+79013897307	8000	-	-	2021-11-03 12:18:00	315	455	ЛОЯЛЬНЫЙ	13	15227,08	Покупал в течение 1 года
22	Стас	-	stascryazhe	+79018005326	20392	Санкт-Пете	2022-05-06 08:40:49	132	545	ЛОЯЛЬНЫЙ	2	28762	Покупал в течение 1 года
23	Стас	-	stascryazhe	+79018005326	37132	Санкт-Пете	2022-05-05 08:20:43	133	545	ЛОЯЛЬНЫЙ	2	28762	Покупал в течение 1 года
24	Борисов Ог	-	+79030084884	21876	-	-	2022-07-20 11:35:52	57	544	ЛОЯЛЬНЫЙ	2	13672,5	Покупал в течение 1 года
25	Борисов Ог	-	+79030084884	5469	-	-	2022-07-20 11:30:33	57	544	ЛОЯЛЬНЫЙ	2	13672,5	Покупал в течение 1 года

Эту группу стоит рассмотреть очень внимательно, чтобы правильно интерпретировать ее содержание. Есть немного тех, кто возвращался в течение нескольких лет – эта информация указана в поле «Интервал покупок». Поверхностно можно заключить, что истинно лояльных покупателей нет, скорее всего из-за специфики бизнеса, но возможно стоит обратить внимание, довольны ли клиенты покупками и сервисом.

Таблица для дельнейшей работы экспортирована в Excel, на первом листе все данные - можно использовать фильтр по названию сегмента, дополнительно все сегменты были экспортированы на отдельные листы.

## Выводы и рекомендации

Совершена предобработка:

- все номера телефонов приведены к одному виду, номера, не прошедшие стандартизацию номера, были удалены;
- обнаружены и удалены тестовые заказы;
- удалены дубликаты по полю «Телефон» и «Дата покупки» для отсеечения клиентов, которые могли продублировать заказ в течение одного дня, также удалены заказы по полю «Телефон» и «Сумма покупки», т.к. многие клиенты могли продублировать заказ в течение нескольких дней, возможно потому что не получали ответа на заявку.

Создана RFM-модель:

- рассчитана давность покупки, частота и общая сумма за все время покупок каждого клиента;
- каждому клиенту присвоен определенный сегмент;
- построена столбчатая диаграмма.

Дополнительно рассчитаны:

- средний чек каждого клиента;
- количество покупок каждого клиента;
- интервал, в течение, которого клиент совершал покупки.

Для проработки используйте следующие сегменты:

ПОТЕРЯННЫЕ ПОКУПАЛ 1 РАЗ ДАВНО, ПОТЕРЯННЫЕ ПОКУПАЛ ОТ 2х РАЗ ДАВНО – сформируйте предложение для рассылки, чтобы вернуть этих клиентов.

В ЗОНЕ РИСКА НЕЛЬЗЯ ПОТЕРЯТЬ, ОБРАТИТЬ ВНИМАНИЕ – напомните о себе с помощью рассылки.