**Анализ ухода пользователей в телекоммуникационной компании**

**Введение:**

1. Для целей данного анализа мы взяли набор данных на сайте kaggle.com (telecom\_customer\_churn\_origin.csv)
2. Анализ данных сделан с помощью языка Питон

**Изучение данных:**

1. Проверяем размер датасета. Датасет состоит из 7043 строк и 38 столбцов.



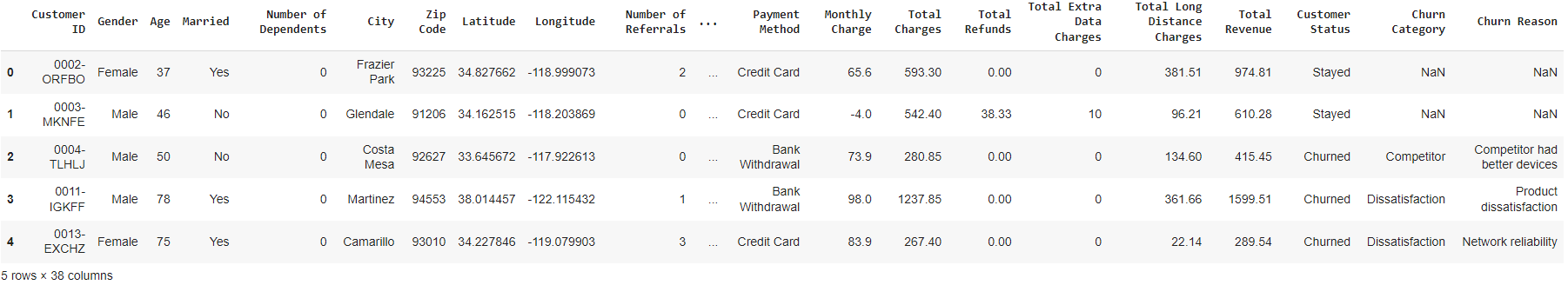
1. Проверяем наличие дублирующихся строк. Таковых нет.



1. Открываем первые 5 строк таблицы, чтобы иметь представление из чего состоят данные.

Каждая строка содержит данные 7043 уникального пользователя по 37 параметрам. 1 столбец содержит идентификационный номер.



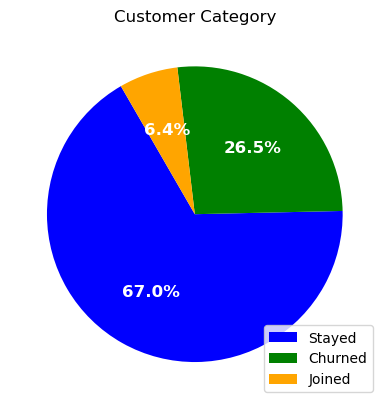
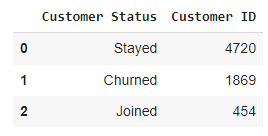


**Анализ данных:**

1. В таблице, среди прочих, заметили столбец Customer Status, который отображает статус пользователя. Компания поделила своих пользователей на 3 категории по статусу.
   * Stayed (далее по тексту - Оставшиеся)
   * Churned (далее по тексту - Ушедшие)
   * Joined (далее по тексту - Присоединившиеся)



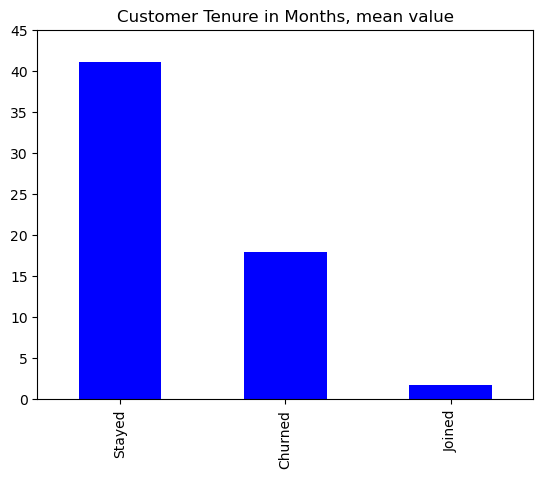
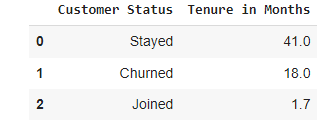
1. Проверяем количество пользователей в каждой категории по статусу. 67% пользователей в категории Оставшиеся, 26.5% - в категории Ушедшие, 6.4% пользователей являются Присоединившимися.



1. Возникает вопрос, когда именно присоединились Присоединившиеся. Заодно, необходимо проверить, сколько уже пользуются услугами компании остальные две категории пользователей.

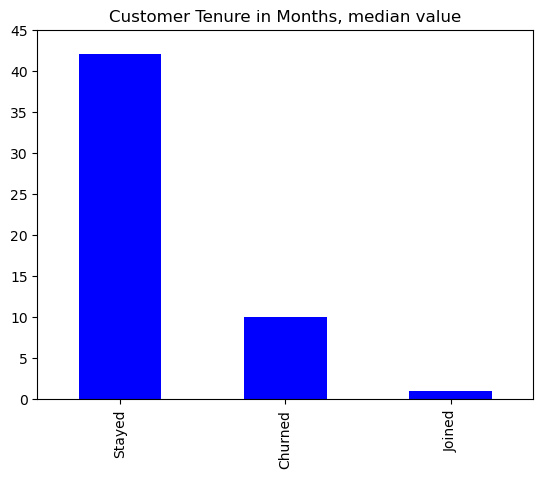
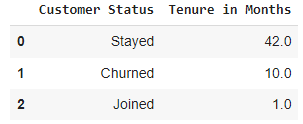
Проверяем среднее и медианное значения по всем трем категориям.

Среднее значение:



Оставшиеся пользователи, в среднем, пользуются услугами компании 41 месяц. Ушедшие – 18 месяцев, тогда как Присоединившиеся только 1.7 месяцев.

Медианное значение:

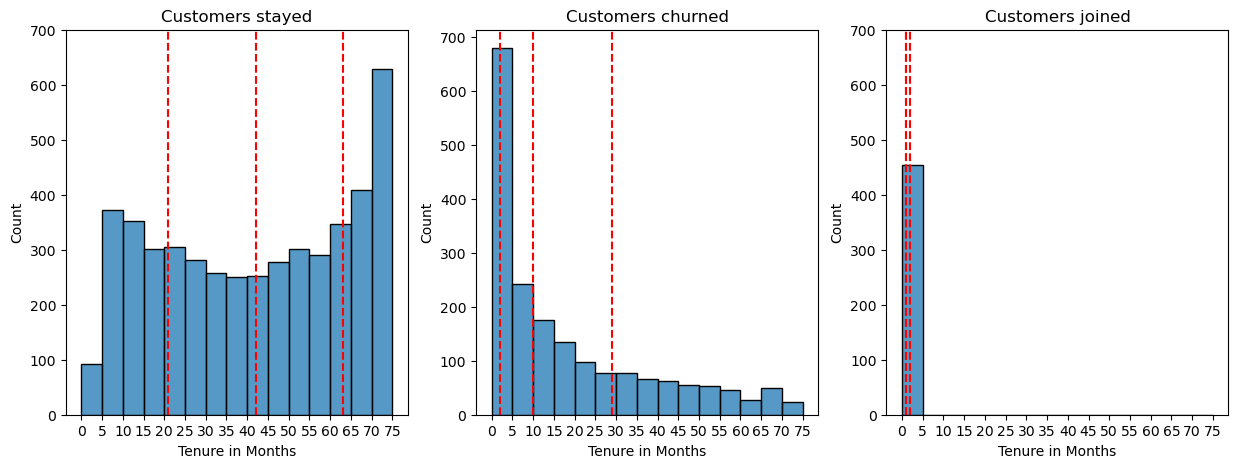


Половина Оставшихся пользователей пользуются услугами компании более 42 месяцев. Половина Ушедших – менее 10 месяцев, тогда как половина Присоединившихся 1 месяц.

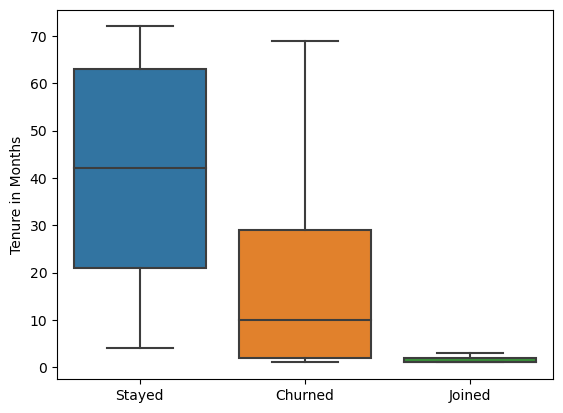
Заметно, что при вычислении средних и медианных значений, картина по Ушедшим пользователям кардинально меняется.

Для полного понимания картины, посмотрим также на квартили.

1. Вид с одного ракурса



1. Вид со второго ракурса



Выяснили, что среднее значение по Ушедшим пользователям, все-таки, нам дает искаженную картину. Оказалось, что целых 25% Ушедших ушли в первые 2 месяца. К этому вопросу мы вернемся чуть позже, сейчас же нас интересует другая категория – Присоединившиеся.

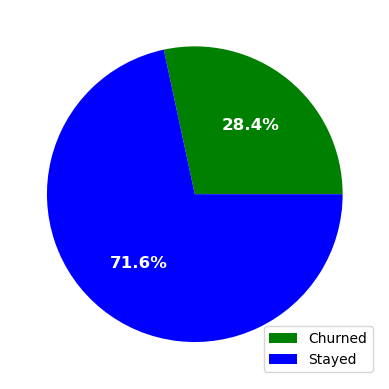
50% данной категории пользуются услугами компании только 1 месяц. Самый долго пользующийся услугами компании Присоединившийся пользователь присоединился только 3 месяца назад.

Основываясь на этих данных, приходим к выводу, что анализе причин ухода пользователей учитывать Присоединившихся пользователей нецелесообразно.

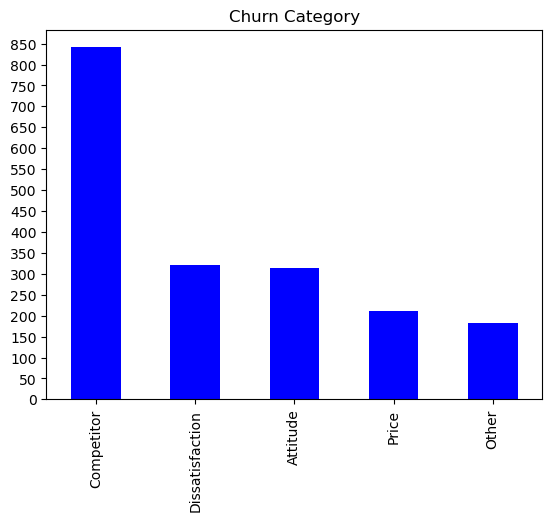
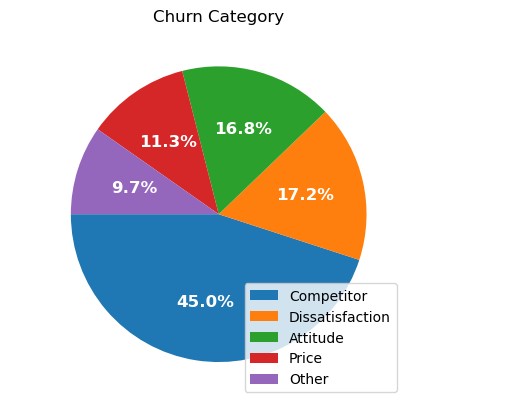
**Поэтому, в дальнейшем, убираем данную категорию пользователей из анализа данных причин ухода.** (*Однако, в последующем, мы можем использовать их данные для другого анализа – прогнозирования их поведения – как долго они будут пользоваться услугами компании.*)

----------------------------------

1. Убираем данные Присоединившихся пользователей из датасета. В таком случае, уже почти три пользователя из десяти уходят из компании. Просто большой и непозволительный отток пользователей.



1. Проверяем, почему пользователи уходят.



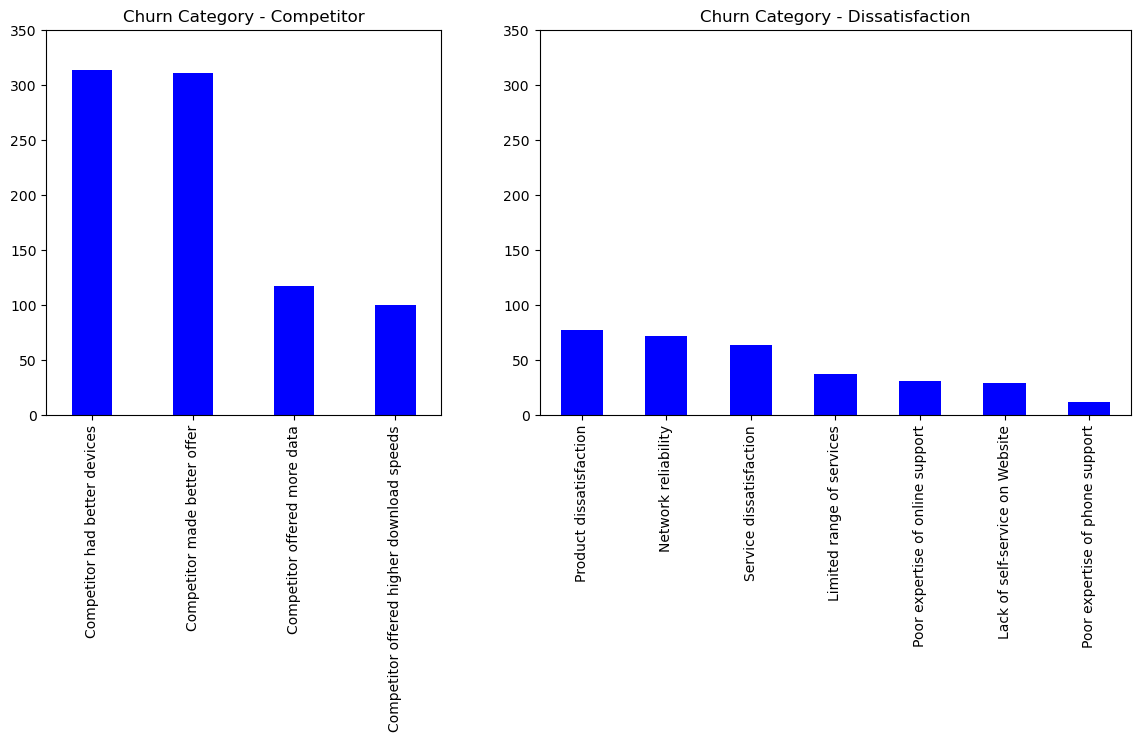
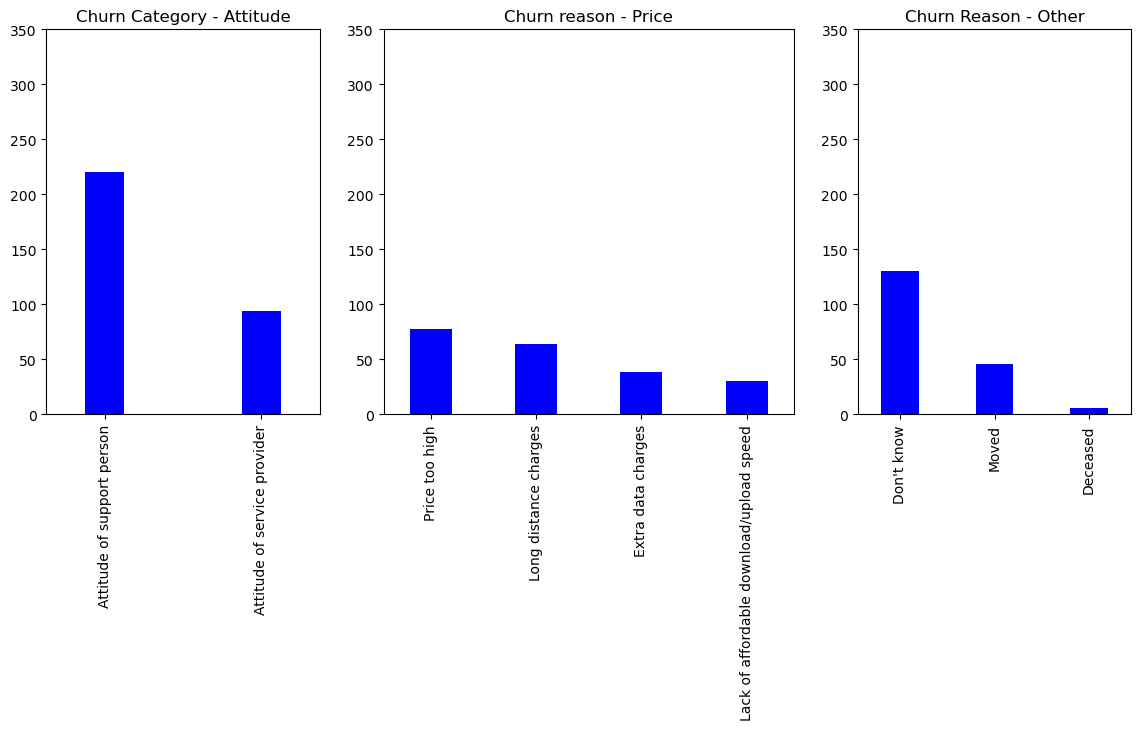
Почти половина (45%) Ушедших ушли по причине конкурирующих компаний. (*В детали углубимся позже)*

По 17% пользователей ушли по причине неудовлетворенности предоставляемыми услугами и … отношением к ним. (*В детали углубимся позже)*

Еще по 10% ушли по причинам, связанным со стоимостью услуг, и другим причинам. (*В детали углубимся позже)*

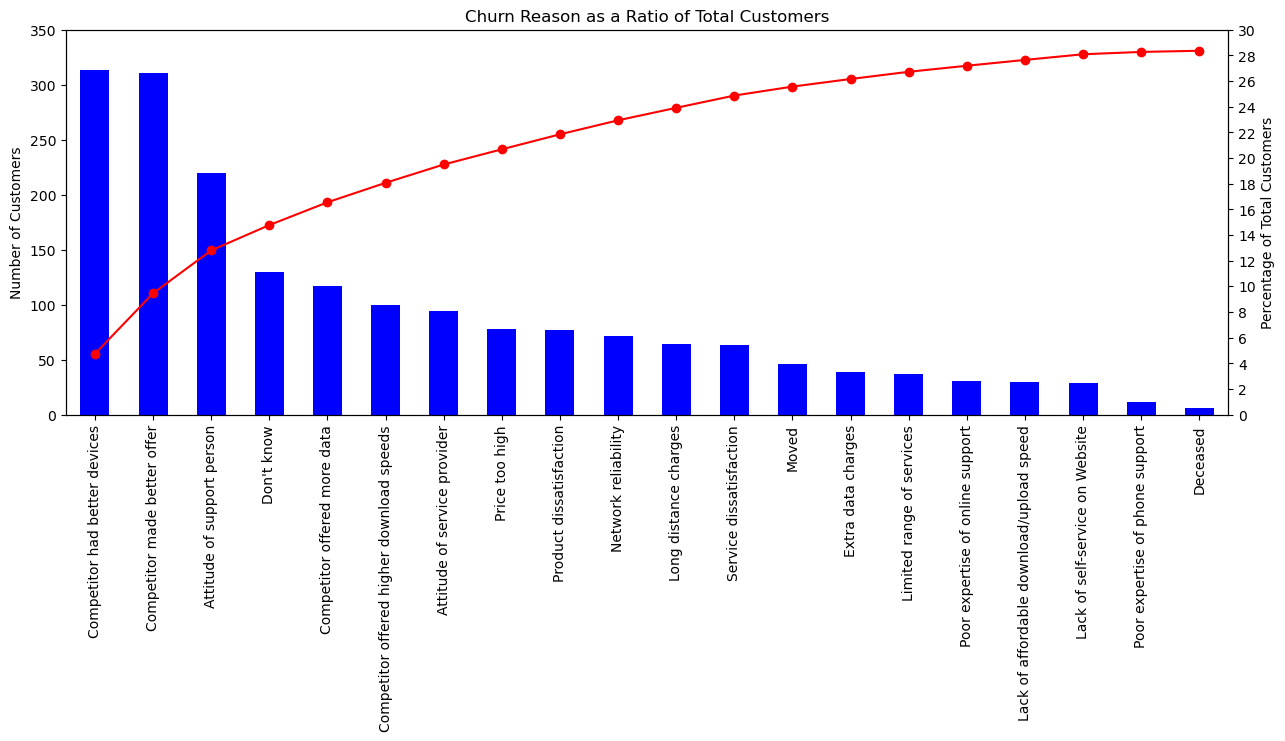
1. Углубляемся в детали причин ухода из компании.

*В графиках ниже представлены конкретные причины ухода пользователей из компании в разрезе 5 категории.*



В данных графиках отображены конкретные причины ухода пользователей в абсолютном значении.

Как было отмечено ранее, 28% пользователей уходят из компании. Ниже посмотрим вклад каждой из этих 20 причин, без разбивки по категориям, в отток пользователей, в процентном выражении.



**То есть, видим, что если компания сможет предоставлять пользователям устройства и предложения не хуже, чем конкурирующие компании, то сможет снизить уход пользователей на целых 10 процентных пунктов. В таком случае, соотношение Оставшихся и Ушедших пользователей будет равно 82% : 18%.**

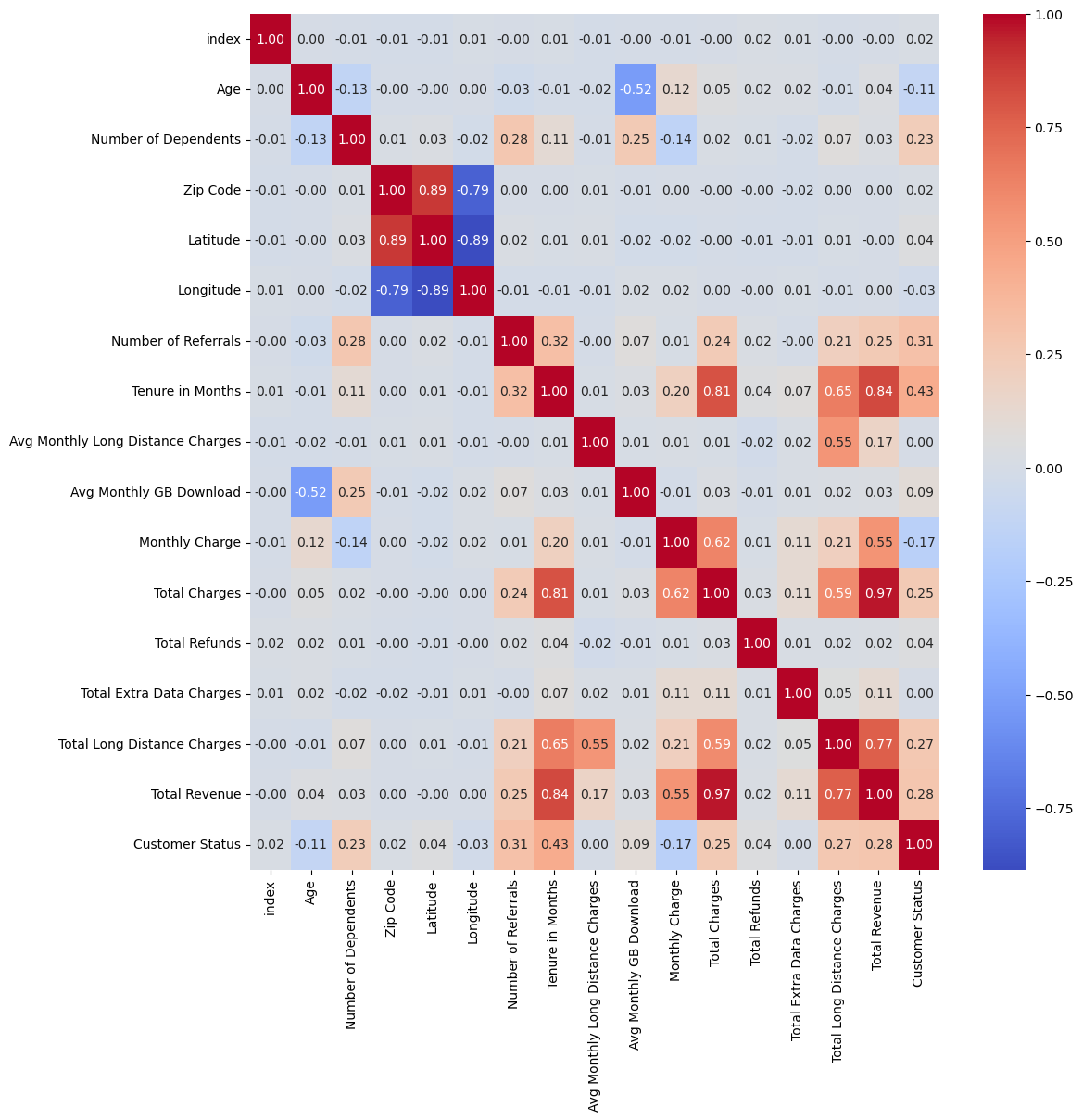
**И так далее по остальным причинам ухода пользователей.**

**Помимо этого, мы видим, что есть причины, на которые компания повлиять не может.** *(Некоторые пользователи умерли, некоторые переехали)*

----------------------------------

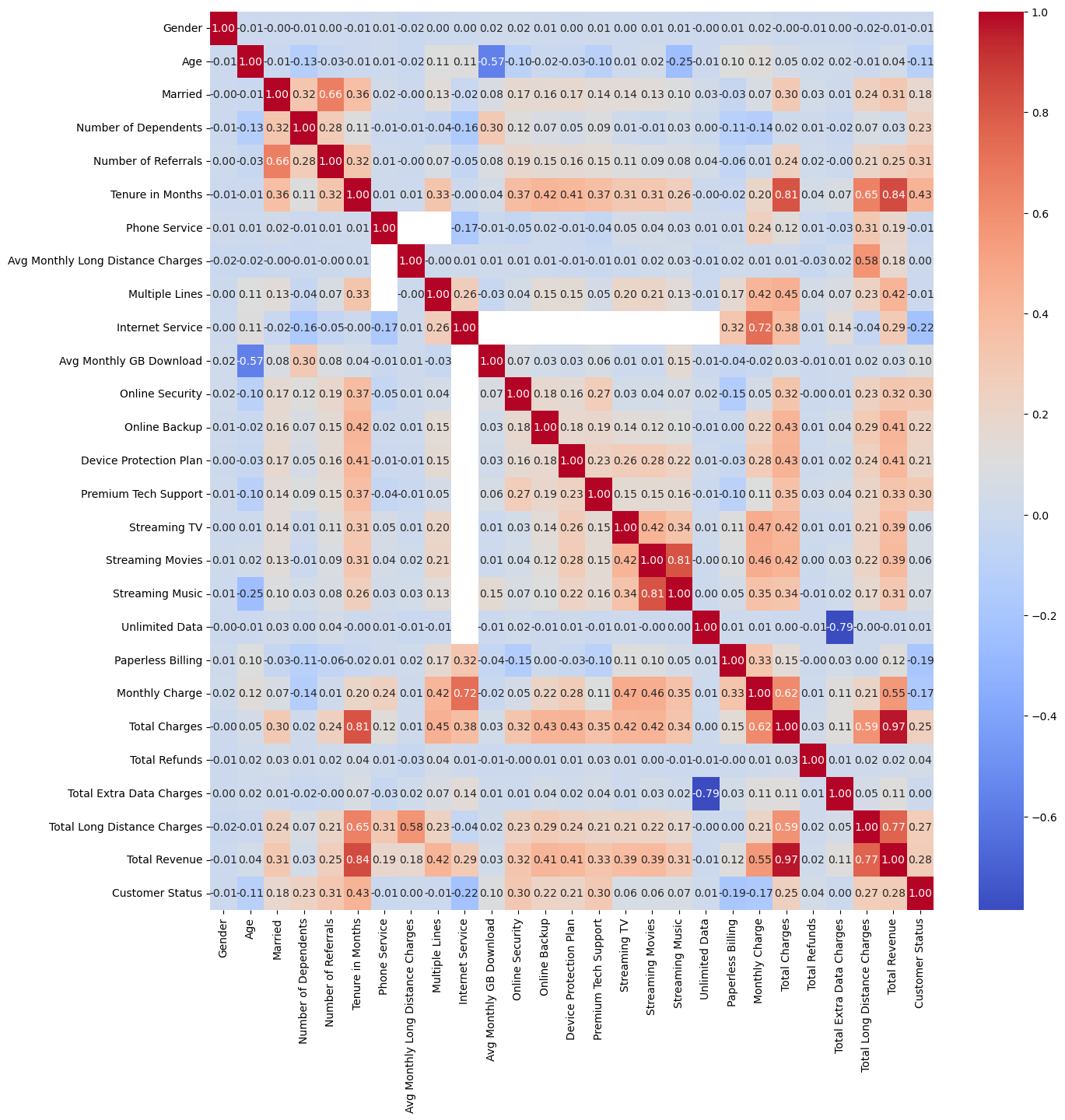
1. Пробуем понять, какие именно пользователи уходят.

Так как параметров, описывающих каждого пользователя очень много (37 столбцов), то попробуем провести корреляцию между пользователями и их числовыми параметрами. *(в таблице есть как числовые параметры, так и текстовые)*



Это корреляция особых инсайтов не дала.

Поэтому, проводим корреляцию между категориями пользователей и их текстовыми параметрами.



Теперь, картина стала более понятной.

**Ниже полученные инсайтов в результате корреляции:**

Факторы, не влияющие на статус пользователя:

1. Пол
2. Наличие услуги мобильной связи
3. Наличие нескольких линий
4. Наличие стриминговых сервисов (ТВ, фильмы, музыка)
5. Среднемесячная плата за междугородную связь
6. Среднемесячный объем загрузки данных

Факторы, отрицательно влияющие на статус пользователя (пользователи уходят):

1. Возраст – возрастные пользователи больше уходили.
2. Пользование услугами интернет – пользователи, пользовавшиеся услугами интернет, больше уходили.
3. Безбумажная оплата – пользователи, пользовавшиеся безбумажной оплатой, больше уходили.
4. Ежемесячная оплата – пользователи, у которых ежемесячная оплата была больше, больше уходили.

Факторы, положительно влияющие на статус пользователя (пользователи остаются):

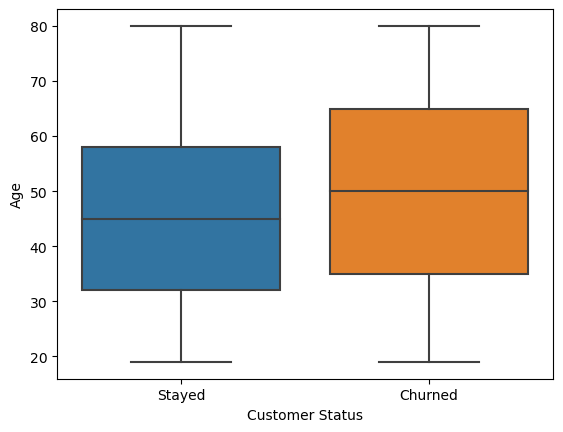
1. Семейное положение – женатые/замужние пользователи склонны остаться.

*\* Не женатые/замужние пользователи, наоборот, склонны уходить. Применимо ко всем фактором ниже.*

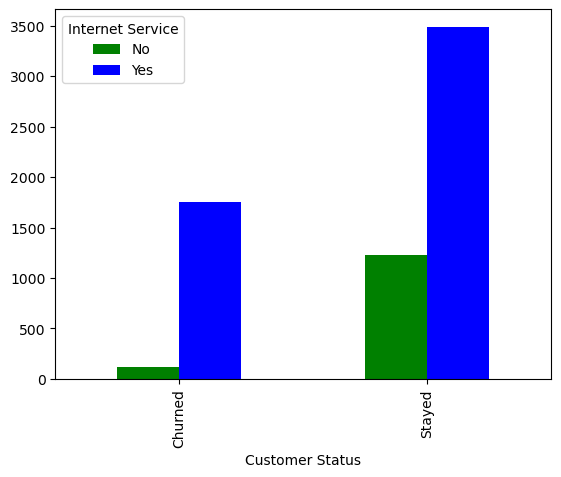
1. Количество иждивенцев – пользователи, имеющие иждивенцев, склонны остаться.
2. Количество рефералов – пользователи, имеющие больше рефералов, склонны остаться.
3. Срок пользования в месяцах – чем дольше пользуется пользователи, тем больше склонны остаться.
4. Онлайн безопасность – пользователи с функцией онлайн безопасности склонны остаться
5. Онлайн резервное копирование – пользователи с функцией онлайн резервного копирования склонны остаться
6. Защита устройства – пользователи с функцией защиты устройства склонны остаться
7. Премиальная техническая поддержка – пользователи с функцией премиальной технической поддержки склонны остаться

----------------------------------

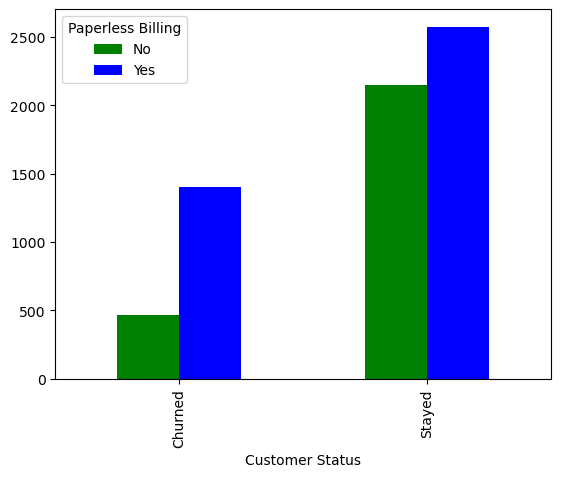
8. Проверяем полученные инсайты.



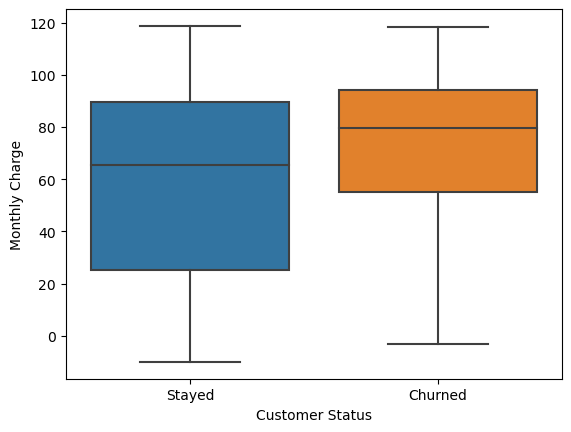
Действительно, большинство Ушедших пользователей старше по возрасту чем Оставшиеся.



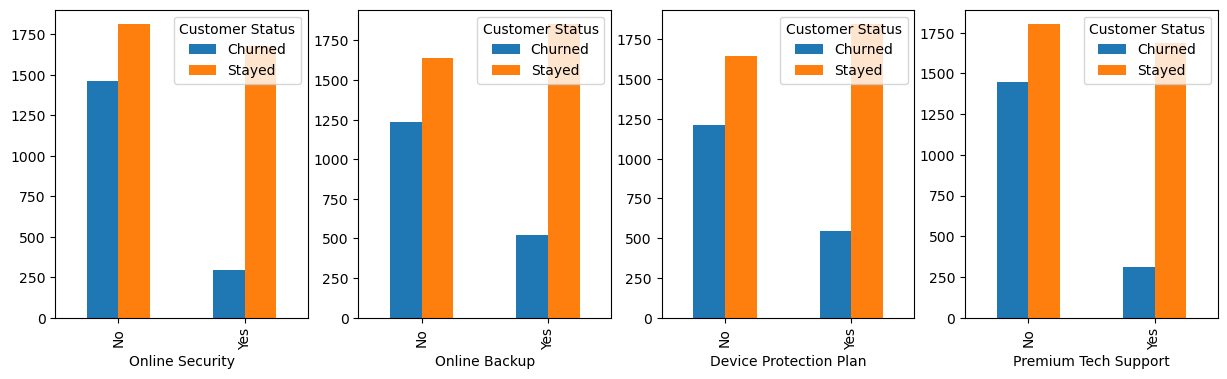
Действительно, среди Ушедших пользователей абсолютное большинство пользовалось интернетом.



Действительно, среди Ушедших пользователей большинство пользовались безбумажной оплатой.



Действительно, Ушедшие пользователи ежемесячно платили больше, чем Оставшиеся.



Действительно, Ушедшие пользователи ежемесячно платили больше, чем Оставшиеся.

----------------------------------

В результате анализа вырисовываются портреты двух сегментов пользователей:

1. Какой тип пользователей склонен остаться

* Женатый/замужний человек среднего возраста, имеющий детей.
* Современный, пользующийся всеми современными услугами, такие как: онлайн безопасность, онлайн резервное копирование, защита устройства, премиальная техническая поддержка.
* Также, больше пользуется услугами междугородней связи, и остается с компанией надолго, тем самым увеличивая LTV.
* Также, имеет много друзей/знакомых, которым активно рекомендует услуги компании, тем самым помогая значительно увеличить количество постоянных пользователей.

1. Какой тип пользователей склонен уходить

* Одинокий мужчина/женщина без детей.
* Старше среднего возраста, без друзей/знакомых.
* В силу возраста, мало пользующийся интернетом и другими дополнительными услугами компании.
* Также, в силу возраста, не привыкший к современным удобным способам оплаты, и предпочитающий оплачивать по старинке.

----------------------------------