

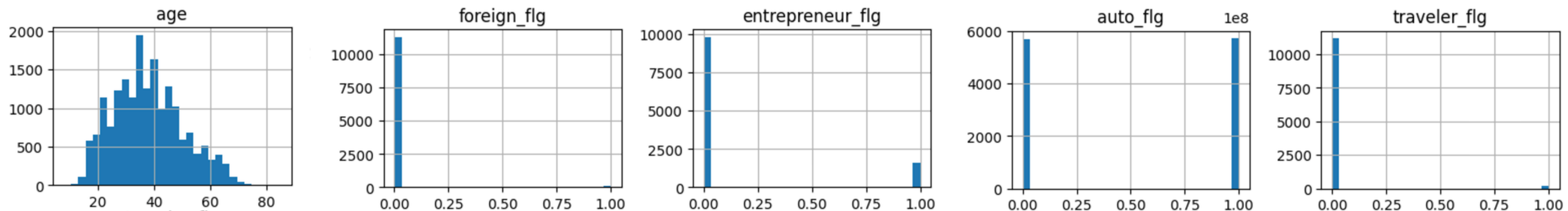
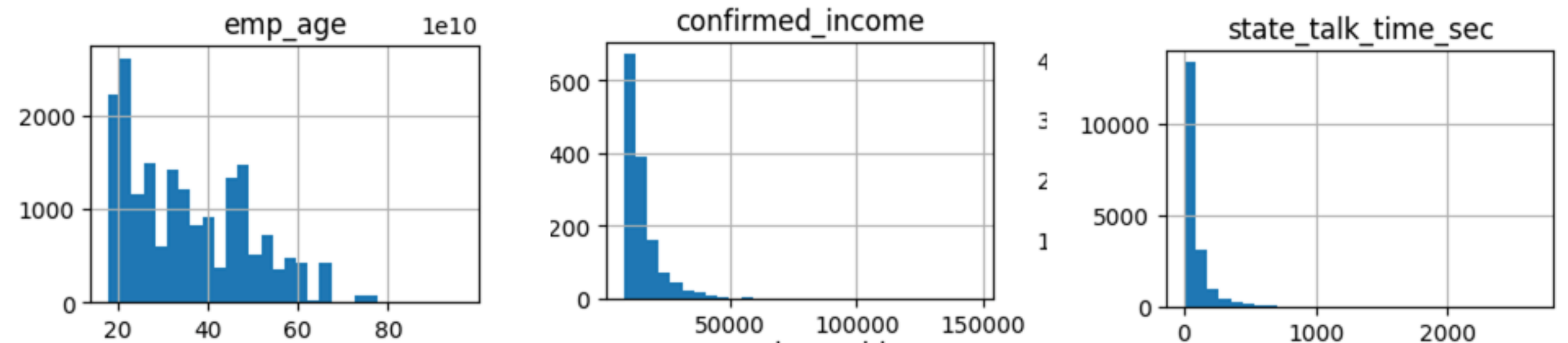
Разведочный анализ данных (EDA)

Пропущенные значения

Был найден процент пропущенных значений в столбцах дата-сета. По этим значениям можно сделать вывод о том, какие столбцы являются релевантными для анализа, а какие - нет в силу большого количества пропущенных данных. Также были произведены изменения некоторых типов данных полей для дальнейшего удобства работы (с датой и временем) и корректного разбиения на типы переменных (количественные, категориальные и ранговые)

Наблюдаемые распределения

У числовых признаков наблюдаются нормальное распределение (возраст клиента), биномиальное (все флаги), скошенное (для возраста сотрудника, времени диалога и подтвержденного дохода)





Исследование категориальных переменных

Удалось обнаружить следующие закономерности:

Основным сегментом пользователей, которым поступают звонки, являются клиенты, проживающие в городах (не миллионниках)

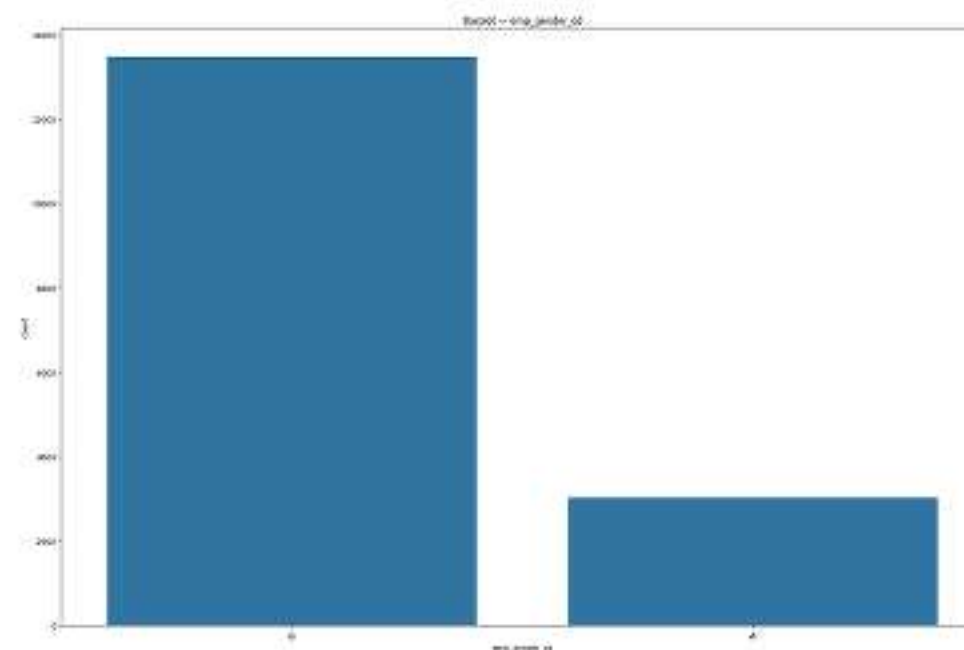
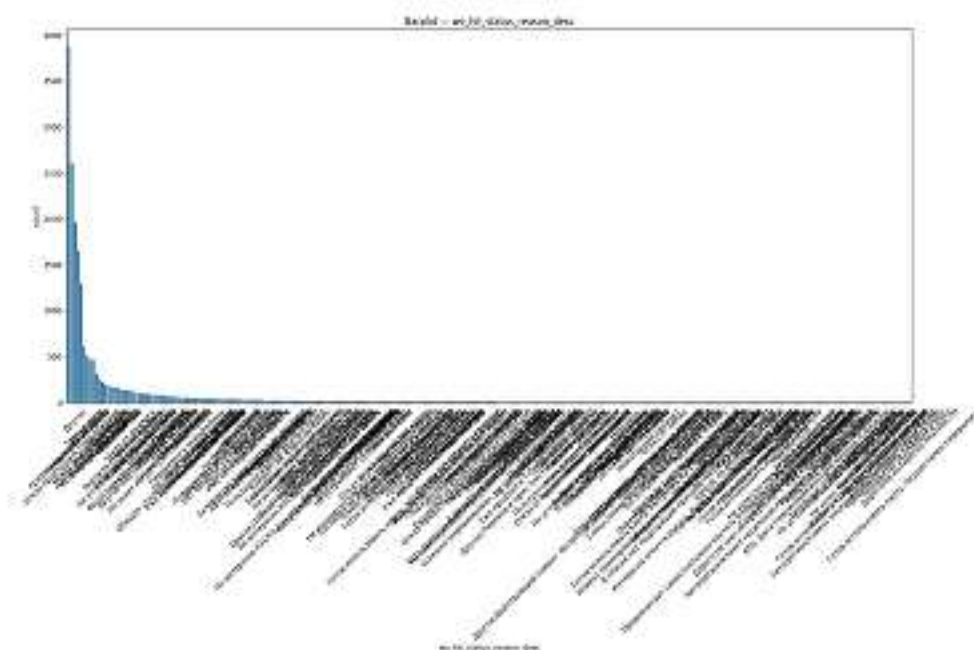
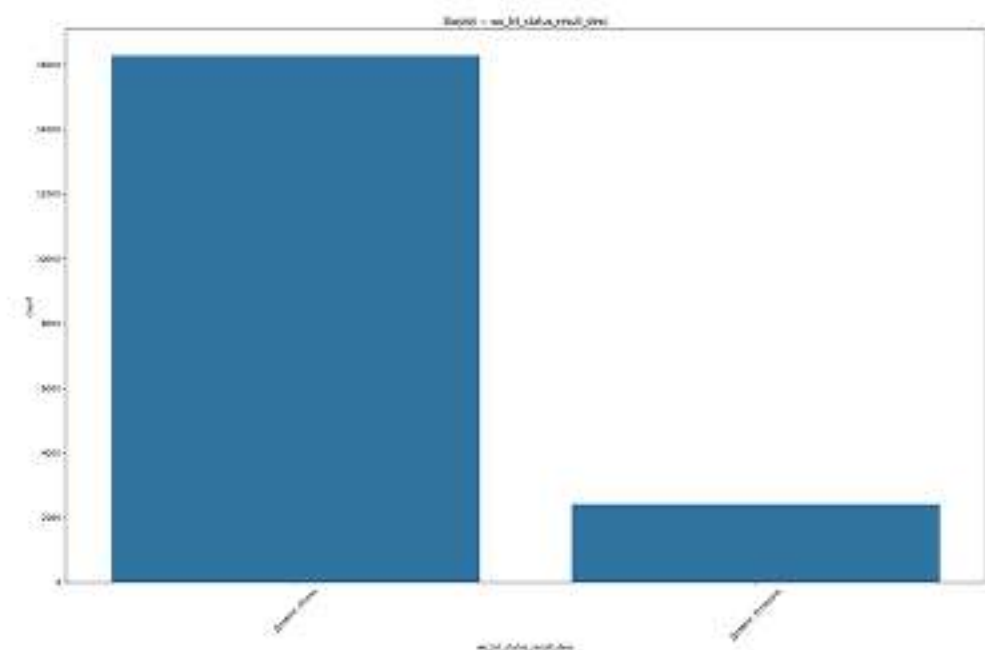
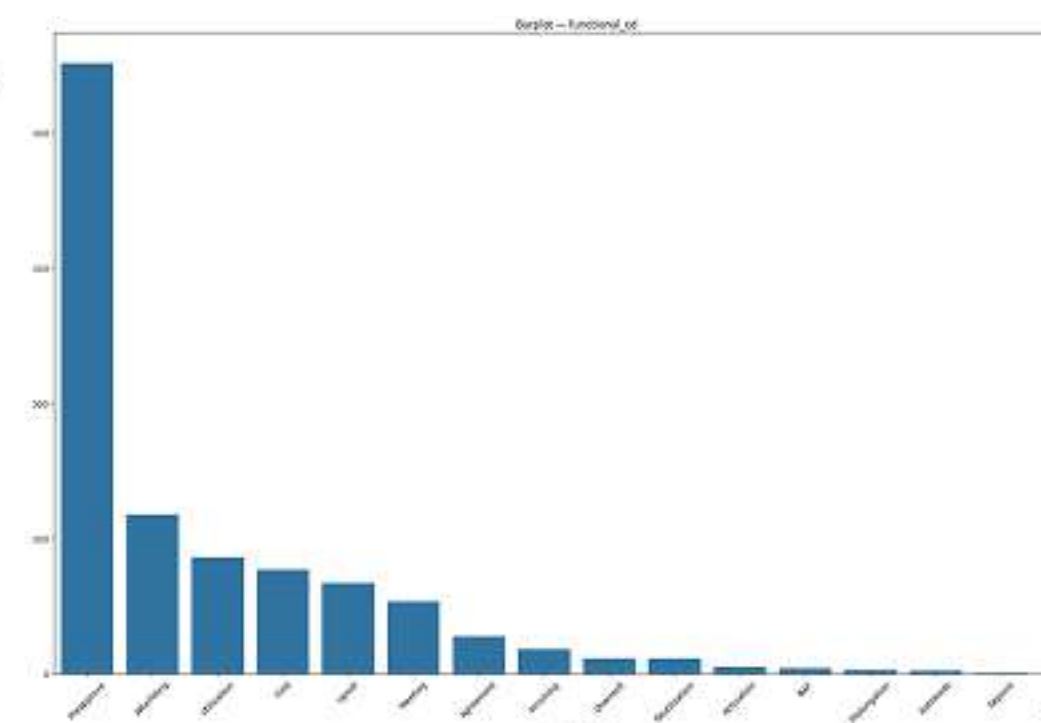
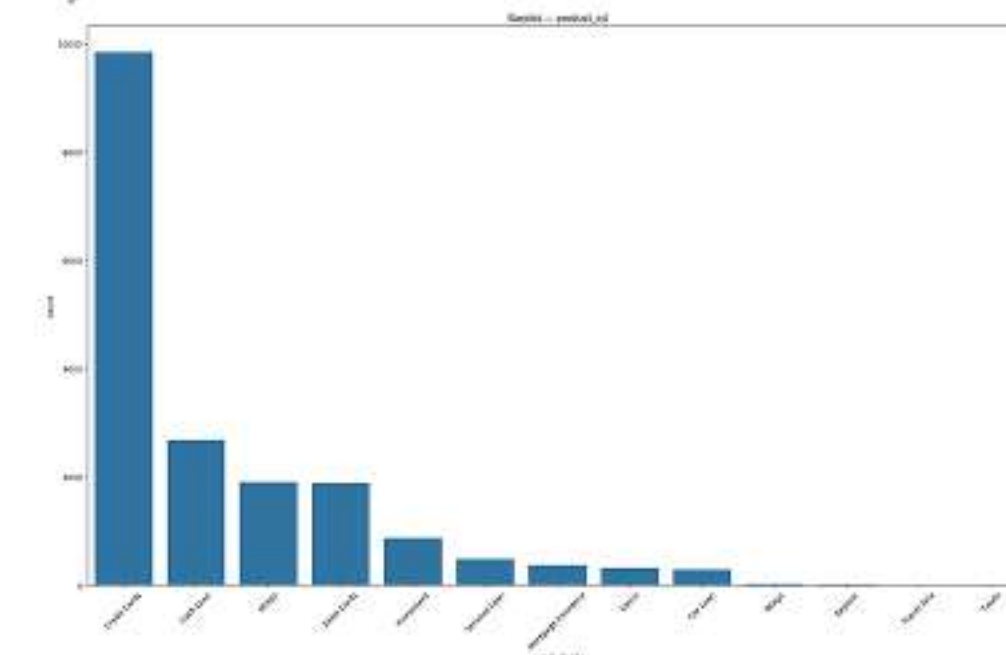
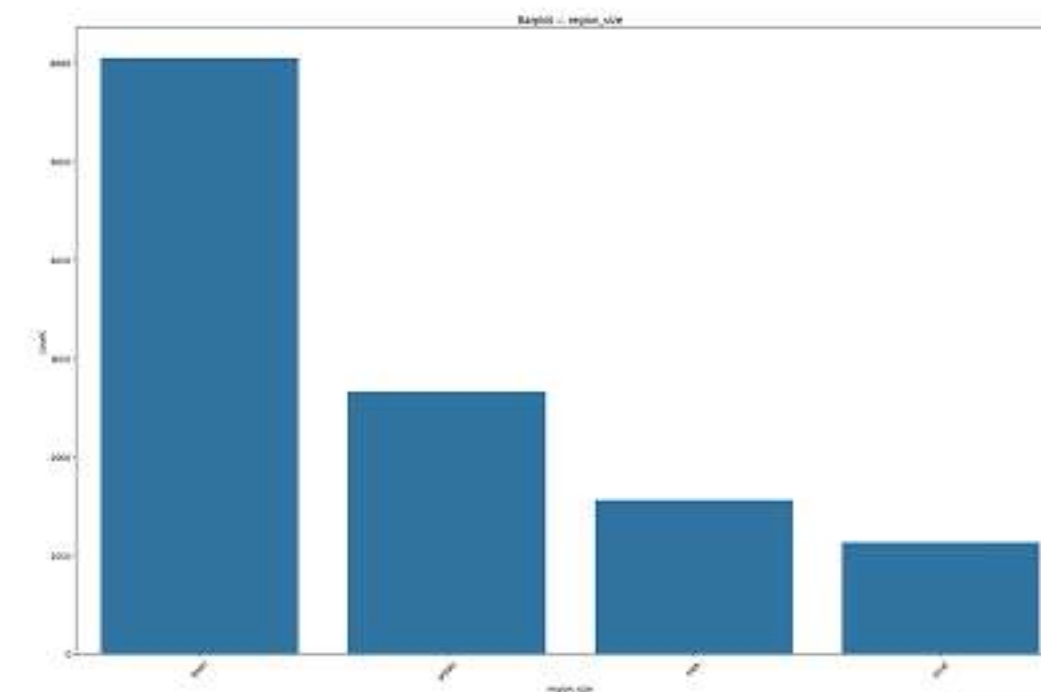
Наблюдается скошенное распределение у segment_cd

Самые популярные продукты при обзвоне - КК, КН, MVNO, ДК

Самые популярные тематики при обзвоне - preapprove, afterfilling, utilization, cold, upsell

Большинство исходящих звонков заканчиваются отказом (топ причин - кредитная карта не нужна, отказ прослушать предложение, другое, кредит уже не нужен)

Наибольший процент работников в штате составляют женщины



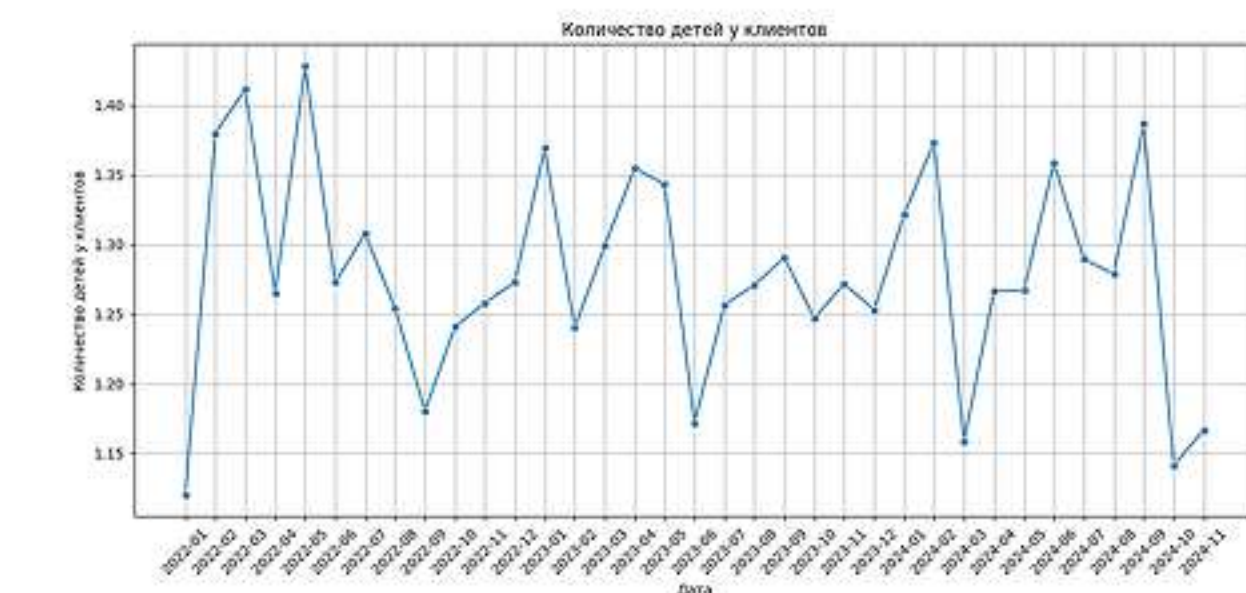
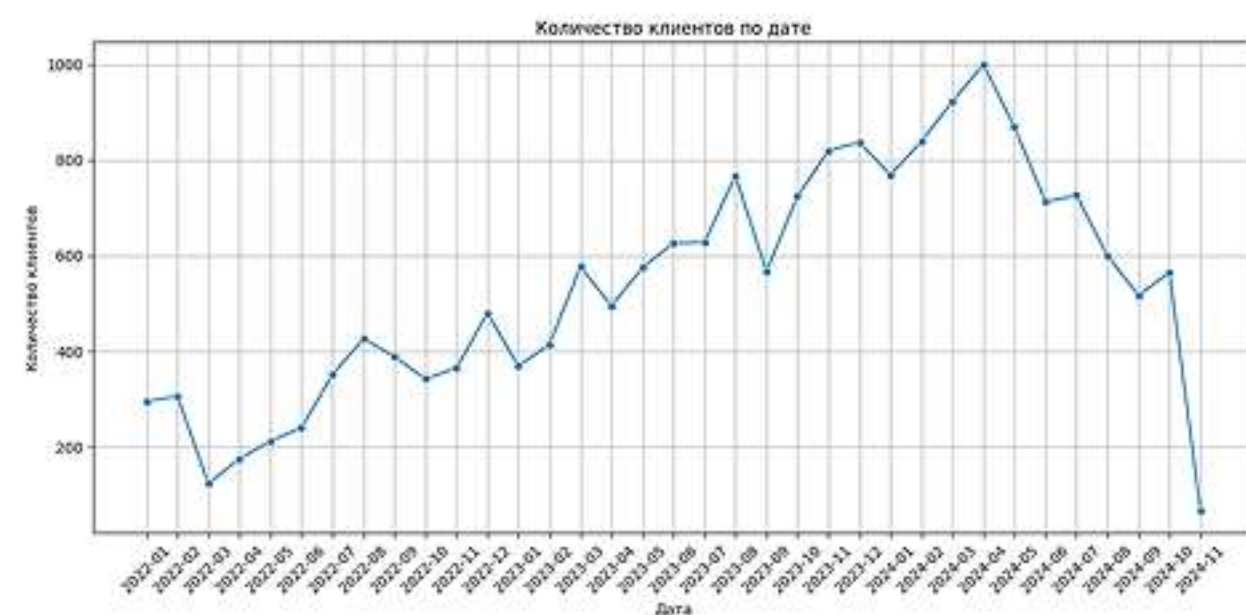
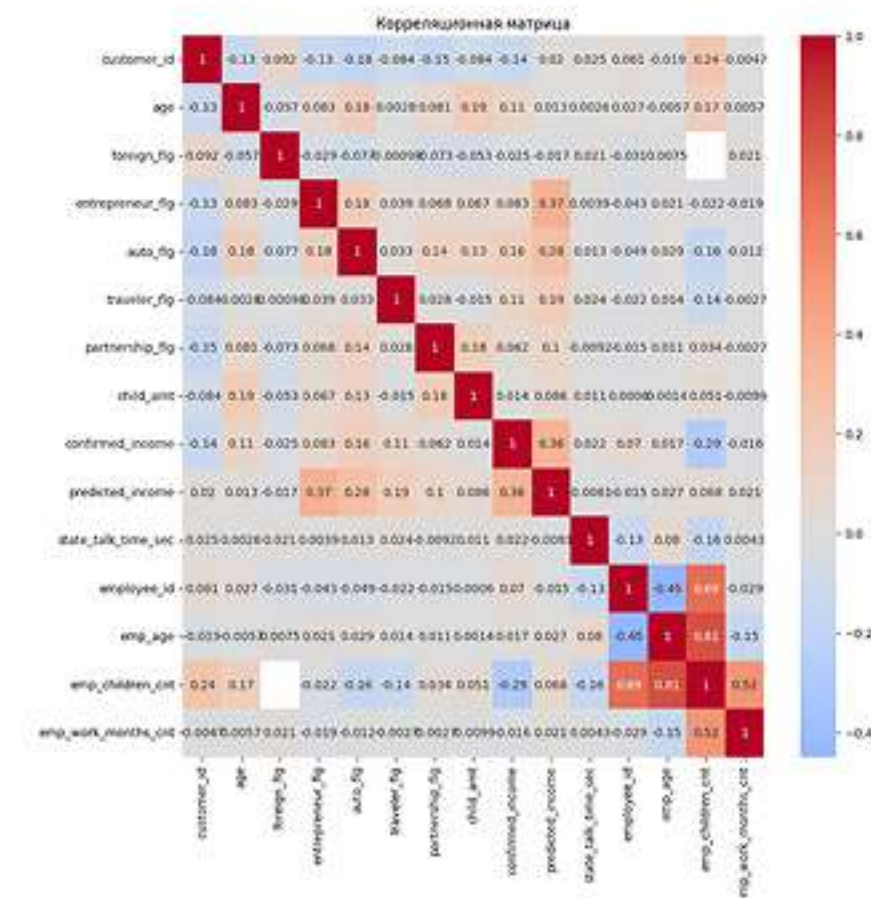


Корреляционная матрица

Интересных зависимостей для генерации продуктовых гипотез не обнаружено

Динамика показателей

Были построены графики количества клиентов, количества клиентов предпринимателей, количества клиентов с партнерами, количества детей у клиентов, средний доход клиентов, средний доход клиентов с усеченными выбросами, среднее время диалога в группировке по месяцам





Предлагаемые гипотезы

Второй по количеству звонков оператора причиной является Afterfilling (дозаполнение заявки). Так как клиент уже заполнял заявку, то он заинтересован в продукте. Уменьшить количество звонков оператора и, соответственно, сохранить затраты на обслуживание клиента позволит организация пуш-уведомлений с предложением дозаполнения заявки в селф-сервисе.

Отправка пуш-уведомлений с предложением самостоятельного дозаполнения заявки в селф-сервисе уменьшит количество звонков оператора.

Средний доход пользователей (если усечь выбросы) сохраняет тенденцию роста, начиная с конца 2023 года. Можно увеличить стоимость обслуживания продуктов.

Незначительное увеличение стоимости продуктов на 3 п.п. увеличит прибыль банка

Можно было бы предположить, что введение “семейных” предложений: промокодов, специальных условий для клиентов с детьми и находящихся в отношениях увеличило бы количество клиентов, однако, данная гипотеза была отвергнута в связи с нестабильностью метрики количества клиентов с детьми (не наблюдается статистическая зависимость) и со спадом количества клиентов с партнерами.



Предлагаемые гипотезы

Дисклеймер: данная гипотеза не несет цели оскорбить чьи-то чувства :)

Существует гипотеза о том, что мужчинам в банковской сфере доверяют больше, чем девушкам. И в колл-центре Т-банка подавляющее большинство сотрудников составляют женщины. Дальше существует гипотеза о том, что женский голос больше располагает.

Увеличение числа мужчин в штате увеличит число успешных сделок

В рамках данного дата-сета нельзя ни подтвердить, ни опровергнуть данную гипотезу. Требуется дополнительное исследование с помощью АВ-теста

Была рассчитана доля успешных сделок среди мужчин и женщин и их количество, доля одинаковая, а количество разное

Стоит увеличить скорость обработки тасок, чтобы быстрее выдавать/одобрять кредит, т.к. Одной из основных тематик, по которой клиент отказывает - неактуальность выдачи кредита

Увеличение скорости обработок заявок (дополнительный найм сотрудников или дообучение уже действующих, объединение групп сотрудников, или наоборот, их декомпозиция) позволят увеличить число успешных сделок



Предлагаемые гипотезы

Существует гипотеза о том, что мужчинам в банковской сфере доверяют больше, чем девушкам. В колл-центре т-банка подавляющее большинство сотрудников составляют женщины. Дальше существует гипотеза о том, что женский голос больше располагает.

Увеличение числа мужчин в штате увеличит число успешных сделок

В рамках данного дата-сета нельзя ни подтвердить, ни опровергнуть данную гипотезу. Требуется дополнительное исследование с помощью АВ-теста



Выбор гипотезы

На мой взгляд, самой перспективной гипотезой является самая первая - она может быть проверена при помощи данного дата-сета при внедрении новой фичи.