

В Тинькофф есть множество продуктов для юридических лиц. Прежде чем начинать сотрудничать с какой-либо из компаний нам хочется понимать, насколько компания честная. Так как, во-первых, мы можем потерять деньги из-за работы с нечестными компаниями, во-вторых, наши клиенты могут быть обмануты и нам не хочется такого. Для этого мы собираем данные о компаниях из различных открытых и не только источников и пытаемся предсказать, с кем можно сотрудничать, а с кем не стоит. Вам представлена часть данных о 15 000 юридических лицах, которые отвечают за их описание и «репутацию».

Описание полей:

id	id компании
date	дата скоринга компании
phones_num	число телефонов компании, которые у нас есть
found_phones	число телефонов компании, найденных в тел книжках наших клиентов
phb_fraud	кол-во телефонов компании, записанных в тел книжках наших клиентов с фродовыми словами (типа мошенники, кидалы и т.д.)
phb_dirty	кол-во телефонов компании, записанных в тел книжках наших клиентов с матерными словами
phb_all_bad	общее число записей телефонов компании ТК с негативными словами
phb_all	общее число записей в ТК телефонов компании
found_phones_dir	число телефонов руководителей компании, найденных в ТК наших клиентов
phb_bad_perc	доля всех негативных записей от общего количества записей телефонов компании
phb_fraud_dir	кол-во телефонов руководителей компании, записанных в тел книжках наших клиентов с фродовыми словами

phb_dirty_dir	кол-во телефонов руководителей компании, записанных в тел книжках наших клиентов с матерными словами
phb_all_bad_dir	общее число записей телефонов руководителей компании ТК с негативными словами
phb_all_dir	общее число записей в ТК телефонов руководителей компании
phb_bad_perc_dir	доля всех негативных записей от общего количества записей телефонов руководителей компании
appl_num	число заявок на кредит от данной компании
claims_percent	доля претензий (от числа заявок), поступивших на компанию от клиентов кредитования
complaints_num(_dir)	количество жалоб, поступивших на компанию (на руководителя)
mvd_request_num	количество запросов МВД, поступивших на компанию или ее руководителя
total_case_cnt	общее количество дел
total_defendant_case_cnt	количество дел в качестве ответчика
defendant_case_perc	доля дел в качестве ответчика
total_defendant_case_amt	сумма исков в качестве ответчика
bad_rep_tcb	флаг плохой репутации партнеров ТСВ
rating_ya	рейтинг компании в Яндекс Картах
ratings_cnt_ya	кол-во оценок у компании в Яндекс Картах
reviews_cnt_ya	кол-во отзывов у компании в Яндекс Картах
rating_yell	рейтинг компании на сайте Yell
reviews_cnt_yell	количество отзывов у компании на сайте Yell
reviews_cnt_yell	количество отзывов у компании на сайте Yell
reviews_cnt_2gis	количество отзывов у компании в картах 2GIS
rating_avito	рейтинг компании на Avito
reviews_cnt_avito	количество отзывов у компании на Avito

Варианты заданий:

1. Выявить закономерности в плохих отзывах
2. Как меняло положение в зависимости от времени
3. Научиться предсказывать негативный рейтинг компании