

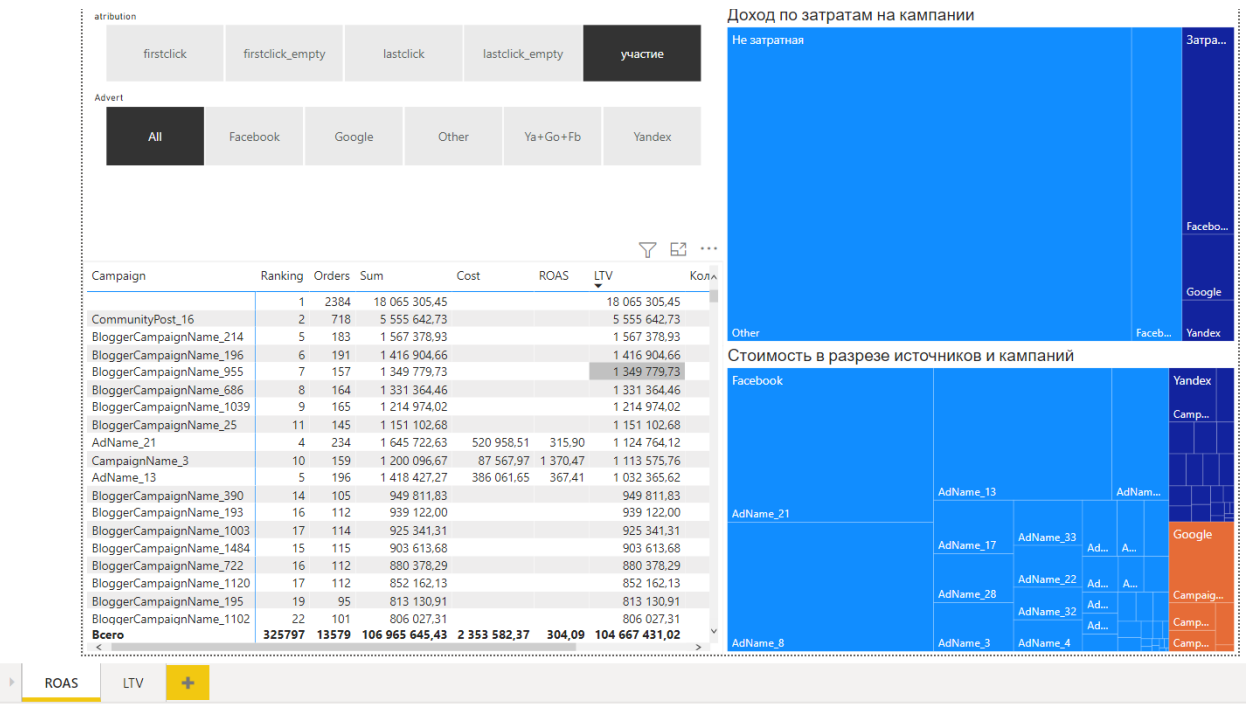
# Решение.

1. Опишите как определяете эффективность рекламы

По LTV в разрезе участия. LTV рассчитываю как сумму взвешенных по кликам разностей доходов и расходов на взаимодействие.

В антирейтинг не включаю кампании, которые попадают в 10-процентную группу минимальных по доходу кампаний (доход менее 100 000 рублей или до 14 заказов) при затратах менее 40000 рублей (исходя из ROAS нижней трети оплачиваемых кампаний от 240).

2. Постройте отчёт для оценки эффективности рекламы с максимально доступной детализацией источников.



3. Выделите 20 самых эффективных и 20 самых неэффективных рекламных активностей (исходя из вашего определения). Используйте три модели атрибуции:

- Последнее касание (lastclick)
- Первое касание (firstclick)
- Участие (по касаниям до заказа, равный вес для каждого касания).

20 самых эффективных кампаний по участию:

	Campaign	Orders	Sum	Cost	ROAS	LTV
1	CommunityPost_16	718	5555643			5555643
2	BloggerCampaignName_214	183	1567379			1567379
3	BloggerCampaignName_196	191	1416905			1416905
4	BloggerCampaignName_955	157	1349780			1349780
5	BloggerCampaignName_686	164	1331364			1331364
6	BloggerCampaignName_1039	165	1214974			1214974
7	BloggerCampaignName_25	145	1151103			1151103
8	AdName_21	234	1645723	520958,5	315,9028	1124764
9	CampaignName_3	159	1200097	87567,97	1370,474	1113576
10	AdName_13	196	1418427	386061,7	367,4095	1032366
11	BloggerCampaignName_390	105	949811,8			949811,8
12	BloggerCampaignName_193	112	939122			939122
13	BloggerCampaignName_1003	114	925341,3			925341,3
14	BloggerCampaignName_1484	115	903613,7			903613,7
15	BloggerCampaignName_722	112	880378,3			880378,3
16	BloggerCampaignName_1120	112	852162,1			852162,1
17	BloggerCampaignName_195	95	813130,9			813130,9
18	BloggerCampaignName_1102	101	806027,3			806027,3
19	BloggerCampaignName_1481	101	794971,7			794971,7
20	BloggerCampaignName_1004	98	791085,7			791085,7

20 самых неэффективных кампаний:

	Campaign	Orders	Sum	Cost	ROAS	LTV
1	AdName_4	13	102055,1	35790,34	285,1471	66264,78
2	AdName_3	22	143289,6	64060,29	223,6793	79229,33
3	AdName_32	16	117301,1	36711,7	319,5197	80589,42
4	BloggerCampaignName_1238	14	100296,6			100296,6
5	EmailCampaignName_232	14	100902,8			100902,8
6	EmailCampaignName_111	14	102352,3			102352,3
7	BloggerCampaignName_869	11	103130,3			103130,3
8	EmailCampaignName_80	12	103423,1			103423,1
9	BloggerCampaignName_918	15	105455			105455
10	BloggerCampaignName_605	15	106367,5			106367,5
11	BloggerCampaignName_523	13	106914,4			106914,4
12	BloggerCampaignName_1041	16	108054,8			108054,8
13	BloggerCampaignName_407	14	108220,3			108220,3
14	EmailCampaignName_203	14	108286			108286
15	BloggerCampaignName_497	16	108572,8			108572,8
16	BloggerCampaignName_1316	13	110090,2			110090,2
17	BloggerCampaignName_1270	15	113098,7			113098,7
18	EmailCampaignName_161	14	113830,6			113830,6
19	BloggerCampaignName_563	15	114023,2			114023,2
20	BloggerCampaignName_122	13	116470,4			116470,4

- Для фейсбука использован уровень объявлений, для Google Ads и Яндекс Директа уровень рекламных кампаний.

Ссылка на интерактивный отчёт: <https://www.dropbox.com/sh/gt4z6j5ndsibbea/AADUGduxKm-r7SPmiWYkxXHba?dl=0>

- Показываю код и даю ссылку на юпитер ноутбук  
<https://www.dropbox.com/sh/gt4z6j5ndsibbea/AADUGduxKm-r7SPmiWYkxXHba?dl=0>

3. Сделайте когортный анализ, по неделям и(или) месяцам.

Сделано:

advert

facebook

google

nan

yandex

Advert, Campaign

☐ Facebook

☐ Google+Yandex

☐ Other

☐ Yandex

Campaign	21	28	35	42	49	56	63	70	77	84	91	Всего	
CommunityPost_16	403,72	381 562,12	245 584,43	226 152,91	201 265,08	168 371,47	93 671,13	53 544,85	63 710,23	21 162,69	7 652,62	3 997,79	18 065 305,45
BloggerCampaignName_214	961,58	213 242,36	211 049,84	130 679,08	140 260,25	132 897,38	82 829,81	62 537,76	50 034,26	17 538,58	4 152,81	1 155,52	5 555 642,73
BloggerCampaignName_196	329,99		86 185,03	3 853,92	10 232,57		6 065,64	3 373,38	4 221,23	25 433,91			1 567 378,93
BloggerCampaignName_955	5 236,71	6 310,04				2 211,12							1 416 904,66
BloggerCampaignName_686	7 718,94	48 184,80	5 854,70	28 086,53	57 527,76	15 581,57	14 737,44						1 349 779,73
BloggerCampaignName_1039	3 342,28	25 081,96	69 743,91	7 604,55	45 477,85	4 743,97	7 010,82						1 331 364,46
BloggerCampaignName_25	9 909,09	9 729,96	524,98										1 214 974,02
AdName_21	2 262,19	6 633,36	11 055,60				47 007,30	6 569,96	4 643,35	12 382,27			1 151 102,68
CampaignName_3	3 396,67	36 614,75	-2 426,58	9 535,13	-10 612,69	-48 179,86	-8 764,71	3 818,53					1 124 764,12
AdName_13	3 395,31	37 347,27	19 988,19	8 471,17	14 771,50	3 965,72	14 316,51	610,84		218,69	508,32		1 113 575,76
BloggerCampaignName_390	5 952,27	-24 685,55	380,79										1 032 365,62
BloggerCampaignName_193	8 662,12	29 405,55	8 291,19	34 514,74		661,45	16 167,50	5 626,86	11 577,69				949 811,83
BloggerCampaignName_1003	3 132,29	53 964,93	11 373,73	14 385,54	6 766,56	113 035,31		760,05	1 013,40	5 827,03			939 122,00
BloggerCampaignName_1484	9 996,11	2 339,67	74 304,44	2 890,44		6 592,41	1 266,88	40 196,26	8 694,14	9 037,38			925 341,31
BloggerCampaignName_722	1 221,50	1 607,56	2 678,26	1 628,67									903 613,68
BloggerCampaignName_1120	5 216,92	1 888,65	9 926,30	5 613,48	613,42								880 378,29
BloggerCampaignName_195		1 365,69	7 765,82	56 153,37	3 729,60	1 564,40							852 162,13
BloggerCampaignName_1102	3 537,39	1 427,74	5 173,09	82 421,99	979,61								813 130,91
BloggerCampaignName_1481	4 335,66	24 929,82	14 159,01		2 804,69	9 603,96	985,35						806 027,31
BloggerCampaignName_1004		4 757,79		6 527,62	696,57	4 643,35				6 331,84			794 971,66
BloggerCampaignName_103	3 370,09	7 865,16	18 749,02	33 006,69	15 572,71	11 764,57	128,47	4 908,50					791 085,73
Всего	2 204,12	13 892,06	6 930,87	3 612,37									780 277,14
Всего	509,77	2 749 308,08	2 311 628,23	1 657 551,97	1 282 067,97	1 223 421,57	932 284,04	413 267,30	479 800,92	278 836,13	94 098,50	25 291,30	104 667 431,02

ROAS LTV +

4. Если есть, дайте рекомендации, как можно улучшить сбор исходных данных для отчёта.

Необходимо доработать CRM для ведения списка устройств, под которыми пользователи входят в аккаунты для их отслеживания в аналитических отчётах. Также необходимо фиксировать данные устройства и utm-метки в ключевые моменты (первый заход, создание заказа, оплата заказа на сайте).

5. Напишите свои рекомендации для директора по маркетингу.

Затраты на рекламу очень низкие, на уровне 2% от дохода, необходимо повышать расходы. Выводы о платных кампаниях и источниках на малых долях могут не отражать реальность.

## Подробнее о решении.

Задача состояла из двух подзадач:

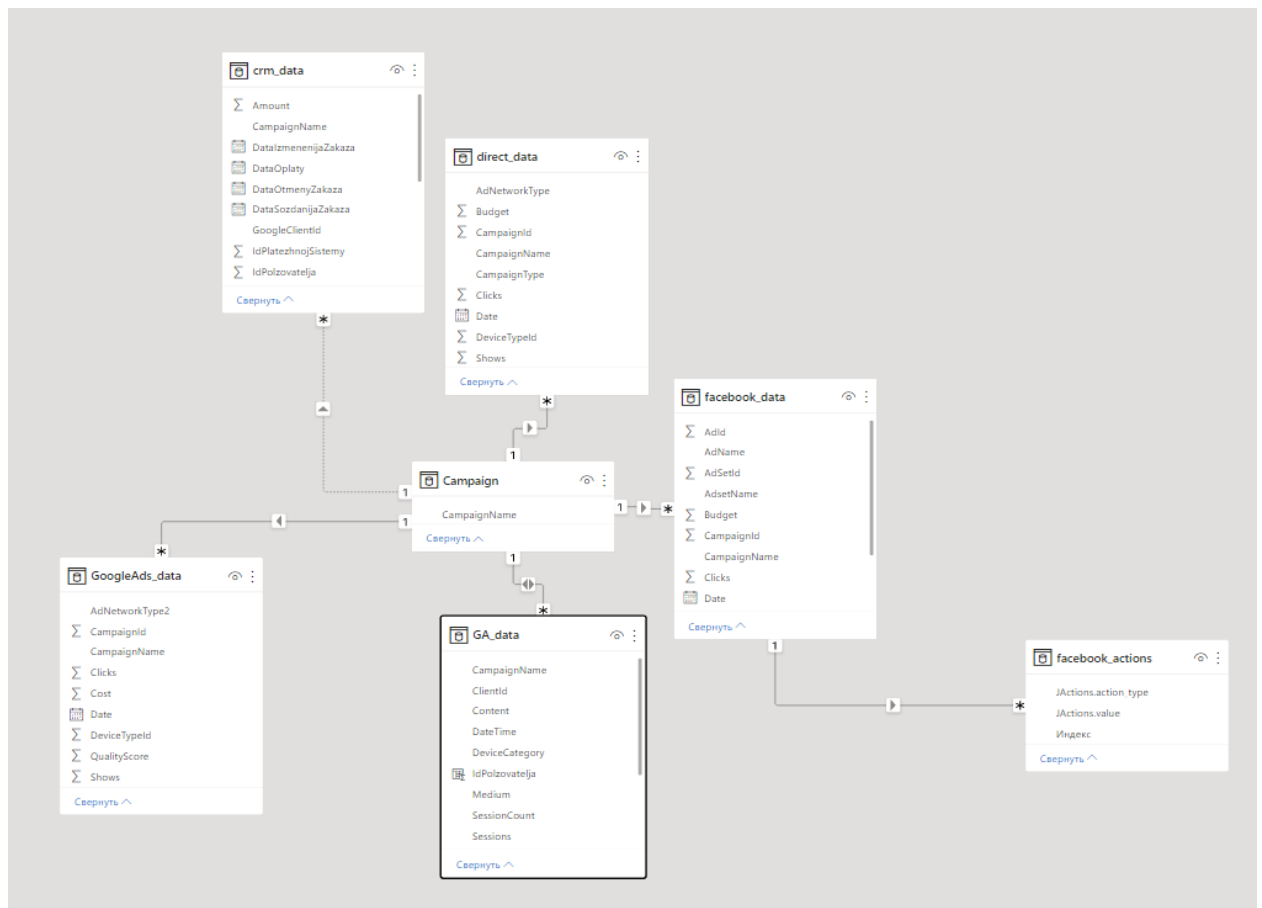
- очистка и сведение разрозненных данных программно и логически;
- выбор метрик и анализ данных.

Особенности решения:

1. Для связи заказов и данных гугл-аналитики применена группировка Google Client ID по трек-номерам на основе повторного использования в заказе и совпадения дат оплат.
2. Данные гугл аналитики заменены на производную таблицу: список всех действий пользователей перед заказом. При множественных заказах каждое действие продублировано по количеству последующих заказов и учитывается как влияющее на все последующие заказы.

В задании были даны данные CRM, Гугл-аналитики и рекламных кабинетов.

Изначальная структура данных не позволяла увязывать данные между собой.



**Данные CRM это список заказов** с датой создания строго в третьем квартале 2020 года. При этом даты оплаты, изменения и отмены встречаются до середины 2021 года. То есть, данные взяты по состоянию не ранее, чем через полгода после создания последнего заказа, следовательно:

1. при среднем времени до оплаты или отмены заказа равном 5 дней **статусы всех заказов следует считать финальными**
2. **данные об оплатах, изменениях и отменах заказов в третьем квартале по более ранним заказам следует считать существующими, но отсутствующими в этой таблице.**

Заказы в статусе StatusZakaza "Возврат заказа: ожидает отправки" без даты оплаты имеют только IdPlatezhnojSistemy равный 1 (из двух возможных), а с датой оплаты только

IdPlatezhnojSistemy равный 17. Следовательно, **платёжная система 17 означает оплату на сайте, а платёжная система 1 означает оплату при получении.**

Заказы с признаком оплаты всегда имеют сумму и дату оплаты, но никогда не имеют признака отмены заказа. Заказы с признаком оплаты имеют только статусы Выполнен, "Возврат заказа: ожидает отправки" (с оплатой на сайте - следует считать, что по факту возврата товара так и не произошло при наличии оплаты) и "Отказ от службы доставки" (30 случаев, даже в случае оплаты при получении - вероятно это связано с частным отказом от заказа). В 45 случаях заказы с признаком оплаты имеют даты отмены (без признака отмены заказа), но во всех этих случаях дата оплаты позже даты отмены, то есть решение об отмене в итоге было пересмотрено. Заказы без признака оплаты имеют суммы только у трёх заказов на сумму около 19000 рублей - по неизвестным причинам, в частном случае этим расхождением признаков можно пренебречь списав на техническую ошибку (в общем случае необходимо выяснение деталей заказов у техподдержки crm). Следовательно, **все заказы не тестовых аккаунтов с признаком оплаты следует считать оплаченными и успешно завершёнными.**

Заказы без признака оплаты в 7,7% случаев имеют статус "Выполнен", в подавляющем большинстве случаев у таких заказов оплата при получении. Вероятно, **с заказами без признака оплаты в статусе "Выполнен" возникли проблемы, но их статус не был актуализован.**

Для связи данных CRM составил процедуру, составляющую таблицу отслеживания пользователей на основе связей с аккаунтами и временем оплаты, которая позволила разметить заказы и данные аналитики увязав их между собой.

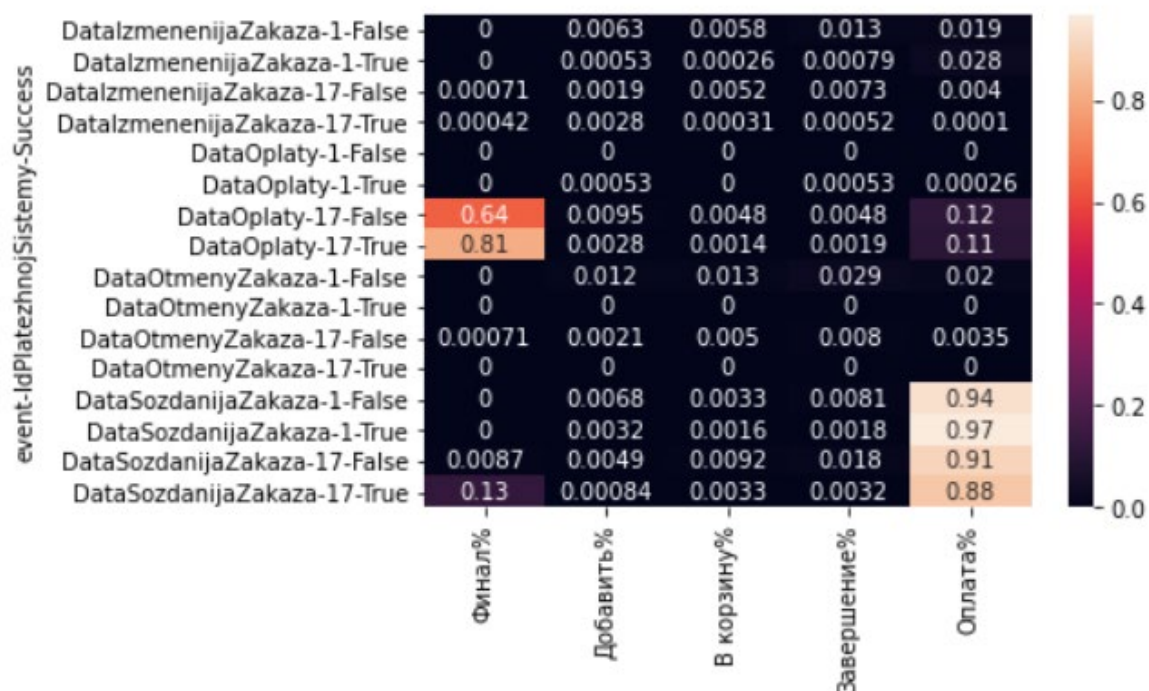
Создаёт таблицу отслеживания пользовательских устройств по аккаунтам (группирует пользовательские устройства). Одно или несколько устройств, использующих только один аккаунт получают в качестве id отслеживания (tid - track id) номер аккаунта. Устройства, использующие несколько аккаунтов, получают в качестве id отслеживания (tid) номер устройства (gcid).

Получает на вход dataset (crm\_data) содержащий соответствия устройств пользователей (gcid) и пользователей (pid), у которых они были замечены, а также перечень тестовых пользователей (test\_users) и перечень дополнительных gcid (ga\_data) без аккаунтов.

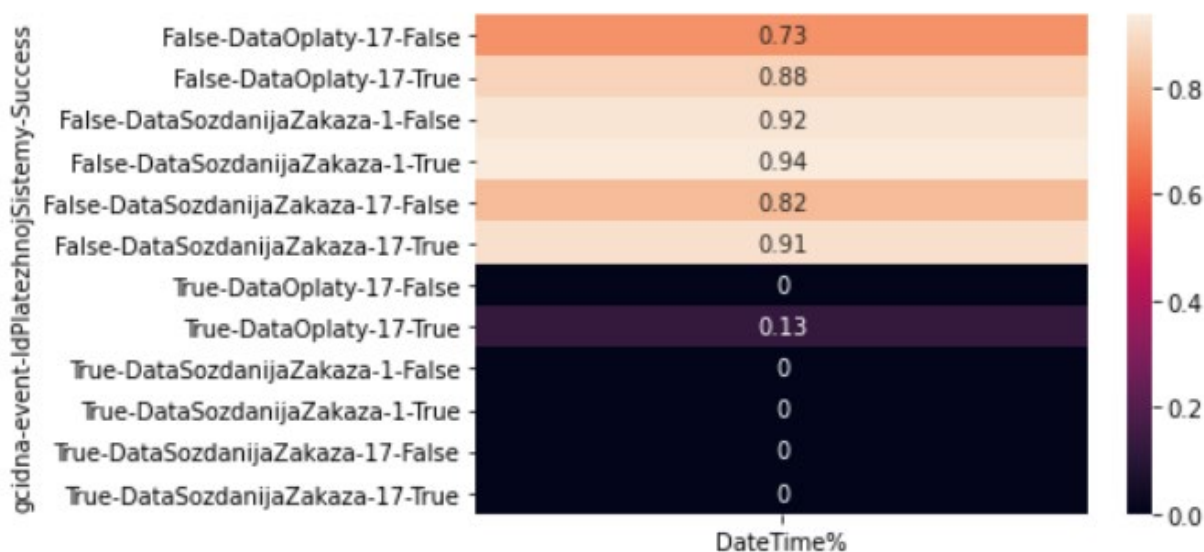
Возвращает соответствие уникальных устройств пользователей (gcid) и id отслеживания (tid - track id) с пометкой устройств, которые встречались только в тестовых аккаунтах (IsTestGCID).

Опция gamerge=True дополняет соответствия устройствами, которые совершили оплату в похожее на зафиксированное по заказу время и не подходящее под другие заказы.

Уточнение данных пользователей основано на совпадении времени оплаты на сайте с неопознанными пользователемскими устройствами:



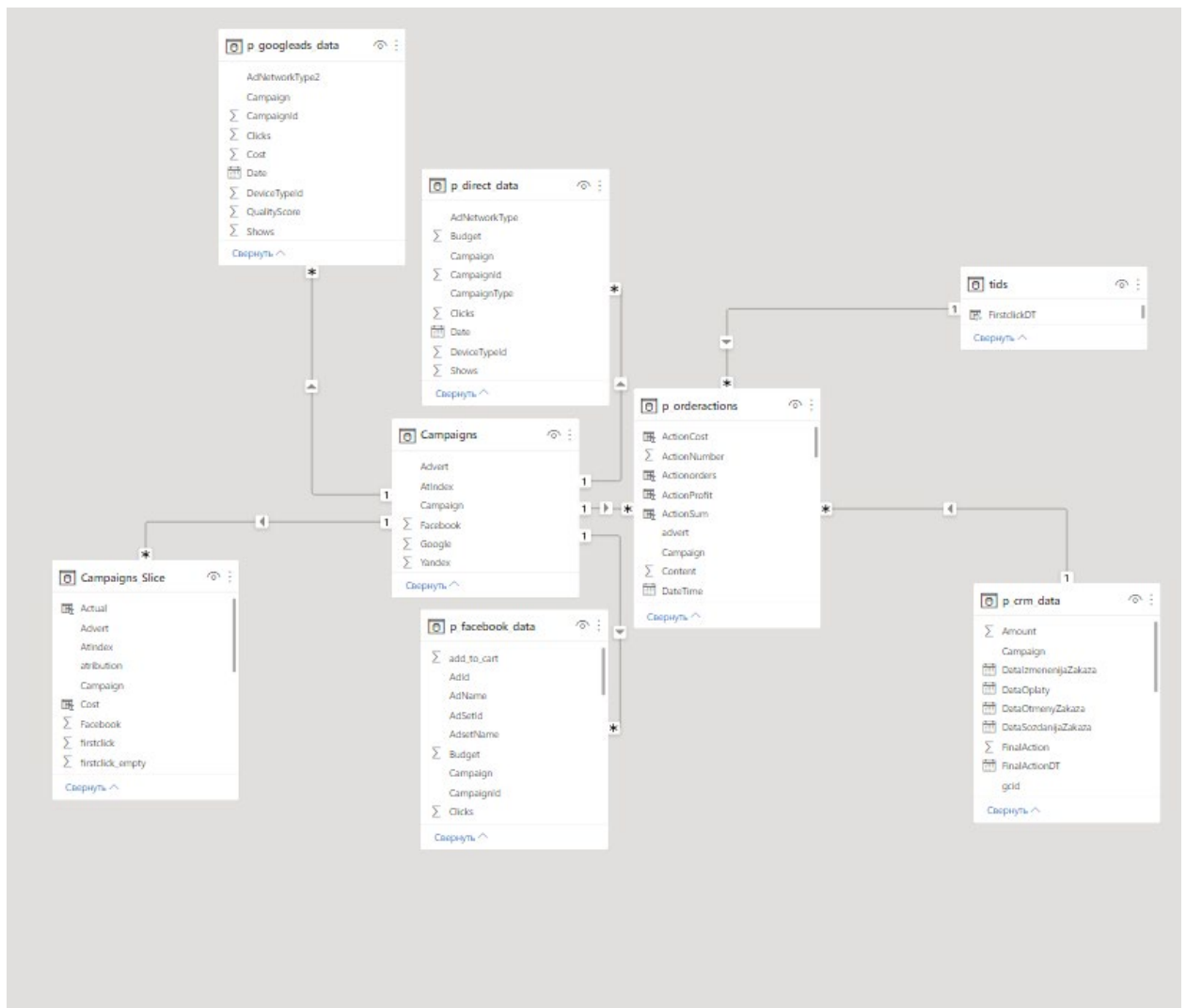
Количество обнаруживаемых оплат (действий пользователей) повышено с 81 до 86%:



Действия пользователей связаны с затратами в рекламных кабинетах Гугла и Яндексa через имя кампании.

Айди объявлений фейсбука обнаружены в данных Content, а в Campaign айди кампании. Айди кампании в Campaign заменён на имя объявления. Однозначное соответствие объявления и действия пользователя установлено через сочетание даты и имени/айди объявления. Информация из Iactions данных фейсбука оказалась бесполезной. Наиболее вероятная причина неудачи тот факт, что фейсбук по документации фиксирует все события на стороне сайта без привязки к конкретному времени и событию перехода, который может отсутствовать. То есть эти события включают переходы по другим рекламным кампаниям.

Итоговая структура данных:



Ссылки на инструменты просмотра:

1. anaconda/Jupyter Lab: <https://www.anaconda.com/products/individual>
2. Power BI: <https://www.microsoft.com/en-us/download/details.aspx?id=58494>