



# Анализ CRM-данных онлайн-школы программирования X

Цель исследования – повысить эффективность работы школы за счёт анализа данных, выявления слабых мест и поиска точек роста.

Anastasiia Radoveliuk

# 1. Подготовка и очистка данных

- Удалите дубликаты и неактуальные столбцы.
- Соответствующим образом обработайте отсутствующие значения.
- Преобразуйте типы данных для таких столбцов, как даты и числовые значения.

```
contacts.info()

class 'pandas.core.frame.DataFrame'
RangeIndex: 18548 entries, 0 to 18547
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----  
 0   Id          18548 non-null   object  
 1   Contact Owner Name 18548 non-null   category
 2   Created Time    18548 non-null   datetime64[ns]
 3   Modified Time   18548 non-null   datetime64[ns]
dtypes: category(1), datetime64[ns](2), object(1)
memory usage: 454.2+ KB
```

```
calls.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 95874 entries, 0 to 95873
Data columns (total 10 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----  
 0   Id          95874 non-null   object  
 1   Call Start Time 95874 non-null   datetime64[ns]
 2   Call Owner Name 95874 non-null   category
 3   CONTACTID    91941 non-null   object  
 4   Call Type     95874 non-null   category
 5   Call Duration (in seconds) 95874 non-null   int64  
 6   Call Status    95874 non-null   category
 7   Outgoing Call Status 95874 non-null   category
 8   Scheduled in CRM 95874 non-null   category
 9   _missing_contact 95874 non-null   bool  
dtypes: bool(1), category(5), datetime64[ns](1), int64(1), object(2)
memory usage: 3.5+ MB
```

```
deals.info()

<class 'pandas.core.frame.DataFrame'>
Index: 21593 entries, 0 to 21592
Data columns (total 26 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----  
 0   Id          21593 non-null   object  
 1   Deal Owner Name 21593 non-null   category
 2   Closing Date  14645 non-null   datetime64[ns]
 3   Quality       21593 non-null   category
 4   Stage         21593 non-null   category
 5   Lost Reason   21593 non-null   category
 6   Campaign      21593 non-null   category
 7   SLA          21593 non-null   timedelta64[ns]
 8   Content        21593 non-null   category
 9   Term          21593 non-null   category
 10  Source        21593 non-null   category
 11  Payment Type  21593 non-null   category
 12  Product       21593 non-null   category
 13  Education Type 21593 non-null   category
 14  Created Time  21593 non-null   datetime64[ns]
 15  Course duration 21593 non-null   Int8  
 16  Months of study 21593 non-null   Int8  
 17  Initial Amount Paid 4165 non-null   float64 
 18  Offer Total Amount 4185 non-null   float64 
 19  Contact Name   21532 non-null   object  
 20  City          21593 non-null   category
 21  Level of Deutsch 21593 non-null   category
 22  _open_deal    21593 non-null   bool  
 23  Payment Category 21593 non-null   object  
 24  Offer Category 21593 non-null   object  
 25  _SLA hours    21593 non-null   float64 
dtypes: Int8(2), bool(1), category(13), datetime64[ns](2), float64(3), object(4), timedelta64[ns](1)
memory usage: 2.3+ MB
```

```
spend.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 19862 entries, 0 to 19861
Data columns (total 8 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----  
 0   Date        19862 non-null   datetime64[ns]
 1   Source      19862 non-null   category
 2   Campaign    19862 non-null   category
 3   Impressions 19862 non-null   int64  
 4   Spend       19862 non-null   float64 
 5   Clicks      19862 non-null   int64  
 6   AdGroup     19862 non-null   category
 7   Ad          19862 non-null   category
dtypes: category(4), datetime64[ns](1), float64(1), int64(2)
memory usage: 727.1 KB
```

## Итог подготовленного датасета

- Calls:** 95,874 строк, 9 столбцов.
- Spend:** 19,862 строк, 6 столбцов.
- Deals:** 21,593 строк, 20 столбцов.
- Contacts:** 18,548 строк, 4 столбца.

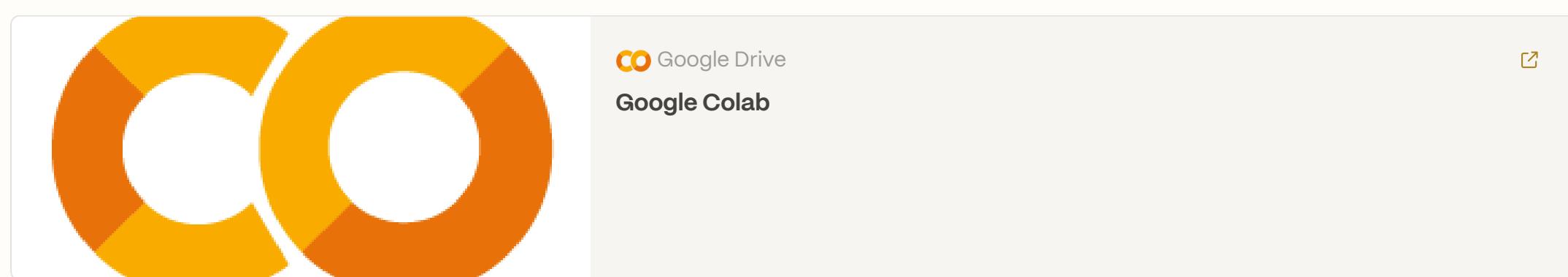
### Основные шаги:

- Удалены дубликаты и неактуальные поля.
- Пропуски → заполнены Unknown или 0/медианой.
- Даты → datetime, категории → category.
- Очистка City, Level of Deutsch, Payment Type.

### Выводы:

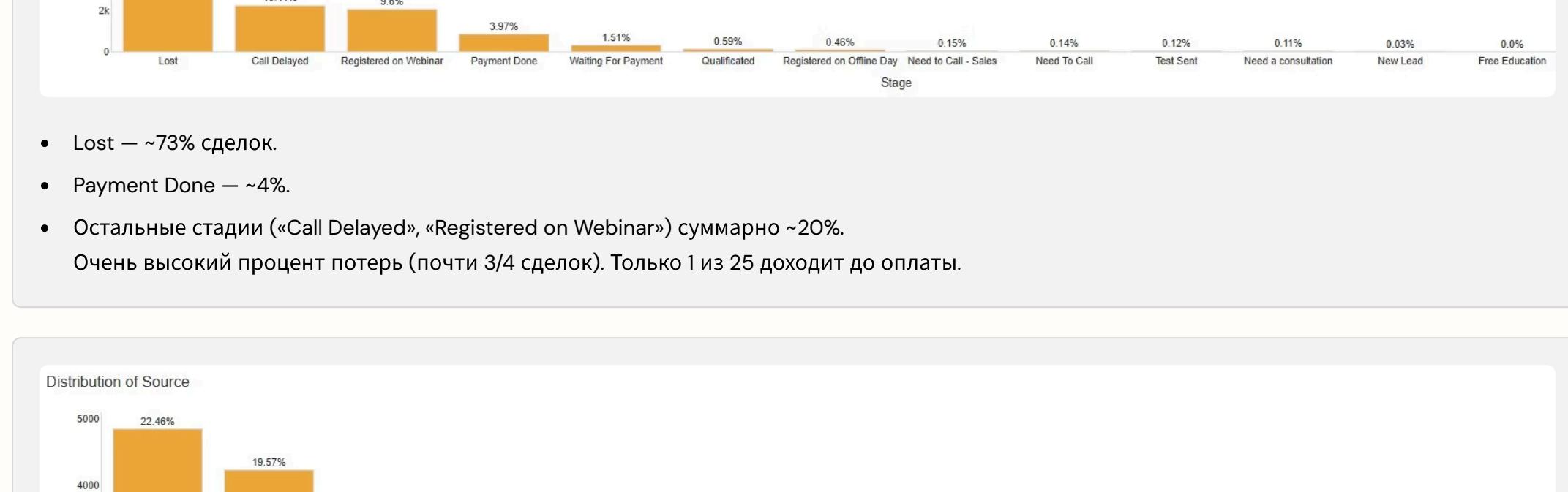
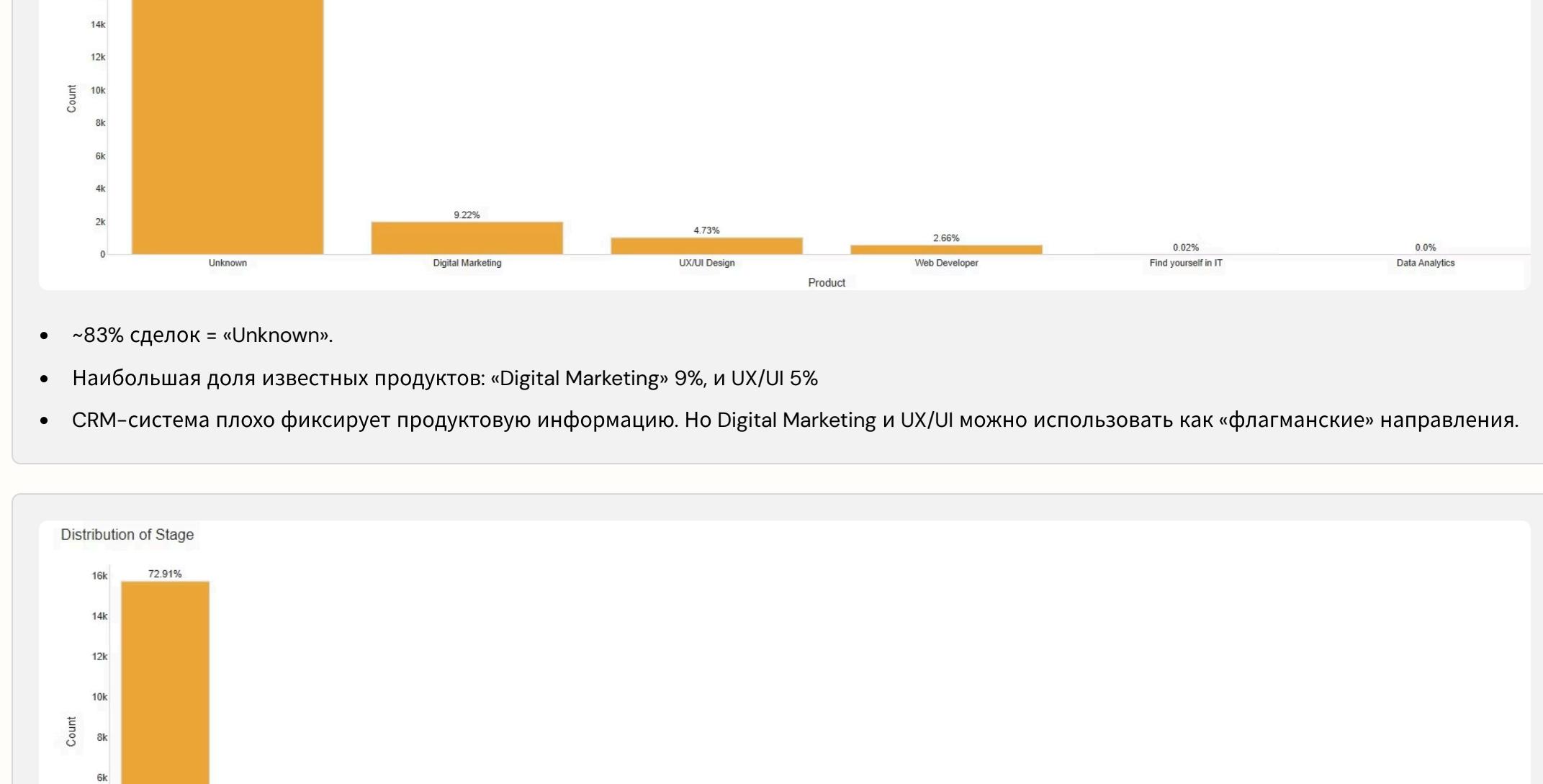
- Данные стандартизированы и готовы к анализу.
- Сокращены ошибки CRM, снижено влияние выбросов.

\* Подробное ознакомление с выполнением задания

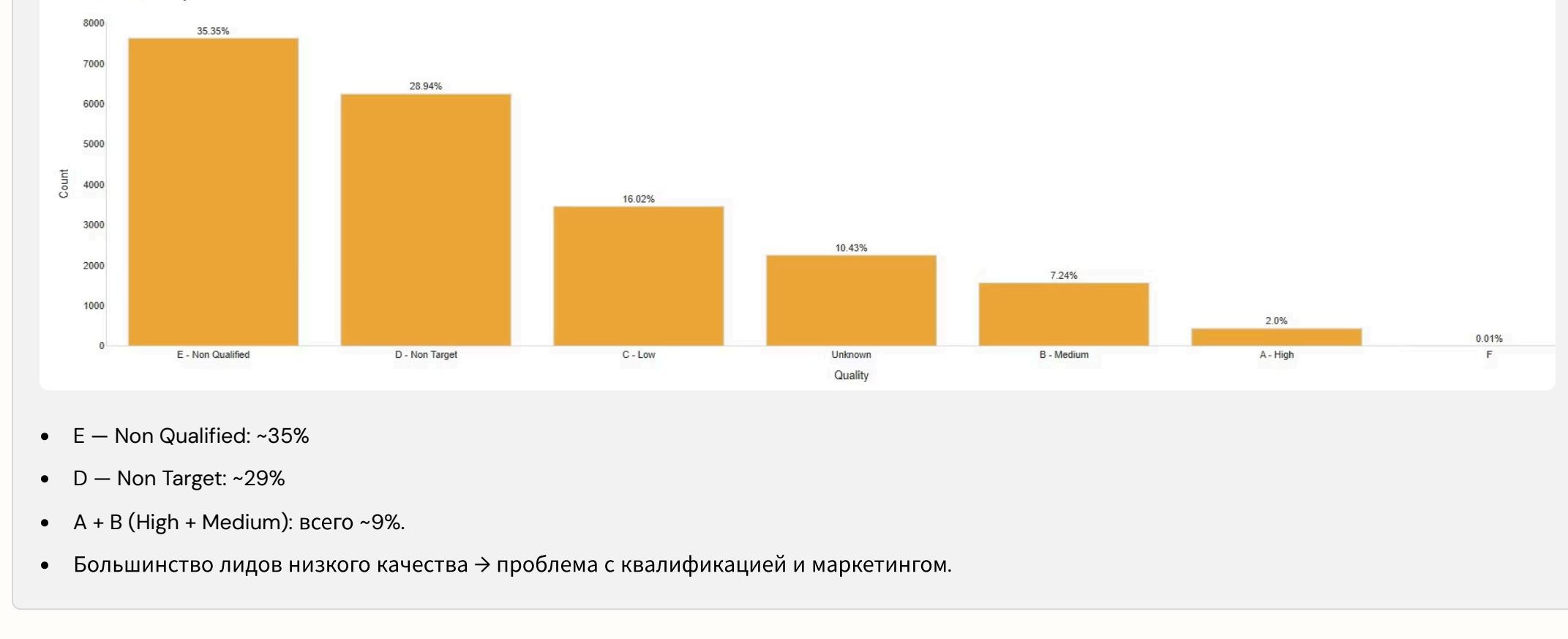


## 2. Описательная статистика

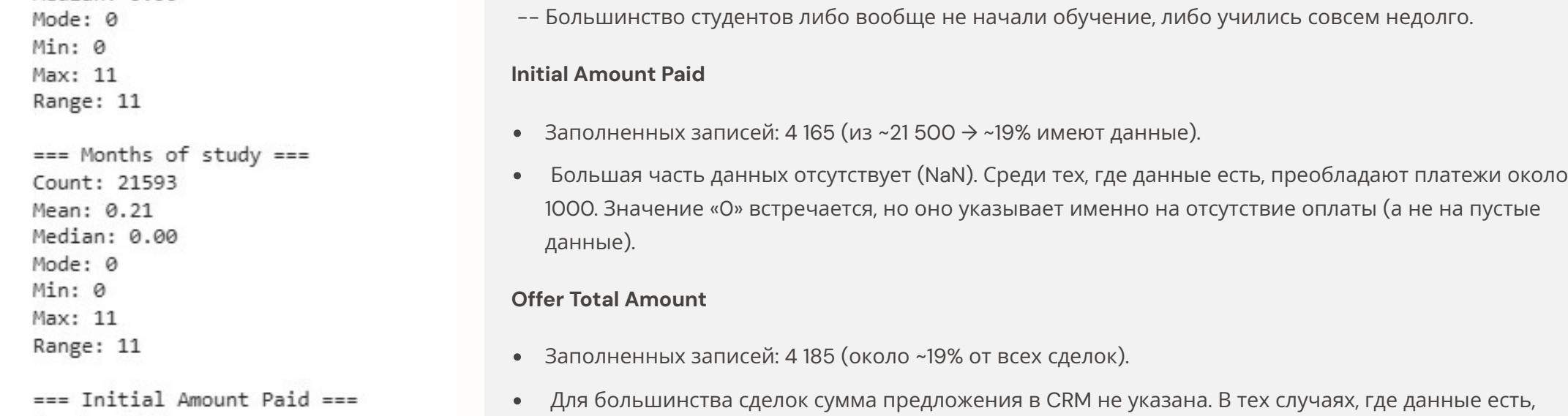
Цель – провести базовый исследовательский анализ (EDA), чтобы понять общие характеристики данных, выявить распределения, а также заметить аномалии и потенциальные проблемы.



- Lost – ~73% сделок.
- Payment Done – ~4%.
- Остальные стадии («Call Delayed», «Registered on Webinar») суммарно ~20%.
- Очень высокий процент потерь (почти 3/4 сделок). Только 1 из 25 доходит до оплаты.



- Основные источники распределены неравномерно, часто встречается «Unknown».
- Это снижает качество аналитики → необходимо улучшать заполнение в CRM.



- E – Non Qualified: ~35%
- D – Non Target: ~29%
- A + B (High + Medium): всего ~9%.
- Большинство лидов низкого качества → проблема с квалификацией и маркетингом.

### Course Duration

-- Большинство сделок имеют значение «0» (курс не стартовал), но есть сделки с полным циклом (11 мес). Это подтверждает высокую долю несостоявшихся студентов.

### Months of Study

-- Большинство студентов либо вообще не начали обучение, либо учились совсем недолго.

### Initial Amount Paid

- Заполненных записей: 4 165 (из ~21 500 → ~19% имеют данные).
- Большая часть данных отсутствует (NaN). Среди тех, где данные есть, преобладают платежи около 1000. Значение «0» встречается, но оно указывает именно на отсутствие оплаты (а не на пустые данные).

### Offer Total Amount

- Заполненных записей: 4 185 (около ~19% от всех сделок).
- Для большинства сделок сумма предложения в CRM не указана. В тех случаях, где данные есть, стандартное значение предложения фиксируется как 11 000. Значение «0» встречается, что можно трактовать как отсутствие финального оффера.

### Общий вывод

Важная находка: около 80% строк не содержат информации о суммах (NaN), что указывает на низкое качество заполнения CRM.

Там, где данные внесены, структура выглядит логично:

-- первый взнос обычно около 1000,

-- полный оффер – около 11 000.

В дальнейшем это позволит выделять сегменты: без оплаты / демо-доступ / полноценные платежи / отсутствуют данные.

### Группы причин

- Коммуникация (~35.8%):** Doesn't Answer + Stopped Answering → проблема взаимодействия с лидерами.
- Качество лидов (~34%):** Duplicate + Non target + Invalid number → слабый таргетинг и некачественные контакты.
- Оффер и цена (~11.1%):** Expensive + Conditions not suitable + Needs time to think → необходимость доработки ценностного предложения.
- Барьеры (~1–2%):** Does not speak English, Does not know how to use a computer → отдельные сегменты требуют иной коммуникации.

## Итог

Этап описательной статистики показал:

- Данные содержат выбросы и перекосы в распределениях.
- Наиболее проблемные места:
  - низкая конверсия воронки, потери на ранних стадиях
  - высокие доли Unknown в продуктах,
  - слабая валидация контактов,
  - низкий CTR кампаний
- Основные точки роста: улучшение качества данных, оптимизация рекламных каналов и работа над коммуникацией с клиентами.

Все эти выводы будут использованы в дальнейшем анализе временных рядов и эффективности кампаний.

\* Подробное ознакомление с выполнением задания



### 3. Анализ временных рядов

Цель – изучить динамику сделок и звонков во времени, а также изучить длительность сделок от момента создания до их закрытия.



#### Выводы

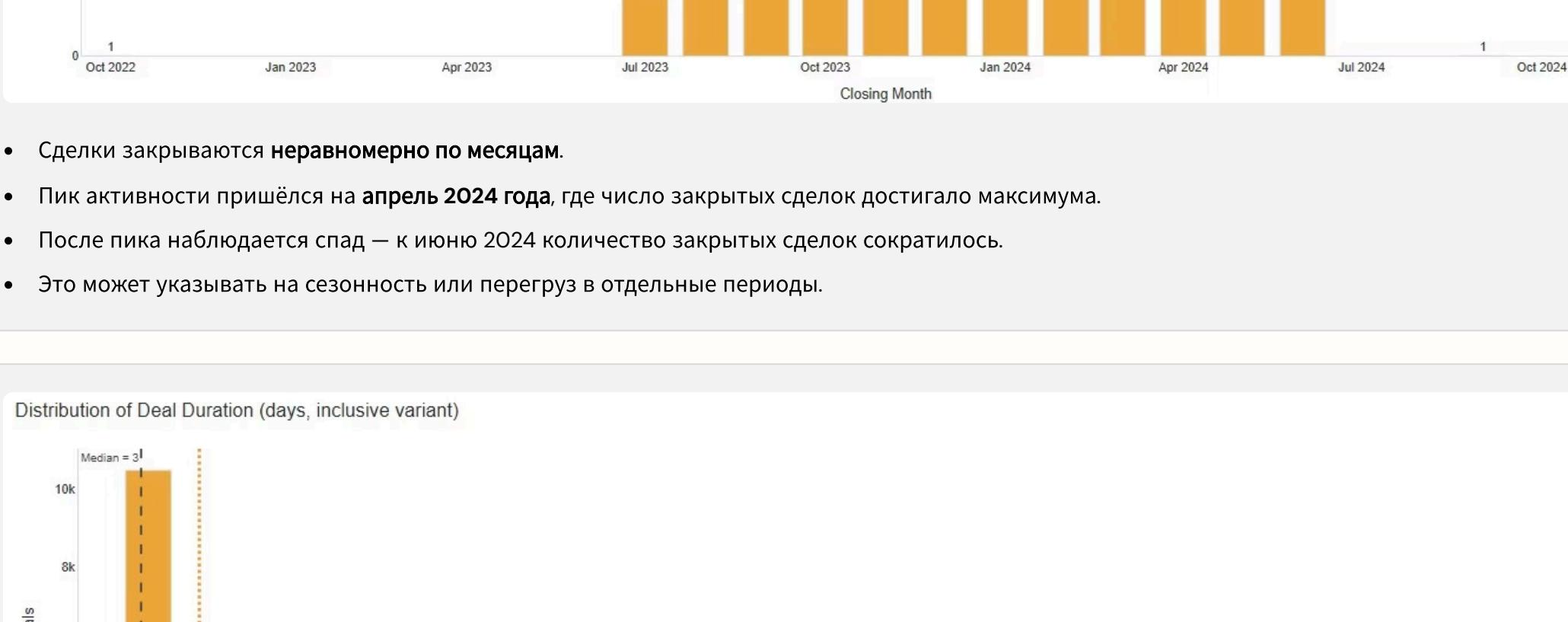
- Количество звонков и сделок активно росло до весны 2024 года, после чего пошёл спад.
- Пики сделок и звонков совпали в апреле 2024.
- Конверсия за год снизилась почти в 2 раза (с 0.34 до ~0.17).
- Несмотря на рост активности, эффективность обработки лидов падает.

#### Возможные причины

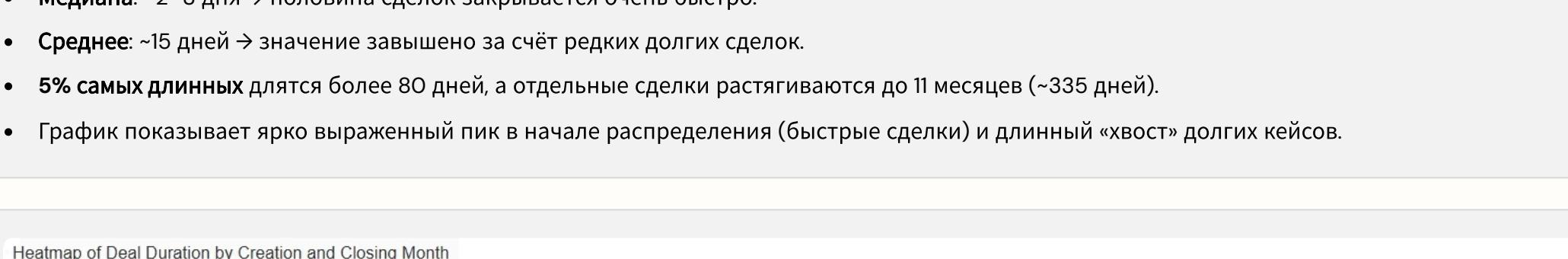
- Ухудшение качества лидов (особенно во 2-й половине периода).
- Перегруз менеджеров звонками – больше лидов, но меньше глубины обработки.
- Недостаточно эффективные скрипты и последующая работа с клиентами.

#### Рекомендации

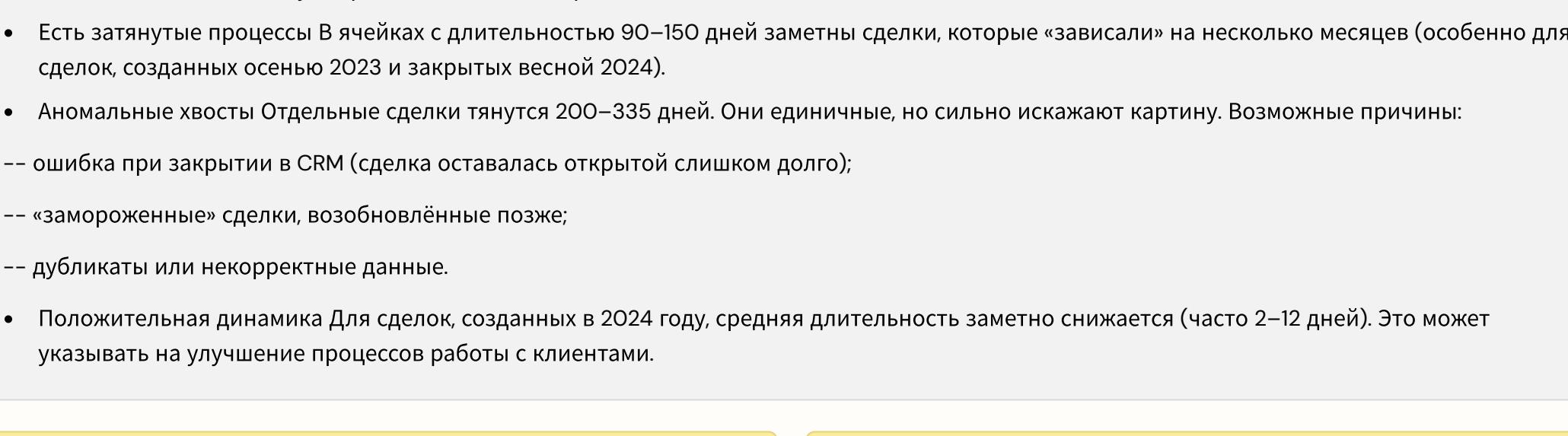
- Проанализировать источники лидов: какие каналы дают звонки без конверсии.
- Усилить обучение менеджеров и оптимизировать скрипты продаж.
- Добавить метрики качества: средняя длительность звонка, конверсия по источникам.
- В периоды пиковых нагрузок распределять звонки равномерно (например, автоматизация или доп. сотрудники).



- Сделки закрываются неравномерно по месяцам.
- Пик активности пришёлся на **апрель 2024 года**, где число закрытых сделок достигало максимума.
- После пика наблюдается спад – к июлю 2024 количество закрытых сделок сократилось.
- Это может указывать на сезонность или перегруз в отдельные периоды.



- Для анализа использован показатель Duration\_1day (same-day сделки считаются как 1 день).
- Медиана:** ~2–3 дня → половина сделок закрывается очень быстро.
- Среднее:** ~15 дней → значение завышено за счёт редких долгих сделок.
- 5% самых длинных** делятся более 80 дней, а отдельные сделки растягиваются до 11 месяцев (~335 дней).
- График показывает ярко выраженный пик в начале распределения (быстрые сделки) и длинный «хвост» долгих кейсов.



- Большинство сделок закрываются быстро. Основная масса сделок завершается в срок 3–30 дней. Это видно на диагонали карты: чем ближе месяц создания к месяцу закрытия, тем меньше средняя длительность.
- Есть затянутые процессы. В ячейках с длительностью 90–150 дней заметны сделки, которые «зависали» на несколько месяцев (особенно для сделок, созданных осенью 2023 и закрытых весной 2024).
- Аномальные хвосты. Отдельные сделки тянутся 200–335 дней. Они единичные, но сильно искажают картину. Возможные причины:
  - ошибка при закрытии в CRM (сделка оставалась открытой слишком долго);
  - «замороженные» сделки, возобновлённые позже;
  - дубликаты или некорректные данные.
- Положительная динамика. Для сделок, созданных в 2024 году, средняя длительность заметно снижается (часто 2–12 дней). Это может указывать на улучшение процессов работы с клиентами.

#### Insights

- Большинство сделок закрываются оперативно (в течение первой недели).
- Среднее сильно искажено выбросами, поэтому **для бизнес-анализа лучше ориентироваться на медиану**.
- Длинный хвост сделок указывает либо на **особо сложных клиентов**, либо на **ошибки в заполнении CRM**.
- Пики по месяцам показывают, когда команда работала с максимальной нагрузкой (апрель 2024) → это может влиять на рост средней длительности.
- Method sensitivity: выбор подсчёта same-day влияет на медиану (+1 день), но картина не меняется — быстрые сделки доминируют.

#### Recommendations

- Мониторинг SLA:** внедрить контроль по длительным сделкам (например, >30 дней).
- Управление нагрузкой:** пики апреля могут означать нехватку ресурсов — стоит распределять поток клиентов равномернее.
- Data quality:** проверить самые длинные сделки (>200 дней), чтобы убедиться, что это не ошибки в данных.
- Рассмотреть возможность увеличения штата/автоматизации в период высокого спроса.

\* Подробное ознакомление с выполнением задания



colab.research.google.com

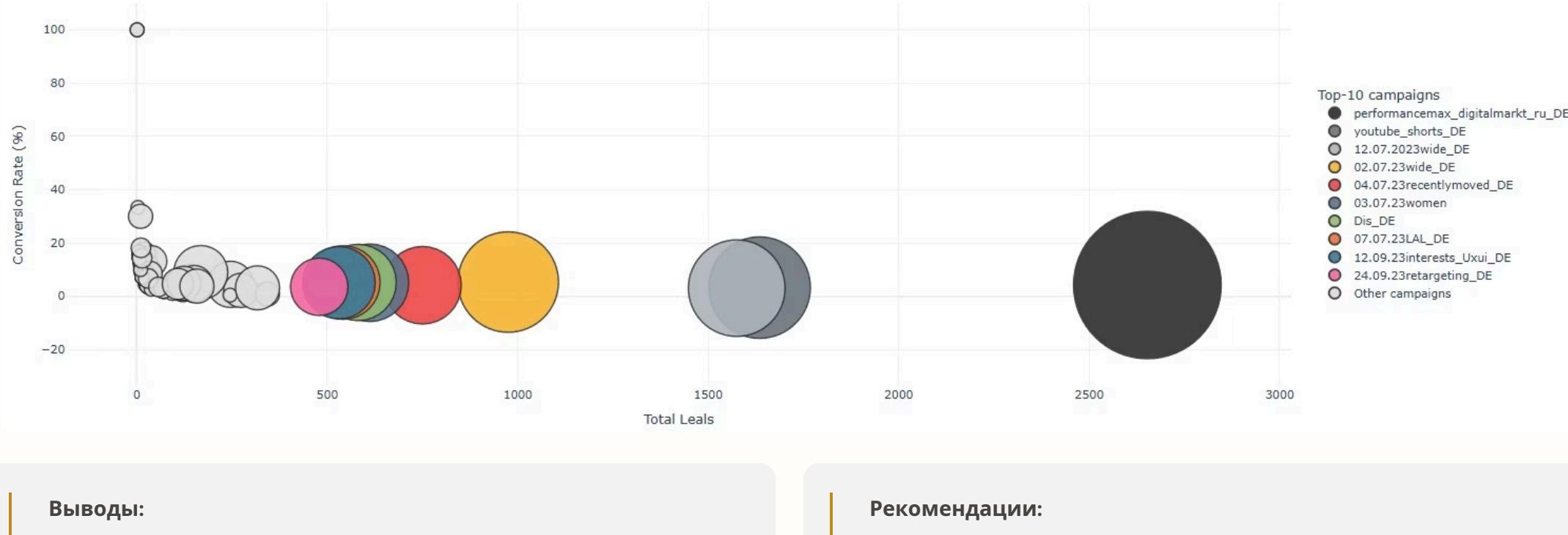
Google Colab



Made with GAMMA

## 4. Анализ кампаний и источников

**Цель:** оценить результативность маркетинговых кампаний и источников трафика с точки зрения генерации лидов и успешных сделок, а также определить наиболее эффективные каналы.



### Выводы:

- Кампания performancemax\_digitalmarkt\_ru\_DE – лидер по объёму лидов (>2600), но конверсия низкая (4.2%). Это «лидогенератор», но с большими потерями.
- Кампании youtube\_shorts\_DE и wide\_DE также показали высокий объём (1600+), но конверсия ещё ниже (3–3.2%).
- Есть кампании с малым объёмом, но более высокой конверсией (>5%), например brand\_search\_eng\_DE, что указывает на более точное попадание в аудиторию.

### Рекомендации:

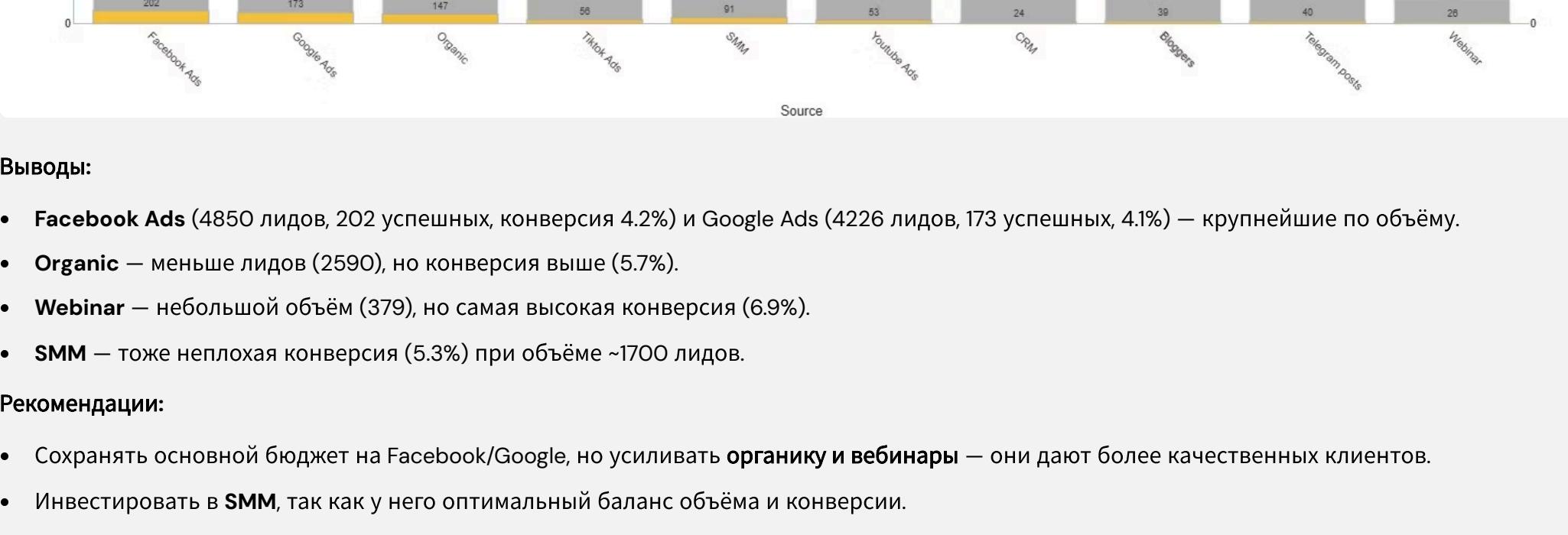
- Масштабировать кампании с высокой конверсией даже при меньших объёмах (ниша, точное попадание).
- Оптимизировать лидогенераторы с низкой конверсией: улучшить сегментацию, корректировать креативы и работу с лидами.
- Убрать/заморозить кампании с крайне низкой конверсией (<3%) и низким объёмом.



Некоторые крупные кампании работают как генераторы трафика, но почти не конвертируют.

### Рекомендации:

- Поддерживать баланс: часть бюджета направлять на широкие кампании ради объёма, часть – на точные ради качества.
- Проводить А/В тесты креативов для крупных кампаний с низкой конверсией.



### Выводы:

- Facebook Ads** (4850 лидов, 202 успешных, конверсия 4.2%) и **Google Ads** (4226 лидов, 173 успешных, 4.1%) – крупнейшие по объёму.
- Organic** – меньше лидов (2590), но конверсия выше (5.7%).
- Webinar** – небольшой объём (379), но самая высокая конверсия (6.9%).
- SMM** – тоже неплохая конверсия (5.3%) при объёме ~1700 лидов.

### Рекомендации:

- Сохранять основной бюджет на Facebook/Google, но усиливать **органику и вебинары** – они дают более качественных клиентов.
- Инвестировать в **SMM**, так как у него оптимальный баланс объёма и конверсии.
- Источники с конверсией <3% анализировать и при необходимости отключать.

## Итоговые рекомендации

### Кампании:

- Оптимизировать крупные низкоконверсионные кампании.
- Масштабировать точечные кампании с >5% конверсии.
- Постоянно мониторить и пересматривать стратегию «широкие vs нишевые».

### Источники:

- Основные лидогенераторы – Facebook и Google.
- Лидеры по качеству – Organic, Webinar, SMM.
- Оптимизировать бюджет в пользу более эффективных источников.

\* Подробное ознакомление с выполнением задания



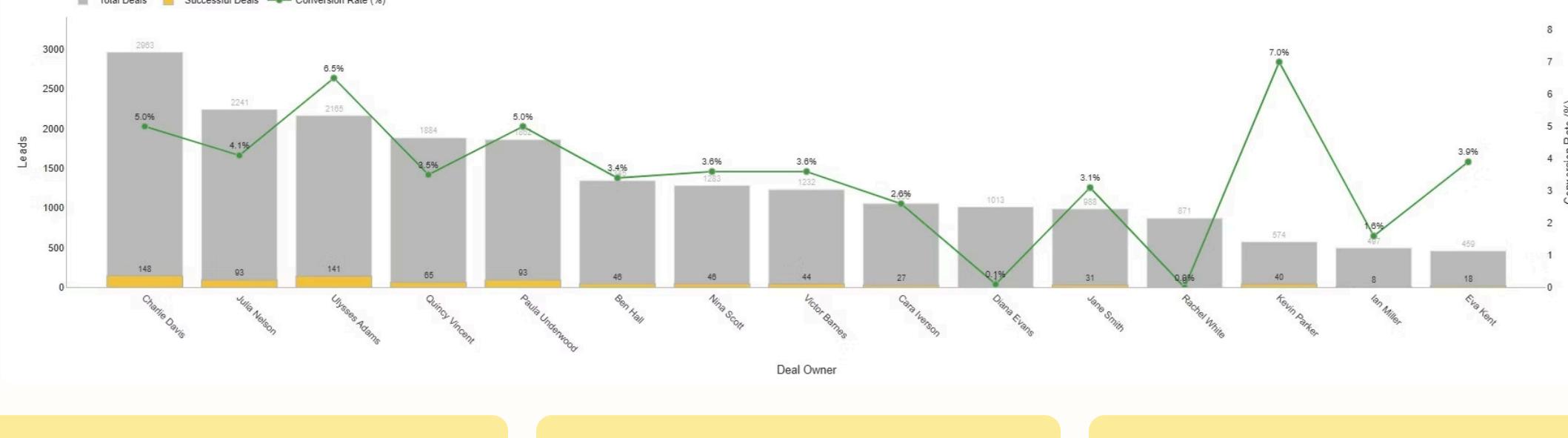
colab.research.google.com



Made with GAMMA

## 5. Работа отдела продаж

Цель: оценить эффективность сотрудников (владельцев сделок) и кампаний с точки зрения количества обработанных сделок, коэффициента конверсии и общей суммы продаж.



### Leads volume:

- Charlie Davis (2,963 leads) and Julia Nelson (2,241 leads) лидеры по количеству обработанных сделок.
- Однако их conversion rate сравнительно низкий (4–5%), что снижает общую эффективность.

### Revenue leaders:

- Ulysses Adams** — \$5.1M total sales.  
**Charlie Davis** — \$3.8M total sales.  
**Julia Nelson** — \$3.6M total sales.

### High conversion:

- Oliver Taylor показывает лучший результат по конверсии (30.7%) при сравнительно небольшом объеме (163 leads).
- Kevin Parker (6.9%) и Ulysses Adams (6.5%) также выше среднего уровня (~4–5%).

Rank	Deal Owner Name	total_leads	completed_leads	failed_leads	total_sales	total_offer_amount	conversion_rate	avg_ticket
5	Charlie Davis	2963	148	2126	445600.0	3822500.0	4.994938	3010.810811
12	Julia Nelson	2241	93	1536	382961.0	3575311.0	4.149933	4117.860215
21	Ulysses Adams	2165	141	1411	541050.0	5117800.0	6.512702	3837.234043
18	Quincy Vincent	1884	65	1379	221601.0	1788900.0	3.450106	3409.246154
17	Paula Underwood	1862	93	1318	326750.0	2842000.0	4.994629	3513.440860
2	Ben Hall	1345	46	887	241700.0	2030800.0	3.420074	5254.347826
15	Nina Scott	1283	46	865	206150.0	1951400.0	3.585347	4481.521739
23	Victor Barnes	1232	44	821	348900.0	2196800.0	3.571429	7929.545455
4	Cara Iverson	1056	27	710	688400.0	912000.0	2.556818	25496.296296
6	Diana Evans	1013	1	884	50450.0	562500.0	0.098717	50450.000000
10	Jane Smith	988	31	892	140050.0	1468500.0	3.137652	4517.741935
19	Rachel White	871	0	813	14000.0	44500.0	0.000000	NaN
13	Kevin Parker	574	40	492	72850.0	741400.0	6.968641	1821.250000
9	Ian Miller	497	8	451	33050.0	299000.0	1.609658	4131.250000
7	Eva Kent	459	18	221	65200.0	554300.0	3.921569	3622.222222
3	Bob Brown	337	0	337	950.0	13500.0	0.000000	NaN
14	Mason Roberts	268	3	229	19300.0	222000.0	1.119403	6433.333333
16	Oliver Taylor	163	50	100	152650.0	1660500.0	30.674847	3053.000000
8	George King	94	2	87	2900.0	33500.0	2.127660	1450.000000
26	Yara Edwards	85	0	51	0.0	0.0	0.000000	NaN
20	Sam Young	67	0	58	0.0	0.0	0.000000	NaN
1	Amy Green	66	0	32	0.0	0.0	0.000000	NaN
22	Unknown	29	1	6	1000.0	11000.0	3.448276	1000.000000
0	Alice Johnson	25	0	25	0.0	0.0	0.000000	NaN
11	John Doe	20	1	7	4600.0	18500.0	5.000000	4600.000000
25	Xander Dean	3	0	3	0.0	0.0	0.000000	NaN
24	Wendy Clark	2	0	2	0.0	0.0	0.000000	NaN
27	Zachary Foster	1	0	0	0.0	0.0	0.000000	NaN

- Несколько сотрудников имеют очень низкую конверсию 1%.

== Campaigns – Top by Leads ==

Campaign	total_leads	completed_leads	conversion_rate	total_sales
performancemax_digitalmarkt_ru_DE	2653	112	4.2	422900.0
youtube_shorts_DE	1635	53	3.2	253300.0
12.07.2023wide_DE	1575	48	3.0	277000.0

== Campaigns – Top by Conversion (volume ≥ 100) ==

Campaign	total_leads	completed_leads	conversion_rate	total_sales
brand_search_eng_DE	168	15	8.9	114050.0
02.07.23wide_DE	975	52	5.3	195950.0
07.07.23LAL_DE	542	28	5.2	126800.0

### By leads volume:

- performancemax\_digitalmarkt\_ru\_DE (2,653 leads) – главный генератор лидов, но низкая конверсия (4.2%).
- youtube\_shorts\_DE (1,635 leads) и wide\_DE (1,575 leads) – также лидеры по количеству, но конверсия лишь 3–3.2%.

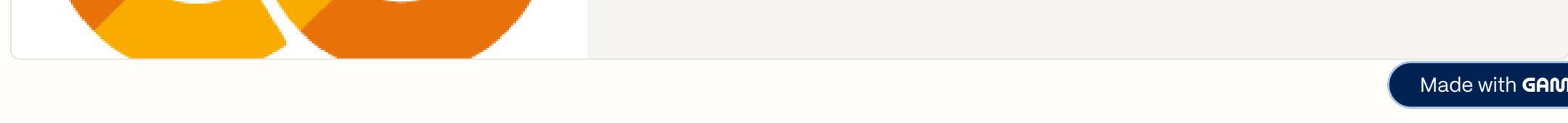
### By conversion rate (≥100 leads):

- brand\_search\_eng\_DE (8.9%), wide\_DE (02.07.23) (5.3%), LAL\_DE (5.2%).
- Эти кампании привлекают меньше лидов, но качество выше.

## Рекомендации

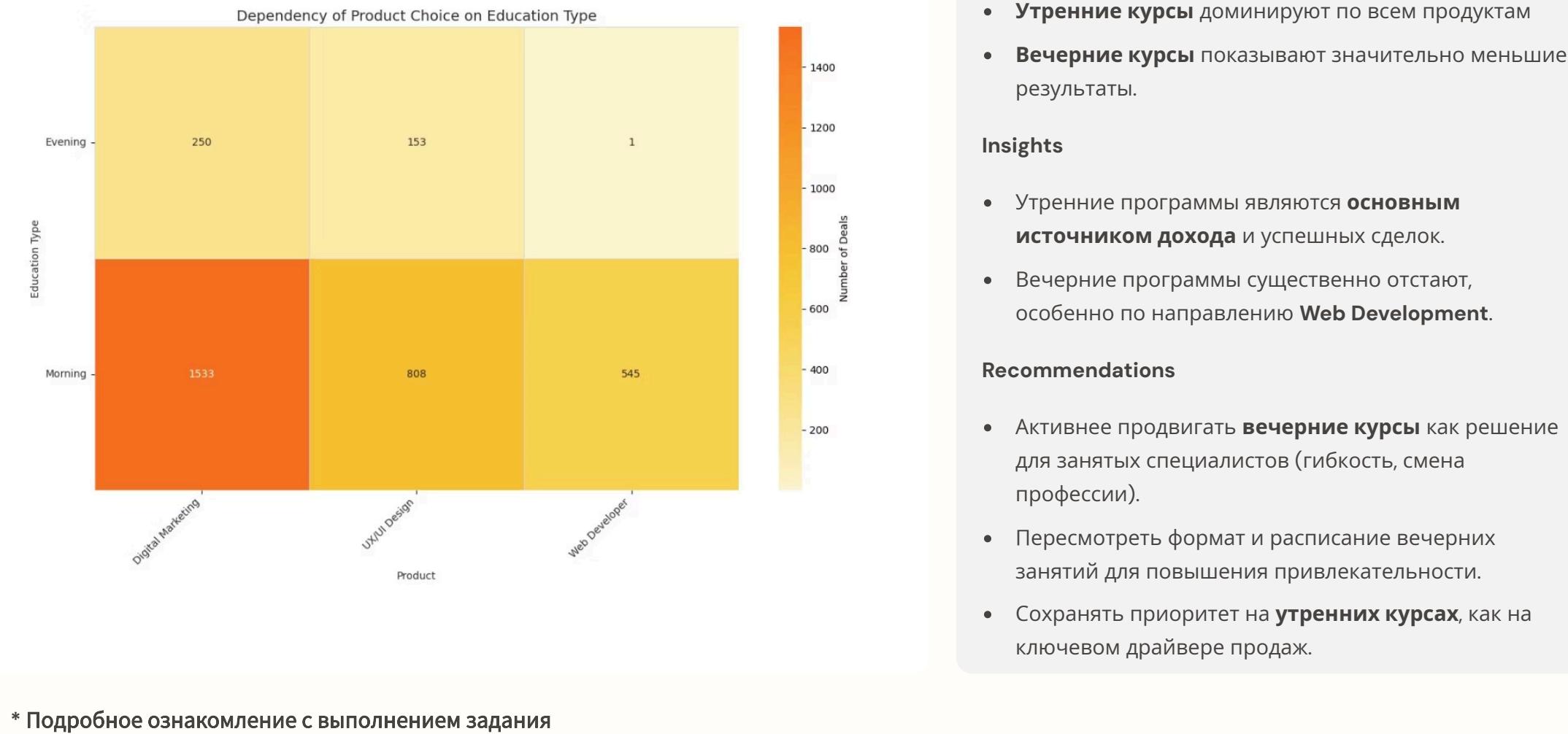
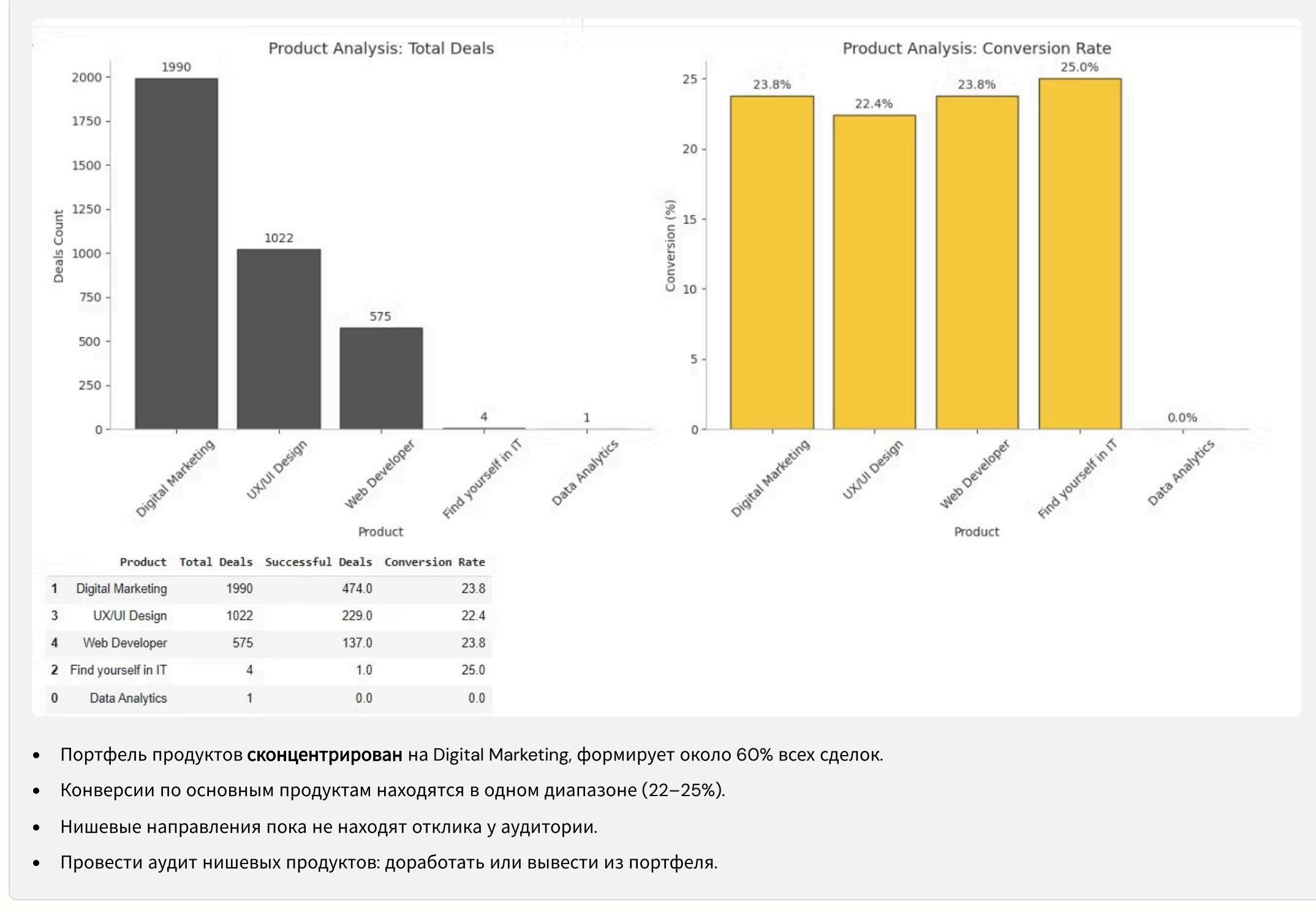
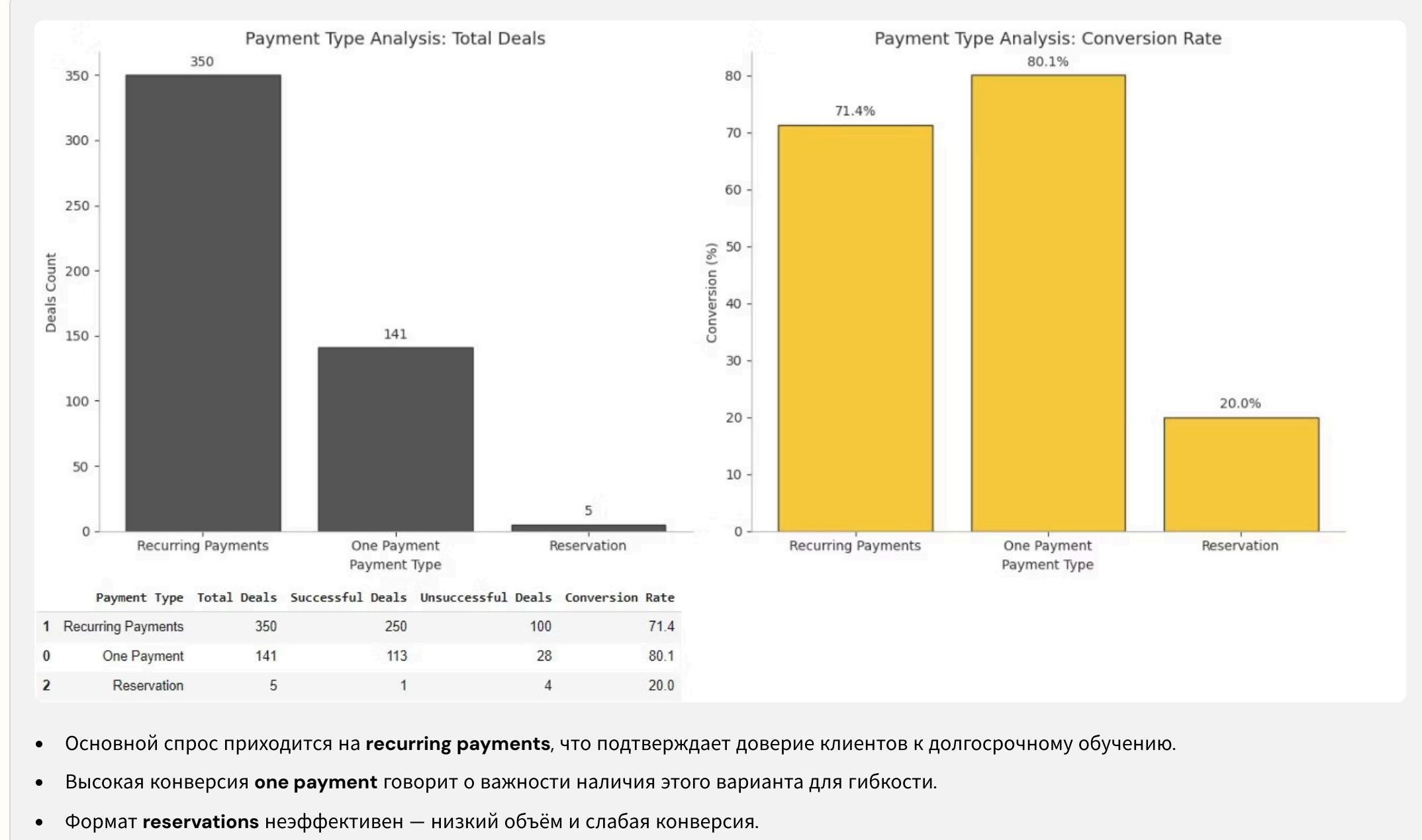
- Провести внутренний аудит отдела продаж: выявить лучших сотрудников и перенять их практики.
- Обучить менеджеров с низкой результативностью или перераспределить их задачи.
- Внедрить более равномерное распределение сделок между сотрудниками, чтобы снизить перегрузку отдельных менеджеров.
- Сосредоточить маркетинговый бюджет на кампаниях, которые показывают высокую конверсию и выручку.
- Использовать дашборды для регулярного мониторинга KPI каждого сотрудника и кампании.
- Увеличить бюджеты для brand\_search\_eng\_DE, LAL\_DE.
- Снять нагрузку с Davis/Nelson и передать часть лидов более эффективным сотрудникам.
- Провести обучение для менеджеров с низкой конверсией.
- Использовать успешные кейсы для увеличения среднего чека в других командах.

\* Подробное ознакомление с выполнением задания



## 6. Анализ платежей и продуктов

Цель: изучить распределение типов оплаты, их влияние на успешность сделок, а также проанализировать популярность и результативность различных продуктов и форматов обучения.



\* Подробное ознакомление с выполнением задания



colab.research.google.com

Google Colab



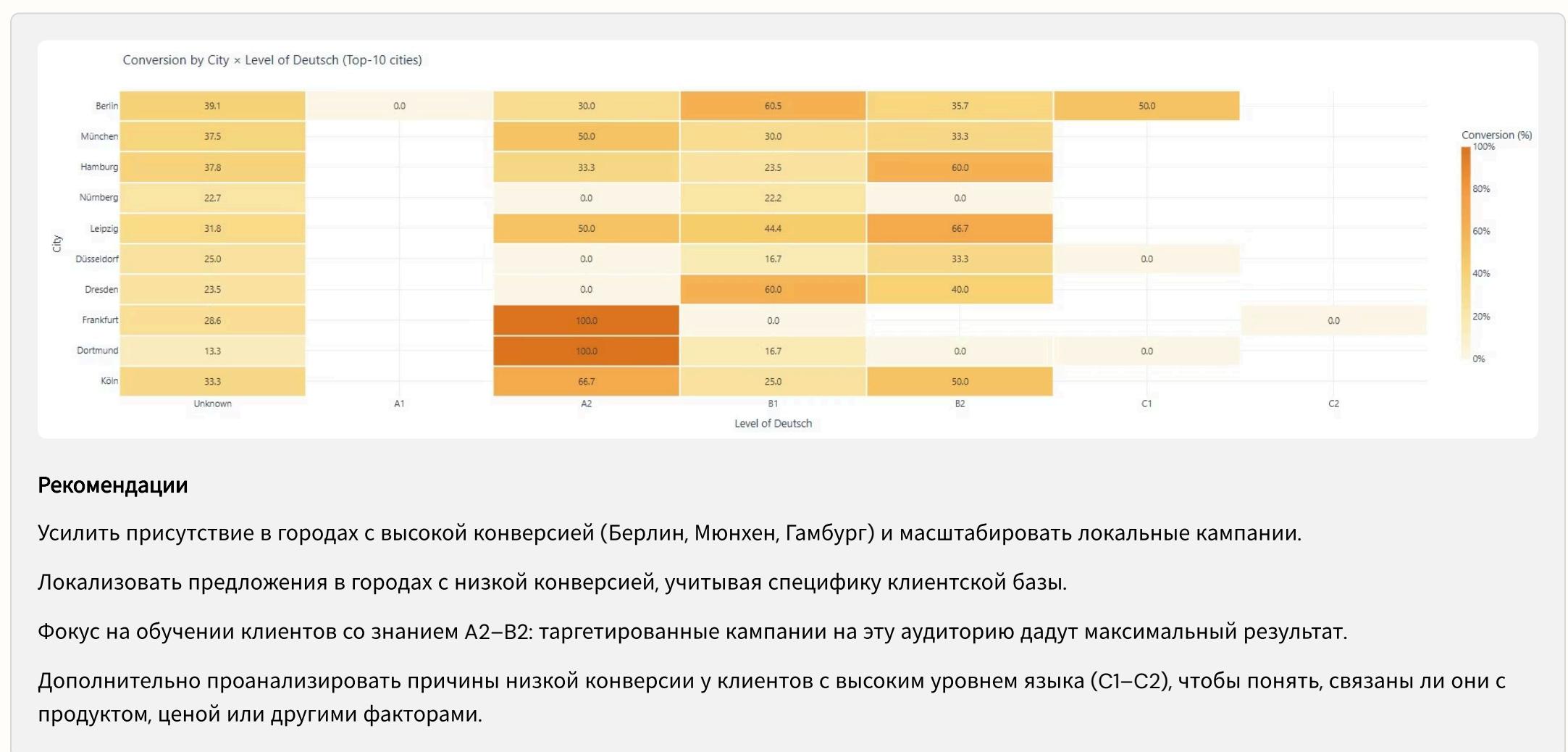
# 7. Географический анализ

Цель: изучить распределение сделок по городам, а также влияние уровня владения немецким языком на успешность сделок в разных регионах.



## Распределение сделок по городам

- Топ-10 городов формируют основную часть сделок.
- Берлин лидирует по объему: 182 сделки (43% конверсии).
- Мюнхен, Гамбург, Лейпциг также показывают значительные объемы и конверсии выше среднего уровня.
- Есть города с низкой конверсией (Дюссельдорф – 21%, Дортмунд).



## Рекомендации

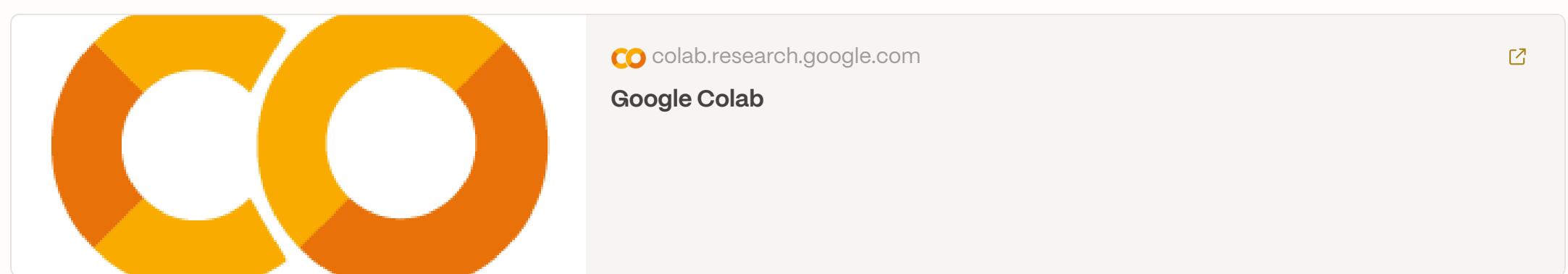
Усилить присутствие в городах с высокой конверсией (Берлин, Мюнхен, Гамбург) и масштабировать локальные кампании.

Локализовать предложения в городах с низкой конверсией, учитывая специфику клиентской базы.

Фокус на обучении клиентов со знанием A2–B2: таргетированные кампании на эту аудиторию дадут максимальный результат.

Дополнительно проанализировать причины низкой конверсии у клиентов с высоким уровнем языка (C1–C2), чтобы понять, связаны ли они с продуктом, ценой или другими факторами.

\* Подробное ознакомление с выполнением задания



# Продуктовая аналитика

Цель: найти точки роста бизнеса онлайн-школы на основе расчёта юнит-экономики и анализа продуктовых метрик, сформулировать гипотезы для улучшения ключевых показателей и описать механику их тестирования.

## 1. Расчёт юнит-экономики по продуктам

	number of acquired potential customers	converting potential customers into paying customers	number of customers received	average order value			average payment count per one customer	gross profit per customer (customer lifetime value)	gross profit per lead	customer acquisition cost (UA acquisition)	acquisition buget	contribution margin
	c=b/ua		aov=revenue/t				apc=t/b	cltv=aovxapc	ltv=cltvxc1	cpa=ac/ua		cm=uax(ltv-cpa)
Product	UA	C1	B	AOV	Revenue	T	APC	CLTV	LTV	LTC/CPA	AC	CM
Web Developer	18548	0.74%	137	€660.51	€348,746.67	528	3.85	€2,545.60	€18.80	€8.06	€149,523.45	€199,223.22
Digital Marketing	18548	2.55%	473	€1,180.56	€2,148,610.91	1820	3.85	€4,542.52	€115.84	€8.06	€149,523.45	€1,999,087.46
UX/UI Design	18548	1.23%	229	€944.03	€876,999.55	929	4.06	€3,829.69	€47.28	€8.06	€149,523.45	€727,476.10
Find yourself in IT	18548	0.01%	1	€0.00	€0.00	2	2.00	€0.00	€0.00	€8.06	€149,523.45	-€149,523.45
Data Analytics	18548	0.00%	0	€0.00	€0.00	0	0.00	€0.00	€0.00	€8.06	€149,523.45	-€149,523.45

- Продукты **Digital Marketing, Web Developer** и **UX/UI Design** показали наибольшую выручку и положительную маржинальность, являются «Локомотивами» бизнеса
- Data Analytics** и **Find yourself in IT** почти не генерируют выручки и показывают отрицательную маржинальность.

## 2. Определение точки роста бизнеса

	UA	C1	B	AOV	Revenue	T	APC	CLTV	LTV	LTC/CPA	AC	CM
	18548	4.53%	840	€1,029.08	€3,374,357.12	3279	3.90	€4,017.09	€181.93	€8.06	€149,523.45	€3,224,833.67
Для поиска точки роста - изменим метрики на 10%												
UA	20403	4.53%	924	€1,029.08	€3,711,792.83	3607	3.90	€4,017.09	€181.93	€8.06	€164,475.80	€3,547,317.04
C1	18548	4.98%	924	€1,029.08	€3,711,792.83	3607	3.90	€4,017.09	€200.12	€8.06	€149,523.45	€3,562,269.38
AOV	18548	4.53%	840	€1,131.99	€3,711,792.83	3279	3.90	€4,418.80	€200.12	€8.06	€149,523.45	€3,562,269.38
APC	18548	4.53%	840	€1,029.08	€3,711,792.83	3607	4.29	€4,418.80	€200.12	€8.06	€149,523.45	€3,562,269.38
CPA	18548	4.53%	840	€1,029.08	€3,374,357.12	3279	3.90	€4,017.09	€181.93	€8.87	€164,475.80	€3,209,881.33
Точка роста - C1 (напрямую влияет на Revenue и CM, без роста AC)												

### 3. Дерево метрик

Revenue	
<b>Продуктовые метрики:</b>	
- AC	
- Т (общее количество транзакций)	
- R_I (выручка на клиента)	
- CLTV (средняя валовая прибыль на клиента)	
- LTV (средняя валовая прибыль на юнит масштабирования)	
- В (клиенты, совершившие покупку)	



Google Docs

CRM\_Data\_Final\_Project\_Product\_Analysis



### 4. Формулировка гипотез

) за ≤ 2 недели)	
HAD #1 ый контакт Lead ≤15 min)	HAD #2 Welcome-цепочка за 24 часа
1: лидом в первые 15 минут,	Серия сообщений лицу - сразу, через 6 ч, через 22 ч - повышает вовлечение (CTR) и оплаты. → рост C1
ти звонок ≤15 мин, шаблон WA/email: 5 min	Настроить welcome-цепочку из четких сообщений "запись на демо/получи пробный урок"
1,	Сделать вариант страницы с блоком доверия
бы на +1 п.п. (с 4.5% до 5.5%),	- переходы → заявки → оплаты. - AOV (средний чек)
Если бронирование демо выросло хотя бы на 10% относительно, и C1 тоже вырос, гипотеза подтверждается	Если конверсия выросла ≥ +1 п.п. и при этом гипотеза верна.
н не дала роста конверсии	Новые креативы не работают, C1 остается прежней
)	Welcome-цепочка не работает, C1 остается прежней



Google Docs

CRM\_Data\_Final\_Project\_Product\_Analysis



### 5. Условия проведения тестов

данные:	
роль всех новых людей	
участие делах их в 2 группах + сладки за разнообразие	
(контроль - все по старому, тест - привлеченные (быстрый контакт, welcome, offer, соц-доказательства))	
изменение C1 = ВЧА между группами	
данные статистически значимы используются в тест данных ( $\alpha = 0.05$ )	
условия проведения гипотезы:	
тест считается подтверждённой, если одновременно выполняются условия:	
ст конверсии (C1) не меньше чем на 1 п.п. ( $MD = 0.01$ ) (4.5% → 5.5%) в тестовой группе относительно контрольной.	
разница в AOV не значима	
если (математическая прибыль) не снизилась	
и стоимость привлечения не выросла	
или бы доля условие не выполнена - гипотеза считается неподтверждённой	
четыре для проведения эксперимента:	
только лидов нужно: $n = (16(1-\rho))^2$	
изделяя конверсия (4.5% + 0.0455)	
минимальный рост, который хотим заметить (+1 п.п. = 0.01)	
16 * 0.0455 * (1 - 0.0455) / 0.012 = 6 920 лидов в каждую группу (13.8k всего)	
Срок до 14 дней (нулю - N лидов/день)	



Google Docs

CRM\_Data\_Final\_Project\_Product\_Analysis



\*Ссылка на детальный, развернутый аналитический отчет

Описание проекта	
Данный проект выполнен в роли аналитика данных для онлайн-школы программирования X.	
Школа ведёт учёт всех клиентов, лидов и сделок в CRM-системе, где хранятся данные о контактах, звонках, рекламных кампаниях, источниках трафика, продажах и платежах.	
Задача: провести очистку и анализ этих данных, выявить слабые и сильные стороны бизнеса, определить эффективность маркетинга и продаж, а также предложить тарифные	
Google Docs	
<a href="#">Report_Final_Project</a>	
Финальный аналитический отчёт по данным CRM онлайн-школы программирования X Описание проекта	
Данный проект выполнен в роли...	