

Project. Metrocar Data Analysis

Klymenko Yuliia

Metrocar

weekday

All

age_range

All

platform

All

Numbers of Active Users

6233

Numbers of Successful Trips

212628

Profit from Trips

\$4,251,667.61

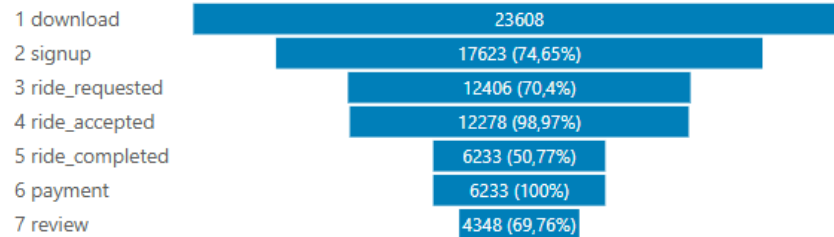
Average Trip Waiting Time

14 minutes

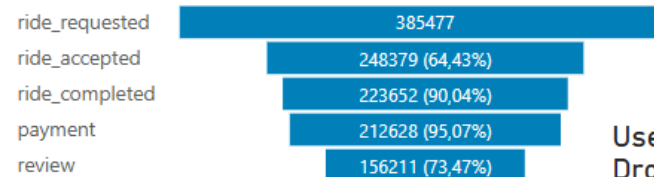
Percentage of Requests Completed

58,02%

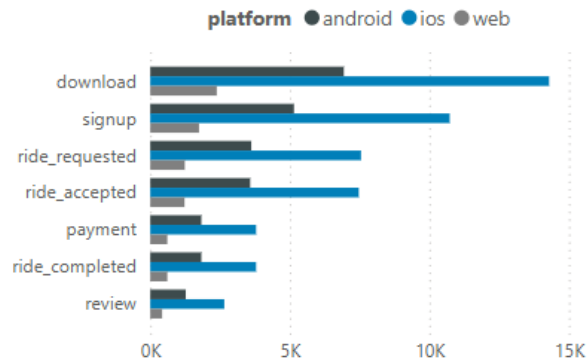
Users funnel



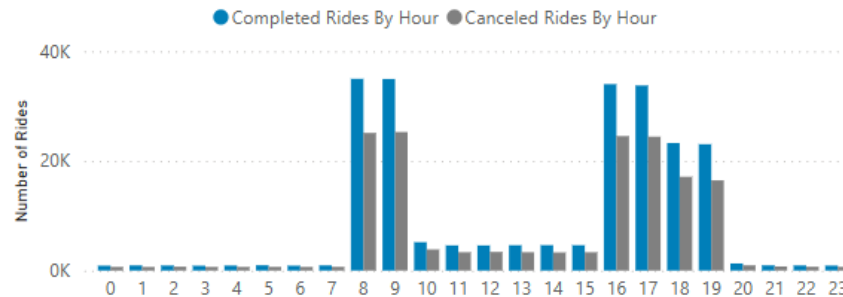
Rides Funnel



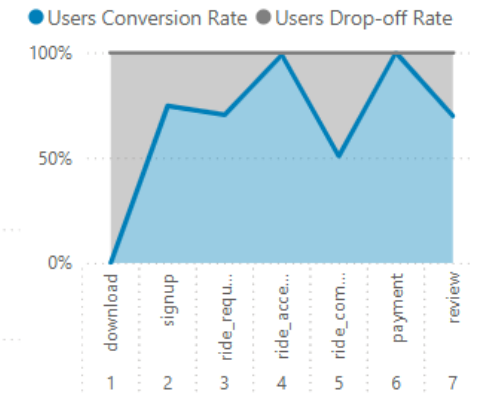
Number of Users by funnel_name and platform



Distribution of Completed Rides and Canceled Rides by Hours



Users Conversion Rate and Drop-off Rate by funnel_step



Про проєкт

Metrocar — це платформа для виклику авто (аналог Uber/Bolt), яка надає можливість користувачам завантажити додаток, зареєструватися, викликати авто, здійснити поїздку, оплатити її та залишити відгук.

В ролі аналітика компанії **Metrocar** я досліджувала, чому багато користувачів не доходять до першої завершеної поїздки, як оптимізувати досвід користувача та покращити конверсію на кожному етапі.

Короткий опис даних

Дані для аналізу містять інформацію про:

- завантаження додатку, включно з платформою та віковою групою користувача;
- реєстрацію нових користувачів і зв'язок із сесією завантаження;
- усі події, пов'язані з поїздками: від запиту до завершення або скасування;
- транзакції, включно з сумами оплат та статусами операцій;
- відгуки користувачів, включно з текстами та оцінками.

Ключові задачі

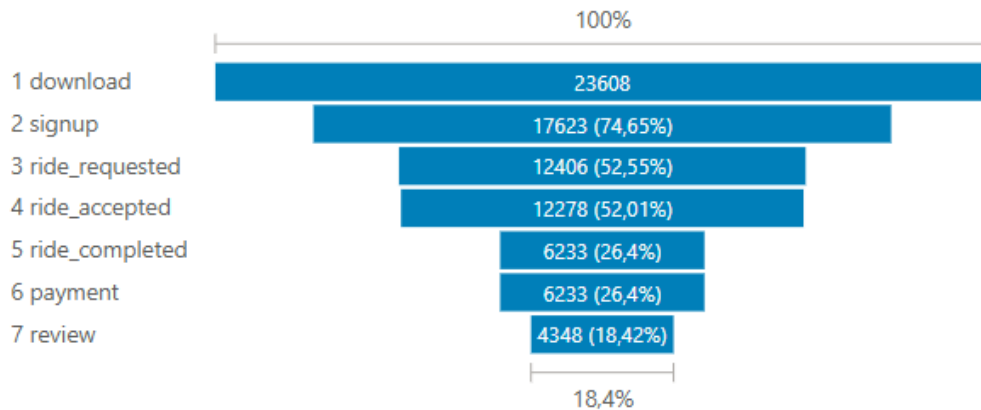
Переді мною були поставлені основні задачі:

- побудувати воронку користувача;
- знайти вузькі місця (drop-off points);
- проаналізувати ефективність платформ;
- оцінити вікові групи користувачів;
- створити інтерактивний дашборд.

Опис етапів аналізу

1. Воронка користувача. На цьому етапі було визначено етапи: завантаження → реєстрація → запит поїздки → прийняття → поїздка → оплата → відгук.

Users funnel

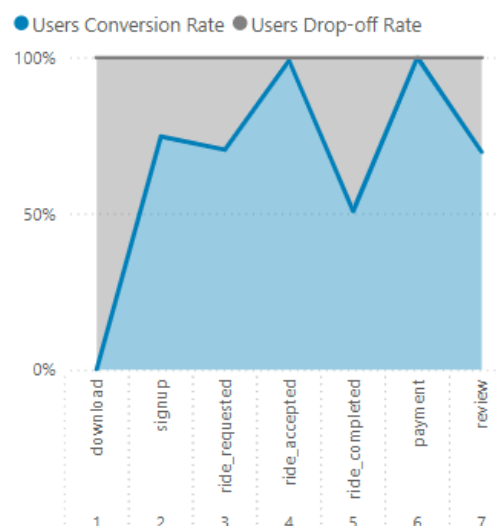


Найбільша втрата користувачів відбувається на етапі між підтвердженням та завершенням поїздки. Робимо висновок, що після підтвердження запиту не завжди відбувається поїздка, тільки 26,4% від загальної кількості завантажень завершують поїздку. При аналізі тут і було сфокусовано найбільше уваги.

Нижче наведено конверсії між етапами: відсоток від попереднього етапу та який відсоток користувачів втрачається на кожному етапі.

funnel_step	Users Conversion Rate	Users Drop-off Rate
1	0,00	100,00
2	74,65	25,35
3	70,40	29,60
4	98,97	1,03
5	50,77	49,23
6	100,00	0,00
7	69,76	30,24

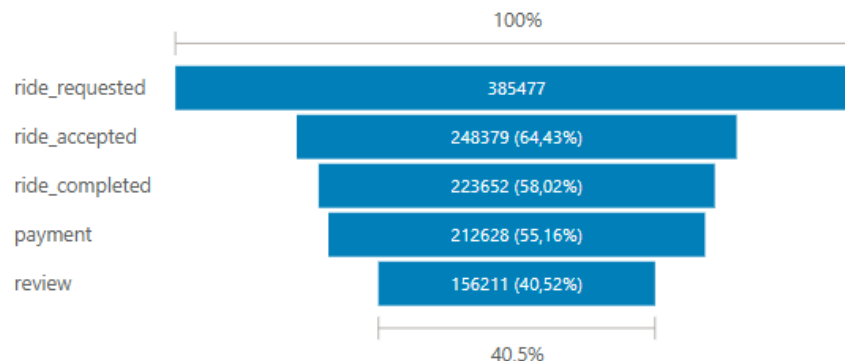
Users Conversion Rate and Drop-off Rate by funnel_step



Від загальної кількості запитів на поїздку 49,23% скасовуються.

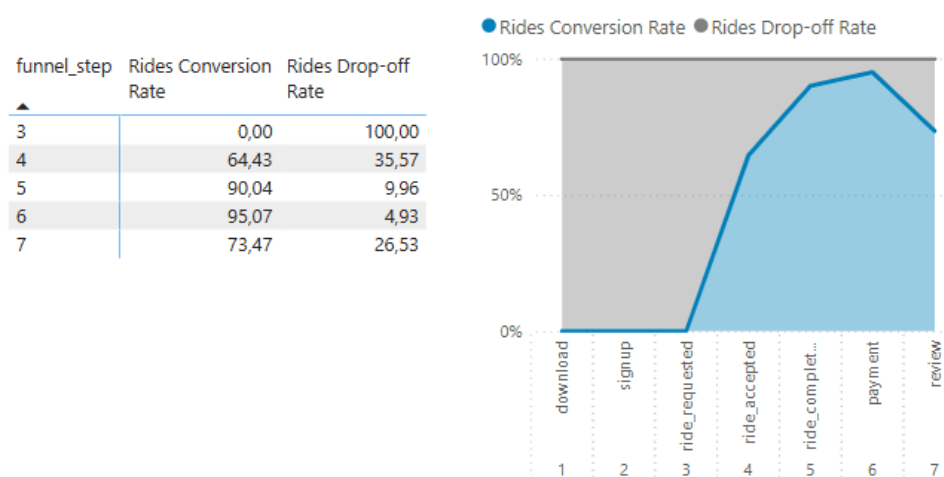
2. Воронка поїздок. На цьому етапі було визначено етапи: запит поїздки → прийняття → поїздка → оплата → відгук.

Rides Funnel



Найбільша втрата відбувається на етапі між запитом поїздки та підтвердженням. Від загальної кількості запитів на поїздку 64,43% підтверджуються та 58,02% закінчуються завершенням поїздки.

Rides Conversion Rate and Drop-off Rate by funnel_step

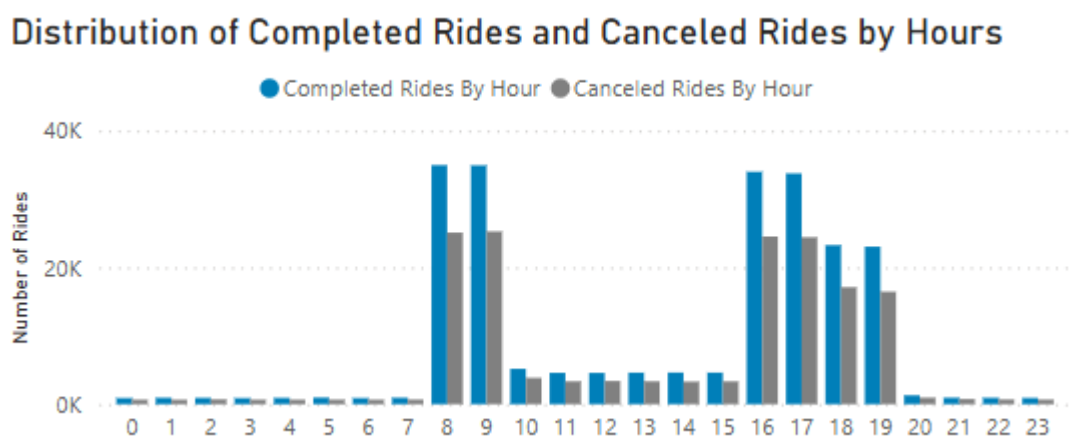


Проаналізувавши конверсії між етапами: відсоток від попереднього етапу та який відсоток користувачів втрачається на кожному етапі, робимо висновок, що 35,57% від кількості запитів на поїздку не підтверджуються, 9,96% та 4,93% від завершених поїздок вважаються не сплаченими (відхилення платежу, тут потрібно дивитись як дані записуються до бази).

3. Залежність від часу доби. Обчислимо середню кількість запитів, підтверджень та скасуванняв по годинно, попередньо обчисливши загальну кількість запитів, підтверджень та скасуванняв по годинах і днях.

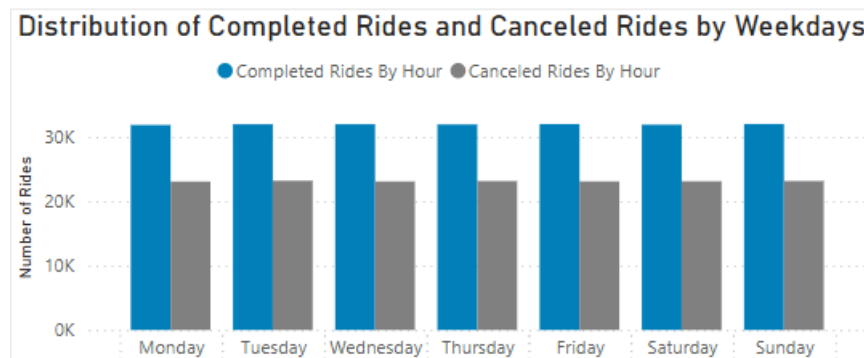
request_hour	avg_requests_per_hour	avg_accepted_per_hour	avg_cancelled_per_hour	cancel_rate_from_accepted	cancel_rate_from_requested
0	5.36	3.57	2.14	59.82	39.8
1	5.54	3.74	2.12	56.63	38.27
2	5.66	3.72	2.32	62.32	40.94
3	5.33	3.56	2.14	60.01	40.05
4	5.5	3.66	2.18	59.44	39.59
5	5.63	3.85	2.1	54.5	37.29
6	5.55	3.7	2.24	60.4	40.25
7	5.6	3.77	2.21	58.61	39.4
8	172.53	116.53	66.22	56.82	38.38
9	173.04	116.57	66.7	57.22	38.54
10	26.42	17.54	10.46	59.62	39.58
11	23.98	15.99	9.38	58.7	39.14
12	23.53	15.79	9.25	58.58	39.31
13	23.94	16.04	9.32	58.11	38.93
14	23.13	15.6	8.91	57.1	38.51
15	23.87	15.92	9.31	58.48	38.99
16	167.78	112.89	64.74	57.35	38.59
17	166.67	112.54	64.29	57.12	38.57
18	115.69	77.6	45.17	58.2	39.04
19	113.81	76.91	43.57	56.66	38.28
20	7.3	4.82	2.91	60.37	39.87
21	5.7	3.72	2.37	63.76	41.64
22	5.52	3.65	2.2	60.47	39.95
23	5.56	3.62	2.25	62.01	40.43

Розглянувши розподіл погодинних завершених поїздок та скасованих, робимо висновок, що найбільш піковими годинами є 8-9, 16-17 години та 18-19 години з трохи меншим попитом.



Кількість скасуванняв замовлень у відсотковому відношенні від загальної кількості запитів на поїздку та від вже підтверджених запитів в пікові години суттєво не відрізняється від аналогічних відношенні в непікові години. Тому це не є основною причиною скасуванняв.

4. Залежність від дня тижня. Розподіл загальної кількості завершених поїздок та скасованих по дням тижня майже однаковий.

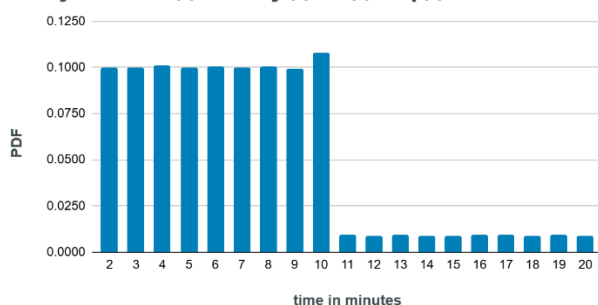


У відсотковому відношенні кількість скасування від загальної кількості запитів по дням тижня також майже рівні.

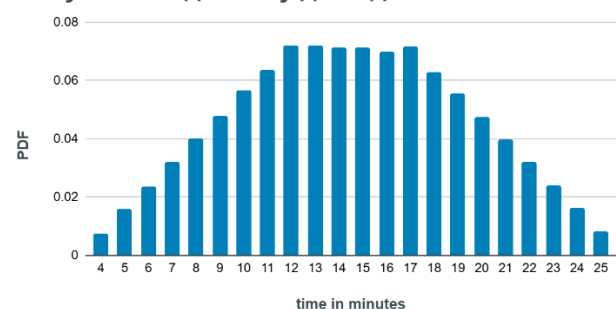
A-Z request_day	123 total_requests	123 cancelled_requests	123 cancel_rate
Monday	54,917	23,063	42
Tuesday	55,168	23,195	42.04
Wednesday	55,057	23,079	41.92
Thursday	55,087	23,146	42.02
Friday	55,080	23,087	41.92
Saturday	55,012	23,108	42.01
Sunday	55,156	23,147	41.97

5. Функції щільності розподілу PDF часу очікування від запиту до підтвердження та від запиту до подачі автомобіля.

Функція щільності розподілу PDF часу очікування від запиту до підтвердження



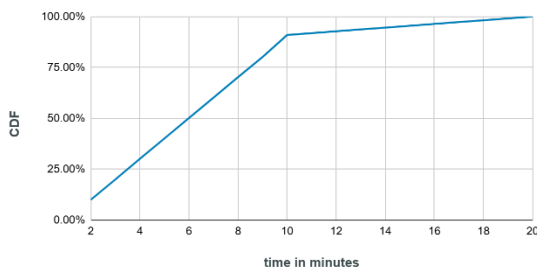
Функція щільності розподілу PDF часу очікування від запиту до подачі автомобіля



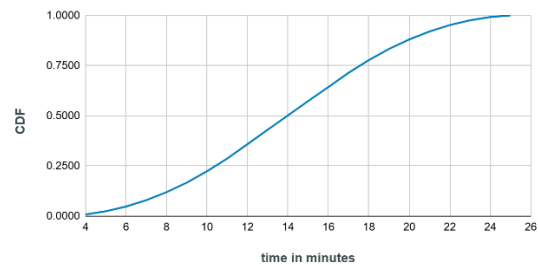
В основному за час до 11 хвилин відбувається підтвердження запиту та в більшості випадків від 11 до 18 хвилин відбувається подача авто.

Також побудовано функції розподілу CDF часу очікування від запиту до підтвердження та від запиту до подачі автомобіля.

Функція розподілу CDF часу очікування від запиту до підтвердження



Функція розподілу CDF часу очікування від запиту до подачі автомобіля



З ймовірністю 90.93% можна сказати, що відбудеться підтвердження запиту за час не більший 10 хвилин.

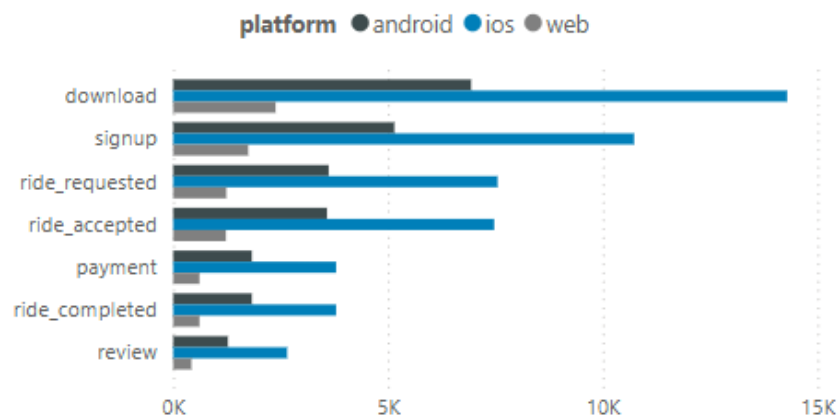
6. Аналіз етапу між погодженням запиту на поїздку та скасуванням запиту.

COUNT of accept_to_cancel_time_minutes										
request_to_accept_time_minutes	accept_to_cancel_time_minutes									
	2	3	4	5	6	7	8	9	10	Grand Total
10	251	241	246	251	245	227	239	259	233	2192
11	265	258	291	264	251	255	243	249	248	2324
12	249	238	242	255	238	240	253	267	237	2219
13	257	266	257	267	258	229	260	245	261	2300
14	253	265	240	251	216	247	244	251	227	2194
15	266	257	272	214	247	256	240	241	242	2235
16	268	233	224	256	237	264	271	252	244	2249
17	258	254	281	209	255	255	271	249	255	2287
18	261	259	253	269	253	232	215	253	235	2230
19	255	253	255	264	260	262	240	254	231	2274
20	261	244	223	267	242	226	260	237	263	2223
Grand Total	2844	2768	2784	2767	2702	2693	2736	2757	2676	24727

Скасування вже погодженого запиту відбувається протягом 2-10 хвилин. Робимо висновок, що довге очікування (12-30 хвилин) призводить до скасування. Цей етап просідає найбільше.

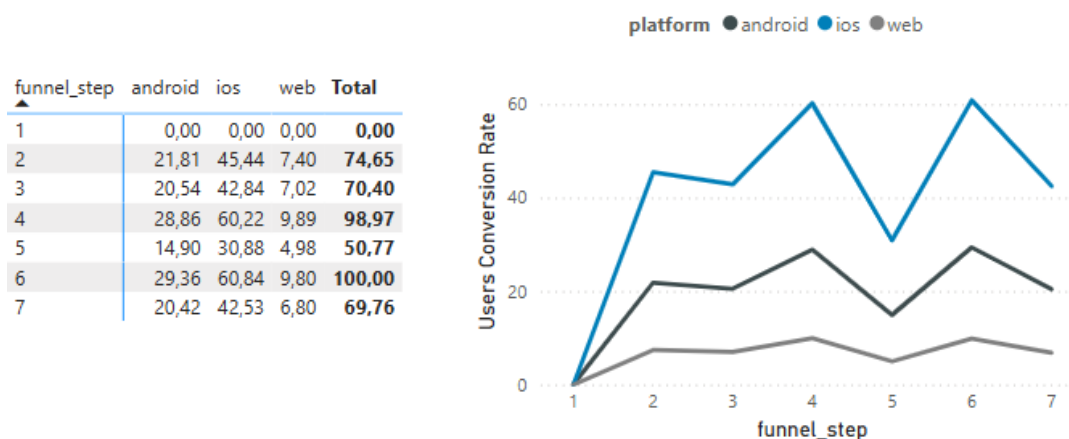
7. Ефективність платформ. Нижче показано розподіл кількості користувачів по платформах на кожному з етапів. Лідером по кількості користувачів на кожному етапі є платформа *ios*.

Number of Users by funnel_name and platform



Розглянемо розподіл конверсії між етапами (відсоток від попереднього етапу) по платформах.

Users Conversion Rate by funnel_step and platform



Також проаналізуємо конверсію між етапами по платформах від загальної кількості першого етапу завантаження. Відсоток втрати користувачів на кожному етапі по платформах майже однаковий.

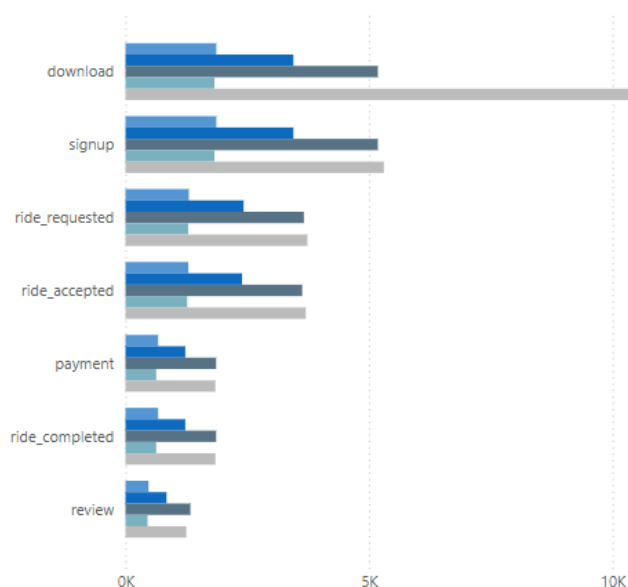
platform	android		ios		web	
funnel_step	Users Conversion Rate By Platform From First Step	Users Drop-off Rate By Platform	Users Conversion Rate By Platform From First Step	Users Drop-off Rate By Platform	Users Conversion Rate By Platform From First Step	Users Drop-off Rate By Platform
1	100,00	0,00	100,00	0,00	100,00	0,00
2	74,23	25,77	75,07	24,93	73,31	26,69
3	52,18	47,82	52,83	47,17	51,91	48,09
4	51,62	48,38	52,28	47,72	51,49	48,51
5	26,39	73,61	26,54	73,46	25,64	74,36
6	26,39	73,61	26,54	73,46	25,64	74,36
7	18,36	81,64	18,55	81,45	17,79	82,21

Доходимо до висновку, що маркетинговий бюджет варто спрямувати на платформи *web* та *android*.

8. Оцінка вікових груп користувачів. Користувачі у вікових категоріях 18-24 та 45-54 користуються сервісом з однаковою активністю, та найактивнішою є вікова група 35-44 років. Також є велика частина користувачів, що не вказали вік (або можливо це користувачі, вік яких поза межами 18-54).

Number_of_users by funnel_name and age_range

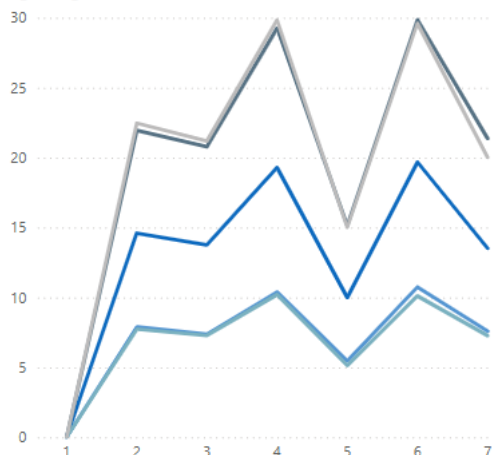
age_range ● 18-24 ● 25-34 ● 35-44 ● 45-54 ● Unknown



Розглянемо розподіл конверсії між етапами (відсоток від попереднього етапу) по віку користувачів.

Users Conversion Rate by funnel_step and age_range

age_range ● 18-24 ● 25-34 ● 35-44 ● 45-54 ● Unknown



funnel_step	18-24	25-34	35-44	45-54	Unknown	Total
1	0,00	0,00	0,00	0,00	0,00	0,00
2	7,90	14,60	21,95	7,73	22,47	74,65
3	7,38	13,76	20,78	7,29	21,19	70,40
4	10,39	19,29	29,24	10,21	29,83	98,97
5	5,46	9,99	15,16	5,13	15,03	50,77
6	10,75	19,69	29,86	10,11	29,60	100,00
7	7,59	13,51	21,37	7,27	20,02	69,76

Обчислено конверсії між етапами по віковим групам від загальної кількості першого етапу завантаження. Відсоток втрати користувачів на кожному етапі по платформах майже однаковий.

age_range	18-24		25-34		35-44		45-54		Unknown	
funnel_step	Users Conversion Rate By Age Range From First Step	Users Drop-off Rate By Age Range	Users Conversion Rate By Age Range From First Step	Users Drop-off Rate By Age Range	Users Conversion Rate By Age Range From First Step	Users Drop-off Rate By Age Range	Users Conversion Rate By Age Range From First Step	Users Drop-off Rate By Age Range	Users Conversion Rate By Age Range From First Step	Users Drop-off Rate By Age Range
1	100,00	0,00	100,00	0,00	100,00	0,00	100,00	0,00	100,00	0,00
2	100,00	0,00	100,00	0,00	100,00	0,00	100,00	0,00	46,98	53,02
3	69,71	30,29	70,35	29,65	70,68	29,32	70,37	29,63	33,08	66,92
4	69,12	30,88	69,42	30,58	70,03	29,97	69,39	30,61	32,78	67,22
5	35,92	64,08	35,60	64,40	35,92	64,08	34,50	65,50	16,34	83,66
6	35,92	64,08	35,60	64,40	35,92	64,08	34,50	65,50	16,34	83,66
7	25,36	74,64	24,43	75,57	25,71	74,29	24,81	75,19	11,06	88,94

Якщо порівнювати втрату користувачів на кожному етапі від загальної кількості користувачів в даній віковій категорії, робимо висновок, що 64-65% від початкової кількості користувачів в віковій категорії не доходить до завершення поїздки.

Найчастіше доходить до завершення поїздки вікова категорія 35-44 років.

Інструменти

Я використовувала MS Power BI для аналізу, включаючи Power Query для перетворення даних і формули DAX для створення нових показників і взаємозв'язків, створювала кастомні стовпці, необхідні для візуалізацій.

Також деяку частину аналізу виконала в Google Sheets.

Висновки та рекомендації

Найбільша втрата користувачів відбувається на етапі між підтвердженням та завершенням поїздки. Значна частина скасувань відбувається через довге очікування, тому можна запропонувати:

- змінити механіку пошуку авто;
- зменшити допустимий час пошуку водія;

- ввести штрафи або бонуси за швидке прийняття, оптимізувати пропозиції та бонуси для годин із низьким попитом;
- сповіщати користувачів про прогнозований час очікування.

Також можна запропонувати:

- збільшити кількість авто в пікові години, і тим самим збільшити кількість завершених поїздок, а отже конверсію і загальний прибуток;
- проаналізувати геолокації, і можливо, найбільша кількість скасувань буде в окремих районах. Тоді проаналізувати їх окремо і робити по ним висновки;
- знайти неактивних користувачів і стимулювати їх цікавими пропозиціями.