**CHANGERS**

**РЕТ – проєкт**

**на тему:** Аналіз клієнтів, побудова персоналізованих комунікацій, які допоможуть посилити взаємодію з клієнтами на основі їхньої поведінки та потреб.

**Виконала**: Мацан Валентина

**2024**ЗМІСТ

ВСТУП……………………………………………………………….……………..…3

1. МЕТОДОЛОГІЯ ДОСЛІДЖЕННЯ …………………………………………….…4

1.1. Вибір теми………………………...……...………………………………….…....4

1.2. Вибір технологій………………………………………………….………………4

2. ЕТАПИ ПРОЄКТУ …………………………………………………………...……5

2.1. Збір та очищення даних………………………………………………………….5

2.2. Аналіз даних RFM-аналіз ………………………………………….…………....7

2.3. Аналіз ключових сегментів ……………………………………………………10

2.4. Аналіз даних ARPU…………………………………………………………….12

2.5. Аналіз даних NPS……………………….………………………………………13

3. ВІЗУАЛІЗАЦІЇ ТА ДАШБОРДИ ………………………………………………...16

3.1. Створені візуалізації……………………………………………………………16

ВИСНОВКИ ………………….……………………………………………………...19

**Вступ**

Цей проєкт зосереджений на вивченні клієнтської бази для збільшення продажів та прибутковості компанії. Ця документація містить важливі деталі, які розкривають методологію, використані інструменти та отримані результати.

Під час цього дослідження, я буду аналізувати дані про клієнтів, проводити їх сегментацію, аналізувати їх поведінку і визначити ключові можливості для бізнесу. Для цього буде проаналізовано Recency (давність останньої покупки), Frequency (частота покупок), і Monetary (грошові витрати), щоб виявити ключові фактори, які впливають на успішність кампаній та підвищення продажів. В роботі будуть застосовуватися метрики, такі як ARPU та NPS, завдяки яким зможемо проаналізувати середній чек, поведінку та їх задоволеність, що зможе призвести до вірних висновків для створенню цільової пропозиції, що має стимулювати повторні покупки та підвищити середній чек.

Через візуалізацію даних, проєкт надасть користувачам зручний інструмент для аналізу та моніторингу ключових показників. Це дозволить керівникам приймати обґрунтовані управлінські рішення на основі актуальної інформації.

**РОЗДІЛ 1**

**Методологія**

* 1. **Вибір теми**

**Тема проєкту:** Аналіз клієнтів, побудова персоналізованих комунікацій, які допоможуть посилити взаємодію з клієнтами на основі їхньої поведінки та потреб.

Тема проєкту обрана на основі інтересів та актуальності для даного бізнесу. Мій PET-проєкт буде спрямований на клієнтську базу, ідентифікувати найприбутковіших клієнтів і створювати цільові пропозиції для кожної групи, що стимулює повторні покупки та підвищує середній чек. Завдяки розумінню того, коли і як часто клієнти купують, можна вибудувати стратегії для підвищення їхньої лояльності, уникнення "відтоку" та заохочення їх до повторних покупок.

Моя робота включає RFM-аналіз, за допомогою якого хочу отримати інформацію, що дозволяє розробляти цільові стратегії для різних груп клієнтів (наприклад, стимулювання сплячих клієнтів, підтримка лояльних та залучення нових клієнтів).

Моя мета - не лише дослідити, але й надати корисні рекомендації для бізнесу щодо покращення стратегій взаємодії з клієнтами для того, щоб посприяє зростанню продажів та покращення загальної ефективність бізнесу, підвищуючи лояльність клієнтів.

* 1. **Вибір технологій**

Для аналізу використовувались інструменти, які вивчалися протягом курсу:

SQL для роботи з базою даних та здійснення потрібних запитів, Excell для обробки та первинного аналізу даних, Power BI, для візуалізації.

**РОЗДІЛ 2**

**Етапи проєкту**

* 1. **Збір та очищення даних**

Проведено збір даних з джерела Power Query.

Основні етапи очищення та перетворення у правильний формат зроблено за допомогою Excell.

Етапи очищення та підготовки представлені нижче:

● Вивантаження даних з BigQuery з датасету logistics.

● Файли Excel були вивантажені та очищені від аномалій та пропусків.

● Проведена перевірка структури даних та виявлено, що деякі значення даних потребують корекції, а також були виявлені пусті значення.

● Для виправлення даних, був використаний Excel.

● Виконана операція видалення рядків, які містять пусті значення. Це було зроблено для забезпечення чистоти даних та уникнення виникнення помилок під час подальшого аналізу.

● Після внесення змін було перевірено результат, щоб переконатися, що дані були очищені та підготовлені коректно.

В результаті всіх вище описаних дій отримано 6 таблиць, а саме:

* **category\_table** – дані про категорії продуктів;
* **customers** – дані про клієнтів;
* **orders** – дані про покупки;
* **orders\_products\_date** – дані про товари та покупки;
* **product\_table** – дані про товари;
* **rating** – рейтинг компанії, виставлений клієнтами «Чи готові клієнти рекомендувати компанію?».

Опис кожної таблиці представлено нижче:

|  |  |
| --- | --- |
| **category\_table** | |
| *Назва атрибута* | *Опис* |
| category\_id | Унікальний ідентифікатор для категорії товару |
| category\_name | Назва категорії |

|  |  |
| --- | --- |
| **customers** | |
| *Назва атрибута* | *Опис* |
| customer\_id | Унікальний ідентифікатор клієнта |
| customer\_name | Ім'я клієнта |
| customer\_email | email клієнта |
| gender | Гендер |
| age | Вік клієнта |
| annual\_income | Річний дохід |
| marital\_status | Статус одруження |

|  |  |
| --- | --- |
| **orders** | |
| *Назва атрибута* | *Опис* |
| customer\_id | Унікальний ідентифікатор клієнта |
| order\_id | Унікальний ідентифікатор покупки |

|  |  |
| --- | --- |
| **orders\_products\_date** | |
| *Назва атрибута* | *Опис* |
| product\_id | Унікальний ідентифікатор товара |
| quantity | Кількість товара |
| orders\_date | Дата покупки |

|  |  |
| --- | --- |
| **product\_table** | |
| *Назва атрибута* | *Опис* |
| product\_name | Назва товара |
| category\_id | Унікальний ідентифікатор для категорії товару |
| price | Ціна товар |

|  |  |
| --- | --- |
| **rating** | |
| *Назва атрибута* | *Опис* |
| order\_id | Унікальний ідентифікатор покупки |
| rating | Рейтинг |

Всі дані було завантажено у BigQuery, при цьому я створила датасет **company-ecom**. Після цього, з ними можна працювати та аналізувати. Наступним етапом було згрупування та вивантаження даних для RFM-аналізу. Нижче представлено код SQL для збору даних.

select c.customer\_id,

        c.customer\_name,

        IFNULL(c.customer\_email, 'Невідомо') as email,

        c.gender,

        c.age,

        c.marital\_status,

        max(op.orders\_date) as last\_purchase\_day,

        DATE\_DIFF(CURRENT\_DATE, MAX(op.orders\_date), DAY) AS days\_since\_last\_purchase,

        sum(op.quantity\*pt.price) as total\_revenue,

        ROUND(AVG(op.quantity \* pt.price), 2) as avg\_total\_revenue,

        count (o.order\_id) as count\_orders

from `pet-pro-vm.company\_ecom.customers`  as c

left join `pet-pro-vm.company\_ecom.orders` as o on c.customer\_id=o.customer\_id

join `pet-pro-vm.company\_ecom.orders\_products\_date` as op on o.order\_id=op.order\_id

left join `pet-pro-vm.company\_ecom.product\_table` as pt on op.product\_id=pt.product\_id

    group by c.customer\_id,

        c.customer\_name,

        email,

        c.gender,

        c.age,

        c.marital\_status;

**2.2. Аналіз даних. RFM-аналіз**

Використані методи статистики та аналізу даних для отримання інсайтів. Проведено RFM- аналіз. Розрахунки проводились в Excell. На основі результатів RFM-сегментації, можемо проаналізувати поведінку клієнтів і визначити ключові можливості для бізнесу. Сегментація дозволяє класифікувати клієнтів за трьома критеріями: Recency (давність останньої покупки), Frequency (частота покупок), і Monetary (грошові витрати). Кожен із трьох показників поділений на кілька рівнів:

* **Recency (давність):**
  + 1 (менше 90 днів) — найбільш активні клієнти.
  + 2 (90-180 днів) — клієнти, що робили покупку досить давно, але ще можуть бути зацікавлені в повторних покупках.
  + 3 (більше 180 днів) — клієнти, що давно не здійснювали покупок, і ймовірність їх повернення нижча.
* **Frequency (частота):**
  + 1 (менше 3 покупок) — нерегулярні покупці.
  + 2 (3-10 покупок) — помірно активні покупці.
  + 3 (більше 10 покупок) — найактивніші, лояльні клієнти.
* **Monetary (грошові витрати):**
  + 1 (менше 1 тис. дол.) — клієнти з невеликими витратами.
  + 2 (1-5 тис. дол.) — клієнти із середніми витратами.
  + 3 (більше 5 тис. дол.) — клієнти з високими витратами.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Recency | | |
| 1 | <90 | 1305 | 70,89% |
| 2 | 90-180 | 376 | 20,42% |
| 3 | >180 | 160 | 8,69% |
|  | Total | 1841 | 1 |
|  |  |  |  |
|  |  |  |  |
|  | Frequency |  |  |
| 1 | <3 | 497 | 27,00% |
| 2 | 3-10 | 1300 | 70,61% |
| 3 | >10 | 44 | 2,39% |
|  | Total | 1841 | 1 |
|  |  |  |  |
|  |  |  |  |
|  | Monetary | | |
| 1 | <1 тис. | 1158 | 62,90% |
| 2 | 1-5 тис. | 655 | 35,58% |
| 3 | >5 | 28 | 1,52% |
|  | Total | 1841 | 1 |

Отримано 21 сегмент:

|  |  |  |  |
| --- | --- | --- | --- |
| **Сегменти** | **%** | **Count\_customers** | **Сегмент** |
| 111 | 10,16% | 187 | **Перспективні** |
| 112 | 1,52% | 28 | **Нові** |
| 113 | 0,05% | 1 | **Нові** |
| 121 | 29,49% | 543 | **Лідери** |
| 122 | 26,13% | 481 | **Лідери** |
| 123 | 1,20% | 22 | **Постійні клієнти** |
| 131 | 0,11% | 2 | **Постійні клієнти** |
| 132 | 2,12% | 39 | **Постійні клієнти** |
| 133 | 0,11% | 2 | **Постійні клієнти** |
| 211 | 7,44% | 137 | **Заснулі** |
| 212 | 0,81% | 15 | **Втрачені** |
| 213 | 0,05% | 1 | **Втрачені** |
| 221 | 7,77% | 143 | **Постійні клієнти** |
| 222 | 4,24% | 78 | **Постійні клієнти** |
| 223 | 0,05% | 1 | **Ті, яких не можна втратити** |
| 232 | 0,05% | 1 | **Ті, яких не можна втратити** |
| 311 | 6,63% | 122 | **Заснулі** |
| 312 | 0,27% | 5 | **Втрачені** |
| 313 | 0,05% | 1 | **Втрачені** |
| 321 | 1,30% | 24 | **Клієнти під загрозою** |
| 322 | 0,43% | 8 | **Ті, хто потребує уваги** |

* 1. **Аналіз ключових сегментів**

1. **Перспективні 111 (10,2%)** – це клієнти з високою давністю покупок (менше 90 днів), низькою частотою покупок (менше 3), і низькими витратами (менше 1 тис.). Це активні, але нерегулярні покупці з низьким бюджетом.

**Висновок:** Цей сегмент може бути перспективним для збільшення лояльності та стимулювання повторних покупок через спеціальні пропозиції або знижки.

1. **Лідери 121 (29,5%) і 122 (26,1%)** – клієнти з недавньою покупкою, помірною частотою покупок, але з різними рівнями витрат. Ці клієнти становлять майже 56% усієї бази.

**Висновок:** Це основний сегмент для бізнесу, тому варто сфокусувати маркетингові кампанії на підтримці лояльності цього сегменту, пропонуючи персоналізовані рекомендації та спеціальні програми, можна просувати нові продукти.

1. **Постійні клієнти (15,54%) -** ця категорія клієнтів може не купувала зовсім недавно, але вони напевно купують часто і дорогі продукти.

**Висновок:** Для сегмента «постійних клієнтів» зосередження на лояльності, персоналізованому досвіді та ексклюзивних перевагах може посилити їхню прихильність і збільшити кількість покупок. Ось кілька ідей:

**програма винагороди за лояльність**: пропонуйте бали за кожну покупку, які можна обміняти на знижки чи подарунки. Постійні клієнти цінують визнання, тому багаторівнева система (наприклад, Silver, Gold, Platinum) може підтримувати їхню зацікавленість.

* **персоналізовані рекомендації:** на основі попередніх покупок або поведінки веб-переглядача можна рекомендувати продукти, які можуть їм сподобатися. Індивідуальні рекомендації можуть зробити покупки зручнішими та приємнішими.
* **ювілейні або значні знижки:** річниця першої покупки або досягнення певної кількості замовлень. Можна надіслати їм спеціальну знижку або подарунок.

1. **Заснулі (14.07%) –** ці клієнти купували тільки недорогі товари, робили покупки один або два рази і ніколи не поверталися.

**Висновок:** Досить великий процент з усіх, тому вважаю, що можна було б нагадати про себе та запропонувати:

**–**  **знижка за повернення**: Надіслати персоналізований лист з пропозицією знижки на наступну покупку. Це може бути спеціальна пропозиція, наприклад: "Ми скучили за вами! Знижка 10% на ваше наступне замовлення".

- **персоналізовані рекомендації**: Запропонувати товари, схожі на ті, які вони купували раніше. Це може бути нагадування про те, що вони купували в минулому, або нові товари, які можуть їх зацікавити.

**- пакетні пропозиції або безкоштовна доставка**: Запропонвати спеціальні пакети товарів зі знижкою або безкоштовну доставку, щоб зменшити додаткові витрати для клієнтів, які можуть вагатися через ціну.

- **реєстрація в програмі лояльності**: наголосити на перевагах, які вони можуть отримати, якщо повернуться: бали, бонуси або знижки за повторні покупки. Для тих, хто ще не зареєстрований, можна запропонувати бонус за реєстрацію.

1. **Сегменти «Ті, хто потребує уваги» та «Ті, яких не можна втратити» (0,43% та 0,11%)** – Ці клієнти, можливо, не купували останнім часом, але вони витратили пристойну суму грошей досить багато разів.

**Висновок:** Вони витрачають більше 5 тис., але їх кількість незначна. Для них можна розробити спеціальні пропозиції або ексклюзивні товари, щоб підвищити їхнє задоволення і утримати високий рівень витрат.

1. **Нові (1,58%) –** це клієнти, щоробили нещодавні покупці, купили товари високою ціною, і лише один або два рази.

**Висновок:** Для стимулювання повторних покупок від нових клієнтів можна запропонувати:

* знижку на другу покупку, наприклад, «Дякуємо, що обрали нас! Отримайте 10% знижки на наступне замовлення протягом 30 днів»
* розказати клієнтам про програму лояльності та бонуси за повторні покупки. Наприклад, заохотити накопичення балів або надати їм додаткові переваги для швидшого досягнення статусу постійного клієнта.
* стимулювати клієнта повернутися за допомогою пропозиції безкоштовної доставки на наступне замовлення, що зижує загальні витрати
* можна запропонувати нещодавнім клієнтам бонус за рекомендацію вашого магазину друзям. Вони отримають винагороду за кожного нового клієнта, якого вони приведуть.
* Зробіть пропозицію, яка обіцяє невеликий сюрприз або бонусний товар для клієнтів, які оформлять замовлення вдруге.
  1. **Аналіз даних ARPU**

Наступним етапом проєкту є розрахунок метрики ARPU. (Середній дохід на користувача). Вказує на середню кількість грошей, яку клієнт приносить компанії протягом кожного місяця за 2024 рік. За допомогою BigQuery виконано розрахунки. Код SQL представлено нижче:

select

      orders\_date,

      extract (Year from orders\_date)  as Year,

      extract (month from orders\_date)  as Month,

      count (distinct c.customer\_id) as count\_customers,

      sum(op.quantity\*pt.price) as total\_revenue,

      round (sum(op.quantity\*pt.price)/count (distinct c.customer\_id),0) as ARPU\_Month

from `pet-pro-vm.company\_ecom.customers`  as c

left join `pet-pro-vm.company\_ecom.orders` as o on c.customer\_id=o.customer\_id

join `pet-pro-vm.company\_ecom.orders\_products\_date` as op on o.order\_id=op.order\_id

left join `pet-pro-vm.company\_ecom.product\_table` as pt on op.product\_id=pt.product\_id

    group by orders\_date

    order by Year, Month

Результати були завантажені до Power BI та зроблена візуалізація, на якій видно, що Середній дохід на користувача збільшується, що є гарним показником для компанії. Отже, компанія ефективно монетизує свою користувацьку базу. Потрібно зосередитися на утриманні клієнтів і підвищенні їхньої лояльності, але рости завжди є куди, тому важливо збалансувати стратегії утримання та залучення нових клієнтів для сталого росту.

* 1. **Аналіз даних NPS**

Проведено аналіз рейтингу продуктів та розраховано NPS - індекс споживчої лояльності користувачів до компанії. В основі індексу – відповіді аудиторії на запитання: «Чи готові ви рекомендувати нашу компанію своїм друзям?» .

За допомогою BigQuery виконано розрахунки. Код SQL представлено нижче:

with Full\_customers as (

      select  r.rating, count (distinct customer\_id) as customers

from `pet-pro-vm.company\_ecom.rating` as r

left join `pet-pro-vm.company\_ecom.orders` as o on r.order\_id=o.order\_id

group by  r.rating),

Dectators as (

select r.rating, count (distinct customer\_id) as Dectators\_customers

from `pet-pro-vm.company\_ecom.rating` as r

left join `pet-pro-vm.company\_ecom.orders` as o on r.order\_id=o.order\_id

where r.rating <=6

group by r.rating),

Promouter as (

select r.rating, count (distinct customer\_id) as Promouter\_customers

from `pet-pro-vm.company\_ecom.rating` as r

left join `pet-pro-vm.company\_ecom.orders` as o on r.order\_id=o.order\_id

where r.rating >8

group by r.rating)

select

sum(customers) as Full\_customers,

sum(Dectators\_customers) as Dectators\_customers,

sum(Promouter\_customers) as Promouter\_customers,

round ((sum(Promouter\_customers)-sum(Dectators\_customers))/sum(customers), 4) as NPS

from Full\_customers

full outer join  Dectators on

Full\_customers.rating=Dectators.rating

full outer join Promouter on

Dectators.rating=Promouter.rating

В результаті обрахунків отримано значення NPS 33,23% , яке означає, що компанія має позитивну лояльність, тобто значно більше промоутерів, ніж декракторів. У цьому випадку, 33,23% – це частка клієнтів, які більше схильні рекомендувати компанію, що є хорошим показником, але завжди є можливість для покращення. Для підтримки та покращення цього рівня лояльності можна продовжувати працювати над обслуговуванням клієнтів, поліпшенням продукту, а також мотивувати промоутерів (наприклад, через програми лояльності або рекомендаційні програми) ще активніше підтримувати бренд.

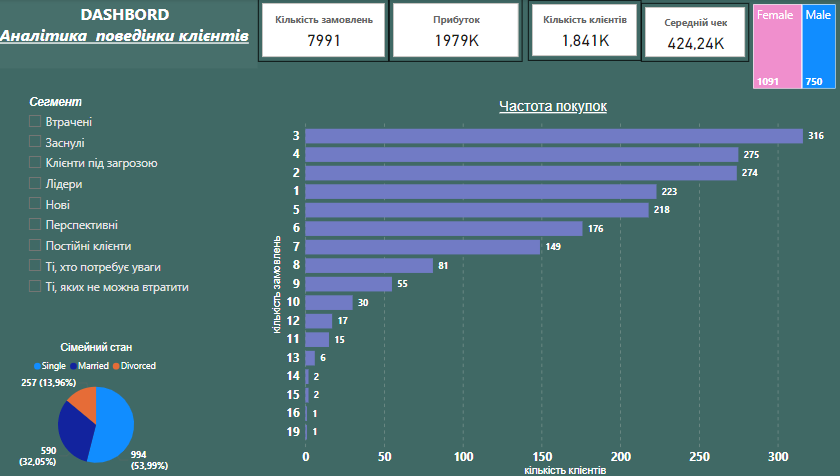
**РОЗДІЛ 3**

**Візуалізація та Дашборд**

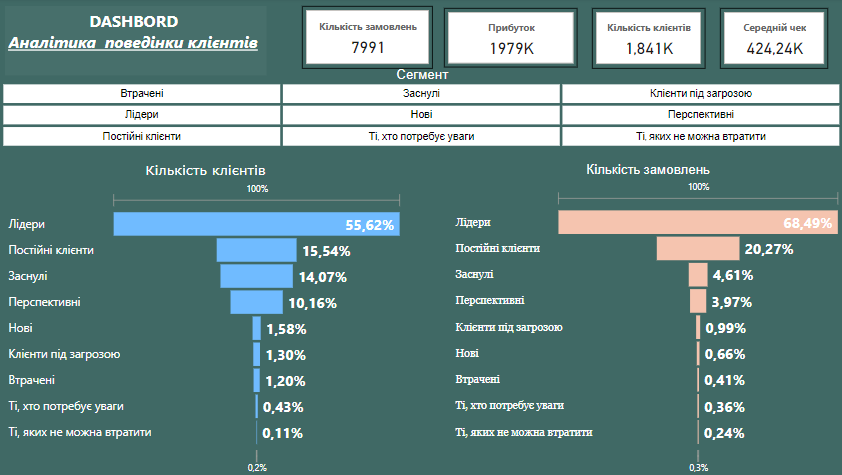
**3.1. Створені візуалізації**

Для візуалізації результатів аналізу використано Power BI. Створено наступні дашборди.

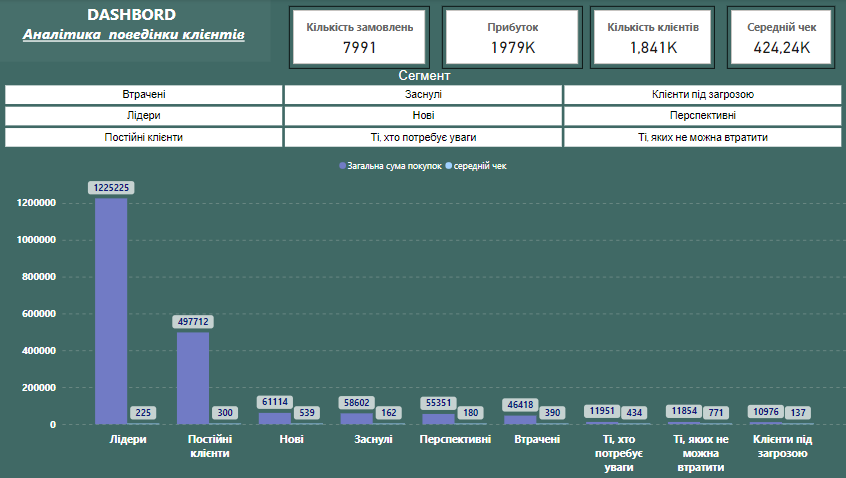
1. Дані про клієнтів, де зображено і сегменти з RFM – аналізу, частота покупок, сімейний стан, гендер та кількість таких клієнтів.



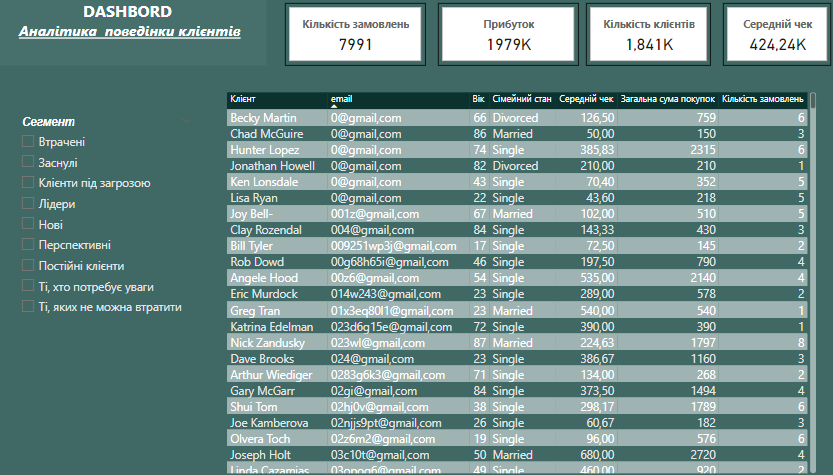
1. Сегменти і їх значення по кількостям та по замовленням



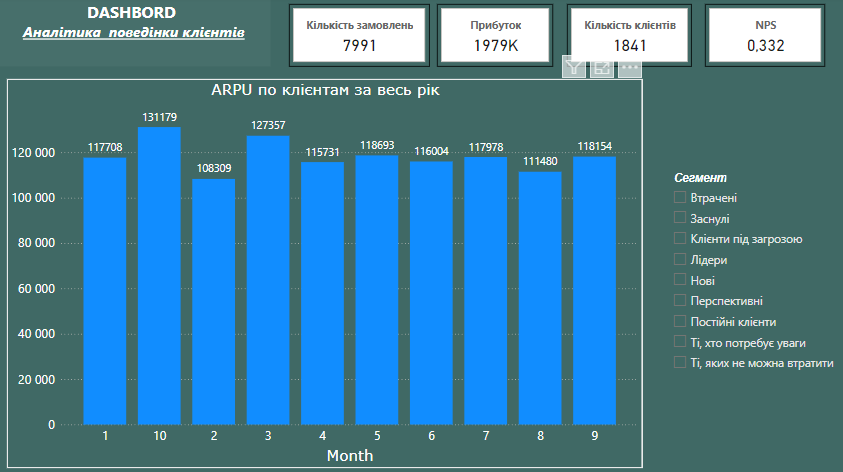
1. Сума покупок та середній чек, видно, що в Лідерів найкращі показники.



1. На дажборді показані всі дані по клієнтам. При виборі необхідного сегмента, можна детально в таблиці побачити всю детальні інформацію по кожному.



1. ARPU по місяцям за весь рік. Середній дохід на користувача збільшується, що є гарним показником. Тут винесено також NPS.



**Висновки**

Проєкт дозволяє зрозуміти, які клієнти є в бізнесі, проаналізовано та виявлено сегменти, на які можна їх поділити.

Ідентифіковано найприбутковіших клієнтів і створено цільові пропозиції для кожної групи, що має стимулювати повторні покупки та підвищить середній чек.

Запропоновано персоналізовані рекомендації та спеціальні програми, завдяки яким можна збільшувати продажі та підвищувати їхню задоволеність та зміцненню зв’язку із брендом. Ці всі методи та пропозиції, що основані на побудові персоналізованих комунікацій, зможуть допомогти посилити взаємодію з клієнтами на основі їхньої поведінки та потреб. Це не тільки посприяє зростанню продажів, але й покращить загальну ефективність бізнесу, підвищить лояльність клієнтів.

Розраховано ARPU (Середній дохід на користувача) він збільшується, що є гарним показником для компанії. Отже, компанія ефективно монетизує свою користувацьку базу. Потрібно зосередитися на утриманні клієнтів і підвищенні їхньої лояльності, але рости завжди є куди, тому важливо збалансувати стратегії утримання та залучення нових клієнтів для сталого росту.

Проведено аналіз рейтингу продуктів та розраховано NPS - індекс споживчої лояльності користувачів до компанії. В основі індексу – відповіді аудиторії на запитання: «Чи готові ви рекомендувати нашу компанію своїм друзям?» . В результаті обрахунків можна зробити висновок, що компанія має позитивну лояльність, тобто значно більше промоутерів, ніж декракторів. У цьому випадку, 33,23% – це частка клієнтів, які більше схильні рекомендувати компанію, що є хорошим показником, але завжди є можливість для покращення. Для підтримки та покращення цього рівня лояльності можна продовжувати працювати над обслуговуванням клієнтів, поліпшенням продукту, а також мотивувати промоутерів (наприклад, через програми лояльності або рекомендаційні програми) ще активніше підтримувати бренд.

Клієнти під сегментом Лідеристановлять майже 56% усієї бази. Варто зосереджуватись саме на цьому сегменту, тому що це основний сегмент для бізнесу. Потрібно збільшити зусилля для їх збереження, пропонуючи персоналізовані рекомендації та спеціальні програми, можна просувати нові продукти. Для сегмента «Перспективні», їх кількість становить більше 10%, можна пропонувати персоналізовані рекомендації та спеціальні програми, можна просувати нові продукти. Для стимулювання повторних покупок від нових клієнтів можна запропонувати знижки або бонуси за повторні покупки.

Додатковий аналіз географічного розподілу клієнтів може допомогти визначити потенційно перспективні ринки для розширення бізнесу та визначення нових цільових аудиторій.

На основі проведеного аналізу, набору метрик можна зробити висновок, що проєкт успішно виконав свою мету та цілі: проаналізовано та створено цільові пропозиції для кожної групи. Через розроблення графіків та дашбордів на основі, проєкт надає користувачам зручний інструмент для аналізу та моніторингу ключових показників. Це дозволяє керівникам приймати обґрунтовані управлінські рішення на основі актуальної інформації.