

Проект: Анализ поведения игроков и внутриигровых транзакций

Цель проекта: изучить влияние характеристик игроков и их игровых персонажей на покупку внутриигровой валюты, а также оценить активность игроков при совершении внутриигровых покупок

Автор: Онищенко Надежда
Дата:

Часть 1. Исследовательский анализ данных
Задача 1. Исследование доли платящих игроков

1.1. Доля платящих пользователей по всем данным:

```
WITH new AS (  
SELECT COUNT(DISTINCT id) AS total_users -- все игроки,  
       (SELECT COUNT(DISTINCT id) FROM fantasy.users WHERE payer = 1) AS  
paid_users --игроки покупающие внутриигровую валюту  
FROM fantasy.users)  
  
SELECT total_users, paid_users, ROUND(paid_users::numeric/total_users,4) AS  
dolya_paid_users  
FROM new;
```

```
/*второй вариант расчета  
SELECT COUNT(payer) AS total_payer,  
       SUM(payer) AS paid_users,  
       ROUND(AVG(payer),4) AS dolya_paid_users  
FROM fantasy.users  
WHERE payer = 0 OR payer = 1;*/
```

1.2. Доля платящих пользователей в разрезе расы персонажа:

```
WITH total AS (  
SELECT COUNT(DISTINCT id) AS total_users, race  
FROM fantasy.users  
JOIN fantasy.race USING(race_id)  
GROUP BY race),  
  
paid AS (  
SELECT COUNT(DISTINCT id) AS paid_users, race  
FROM fantasy.users  
JOIN fantasy.race USING(race_id)  
WHERE payer = 1  
GROUP BY race)  
  
SELECT race, total_users, paid_users,  
       ROUND(paid_users/total_users::NUMERIC,4) AS dolya_paid_users  
FROM total  
FULL JOIN paid USING(race)  
ORDER BY paid_users;
```

Задача 2. Исследование внутриигровых покупок
2.1. Статистические показатели по полю amount:

```
SELECT COUNT(amount) AS total_count_amount --общее количество покупок,  
       SUM(amount) AS sum_amount,  
       MIN(amount) AS min_amount,  
       MAX(amount) AS max_amount,  
       ROUND(AVG(amount::numeric),2) AS avg_amount,  
       PERCENTILE_DISC(0.5) WITHIN GROUP (ORDER BY amount) AS mediana_amount,  
       ROUND(STDDEV(amount::numeric),2) AS std_amount  
FROM fantasy.events
```

2.2: Аномальные нулевые покупки:

```
SELECT (SELECT COUNT(amount) FROM fantasy.events) AS total_amount,  
       COUNT(amount) AS zero_amount, --покупки с нулевой стоимостью  
       ROUND(COUNT(amount)::numeric/(SELECT COUNT(amount) FROM fantasy.events),4)  
AS dolya_zero_amount  
FROM fantasy.events  
WHERE amount = 0.0;
```

```
/* второй вариант расчета  
SELECT COUNT(*) AS total_amount,  
       COUNT(*) FILTER (WHERE amount = 0) AS zero_amount,  
       ROUND(COUNT(*)FILTER (WHERE amount = 0)::numeric / COUNT(*), 6) AS  
dolya_zero_amount  
FROM fantasy.events;*/
```

```
SELECT race, COUNT(amount) AS zero_amount,  
       ROUND(COUNT(amount)::numeric/(SELECT COUNT(amount) FROM fantasy.events),6)  
AS dolya_zero_amount  
FROM fantasy.events  
FULL JOIN fantasy.users USING(id)  
FULL JOIN fantasy.race USING(race_id)  
WHERE amount = 0.0  
GROUP BY race  
ORDER BY zero_amount DESC;
```

2.3: Сравнительный анализ активности платящих и неплатящих игроков:

```
WITH new AS ((
SELECT CASE WHEN payer=1 THEN 'paid_users'
      END AS users_group,
      COUNT(DISTINCT events.id) AS users_count, --совершившие покупки, и
являющиеся покупателями внутриигровой валюты
      COUNT(events.amount) AS count_amount,
      SUM(events.amount) AS sum_amount
FROM fantasy.events
LEFT JOIN fantasy.users USING(id)
WHERE payer = 1 AND amount <> 0.0
GROUP BY payer)
UNION
(SELECT CASE WHEN payer=0 THEN 'not_paid_users'
      END AS users_group,
      COUNT(DISTINCT events.id) AS users_count, --совершившие покупки,
неплатящие
      COUNT(events.amount) AS count_amount,
      SUM(events.amount) AS sum_amount
FROM fantasy.events
LEFT JOIN fantasy.users USING(id)
WHERE payer = 0 AND amount <> 0.0
GROUP BY payer))

(SELECT users_group,
      users_count, count_amount, ROUND(count_amount::numeric/users_count,2) AS
dolya_user,
      sum_amount, ROUND(sum_amount::numeric/users_count,2) AS dolya_user
FROM NEW
GROUP BY users_group, users_count, count_amount, sum_amount)
UNION
(SELECT 'total_users_amount' AS users_group,
      SUM(users_count), SUM(count_amount),
      ROUND(SUM(count_amount)::numeric/SUM(users_count),2) AS dolya_user,
      SUM(sum_amount),
      ROUND(SUM(sum_amount)::numeric/SUM(users_count),2) AS dolya_user
FROM NEW)
ORDER BY users_count;
```

2.4: Популярные эпические предметы:

```
WITH NEW AS (
```

```

SELECT game_items, COUNT(amount) AS count_game_items, --распределение покупок по
эпическим предметам
        ROUND(COUNT(amount)::numeric/(SELECT COUNT(amount) FROM fantasy.events
WHERE amount <> 0),6) AS dolya_count,
        ROUND(COUNT(DISTINCT id)::numeric/(SELECT COUNT(DISTINCT id) FROM
fantasy.events WHERE amount <> 0),6) AS dolya_users_total
FROM fantasy.events
FULL JOIN fantasy.items USING(item_code)
WHERE amount <> 0
GROUP BY game_items
ORDER BY count_game_items DESC)

SELECT *
FROM NEW;

```

```

WITH avg_dolya AS (
SELECT ROUND(COUNT(DISTINCT id)::numeric/(SELECT COUNT(DISTINCT id)
FROM fantasy.events
LEFT JOIN fantasy.users USING(id)
WHERE amount <> 0 AND payer = 1),6) AS
dolya_users_paid,
        ROUND(COUNT(DISTINCT id)::numeric/(SELECT COUNT(DISTINCT id)
FROM fantasy.events
LEFT JOIN fantasy.users USING(id)
WHERE amount <> 0 AND payer = 0),6) AS
dolya_users_not_paid
FROM fantasy.events
FULL JOIN fantasy.items USING(item_code)
WHERE amount <> 0
GROUP BY game_items)

SELECT ROUND(avg(dolya_users_paid),3) AS avg_dolya_users_paid,
        ROUND(avg(dolya_users_not_paid),3) AS avg_dolya_users_not_paid,
        ROUND(avg(dolya_users_paid)/avg(dolya_users_not_paid),3)
FROM avg_dolya;

```

Часть 2. Решение ad hoc-задач

Задача 1. Зависимость активности игроков от расы персонажа:

```

WITH a AS (
SELECT race, COUNT(DISTINCT id) AS total_users, --добавляя условия amount<>0
значения в total=user_amount
        users_amount, paid_users
FROM fantasy.race

```

```

LEFT JOIN fantasy.users USING(race_id)
LEFT JOIN (SELECT race_id, COUNT(DISTINCT events.id) AS users_amount --совершившие
покупки
FROM fantasy.race
LEFT JOIN fantasy.users USING(race_id)
RIGHT JOIN fantasy.events USING(id)
WHERE amount <> 0
GROUP BY race_id) AS amount USING(race_id)
LEFT JOIN (SELECT race_id, COUNT(DISTINCT events.id) AS paid_users --совершившие
покупки, платящие
FROM fantasy.race
LEFT JOIN fantasy.users USING(race_id)
RIGHT JOIN fantasy.events USING(id)
WHERE payer = 1 AND amount <> 0
GROUP BY race_id) AS paid USING(race_id)
GROUP BY race, users_amount, paid_users),

b AS (
SELECT race,
COUNT(amount) AS count_amount,
SUM(amount) AS sum_amount
FROM fantasy.race
LEFT JOIN fantasy.users USING(race_id)
LEFT JOIN fantasy.events USING (id)
WHERE amount <> 0.0
GROUP BY race)

SELECT race, total_users, users_amount,
ROUND(users_amount::numeric/total_users,4) AS dolya_users_amount,
ROUND(paid_users::numeric/users_amount,4) AS dolya_paid_users,
ROUND(AVG(count_amount::numeric/users_amount),2) AS avg_count_amount,
ROUND(AVG(sum_amount::numeric/count_amount),2) AS avg_amount,
ROUND(AVG(sum_amount::numeric/users_amount),2) AS avg_sum_amount
FROM a
FULL JOIN b USING(race)
GROUP BY race, total_users, users_amount, paid_users
ORDER BY avg_count_amount DESC;

```

77.87 | 529.02 | 41194.44 |

Задача 2: Частота покупок

```

WITH time_new AS (--перевод времени в формат date
SELECT *, to_date(date, '%YYYY%MM%DD')
FROM fantasy.events),

plus_duration AS (--добавление столбца периода между покупками
SELECT *, to_date - LAG(to_date) OVER(PARTITION BY id ORDER BY to_date) AS
duration
FROM time_new),

next AS (--вывод нужных столбцов и условий
SELECT plus_duration.id AS users, payer, COUNT(amount) AS count_amount,
ROUND(AVG(duration::numeric),2) AS avg_duration
FROM plus_duration
LEFT JOIN fantasy.users USING(id)
WHERE amount <> 0.0
GROUP BY users, payer
HAVING COUNT(amount) >= 25),

finish AS (--ранжирование таблицы
SELECT *, NTILE(3) OVER(ORDER BY avg_duration) AS group_duration
FROM next),

paid AS (SELECT COUNT(DISTINCT users) AS paid_users, group_duration--подсчет
платящих игроков
FROM finish
WHERE payer = 1
GROUP BY group_duration)

SELECT CASE WHEN group_duration = 1 THEN 'высокая частота'
WHEN group_duration = 2 THEN 'умеренная частота'
WHEN group_duration = 3 THEN 'низкая частота'
END AS rate,
COUNT(DISTINCT users) AS total_users_amount,
paid_users AS paid_users,
ROUND(paid_users/COUNT(DISTINCT users)::numeric,2) AS dolya,
ROUND(AVG(count_amount::numeric),2) AS avg_count_amount_group,
ROUND(AVG(avg_duration::numeric),2) AS avg_duration_group
FROM finish
LEFT JOIN paid USING(group_duration)
GROUP BY group_duration, paid_users
ORDER BY avg_duration_group;

```

Часть 3. Выводы и аналитические комментарии

1. Результаты исследовательского анализа данных:

1.1. Какая доля платящих игроков характерна для всей игры и как раса персонажа влияет на изменение этого показателя?

- общее количество зарегистрированных игроков 22 214 чел.
- около 18% от них игроков являются платящими игроками или игроками совершающими покупки за реальные деньги
- доли платящих игроков в зависимости от расы персонажа близки к значению доли платящих игроков от общего количества игроков

1.2. Сколько было совершено внутриигровых покупок и что можно сказать об их стоимости (минимум и максимум, есть ли различие между средним значением и медианой, какой разброс данных)?

- всего было совершено 1 307 678 покупок
- минимальная покупка (за исключением нулевых) составляет 0,01 игровой валюты, максимальная - 486 615,1 и.в.
- средняя стоимость покупки составляет 526 и.в.
- при этом медиана равна 75, что говорит о том что 50 % покупок совершаются по низкой цене, в том числе 0,00. Скорее всего большая часть из них это наиболее популярные эпические предметы, которые необходимы для базового уровня

1.3. Есть ли аномальные покупки по стоимости? Если есть, то сколько их?

- доля нулевых покупок составляет 0.000694 или 907 шт.
- нулевых покупки в разрезе расы персонажа выявили наибольшую популярность таких покупок расой Elfof и без указания продавца, возможно это обмен с другими игроками

1.4. Сколько игроков совершают внутриигровые покупки и насколько активно? Сравнение поведения платящих и неплатящих игроков.

- количество игроков совершивших покупки, но не за "райские лепестки" превышает в 4,6 раза игроков совершающих покупки за реальную валюту (11348 к 2444 чел.)
- количество покупок совершаемых неплатящим игроком больше в среднем на 16%, в то время как сумма таких покупок меньше на 12%
- суммарная сумма покупок платящих и неплатящих игроков превышает расчетную сумму в общем по данной статье, скорее всего это аномалия могла возникнуть в связи попаданием некоторых игроков в обе категории

1.5. Есть ли среди эпических предметов популярные, которые покупают чаще всего?

- наибольшей популярностью среди эпических предметов пользуется Book of Legends (77 % от количества всех покупок), количество его покупок кратно превышает остальные. На втором и третьем месте Bag of Holding (21%) и Necklace of Wisdom (1%) соответственно. В тоже время доля игроков покупающих такие эпические предметы как Book of Legends и Bag of Holding составляет около 87.5%. Скорее всего эти предметы первой необходимости, для того чтобы разобраться и продвигаться в игре.
- покупки остальных предметов ниже 1% в общей доле
- платящие игроки в среднем в 4,6 раза охотнее покупают эпические предметы

2. Результаты решения ad hoc задач

2.1. Существует ли зависимость активности игроков по совершению внутриигровых покупок от расы персонажа?

- все игроки зарегистрированные в игре распределяются на персонажи следующим образом: Human (28%), Hobbit, Orc, Northman (по 16%), Elf (11%) и Angel и Demon (по 6%).
- в не зависимости от выбранного персонажа покупки совершают около 62% игроков, около 18 % из них являются платящими игроками.
- при этом количество эпических предметов в среднем покупаемые персонажем Human больше (около 120 предметов на игрока), меньше всего предметов покупают игроки выбравшие Demon и Elf в качестве персонажа (около 78 предметов) остальные персонажи покупают в среднем от 81 до 106 эпических предметов.
- в тоже время средняя цена покупки для персонажа Human самая низкая (404 игровой валюты), а у персонажа Northman самая высокая (760 игровой валюты)
- по средней итоговой сумме затрат на эпические предметы лидируют персонажи Northman (62,5 тыс.). Меньше всего тратят персонажи Demon и Orc (около 41,5 тыс.).

2.2. Как часто игроки совершают покупки?

- все игроки совершающие покупки были разделены на 3 равные по количеству участников группы по показателю периода между покупками с учетом только игроков совершивших суммарно более 25 покупок эпических предметов
- доля платящих игроков в каждой группе оказалась примерно равна и составляет 18 % от общего количества игроков совершивших покупки
- игроки попавшие в группу “высокая частота” в разы чаще совершают покупки (в среднем 391 шт. на игрока, к 59 в группе “умеренная частота” и 34 в группе “низкая частота”
- в группе “высокая частота” покупка эпического предмета совершается в среднем раз в 3 дня, в “умеренная частота” раз в 7,5 дней, “низкая частота” раз в 13 дней.

3. Общие выводы и рекомендации

Анализ показывает, что начинающие использовать внутриигровую валюту игроки совершают более дорогие покупки, но реже. Вероятно, такие покупки более обдуманные и направлены на улучшения, которые невозможно осуществить за стандартную валюту. Также высокая доля покупок эпических предметов среди платящих игроков свидетельствует о том, что покупки за реальные деньги увеличивают их активность. Рекомендуется усилить наличие эпических предметов, доступных только за внутриигровую валюту, особенно второго по значимости, так как это может стимулировать дополнительные внутриигровые транзакции и увеличить доход.

В ходе анализа зависимости внутриигровых покупок от расы персонажа выявлена неравномерность распределения. Игроки, выбирающие персонажа Human, совершают больше покупок по количеству предметов, тогда как для персонажа Northman суммарные затраты выше, хотя количество покупок меньше. Самые низкие затраты и меньшие показатели по количеству покупок характерны для игроков с персонажами Demon и Orc, что свидетельствует о меньших затратах как в количестве, так и в денежном выражении.

На основе данных видно, что наиболее активные покупатели совершают значительно большие покупки. В то время как их стимул практически очевиден, продажи среди остальных групп игроков требуют активных мер по стимулированию. Между тем, уровень покупательской активности в этих категориях значительно ниже и требует дополнительного внимания для увеличения объема внутриигровых транзакций.