

## ➤ مقدمه و هدف تسک

در این تست، کاربران به سه گروه مختلف با سطوح متفاوت هزینه عملیاتی تخصیص داده شده‌اند:

گروه A: بدون هزینه عملیاتی (0 تومان)

گروه C: با هزینه عملیاتی 40,000 تومان (بر اساس داده‌های واقعی فایل CSV)

گروه B: با هزینه عملیاتی 80,000 تومان (بر اساس داده‌های واقعی فایل CSV)

توجه: این تحلیل بر اساس داده‌های واقعی موجود در فایل CSV انجام شده است، که در آن تخصیص هزینه‌ها به گروه‌های B و C با توضیحات اولیه تسک متفاوت بود (هزینه گروه B در CSV معادل 80,000 و گروه C معادل 40,000 تومان است). نام‌گذاری گروه‌ها (A, B, C) برای حفظ سادگی مطابق با توضیحات اولیه تسک حفظ شده است.

هدف اصلی این تحلیل، بررسی جامع داده‌های این آزمایش، ارزیابی عملکرد هر یک از گروه‌ها بر اساس متریک‌های کلیدی، شناسایی گروه با عملکرد بهینه (در صورت وجود تفاوت معنی‌دار) از منظر مالی و رفتاری کاربران، و در نهایت ارائه بینش‌ها و پیشنهادات عملی برای کمک به تصمیم‌گیری استراتژیک شرکت ازکی در خصوص سیاست‌گذاری هزینه عملیاتی است.

## 2. مراحل انجام تحلیل

تحلیل حاضر شامل مراحل زیر بوده است:

- ✓ آماده‌سازی و پاکسازی داده‌ها: بارگذاری دیتاست، بررسی اولیه، مدیریت مقادیر گمشده و تبدیل انواع داده‌ها.
- ✓ تحلیل اکتشافی داده‌ها (EDA): بررسی توزیع متغیرهای مختلف و شناسایی الگوهای اولیه.
- ✓ تعریف متریک‌های ارزیابی: انتخاب متریک‌های کلیدی برای سنجش عملکرد.
- ✓ تحلیل نتایج تست و مقایسه گروه‌ها: محاسبه متریک‌ها، انجام آزمون‌های آماری برای بررسی معنی‌داری تفاوت‌ها، و بصری‌سازی نتایج.
- ✓ نتیجه‌گیری و ارائه پیشنهادات: جمع‌بندی یافته‌ها و ارائه توصیه‌های عملی.

### 3. آماده‌سازی و پاکسازی داده‌ها

دیتاست ارائه‌شده در فایل data.csv شامل 5000 رکورد و 10 ستون اولیه بود. اقدامات انجام‌شده در این مرحله:

تجزیه ستون group: ستون group (حاوی رشته JSON) به دو ستون مجزا تفکیک شد:

group\_letter: حرف گروه (A, B, C).

operation\_cost: هزینه عملیاتی (0 برای A، 40000 برای C، و 80000 برای B).

توزیع کاربران در گروه‌ها پس از این مرحله به شرح زیر بود:

- گروه A (هزینه 0): 1688 کاربر
- گروه C (هزینه 40,000): 1668 کاربر
- گروه B (هزینه 80,000): 1644 کاربر

تبدیل انواع داده:

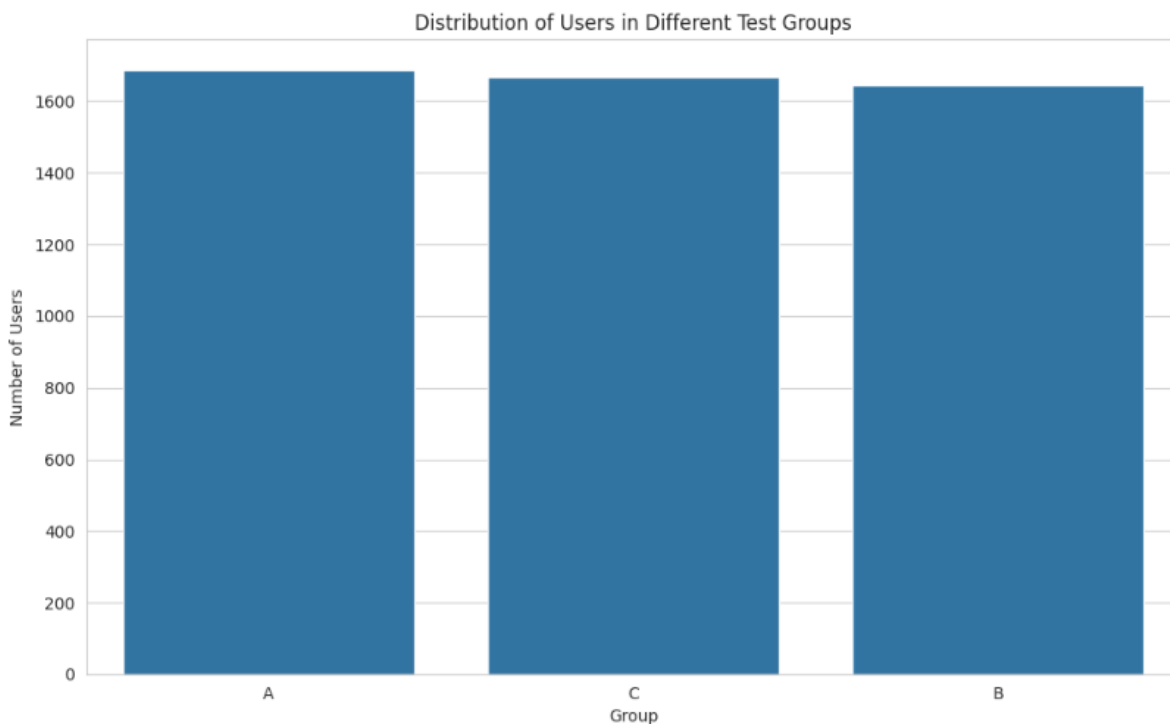
- ✓ ستون date از نوع object به فرمت استاندارد datetime تبدیل شد.
- ✓ ستون price به نوع عددی float تبدیل شد. هیچ ردیفی به دلیل قیمت نامعتبر حذف نشد.
- ✓ مدیریت مقادیر گمشده (Missing Values):
- ✓ ستون user\_device دارای 28 مقدار گمشده بود. این مقادیر با رشته 'unknown' جایگزین شدند. این روش انتخاب شد تا از دست رفتن اطلاعات سایر ستون‌ها در این ردیف‌ها جلوگیری شود و امکان بررسی این دسته به صورت جداگانه فراهم آید. ستون renewal\_status در داده‌های اولیه فاقد مقدار گمشده بود.

**بررسی نهایی:** پس از پاکسازی، دیتاست شامل 5000 رکورد و 11 ستون (با اضافه شدن group\_letter و operation\_cost و حذف group اصلی) و بدون مقدار گمشده بود.

### 4. تحلیل اکتشافی داده‌ها (EDA)

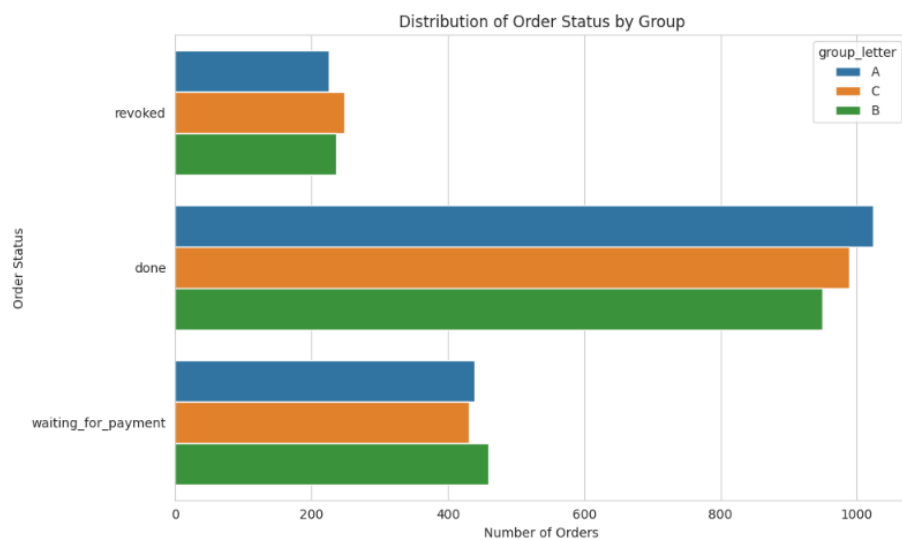
در این مرحله، به بررسی ویژگی‌های داده‌ها و توزیع متغیرهای مختلف پرداخته شد:

توزیع کاربران در گروه‌ها: همانطور که در مرحله پاکسازی ذکر شد، تعداد کاربران در هر سه گروه تقریباً یکسان و متعادل بود (گروه A: ~33.8%، گروه C: ~33.4%، گروه B: ~32.9%).

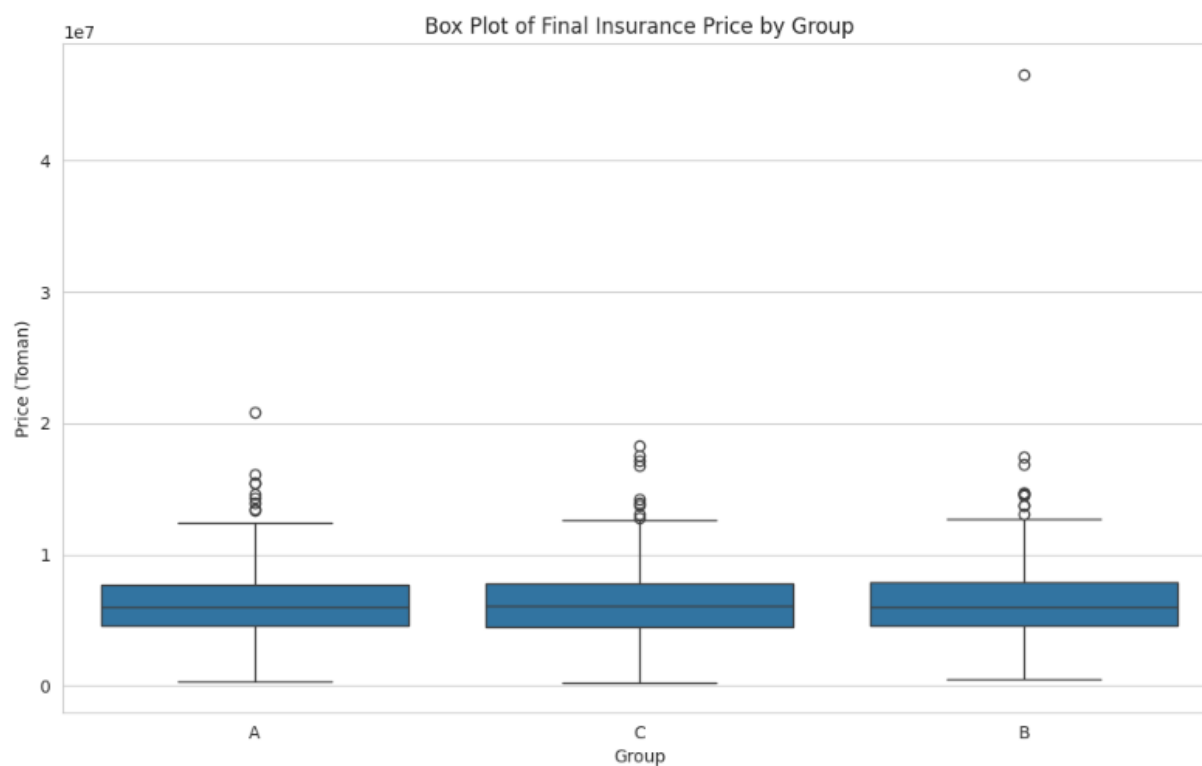
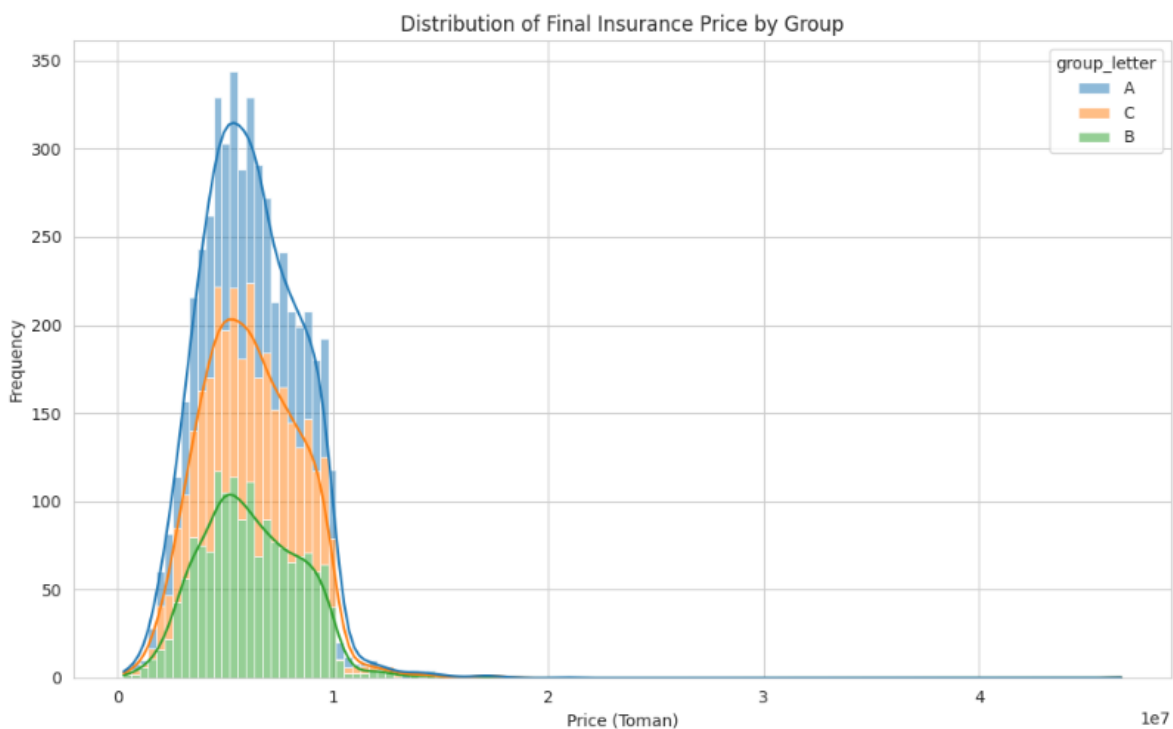


توزیع وضعیت سفارش (order\_status):

اکثریت سفارش‌ها (حدود 59.2%) با وضعیت done (موفق) ثبت شده‌اند. حدود 26.6% در وضعیت waiting\_for\_payment و 14.2% در وضعیت revoked بودند. از نظر تعداد مطلق، گروه A بیشترین سفارش موفق و گروه B کمترین را داشت.



توزیع قیمت نهایی بیمه (price):



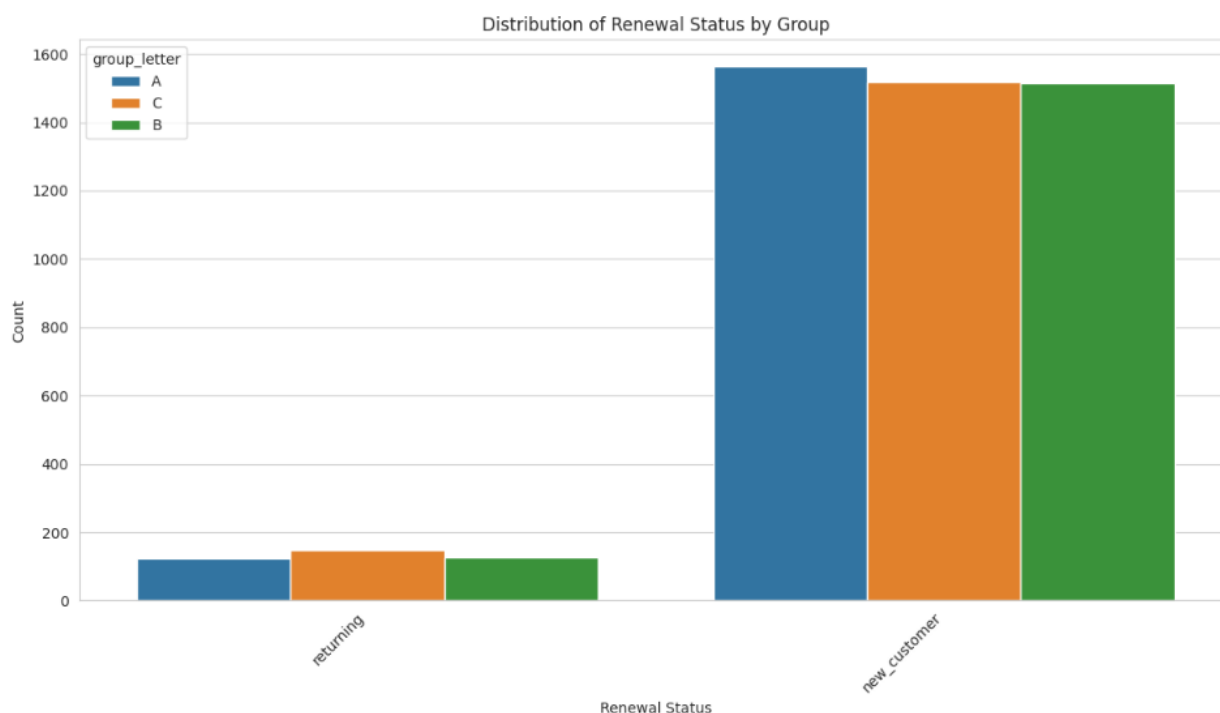
هیستوگرام قیمت‌ها چولگی به راست و وجود یک داده پرت بسیار بزرگ (حدود 4.66 میلیون تومان) را به خصوص در گروه B نشان داد. این موضوع اهمیت استفاده از آزمون‌های آماری مقاوم به داده‌های پرت (مانند **Mann-Whitney U**) را برای مقایسه AOV بیشتر می‌کند.

باکس پلات نشان داد که میانه قیمت‌ها در سه گروه تفاوت فاحشی با یکدیگر ندارد، هرچند پراکندگی و داده‌های پرت در گروه B بیشتر به نظر می‌رسد.

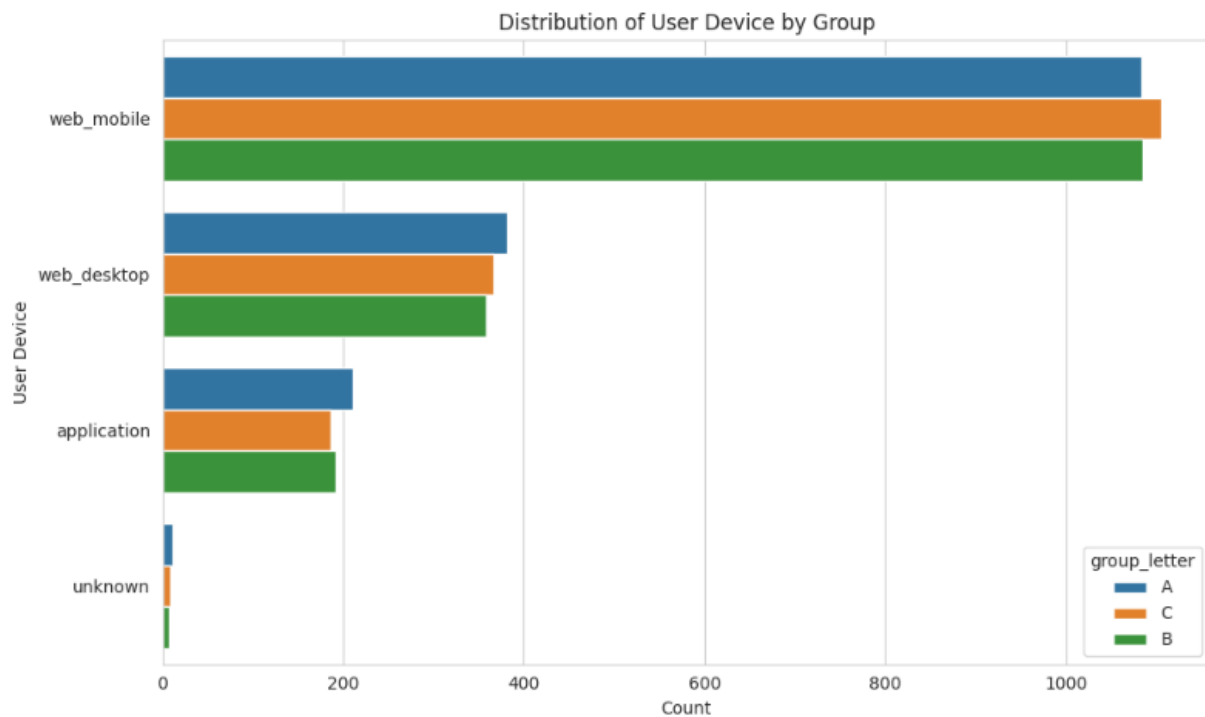
توزیع سایر متغیرهای کلیدی:

نوع بیمه (type): تمامی 5000 رکورد مربوط به vehicle\_body\_insurance بودند.

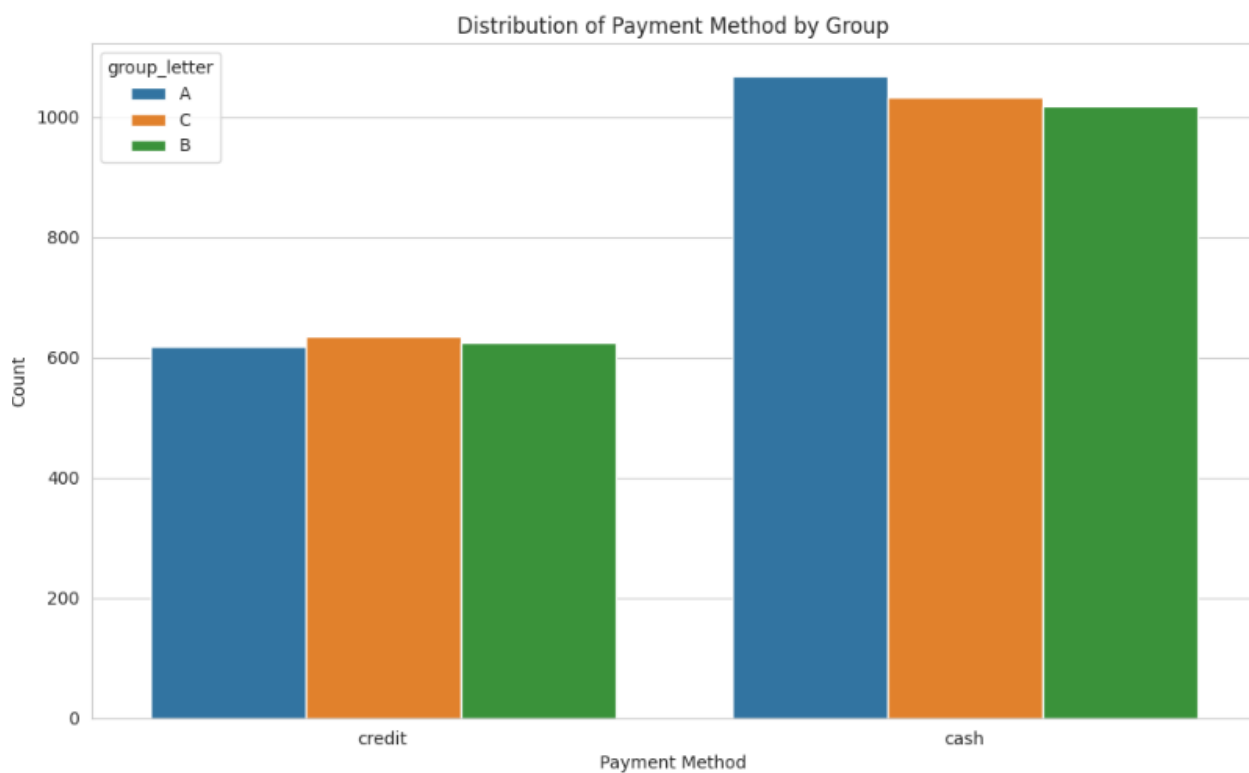
وضعیت تمدید (renewal\_status): اکثریت غالب کاربران (new\_customer، حدود 92%) مشتریان جدید بودند.



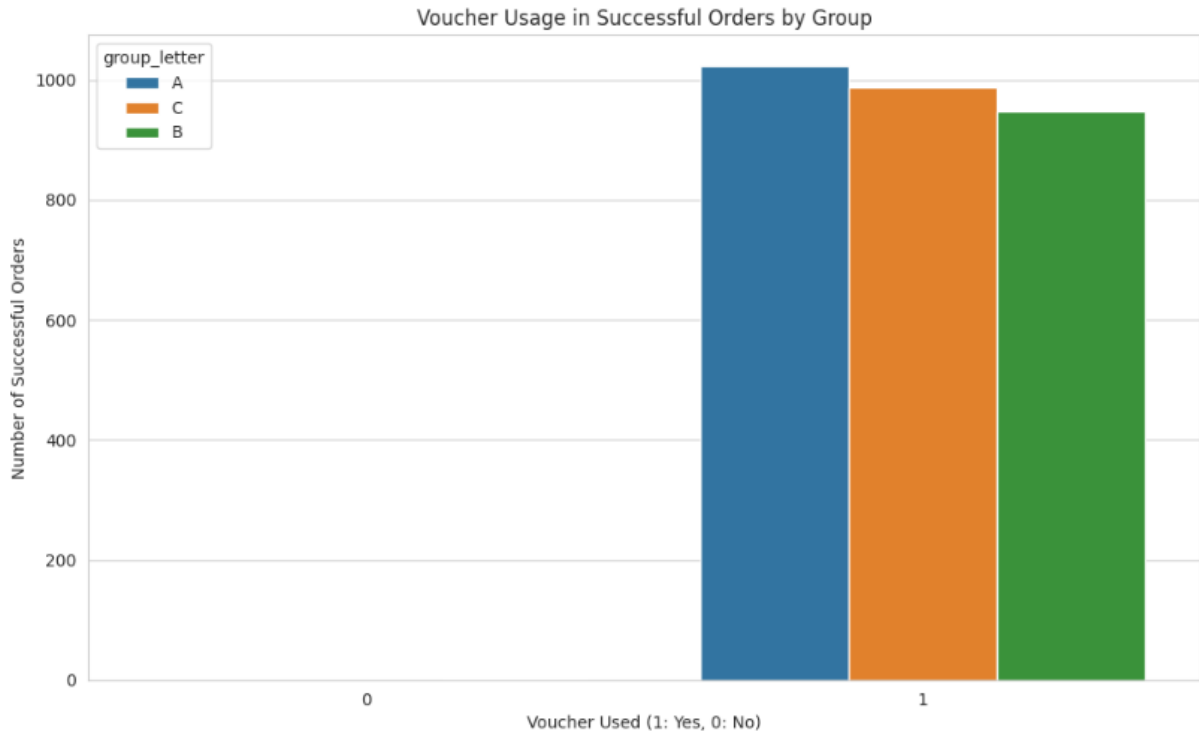
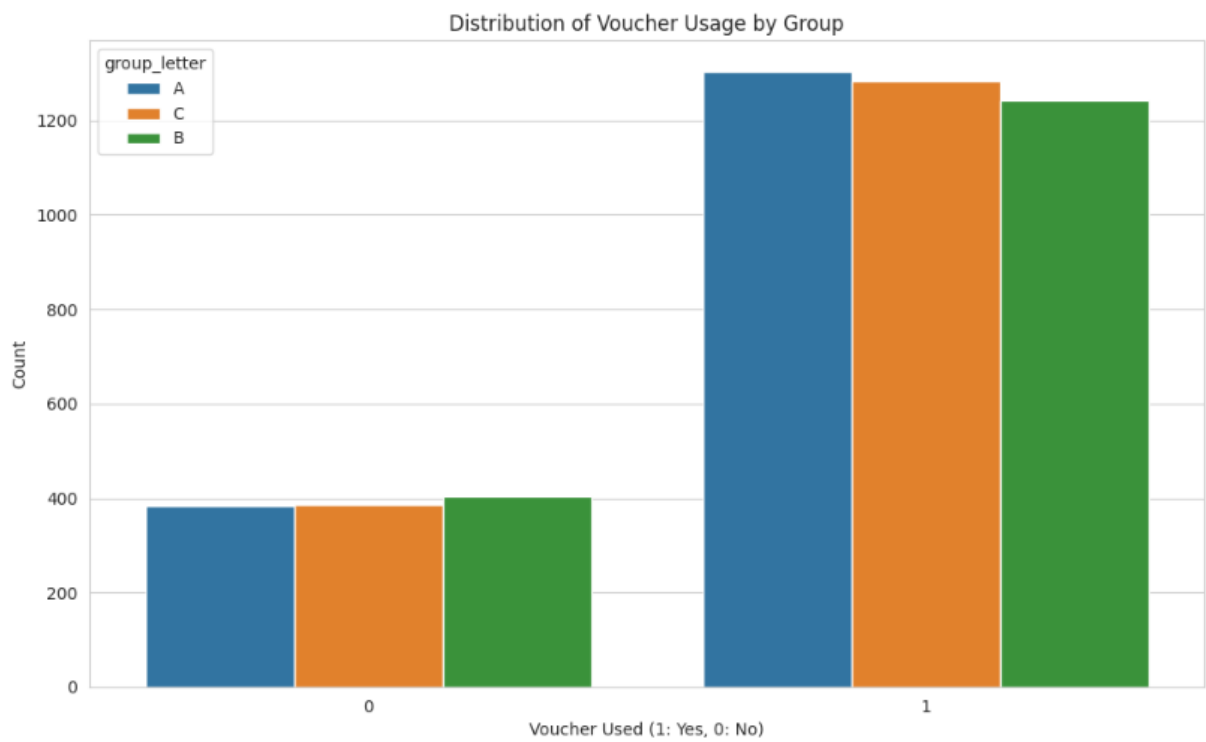
دستگاه کاربر (user\_device): web\_mobile بیشترین استفاده را داشته (بیش از 65%)، سپس web\_desktop (حدود 25%) و application (حدود 8%).



روش پرداخت (payment\_method): پرداخت نقدی (cash) با حدود 62% بیشتر از پرداخت اعتباری/آنلاین (credit) با حدود 38% رایج بوده است.



ووچر (voucher): حدود 76.5% از کاربران از ووچر استفاده کرده‌اند.



## 5. تعریف متریک‌های ارزیابی عملکرد

برای ارزیابی دقیق عملکرد هر گروه، متریک‌های زیر تعریف شدند:

1. نرخ تبدیل (CR): تعداد سفارش‌های موفق / تعداد کل کاربران در گروه
2. میانگین درآمد هر خرید موفق (AOV): میانگین price برای سفارش‌های موفق
3. کل درآمد (Total Revenue): مجموع price برای سفارش‌های موفق
4. درآمد به ازای هر کاربر (RPU): کل درآمد گروه / تعداد کل کاربران گروه (متریک کلیدی برای تصمیم‌گیری)
5. کل هزینه عملیاتی جمع‌آوری شده: مجموع operation\_cost برای سفارش‌های موفق

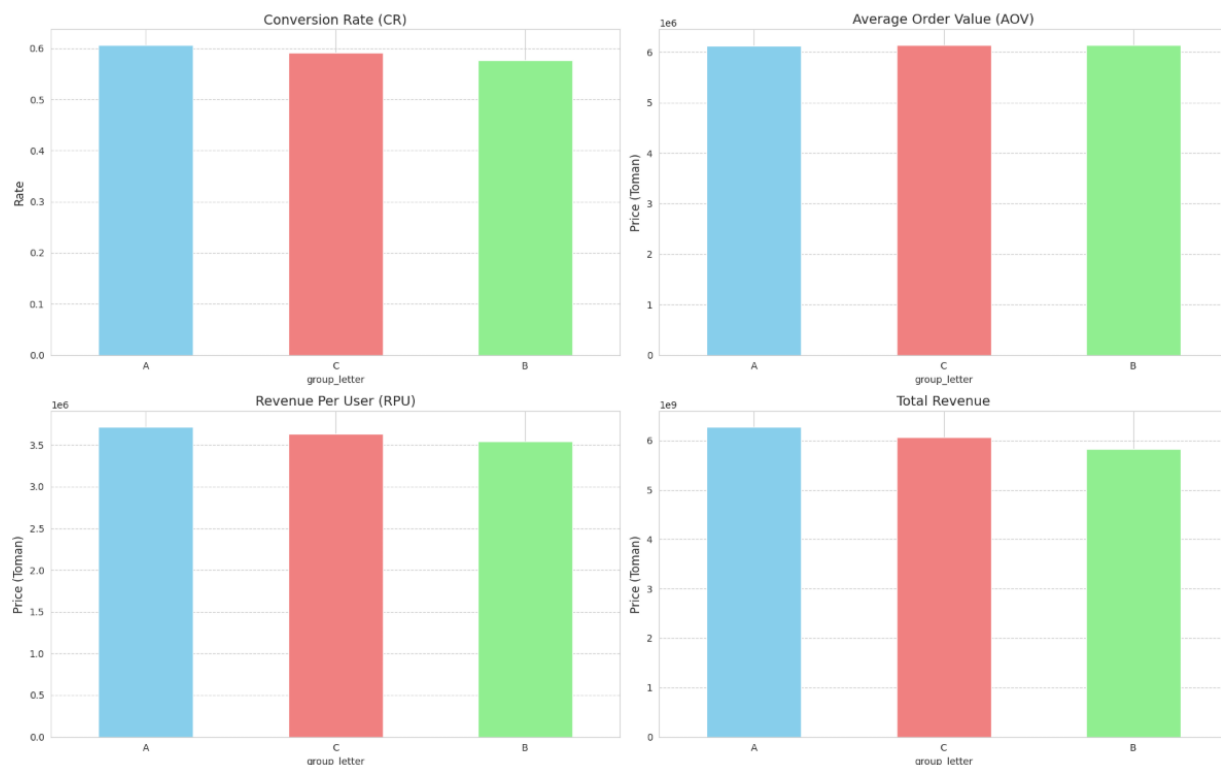
## 6. تحلیل نتایج تست و مقایسه گروه‌ها

6.1. خلاصه نتایج متریک‌های کلیدی

نتایج محاسبه شده برای متریک‌های کلیدی به شرح زیر است:

| متریک                         | گروه A (0 تومان)    | گروه C (40,000 تومان) | گروه B (80,000 تومان) |
|-------------------------------|---------------------|-----------------------|-----------------------|
| تعداد کل کاربران              | 1688                | 1668                  | 1644                  |
| تعداد خرید موفق               | 1024                | 988                   | 949                   |
| نرخ تبدیل (CR)                | 60.66%              | 59.23%                | 57.73%                |
| میانگین درآمد هر خرید (AOV)   | تومان 6,130,148     | تومان 6,141,403       | تومان 6,140,145       |
| کل درآمد                      | تومان 6,277,271,147 | تومان 6,067,706,147   | تومان 5,826,997,433   |
| درآمد به ازای هر کاربر (RPU)  | تومان 3,718,763     | تومان 3,637,713       | تومان 3,544,402       |
| کل هزینه عملیاتی جمع‌آوری شده | 0 تومان             | تومان 39,520,000      | تومان 75,920,000      |





مشاهدات اولیه از جدول:

نرخ تبدیل به صورت عددی از گروه A به C و سپس به B کاهش می‌یابد.  
 AOV در گروه‌های C و B بسیار نزدیک و کمی بالاتر از A است.  
 RPU و کل درآمد برای گروه A به صورت عددی بالاترین مقدار را دارند.

## 6.2. تحلیل معنی‌داری آماری تفاوت‌ها

دلیل انتخاب تست‌های آماری:

برای مقایسه نرخ تبدیل (CR) که یک متغیر نسبتی (دو جمله‌ای) است، از آزمون Z برای دو نسبت (proportions\_ztest) استفاده شد. این آزمون برای مقایسه اینکه آیا تفاوت مشاهده شده در نسبت‌ها بین دو گروه مستقل، از نظر آماری معنی‌دار است یا خیر، مناسب می‌باشد.

برای مقایسه میانگین درآمد هر خرید موفق (AOV)، با توجه به اینکه توزیع قیمت‌ها (همانطور که در EDA مشاهده شد) نرمال نبود و دارای چولگی و داده‌های پرت بود، از آزمون غیرپارامتریک Mann-Whitney U استفاده شد. این آزمون نیازی به فرض نرمال بودن داده‌ها ندارد و برای مقایسه میانه (یا به طور کلی توزیع) دو گروه مستقل مناسب است.

سطح معنی‌داری ( $\alpha$ ) برای تمام آزمون‌ها 0.05 در نظر گرفته شد.

نتایج آزمون‌ها:

نرخ تبدیل (CR):

✓ مقایسه گروه A و گروه C:

Z-statistic: 0.846

P-value: 0.3976

نتیجه: غیر معنی‌دار. تفاوت مشاهده شده در نرخ تبدیل (60.66% برای A و 59.23% برای C) از نظر آماری در سطح اطمینان 95% معنی‌دار نیست.

✓ مقایسه گروه A و گروه B:

Z-statistic: 1.726

P-value: 0.0844

نتیجه: غیر معنی‌دار. اگرچه P-value به 0.05 نزدیک‌تر است، اما همچنان بزرگتر بوده و تفاوت مشاهده شده در نرخ تبدیل (60.66% برای A و 57.73% برای B) از نظر آماری معنی‌دار نیست.

✓ مقایسه گروه C و گروه B:

Z-statistic: 0.880

P-value: 0.3787

نتیجه: غیر معنی‌دار. تفاوت مشاهده شده در نرخ تبدیل (59.23% برای C و 57.73% برای B) از نظر آماری معنی‌دار نیست.

تفسیر CR: نتایج آزمون‌های آماری نشان می‌دهد که با اطمینان 95% نمی‌توان گفت که اعمال هزینه‌های عملیاتی 40,000 یا 80,000 تومان تأثیر واقعی و معنی‌داری بر نرخ تبدیل کاربران داشته است. اگرچه به صورت عددی کاهش مشاهده می‌شود، این کاهش می‌تواند ناشی از نوسانات تصادفی در نمونه باشد.

میانگین درآمد هر خرید موفق (AOV):

✓ مقایسه گروه A و گروه C:

U-statistic: 505677

P-value: 0.9891

نتیجه: غیر معنی‌دار.

✓ مقایسه گروه A و گروه B:

U-statistic: 491260

P-value: 0.6710

نتیجه: غیر معنی‌دار.  
✓ مقایسه گروه C و گروه B:

U-statistic: 473454

P-value: 0.7056

نتیجه: غیر معنی‌دار.

تفسیر AOV: به طور شگفت‌آوری، تفاوت‌های عددی مشاهده شده در AOV بین گروه‌ها (که انتظار می‌رفت حداقل به اندازه هزینه عملیاتی اضافه شده تفاوت داشته باشند) نیز از نظر آماری معنی‌دار نیستند. این می‌تواند به دلایلی مانند واریانس بسیار بالای قیمت‌ها در هر گروه (وجود داده‌های پرت)، یا حجم نمونه ناکافی سفارش‌های موفق برای تشخیص دقیق این تفاوت‌ها با توان آماری بالا باشد.

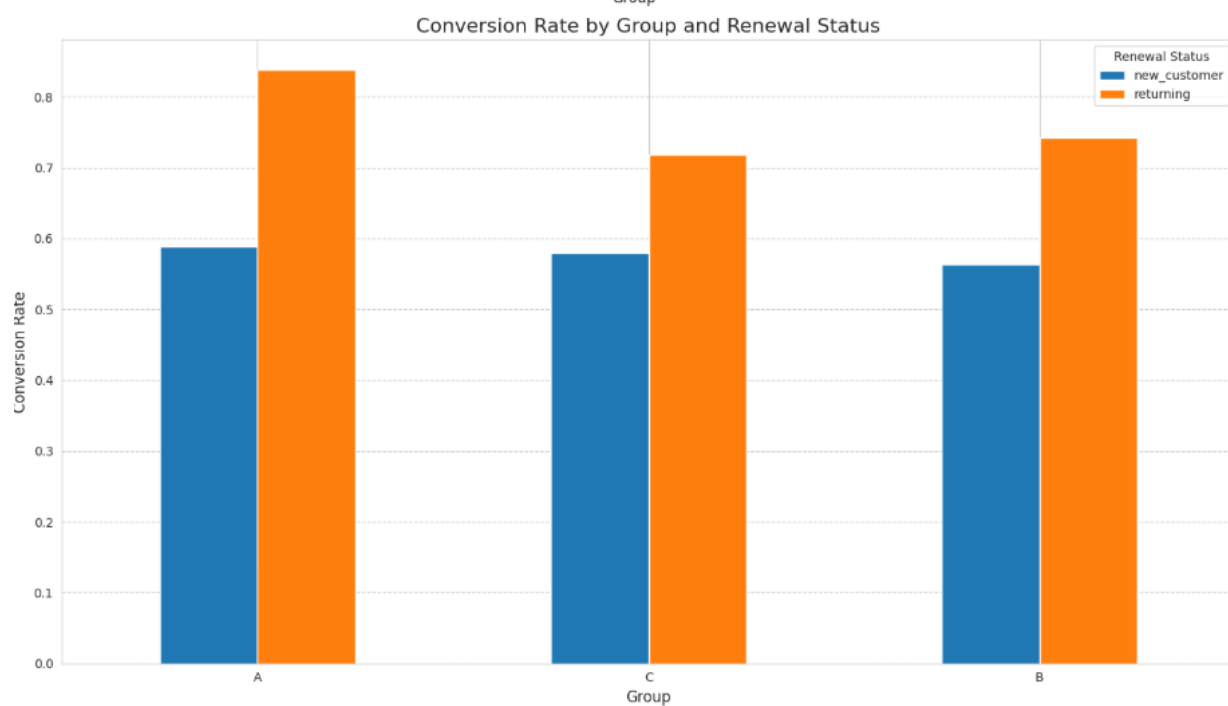
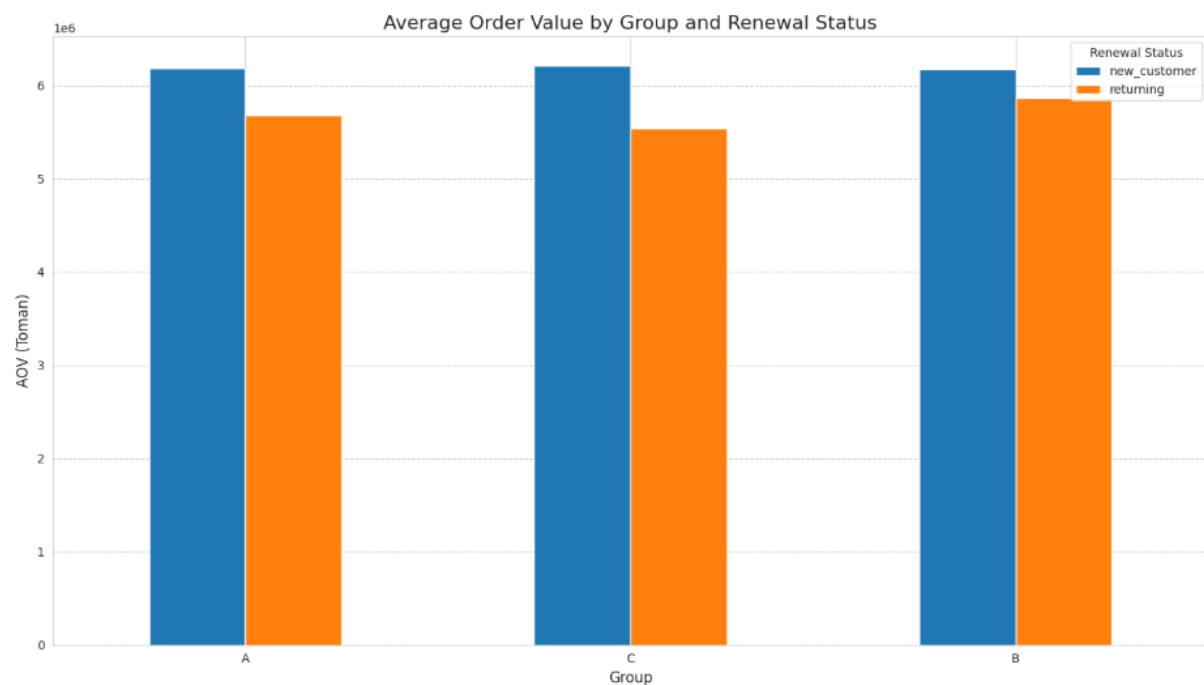
### 6.3. تحلیل درآمد به ازای هر کاربر (RPU)

با وجود عدم معنی‌داری آماری در CR و AOV به صورت مجزا، بررسی RPU همچنان می‌تواند به عنوان یک شاخص ترکیبی مفید باشد، اما باید با احتیاط تفسیر شود:

گروه A (بدون هزینه): 3,718,763 تومان  
گروه C (هزینه 40,000): 3,637,713 تومان  
گروه B (هزینه 80,000): 3,544,402 تومان

تفسیر RPU: به صورت عددی، گروه A بالاترین RPU را دارد. گروه C حدود 2.18% RPU کمتری نسبت به A و گروه B حدود 4.69% RPU کمتری نسبت به A دارد. با این حال، از آنجایی که تفاوت‌های CR و AOV که این RPU را تشکیل می‌دهند از نظر آماری معنی‌دار نبودند، نمی‌توان با قطعیت گفت که این تفاوت‌های عددی در RPU نیز در جامعه آماری بزرگتر پایدار خواهند بود.

### 6.4. تحلیل بر اساس ویژگی‌های کاربران (مثال: وضعیت تمدید - renewal\_status)



### تفسیر CR بر اساس وضعیت تمدید:

مشتریان تمدیدکننده (returning) به طور قابل ملاحظه‌ای نرخ تبدیل بالاتری در تمام گروه‌ها نشان دادند (مثلاً در گروه A، نرخ تبدیل تمدیدکنندگان 83.87٪ در مقابل 58.82٪ برای مشتریان جدید بود).

در بین مشتریان جدید، روند عددی کاهش CR با افزایش هزینه ( $A > C > B$ ) مشاهده شد.

در بین مشتریان تمديدکننده، روند عددی CR کمی متفاوت بود ( $A > B > C$ ).

### تفسير AOV بر اساس وضعیت تمديد:

برای مشتریان جدید، AOV در گروه‌های C و B کمی بالاتر از A بود.

برای مشتریان تمديدکننده، AOV گروه B بالاتر از A و C بود.

نکته مهم: با توجه به عدم معنی‌داری کلی در CR و AOV، تفاوت‌های مشاهده شده در این سگمنت‌ها نیز احتمالاً از نظر آماری معنی‌دار نیستند و باید با احتیاط تفسیر شوند.

### 7. نتیجه‌گیری نهایی

بر اساس تحلیل جامع داده‌های A/B تست و با در نظر گرفتن نتایج آزمون‌های معنی‌داری آماری:

1. **عدم وجود تفاوت آماری معنی‌دار:** مهم‌ترین یافته این است که تفاوت‌های مشاهده شده در نرخ تبدیل (CR) و میانگین ارزش سفارش (AOV) بین سه گروه آزمایشی (با هزینه‌های عملیاتی 0، 40,000 و 80,000 تومان) در سطح اطمینان 95% از نظر آماری معنی‌دار نبودند. این بدان معناست که بر اساس این تست، نمی‌توان با قطعیت نتیجه گرفت که اعمال این سطوح از هزینه عملیاتی تأثیر واقعی و قابل اتکایی بر تمایل کاربران به خرید یا میزان خرید آن‌ها داشته است.

### 2. عملکرد عددی متریک‌ها:

- ✓ با وجود عدم معنی‌داری آماری، به صورت عددی، گروه A (بدون هزینه عملیاتی) بالاترین نرخ تبدیل (60.66%) و بالاترین درآمد به ازای هر کاربر (RPU معادل 3,718,763 تومان) را نشان داد.
- ✓ گروه‌های C و B که هزینه عملیاتی از آن‌ها دریافت شده بود، به صورت عددی CR و RPU کمتری داشتند.
- ✓ AOV در سه گروه تفاوت عددی بسیار کمی داشت که این نیز از نظر آماری معنی‌دار نبود.

### 3. دلایل احتمالی عدم معنی‌داری:

**اثر واقعی کوچک:** ممکن است تأثیر واقعی این سطوح از هزینه عملیاتی بر رفتار کاربران آنقدر کوچک باشد که با حجم نمونه فعلی قابل تشخیص نباشد.

**حجم نمونه یا توان آماری ناکافی:** برای تشخیص اثرات کوچک، به حجم نمونه بزرگتری نیاز است. شاید تعداد کل کاربران یا تعداد سفارشات موفق در هر گروه برای رسیدن به توان آماری مطلوب کافی نبوده است.

**واریانس بالای داده‌ها:** پراکندگی زیاد در متغیرهایی مانند قیمت می‌تواند تشخیص تفاوت‌های کوچک را دشوارتر کند.

در مجموع، با توجه به نتایج آماری این تست، نمی‌توان به طور قطعی یکی از گزینه‌های هزینه عملیاتی را بر دیگری برتر دانست. هرچند گروه A به صورت عددی عملکرد بهتری در RPU داشته، این برتری از نظر آماری پشتیبانی نمی‌شود.

## 8. پیشنهادات عملی

با توجه به عدم قطعیت آماری در نتایج تست، پیشنهادات زیر ارائه می‌گردد:

1. اجرای مجدد تست با حجم نمونه بزرگتر و/یا مدت زمان طولانی‌تر: برای افزایش توان آماری و امکان تشخیص تفاوت‌های واقعی (اگر وجود داشته باشند)، اجرای مجدد تست با تعداد کاربران بیشتر در هر گروه یا برای مدت زمان طولانی‌تر توصیه می‌شود. قبل از اجرای مجدد، محاسبه حجم نمونه مورد نیاز (Sample Size Calculation) بر اساس حداقل اثر قابل تشخیص (Minimum Detectable Effect - MDE) مد نظر، مفید خواهد بود.

2. بررسی بازه‌های هزینه‌ای متفاوت یا تفاوت‌های بزرگتر:

ممکن است هزینه‌های تست شده (40,000 و 80,000 تومان) برای ایجاد تغییر رفتار قابل توجه کافی نبوده باشند یا برعکس، نقطه بهینه در بازه دیگری قرار داشته باشد. تست کردن سطوح هزینه متفاوت (مثلاً 20,000 تومان یا حتی 100,000 تومان، بسته به استراتژی کسب و کار) می‌تواند بینش‌های جدیدی ارائه دهد.

3. حفظ وضعیت موجود (در صورت عدم دریافت هزینه عملیاتی): اگر استراتژی فعلی شرکت عدم دریافت هزینه عملیاتی است و این تست نتوانسته برتری معنی‌داری برای دریافت هزینه نشان دهد، یک گزینه محتاطانه ادامه همین روند تا زمان کسب اطلاعات بیشتر از تست‌های آتی است. گروه A به صورت عددی RPU بالاتری داشته است.

4. تمرکز بر بهینه‌سازی‌های با تأثیر قطعی‌تر: با توجه به عدم قطعیت در مورد تأثیر هزینه عملیاتی، منابع شرکت می‌تواند به سمت بهینه‌سازی‌هایی هدایت شود که پتانسیل تأثیرگذاری بیشتری دارند. همانطور که در EDA مشاهده شد، درصد قابل توجهی از سفارش‌ها لغو شده یا در انتظار پرداخت باقی مانده‌اند. بهبود قیف تبدیل، تجربه کاربری، فرآیند پرداخت و کاهش نقاط اصطکاک در مسیر خرید می‌تواند تأثیر مستقیم و قابل اندازه‌گیری بر افزایش نرخ تبدیل و درآمد داشته باشد.

5. تحلیل عمیق‌تر سگمنت‌ها (در تست‌های آتی): در صورت اجرای تست‌های بعدی با حجم نمونه کافی، تحلیل رفتار سگمنت‌های مختلف مشتریان (مثلاً بر اساس وفاداری، ارزش طول عمر، نوع بیمه خریداری شده، یا ویژگی‌های دموگرافیک) نسبت به سطوح مختلف هزینه عملیاتی می‌تواند به شناسایی فرصت‌های قیمت‌گذاری متمایز کمک کند.

6. در نظر گرفتن "هزینه فرصت": اگر شرکت تصمیم به عدم دریافت هزینه عملیاتی بگیرد (گزینه A)، باید در نظر داشت که درآمد بالقوه از محل جمع‌آوری این هزینه را از دست می‌دهد. این موضوع باید در کنار تأثیر احتمالی (هرچند در این تست غیر معنی‌دار) بر نرخ تبدیل سنجیده شود.

نتیجه نهایی برای تصمیم‌گیری: با داده‌های فعلی، انتخاب قطعی بین گزینه‌ها دشوار است. اما اگر مجبور به انتخاب بر اساس این تست باشیم و با در نظر گرفتن اینکه گروه A به صورت عددی بهترین RPU را داشته و اعمال هزینه عملیاتی ریسک (هرچند در اینجا غیر معنی‌دار) کاهش CR را به همراه دارد، عدم دریافت هزینه عملیاتی (گزینه A) یا دریافت هزینه بسیار کم (که در این تست بررسی نشده) می‌تواند گزینه کم‌ریسک‌تر و با پتانسیل درآمدزایی بالاتر برای کل کاربران باشد. اما تأکید می‌شود که این نتیجه‌گیری با احتیاط و با در نظر گرفتن عدم معنی‌داری آماری است.

## 9. محدودیت‌های تحلیل

- ✓ عدم معنی‌داری آماری نتایج: مهم‌ترین محدودیت این تحلیل، عدم توانایی در اثبات تفاوت‌های معنی‌دار آماری بین گروه‌ها برای CR و AOV است.
- ✓ حجم نمونه و توان آماری: همانطور که ذکر شد، ممکن است حجم نمونه برای تشخیص اثرات کوچک کافی نبوده باشد.
- ✓ دوره زمانی و نوع بیمه: تحلیل محدود به یک دوره زمانی و یک نوع بیمه (vehicle\_body\_insurance) است.
- ✓ عدم دسترسی به هزینه‌های واقعی عملیات: برای تحلیل دقیق سودآوری، اطلاعات هزینه واقعی مورد نیاز است.
- ✓ عوامل خارجی: تأثیر احتمالی عوامل خارجی دیگر لحاظ نشده است.