



کاوش عمیق ، تصمیم هوشمند

نمونه‌ی نمایشی – تحلیل واقعی روی داده‌ی عمومی

پنج یافته از یک کسب‌وکار واقعی

عمق تحلیل ما، این بار روی داده‌ی واقعی فروش یک کسب‌وکار.

کسب‌وکار

یک دستگاه فروش خودکار قهوه

منبع داده

دیتاست عمومی Kaggle · مجوز CC۰

دوره‌ی تحلیل

مارس تا ژوئیه‌ی ۲۰۲۴ (اسفند تا تیر) ۹۷۶۰ تراکنش

این کسب‌وکار مشتری کاوان نیست. داده از دیتاست عمومی «Coffee Sales» با مجوز CC۰ گرفته شده تا عمق تحلیل ما را نشان دهد، بدون افشای اطلاعات هیچ‌کدام از مشتریان واقعی‌مان. واحد پول همان هریونیای اوکراین است؛ برای حفظ صداقت داده، به تومان تبدیل نشده.

خلاصه‌ی مدیریتی

۹۷۶ تراکنش واقعی از یک دستگاه فروش خودکار قهوه، در بازه‌ای پنج‌ماهه بررسی شد. آنچه در ادامه می‌آید، پنج الگویی است که در نگاه اول از میان اعداد خام دیده نمی‌شوند.

سهم فروش کارتی از کل

۹۰/۴%

و در حال افزایش ماه‌به‌ماه

میانگین هر تراکنش

۳۳,۹

هریونیا

فروش کل دوره

۳۳,۰۶۷

هریونیا

۱. فروش در سه ساعت مشخص از روز فشرده می‌شود

بازه‌ی ۹ تا ۱۲ ظهر، تنها سه ساعت از پانزده ساعت کاری روز، به‌تنهایی ۲۴٪ کل فروش دوره را ساخته — و تنها ساعت ۱۰ صبح، ۱۲/۸٪ از کل درآمد را.

۲. یک‌سوم مشتریان، سه‌چهارم درآمد را می‌سازند

از ۳۷۹ مشتری شناسایی‌شده، تنها ۳۶/۷٪ بیش از یک‌بار خرید کرده‌اند، اما همین گروه ۷۳/۱٪ از کل درآمد کارتی را ساخته است.

۳. پرفروش‌ترین نوشیدنی، پرسودترین نیست

«آمریکانو با شیر» با ۲۲۷ فنجان پرفروش‌ترین قلم است، اما «لاته» با ۲۴ فنجان کمتر، بیشتر درآمد ساخته — چون هر فنجانش ۱۵٪ گران‌تر فروخته می‌شود.

هر یافته در صفحه‌ی بعد با نمودار و تفسیر همراه است.

ساعت ۱۰ صبح، اوج بی‌رقیب روز است

فروش هر ساعت از روز، جمع‌شده روی کل دوره‌ی پنج‌ماهه. هیچ ساعت دیگری نزدیک به ساعت ۱۰ نمی‌رسد.



اوج ساعت ۱۰ صبح تصادفی نیست: هفت روز هفته را جدا بررسی کردیم و در شش روز از هفت روز، ساعت ۱۰ یا ۱۱ صبح پرفروش‌ترین ساعت همان روز بود.

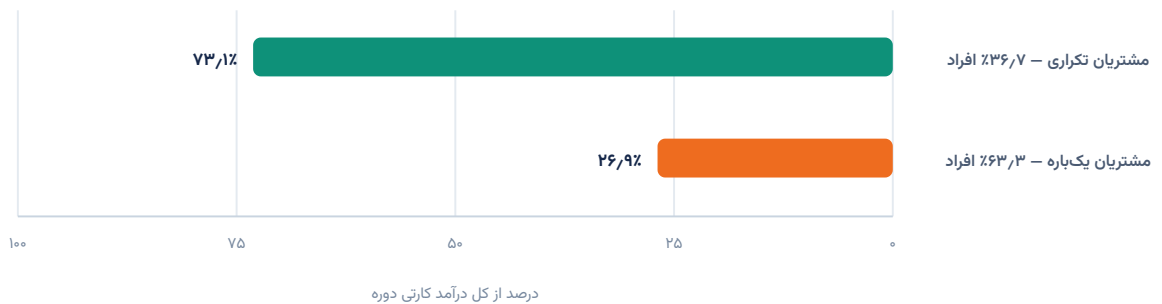


در مقابل، فروش بر حسب روز هفته تقریباً صاف است — بین ۴,۳۶۳ تا ۵,۱۰۲ هریونیا در هر روز، بدون افت معنادار آخر هفته. فشردگی فقط در ساعت است، نه در روز.

یعنی چه: برای این کسب‌وکار، برنامه‌ریزی سرویس و پرکردن دستگاه باید بر پایه‌ی «ساعت» تنظیم شود، نه «روز». رسیدن به دستگاه پیش از ساعت ۹ صبح — در هر روز هفته، بدون استثنا — بیشترین اثر را دارد.

یک هسته‌ی کوچک، اکثر درآمد را می‌سازد

از ۳۷۹ مشتری شناسایی‌شده در این دوره، اکثریت فقط یک‌بار خرید کرده‌اند – ولی درآمد داستان دیگری می‌گوید.



۳۶٫۷٪ مشتریان (۱۳۹ نفر) بیش از یک‌بار خرید کرده‌اند. همین گروه کوچک، ۷۳٫۱٪ از کل درآمد کارتی را ساخته – پرتکرارترین مشتری به‌تنهایی ۸۱ بار در پنج ماه خرید کرده، یعنی تقریباً هر روز کاری.

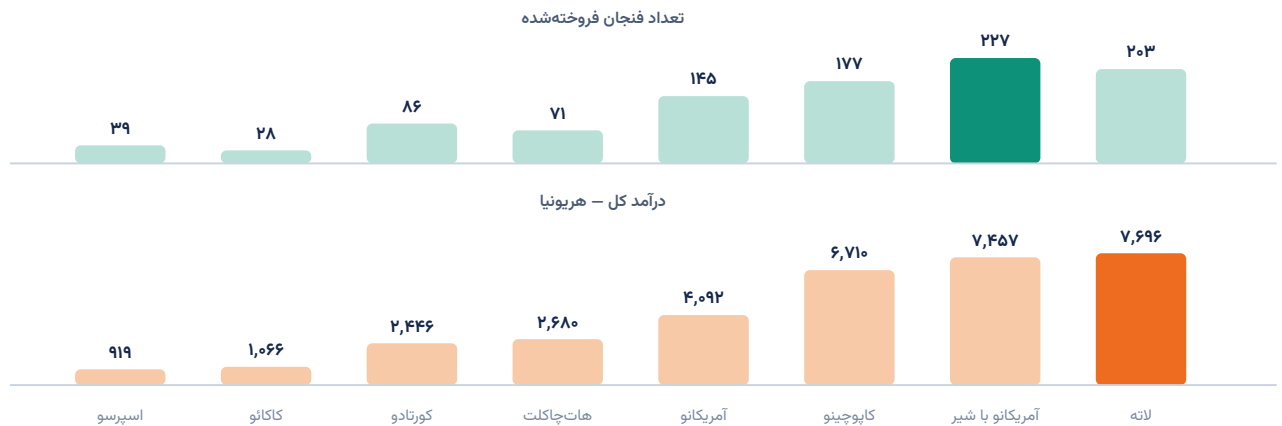


سهم پرداخت کارتی از درآمد، ماه‌به‌ماه و پیوسته بالا رفته – از ۸۳٫۸٪ در اسفند تا ۱۰۰٪ در تیر (ماه ناقص). نقد در حال حذف‌شدن طبیعی از این کسب‌وکار است.

یعنی چه: دو اقدام متفاوت برای دو مسئله: برای مشتریان یک‌باره، یک مشوق ساده برای بازگشت؛ و برای زیرساخت، اولویت با اطمینان و سرعت کارت‌خوان است، نه مدیریت نقد – چون نقد عملاً دارد از این کسب‌وکار حذف می‌شود.

ا پرفروش‌ترین قلم، پرسودترین نیست

اگر اقلام را با تعداد فنجان رتبه‌بندی کنید، ترتیب یک چیز است؛ اگر با درآمد رتبه‌بندی کنید، چیز دیگری.



«لاته» و «کاپوچینو» با قیمت هر فنجان ۳۷/۹۱ هریونیا، حدود ۱۵٪ گران‌تر از «آمریکانو با شیر» (۳۲/۸۵ هریونیا) فروخته می‌شوند. همین اختلاف قیمت کافی است تا با وجود فنجان کمتر، درآمد بیشتری بسازند.

نوشیدنی	تعداد فنجان	درآمد (هریونیا)	قیمت هر فنجان
لاته	۲۰۳	۷,۶۹۶	۳۷/۹۱
آمریکانو با شیر	۲۲۷	۷,۴۵۷	۳۲/۸۵
کاپوچینو	۱۷۷	۶,۷۱۰	۳۷/۹۱
آمریکانو	۱۴۵	۴,۰۹۲	۲۸/۲۲
کورتادو	۸۶	۲,۴۴۶	۲۸/۴۴
هات‌چاکلت	۷۱	۲,۶۸۰	۳۷/۷۵
اسپرسو	۳۹	۹۱۹	۲۳/۵۷
کاکائو	۲۸	۱,۰۶۶	۳۸/۰۸
مجموع	۹۷۶	۳۳,۰۶۷	۳۳/۹ (میانگین)

یعنی چه: این فقط تفاوت درآمد است، نه سود — بهای تمام‌شده‌ی هر فنجان در این داده نیست. برجسته‌کردن لاته در دسترس‌ی دستگاه، بدون هیچ هزینه‌ی اضافه‌ای، می‌تواند درآمد هر بازدید را بالا ببرد.

چرا پیش‌بینی این‌جا سخت است – و ما آن را پنهان نکردیم

مدل روی ۱۵ هفته آموزش دید و روی ۴ هفته‌ی پایانی (که کنار گذاشته شده بود) ارزیابی شد. باند روشن، بازه‌ی واقعی خطاست – نه یک تزئین امیدوارکننده.



فروش هفتگی، هریونیا – زمان از راست به چپ پیش می‌رود

خطای مدل روی داده‌ی کنارگذاشته‌شده ۳۸/۱٪ بود – عددی که در گزارش‌های چندشعبه‌ای مشتریان واقعی ما معمولاً به مراتب پایین‌تر است.

هفته‌ی غیرمنتظره

مه

(اردیبهشت) جهش فروش بدون توضیح
در همین داده

علت اصلی خطای بالا

۱

منبع داده – یک دستگاه تنها، بدون متغیر
بیرونی

خطای مدل روی داده‌ی کنارگذاشته‌شده

۳۸/۱٪

میانگین قدرمطلق خطای نسبی (MAPE)

چرا این عدد را نپوشاندیم: یک باند اطمینان باریک و اطمینان‌بخش، وقتی پشتش چنین خطایی باشد، دقیقاً همان «ادعای جعلی»‌ای است که کاوان از آن پرهیز می‌کند. با داده‌ی بیشتر – چند ماه بیشتر، یا چند شعبه به جای یک واحد – این خطا معمولاً به شکل محسوسی پایین می‌آید؛ این خودش یکی از دلایلی است که برای گزارش جامع، حداقل دوازده ماه داده می‌خواهیم.

چه کاری، بر پایه‌ی کدام یافته

چون این کسب‌وکار مشتری ما نیست، برآورد ریالی برای این اقدام‌ها نساختیم – فقط جهت و دلیل هرکدام را از دل همان یافته‌های واقعی نوشتیم.

۱ سرویس و پرکردن دستگاه پیش از ساعت ۹ صبح

بر پایه‌ی یافته‌ی ۱

هر روز هفته، بدون استثنا – چون ۲۴٪ فروش روز در سه ساعت اول اتفاق می‌افتد.

۲ مشوق بازگشت برای مشتریان یک‌باره

بر پایه‌ی یافته‌ی ۲

۶۳٪ مشتریان فقط یک‌بار خریده‌اند؛ حتی بازگشت بخش کوچکی از آن‌ها اثر بزرگی دارد.

۳ برجسته‌کردن لاتِه در دسترسی دستگاه

بر پایه‌ی یافته‌ی ۳

همان تلاش مشتری، درآمد ۱۵٪ بیشتر – بدون تغییر قیمت یا هزینه‌ی اضافه.

۴ اولویت با اطمینان کارت‌خوان، نه مدیریت نقد

بر پایه‌ی یافته‌ی ۲

نقد در ماه پایانی داده عملاً صفر شد؛ منبع اصلی درآمد اکنون کارت است.

روشناسی

مرحله	آنچه انجام شد
منبع و مجوز داده	دیتاست عمومی «Coffee Sales»، نویسنده Yaroslav Isaienko، مجوز CC۰ (مالکیت عمومی)
پاک‌سازی	نیازی به حذف رکورد نبود؛ ۸۹ تراکنش نقدی به طبیعت فاقد شناسه‌ی کارت‌اند
تحلیل آماری	تفکیک فروش بر حسب ساعت، روز هفته، نوع نوشیدنی، مشتری و روش پرداخت
پیش‌بینی تقاضا	میانگین متحرک چهارهفته‌ای؛ ارزیابی روی ۴ هفته‌ی پایانی کنارگذاشته‌شده، خطا ۳۸/۱٪
برآورد اثر مالی	خارج از دامنه‌ی این نمونه‌ی نمایشی – دلیلش در بالا آمد

پیوست فنی

این صفحه برای خواندن در جلسه نیست؛ برای زمانی است که بخواهید بدانید هر عدد دقیقاً از کجا آمده – یا آن را به کارشناس خودتان نشان دهید.

آنچه در این دامنه ندیدیم

پرسشی که بی‌جواب ماند	برای پاسخ به آن چه لازم است
سود واقعی هر فنجان چقدر است؟	بهای تمام‌شده‌ی مواد در این داده نیست – فقط تحلیل فروش شد، نه سود
چرا مه ۲۰۲۴ (اردیبهشت) جهش فروش داشت؟	تقویم رویداد یا تعطیلات محل نصب دستگاه لازم است
آیا این الگو مخصوص همین یک دستگاه است؟	نیازمند داده‌ی چند واحد برای مقایسه
حجم داده‌ی پردازش‌شده	۹۷۶ تراکنش واقعی، ۱۳۶ روز فروش، بازه‌ی پنج‌ماهه (۱ مارس تا ۱۷ ژوئیه‌ی ۲۰۲۴ - اسفند تا تیر)
منبع	Kaggle، دیتاست «Coffee Sales»، نویسنده Yaroslav Isaienko، مجوز CC۰
پاک‌سازی	بدون حذف رکورد؛ فقط جداسازی تراکنش‌های نقدی (بدون شناسه‌ی کارت) از کارتی
مدل و ارزیابی پیش‌بینی	میانگین متحرک چهارهفته‌ای؛ بک‌تست روی ۴ هفته‌ی پایانی از ۱۹ هفته، خطا (MAPE) ۳۸٫۱٪
تحلیل مشتری	مشتری «تکراری» یعنی بیش از یک تراکنش کارتی با یک شناسه‌ی کارت در کل دوره

چرا خطای پیش‌بینی این‌جا (۳۸٫۱٪) این‌قدر بالاست؟ چون منبع داده فقط یک واحد فروش است، بدون متغیر بیرونی. نوسان یک واحد به‌تنهایی طبیعتاً زیاد است؛ با چند ماه یا چند شعبه‌ی بیشتر برای میانگین‌گیری، این خطا معمولاً پایین‌تر می‌آید. عمداً این عدد را کوچک نشان ندادیم – به همان دلیلی که هیچ عدد ساختگی در این گزارش نمی‌گذاریم.

یادآوری: این گزارش یک نمونه‌ی نمایشی است، نه گزارش یک مشتری کاوان. داده از یک دیتاست عمومی گرفته شده تا عمق واقعی تحلیل ما را نشان دهد، بدون افشای اطلاعات هیچ مشتری واقعی. گزارش واقعی شما بر پایه‌ی داده‌ی خودتان تهیه می‌شود و پیش از دریافت هرگونه داده، توافق‌نامه‌ی محرمانگی امضا می‌شود.