



بوتکمپ تحلیل داده

پروژه: تحلیل داده‌های سرویس دوچرخه اشتراکی شرکت‌های

CitiBike و MiBici

نگارندگان: گروه شمع (Candle)

مهدی آریایی

علیرضا صالحی

عاطفه بنایی

شیوا حمیدی فرد

مهلقا صدوqi

فهرست مطالب

4	بخش اول: مقدمه و کلیات پروژه
5	1. معرفی پروژه و هدف کل
5	2. معرفی منابع داده و بازه زمانی مورد بررسی
6	3. ذینفعان و نیازهای آنها
8	بخش دوم: آماده‌سازی و خلاصه آماری داده‌ها
9	1. منبع و بارگذاری داده‌ها
9	2. پاکسازی داده‌ها
9	1.۱.۲ پاکسازی داده‌های MiBici
11	1.۲.۲ پاکسازی داده‌های CitiBike
12	1.۳.۲ خلاصه‌ی آماری داده‌ها (سوال ۱)
12	1.۳.۲ خلاصه‌ی آماری- تحلیل سال ۲۰۱۹
13	2.۱.۲ آمار جمعیتی و رفتار کاربران
18	3.۱.۲ الگوهای مدت زمان و مسافت سفر
20	4.۱.۲ روندهای زمانی (روزانه، ماهانه، فصلی و سالانه)
23	5.۱.۲ استفاده از ایستگاه و جریان ترافیک
27	6.۱.۲ سهم بازار
29	بخش سوم: پاسخ به سوالات کسب و کار
30	1.۱.۳ تشریح مسئله کسب و کار
30	2.۱.۳ متداول‌وزی و معیارهای استفاده شده
32	3.۱.۳ مدل خوشه‌بندی k_means
33	4.۱.۳ یافته‌ها و بینش‌ها
38	5.۱.۳ توصیه‌ها و اقدامات عملی
38	6.۱.۳ جمع‌بندی و نتیجه‌گیری
39	2.۱.۳ شناسایی و پیش‌بینی دوچرخه‌های مسافت بلند (سوال ۳)
39	1.۱.۳ روش‌های تشخیصی (رویکرد اولیه)
49	2.۲.۳ روش‌های پیش‌بینی‌کننده (رویکرد مکمل)
54	3.۲.۳ نتیجه‌گیری نهایی
55	4.۲.۳ پیشنهادات برای اجرا در سیستم واقعی
56	3.۱.۳ گسترش و نوسانات شرکت (سوال ۴)
56	2.۳.۳ ابزارها و فناوری‌های مورد استفاده
56	3.۳.۳ آماده‌سازی داده‌ها
57	4.۳.۳ سوال اول: گسترش یکسان سرویس در شهر
57	1.۴.۳.۳ متداول‌وزی و معیارهای استفاده شده
59	2.۴.۳.۳ یافته‌ها و بینش‌ها
74	3.۴.۳.۳ خلاصه تحلیل‌های انجام شده
74	4.۴.۳.۳ جمع‌بندی سوال اول
75	5.۱.۳.۳ عوامل موثر بر نوسانات سفرها
75	1.۵.۱.۳ متداول‌وزی و معیارهای استفاده شده
78	2.۵.۱.۳ یافته‌ها و بینش‌ها
84	3.۵.۱.۳ جمع‌بندی سوال دوم
85	6.۳.۳ نتیجه‌گیری نهایی
86	4.۱.۳ بخش‌بندی کاربران و فادر و موردی (سوال ۵)
86	1.۱.۳ تشریح مسئله
86	2.۱.۳ دسته‌بندی کاربران

88	۱.۳.۴.۳ بررسی و تحلیل توزیع ویژگی‌های کاربران وفادار در مقایسه با کاربران مور迪 و آزمون‌های آماری
90	۲.۳.۴.۳ یافته‌ها و توصیه‌هایی برای تیم بازاریابی
91	۳.۴.۳ الگوی استفاده کاربران وفادار از ایستگاه‌ها
92	۱.۳.۴.۳ تحلیل تمرکز بر اساس ضریب جین
93	۲.۳.۴.۳ یافته‌ها و بینش‌ها
94	۵. شناسایی کاربران با کاهش یا توقف استفاده (سوال ۶)
94	۱.۵.۳ تشریح مسئله
94	۲.۵.۳ معرفی داده‌ها و آماده‌سازی اولیه
94	۱.۲.۵.۳ مراحل پردازش و پاک‌سازی داده‌ها
95	۳.۵.۳ متداول‌تری و معیارهای استفاده شده
97	۴.۵.۳ تحلیل همبستگی و انتخاب ویژگی‌های رفتاری برای خوشبندی کاربران
100	۲.۴.۵.۳ ویژگی‌های انتخاب شده برای خوشبندی کاربران
101	۳.۴.۵.۳ روش خوشبندی کاربران بر اساس الگوهای رفتاری
103	۳.۱.۵.۵.۳ تعریف خوشبندی و تفسیر رفتاری
104	۲.۵.۵.۳ تحلیل تصویری دوبعدی ویژگی‌ها بین خوشبندی
104	مشاهدات کلیدی از نمودارهای دوبعدی
106	۳.۵.۵.۳ بررسی رفتار زمانی کاربران نمونه در هر خوشبندی
110	۶. افزایش سود و بهینه‌سازی ناوگان (سوال ۷)
110	۱.۶.۳ ارائه و تحلیل داده‌های استخراج شده از گزارش‌های ماهانه
114	۲.۶.۳ بهبود بازتوزیع دوچرخه‌ها بین ایستگاه‌ها با سرمایه‌گذاری بر روی وسائل نقلیه
114	۱.۲.۶.۳ پارامترها و فرضیات کلیدی
115	۲.۲.۶.۳ محاسبات انجام شده
116	۳.۲.۶.۳ تحلیل حساسیت و نمودارها
120	۴.۲.۶.۳ نتایج تحلیل روش بازتوزیع (الف)
120	۳.۶.۳ گسترش ناوگان دوچرخه‌ها و ساخت ایستگاه‌های جدید
120	۱.۳.۶.۳ پارامترها و فرضیات کلیدی
121	۲.۳.۶.۳ محاسبات انجام شده
122	۳.۳.۶.۳ تحلیل حساسیت و نمودارها
124	۴.۳.۶.۳ ایستگاه‌هایی با بیشترین عدم توازن
124	۵.۳.۶.۳ نتایج تحلیل روش گسترش ناوگان و ایستگاه‌ها (ب)
124	۴.۶.۳ استراتژی ترکیبی (بهبود بازتوزیع و گسترش ناوگان)
125	۲.۴.۶.۳ محاسبات نهایی
125	۳.۴.۶.۳ تحلیل حساسیت و نمودارها
127	۴.۴.۶.۳ نتایج تحلیل روش ترکیبی (ج)
127	۵.۶.۳ نتیجه‌گیری و توصیه‌ها
128	۶.۶.۳ جمع‌بندی نهایی

بخش اول:
مقدمه و کليات پروژه

۱.۱ معرفی پژوهه و هدف کلی

این پژوهه تحلیلی، با تمرکز بر داده‌های واقعی از دو شرکت برجسته در حوزه سرویس دوچرخه‌اشتراکی، یعنی MiBici در گوادالاخارا، مکزیک و CitiBike در نیویورک، آمریکا، طراحی شده است. در عصر حاضر که حجم عظیم داده‌ها در اختیار کسب‌وکارها قرار دارد، توانایی استخراج بینش‌های ارزشمند از این داده‌ها به یک مزیت رقابتی حیاتی تبدیل شده است. هدف اصلی این گزارش، با بهره‌گیری از تکنیک‌های پیشرفته تحلیل داده، ارائه بینش‌های عمیق و کاربردی مبتنی بر داده به مدیران و ذینفعان این شرکت‌ها است. ما قصد داریم با کاوش در الگوهای استفاده از دوچرخه، رفتار کاربران، و عملیات ایستگاه‌ها، به سوالات کلیدی کسب‌وکار در حوزه‌های بازاریابی، مدیریت عملیات، قیمت‌گذاری و استراتژی رشد پاسخ دهیم. نتایج و توصیه‌های حاصل از این تحلیل، به تصمیم‌گیری‌های آگاهانه‌تر برای بهینه‌سازی عملکرد، افزایش رضایت مشتریان و در نهایت، رشد پایدار درآمد کمک خواهد کرد.

۱.۲ معرفی منابع داده و بازه زمانی مورد بررسی

تحلیل‌های این پژوهه براساس دو مجموعه داده اصلی و حجمی از سرویس‌های دوچرخه‌اشتراکی انجام شده است:

- **داده‌های اصلی (شرکت MiBici - گوادالاخارا، مکزیک):**

این مجموعه داده شامل اطلاعات تردد روزانه کاربران شرکت MiBici است که در اوخر سال ۲۰۱۴ تأسیس شده است.

بازه زمانی مورد بررسی: از ابتدای سال ۲۰۱۵ تا پایان سال ۲۰۱۹.

حجم داده‌ها: این مجموعه شامل ۱۱,۹۶۸,۴۶۰ رکورد سفر کاربر است که حجم قابل توجهی از اطلاعات را برای تحلیل‌های جامع فراهم می‌کند.

فیلدهای اطلاعات موجود برای هر سفر:

Trip ID	•
User ID	•
Gender	•
Year of Birth	•
Trip Start Date & Time	•
Trip End Date & Time	•
Origin Station ID	•
Destination Station ID	•

- **داده‌های تکمیلی (شرکت CitiBike - نیویورک، آمریکا):**

این مجموعه داده برای پاسخگویی به برخی سوالات خاص (سوالات ۳ و ۷) استفاده شده و مربوط به شرکت CitiBike است که در اوخر می ۲۰۱۳ تأسیس و داده‌های آن از ابتدای ژوئن ۲۰۱۳ در دسترس قرار گرفته است.

باشه زمانی مورد بررسی: از ابتدای ژوئن ۲۰۱۳ تا پایان دسامبر ۲۰۱۴.

حجم داده‌ها: این مجموعه شامل ۵,۶۱۴,۸۸۸ رکورد سفر می‌باشد.

فیلدهای اطلاعاتی موجود برای هر سفر:

ride_id	•
trip_start_date	•
trip_started_at	•
trip_ended_date	•
trip_ended_at	•
trip_duration_seconds	•
bike_id	•
origin_station_id	•
start_station_name	•
start_station_latitude	•
start_station_longitude	•
destination_station_id	•
destination_station_name	•
end_station_latitude	•
end_station_longitude	•
user_type	•
birth_year	•
gender	•

۱.۳. ذینفعان و نیازهای آنها

این پژوهه تحلیلی به طور مستقیم به نیازهای اطلاعاتی و تصمیم‌گیری چندین ذینفع کلیدی در شرکت‌های سرویس دوچرخه اشتراکی پاسخ می‌دهد. درک صحیح از این نیازها، راهنمای اصلی رویکرد تحلیلی و ارائه توصیه‌ها در این گزارش است. ذینفعان اصلی و مسائل کسب‌وکار مرتبط با آنها عبارتند از:

مدیر بازاریابی:

- نیاز به شناسایی گروه‌های مشتری هدف برای تبلیغ و عرضه دوچرخه‌های جدید با قابلیت طی مسافت‌های بلند.
- لزوم تفکیک کاربران به دسته‌های "وفادر" و "موردی" به منظور طراحی و اجرای کمپین‌های بازاریابی هدفمند و استراتژی‌های جدید قیمت‌گذاری.
- اهمیت شناسایی کاربران با کاهش چشمگیر در میزان استفاده یا توقف کامل خدمات، به منظور ارائه مشوق‌های مالی مناسب و بازگرداندن آنها به چرخه مشتریان فعال.

مدیرعامل و تیم مدیریت ارشد:

- نیاز به دریافت گزارشی جامع در مورد چگونگی رشد و گسترش خدمات شرکت از زمان تأسیس، به همراه تحلیل نوسانات مشاهده شده در تعداد سفرها و عوامل تأثیرگذار بر آنها.
- تردید در مورد بهینه‌ترین مسیر برای افزایش سودآوری، به ویژه در خصوص خرید دوچرخه‌های جدید و افزایش تعداد ایستگاه‌ها؛ این تیم نیازمند تحلیل عمیق در مورد بهره‌وری استفاده از دوچرخه‌های موجود و بررسی اثربخشی گرینه‌های جایگزین است.

مدیر قیمت‌گذاری سفرها:

- نیاز به دسته‌بندی کاربران به دو گروه "وفدار" و "موردی" برای اعمال استراتژی‌های متفاوت و بهینه در قیمت‌گذاری خدمات.

بخش دوم:
آماده‌سازی و خلاصه آماری داده‌ها

۱.۲ منبع و بارگذاری داده‌ها

فرآیند آماده‌سازی داده‌ها برای این پروژه با جمع‌آوری دقیق داده‌های خام از منابع اصلی آغاز شد. داده‌های مربوط به شرکت MiBici (گوادالاخارا، مکزیک) از طریق بخش داده‌های باز (Open Data) وبسایت رسمی این شرکت دریافت شد. این فرآیند شامل دانلود فایل‌های CSV ماهانه برای تمامی سال‌های ۲۰۱۵ تا ۲۰۱۹ بود. پس از دانلود، تمامی این فایل‌های ماهانه با استفاده از کتابخانه pandas در زبان برنامه‌نویسی Python تجمعی و در یک ساختار یکپارچه سازماندهی شدند.

به منظور بهینه‌سازی عملکرد و کارایی در فرآیندهای آنی تحلیل، داده‌های تجمعی شده از فرمت CSV به فرمت Parquet تبدیل شدند. استفاده از فرمت Parquet مزایای قابل توجهی دارد؛ از جمله فشرده‌سازی بالاتر که منجر به کاهش حجم فایل‌ها و در نتیجه، سرعت بارگذاری و پردازش سریع‌تر می‌شود، و همچنین قابلیت ذخیره‌سازی ستونی (Columnar) که برای کوئری‌های تحلیلی بسیار کارآمد است. این فایل‌های Parquet سپس به پلتفرم Kaggle اضافه شدند تا به عنوان منبع اصلی داده برای اجرای پروژه در محیط‌های اشتراکی مورد استفاده قرار گیرند.

مشابه با داده‌های MiBici، مجموعه داده‌های تکمیلی مربوط به شرکت CitiBike (نیویورک، آمریکا) نیز از طریق بخش داده‌های سیستمی (System Data) وبسایت این شرکت جمع‌آوری شد. داده‌های اصلی مورد نیاز از این منبع دانلود و همانند داده‌های MiBici، پس از تجمعی، به فرمت Parquet تبدیل و برای دسترسی آسان تیم در Kaggle بارگذاری شدند.

جزئیات کامل لینک‌های وبسایت‌های MiBici و CitiBike در بخش "منابع و مراجع" در پایان این گزارش ارائه شده است.

این رویکرد جامع در گردآوری و آماده‌سازی اولیه داده‌ها، زیربنای مستحکمی برای مراحل بعدی پاکسازی و تحلیل‌های عمیق‌تر در این پروژه فراهم آورد.

۲.۲ پاکسازی داده‌ها

فرآیند پاکسازی داده‌ها، گامی حیاتی پس از بارگذاری اولیه بود تا از کیفیت و یکپارچگی داده‌ها برای تحلیل‌های آنی اطمینان حاصل شود. این مراحل برای هر دو مجموعه داده CitiBike و MiBici، با توجه به ساختار و ویژگی‌های خاص هر کدام، به صورت مجزا اما با رویکردهای مشابه انجام گرفت:

۱.۲.۲ پاکسازی داده‌های MiBici

مجموعه داده اصلی MiBici شامل ۱۳ ستون و ۱۱,۹۶۸,۴۶۰ ردیف بود. مراحل پاکسازی و آماده‌سازی این داده‌ها به شرح زیر است:

تغییر نام ستون‌ها:

نام ستون‌های اصلی داده‌های MiBici برای سهولت درک و تحلیل، به انگلیسی ترجمه و استانداردسازی شدند:

- Viaje_Id → trip_id •
- Usuario_Id → user_id •

- Genero → gender •
- AÑo_de_nacimiento → birth_year •
- Origen_Id → origin_station_id •
- Destino_Id → destination_station_id •
- Inicio_del_viaje → trip_started_at •
- Fin_del_viaje → trip_ended_at •

تبدیل نوع داده‌ها:

- ستون‌های trip_ended_at و trip_started_at که در ابتدا از نوع object (رشته) بودند، به نوع datetime تبدیل شدند تا امکان محاسبات زمانی (مانند محاسبه مدت زمان سفر) فراهم شود. •
- ستون‌های path_id و gender، year_month به نوع string تبدیل شدند. •
- ستون birth_year به عدد صحیح (int64) تبدیل شد. •

ساخت ویژگی‌های جدید:

- محاسبه مدت‌زمان سفر (trip_duration_seconds): مدت‌زمان سفر بر حسب ثانیه از اختلاف بین زمان شروع و پایان محاسبه شده است. •
- ساخت شناسه مسیر (Path ID): ترکیب شناسه مبدأ و مقصد برای ساخت یک مسیر یکتا با قالب <O<origin_id>D<destination_id> •
- محاسبه مقدار فاصله مستقیم بین ایستگاه‌ها (direct_distance_km): با استفاده از id، path_id، از دیتابریم paths_df اضافه شده است. •
- محاسبه سرعت تقریبی سفر (speed_kmh): سرعت تقریبی بر حسب کیلومتر بر ساعت با تقسیم فاصله مستقیم بر مدت‌زمان سفر. •
- ستون های day، month و year: با استفاده از ستون trip_started_at بر اساس شماره آن روز در آن ماه محاسبه شده است. •

: مدیریت مقادیر از دست رفته (Missing Values)

- ابتدا درصد مقادیر Null برای هر ستون بررسی شد. •
- فقط در ستون‌های gender و birth_year مقادیر گمشده وجود دارد. این مقادیر را بعدا در سوالی که از این دو ستون استفاده می‌شود، با استفاده از تابع dropna() حذف می‌کنیم. این رویکرد برای حفظ یکپارچگی داده‌ها در تحلیل‌های بعدی اتخاذ گردید. •

: رسیدگی به رکوردهای تکراری (Duplicate Records)

- ستون trip_id برای شناسایی داده‌های تکراری بررسی شد و هیچ رکورد تکراری مشاهده نشد. •

: مدیریت داده‌های پرت (Outliers) و بررسی‌های منطقی سفرها

- محدودیت زمان شروع و پایان: با بررسی وبسایت رسمی شرکت MiBici، مشخص شد که سرویس دوچرخه اشتراکی این شرکت از ساعت ۱ صبح تا ۵ صبح تا ۱ بامداد فعال است. با در نظر گرفتن این محدودیت عملیاتی، سفرهایی با زمان شروع و پایان بعد از ۱ بامداد و قبل از ۵ صبح نامعتبر در نظر گرفته شده و حذف شدند.

در مجموع ۴۹۱۳ یعنی ۰.۵٪ درصد سفر نامعتبر در داده‌ها وجود داشت که تصمیم بر حذف آن‌ها گرفته شد.

- ترتیب زمان سفر: بررسی شد که آیا زمان شروع سفر بعد از زمان پایان آن ثبت شده است، و همچنین سفرهایی با زمان شروع برابر با زمان پایان نیز مورد بررسی قرار گرفتند. خوشبختانه این مشکل در دیتابست وجود نداشت و رکوردي حذف نشد.

بررسی ایستگاه‌ها:

تمامی گام‌های انجام شده برای سفرهای، یعنی بررسی نوع داده‌ها، بررسی موارد خالی، صحت داده‌ها از لحاظ منطقی، وجود داده‌های تکراری و ... برای ایستگاه‌ها نیز بررسی شد و خوشبختانه داده‌های ایستگاه‌ها تمیز بود و نیاز به حذف هیچ موردی نبود.

حجم نهایی مجموعه داده پس از پاکسازی: ۱۱,۹۶۳,۵۴۷ ردیف (۹۹.۹۶٪ از داده‌های اولیه).

۲.۲.۲ پاکسازی داده‌های CitiBike

مجموعه داده اصلی CitiBike شامل ۱۵ ستون و ۵,۶۱۴,۸۸۸ ردیف بود. مراحل پاکسازی و آماده‌سازی این داده‌ها به شرح زیر است:

بررسی و مدیریت مقادیر از دست رفته (Missing Values):

- ستون birth_year (سال تولد کاربر) دارای ۳۲۱ مقدار Null بود. با توجه به اینکه این ستون برای تحلیلهای اصلی در سوالات ۳ و ۷ مورد نیاز نیست، هیچ عملیات خاصی برای پر کردن یا حذف این مقادیر انجام نشد.
- ستون‌های destination_station_id, destination_station_name, destination_station_latitude, destination_station_longitude مرتبط با مقصد (شامل ۱۸,۰۵۹ ردیف اصلی مشاهده شد که پس از اعمال سایر مراحل پاکسازی، تعداد ردیف‌های دارای Null در destination_station_id به ۱۶,۴۵۲ ردیف (۹۰.۳٪ از کل داده‌ها) کاهش یافت. با توجه به نقش این ستون‌ها در تحلیلهای مرتبط با ایستگاه‌های پایانی، این مقادیر از دست رفته به منظور بررسی سوال ۷ از داده‌ها حذف شدند.

رسیدگی به رکوردهای تکراری (Duplicate Records):

بررسی‌ها نشان داد که هیچ ردیف کاملاً تکراری در مجموعه داده CitiBike وجود نداشت.

بررسی و مدیریت ناسازگاری در مدت زمان سفر (Trip Duration Inconsistency):

مشاهده شد که ۱۱۸ ردیف دارای ناسازگاری بین مقدار ثبت شده در ستون trip_duration_seconds و مقدار محاسبه شده از اختلاف trip_ended_at و trip_started_at بودند. برای اطمینان از دقت تحلیل‌ها، تصمیم گرفته شد که در محاسبات و تحلیل‌های بعدی، از مقدار محاسبه شده از اختلاف زمان شروع و پایان سفر استفاده شود.

مدیریت داده‌های پرت و سفرهای نامعتبر:

در سرویس CitiBike، دو نوع کاربر اصلی وجود دارد: Subscriber (کاربران با اشتراک) و Customer (کاربران موردي/بدون اشتراک). با بررسی وبسایت رسمی شرکت و قوانین استفاده از خدمات، مشخص شد که حداقل مدت زمان مجاز برای هر سفر برای کاربران ۳۰ دقیقه و برای کاربران ۴۵ دقیقه تعیین شده است.

با در نظر گرفتن این محدودیت‌ها، سفرهایی با مدت زمان فراتر از این آستانه‌ها به عنوان داده پرت یا نامعتبر شناسایی شدند.

علاوه بر این، سفرهایی که تاریخ شروع و پایان آنها در یک روز نبود (نشان‌دهنده مدت زمان‌های بسیار طولانی و غیرعادی تا ۷۰ روز) و یا سفرهایی که مدت زمان آنها بر حسب ثانیه (در ستون trip_duration_seconds) به طور غیرمنطقی بسیار طولانی‌تر از محدوده‌های مجاز بود (حتی در مواردی که تاریخ شروع و پایان در یک روز بود)، شناسایی شدند.

آمار سفرهای نامعتبر:

- سفرهایی با مدت زمان بیش از ۳۰ دقیقه برای کاربران customer: ۱۶۷,۷۲۳ ردیف (%)۲.۹۹
- سفرهایی با مدت زمان بیش از ۴۵ دقیقه برای کاربران subscriber: ۴۵,۵۶۶ ردیف (%)۰.۸۱
- سفرهایی که تاریخ شروع و پایان آنها متفاوت بود: ۲۵,۰۴۸ ردیف (%)۰.۴۵

تصمیم‌گیری برای حذف: مجموعاً ۲۳۱,۴۲۷ ردیف (۴.۱٪ از کل داده‌ها) که شامل موارد فوق بودند، به عنوان داده‌های نامعتبر یا پرت تشخیص داده شده و از مجموعه داده حذف گردیدند. این تصمیم برای تضمین دقت و اعتبار تحلیل‌ها، بهویژه در سوال ۳ و ۷، اتخاذ شد.

حجم نهایی مجموعه داده پس از پاکسازی: ۱,۴۶۱,۳۸۳ ردیف (۹۵.۸٪ از داده‌های اولیه).

۳.۲ خلاصه‌ی آماری داده‌ها (سوال ۱)

۱.۳.۲ خلاصه‌ی آماری- تحلیل سال ۲۰۱۹

برای پیدا کردن دید مناسب و کاملی نسبت به وضعیت شرکت MiBici، در جدول ۱.۲ یک نمای کلی و سریع از مهمترین ابعاد عملیاتی و رفتاری این سرویس در سال ۲۰۱۹ ارائه می‌شود.

جدول (۱-۲) معیارهای کلیدی عملیاتی شرکت MiBici در سال ۲۰۱۹

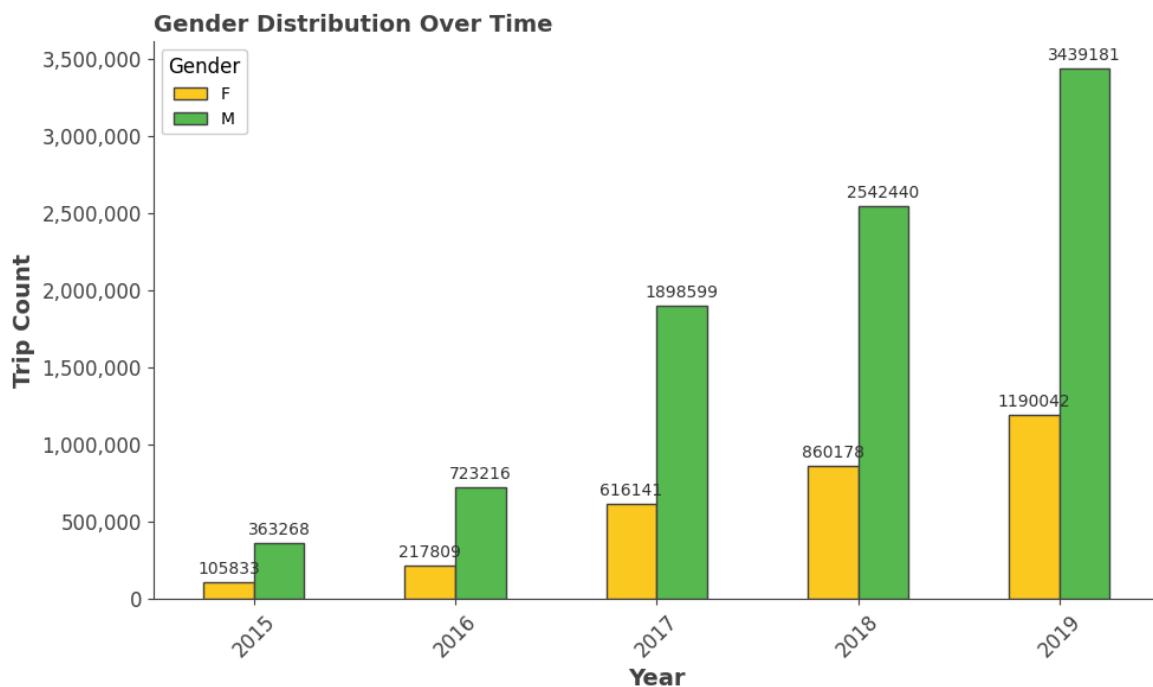
متريک	تعريف	مقدار عددی
تعداد کل سفرها	مجموع کل سفرهای دوچرخه‌سواری ثبت شده در سال ۲۰۱۹.	۴۴۹۷۶۱۵
تعداد کاربران يكتا	تعداد شناسه‌های منحصر به فرد کاربران که در سال ۲۰۱۹ حداقل یک سفر داشته‌اند.	۲۸۳۶
میانگین تعداد سفر به ازای هر کاربر	متوسط تعداد سفرهای انجام شده توسط هر کاربر در سال ۲۰۱۹.	۱۵۸.۸۳
مجموع زمان سفرها (دقیقه)	مجموع کل مدت زمان استفاده از دوچرخه‌ها در سال ۲۰۱۹ بر حسب دقیقه	۸۷۰ هزار ساعت معادل بیش از ۳۶ هزار روز
میانگین مدت زمان سفر (دقیقه)	میانگین مدت زمان هر سفر در سال ۲۰۱۹ بر حسب دقیقه	۱۱.۶۱ دقیقه
میانه مدت زمان سفر (دقیقه)	میانه مدت زمان سفرها در سال ۲۰۱۹ بر حسب دقیقه	۹.۷۷ دقیقه
میانگین سفر در هر روز	متوسط تعداد سفرهای انجام شده در هر روز	۱۲۳۲۱.۶۸
مجموع مسافت طی شده (کیلومتر)	مجموع کل مسافت‌های طی شده توسط تمامی دوچرخه‌ها بر حسب کیلومتر	۶۷۷۰۶۲۷.۸۵ کیلومتر
میانگین مسافت هر سفر (کیلومتر)	میانگین مسافت طی شده در هر سفر بر حسب کیلومتر.	۱.۵۱ کیلومتر
میانگین سرعت سفرها (km/h)	میانگین سرعت کلی سفرهای انجام شده بر حسب کیلومتر بر ساعت.	۸.۵۸ کیلومتر بر ساعت
تعداد ایستگاه‌های فعال	تعداد کل ایستگاه‌های منحصر به فردی که به عنوان مبدأ یا مقصد استفاده شده‌اند.	۲۷۴ ایستگاه
محبوب‌ترین ماه سال	ماهی که بیشترین تعداد سفر در آن ثبت شده است.	اکتبر
محبوب‌ترین روز هفته	روزی از هفته که بیشترین متوسط تعداد سفر را داشته است.	سه‌شنبه
شلوغ‌ترین روز سال	تاریخ خاصی که بیشترین تعداد سفر روزانه در آن ثبت شده است.	۲۲ ژانویه
بیشترین تعداد سفر یک کاربر	حداکثر تعداد سفرهایی که توسط یک کاربر منحصر به فرد انجام شده است.	۱۶ سفر
سهم ۱۰٪ کاربران فعال برتر	درصد کل سفرها که توسط ۱۰ درصد از فعال‌ترین کاربران (از نظر تعداد سفر) انجام شده‌اند.	۴۰.۳۵٪

۲.۳.۲ آمار جمعیتی و رفتار کاربران

توزيع جنسیتی کاربران:

نتایج تحلیل نشان می‌دهد که استفاده از این سیستم به‌طور قابل توجهی میان کاربران مرد غالب بوده است، بطوریکه ۷۰.۳٪ کاربران مرد و ۲۷٪ کاربران زن هستند و جنسیت ۱۲.۷٪ از کاربران نامشخص است. همچنین ۷۴.۹٪ از کل سفرها توسط مردان و ۲۵٪ سفرها توسط کاربران زن انجام شده است. مطابق شکل (۱-۲) اختلاف چشمگیری در تعداد سفرها میان مردان و زنان در تمام سال‌ها دیده می‌شود. در سال ۲۰۱۹، مردان بیش از ۳.۴ میلیون سفر انجام داده‌اند،

در حالی که زنان حدود ۱.۱۹ میلیون سفر ثبت کرده‌اند. با وجود رشد هم‌زمان تعداد سفرها برای هر دو جنسیت، نسبت زنان به مردان ثابت یا با رشد کنترلی همراه بوده است.



شکل (۱-۲) توزیع تعداد سفر کاربران مرد و زن با گذرهای زمان

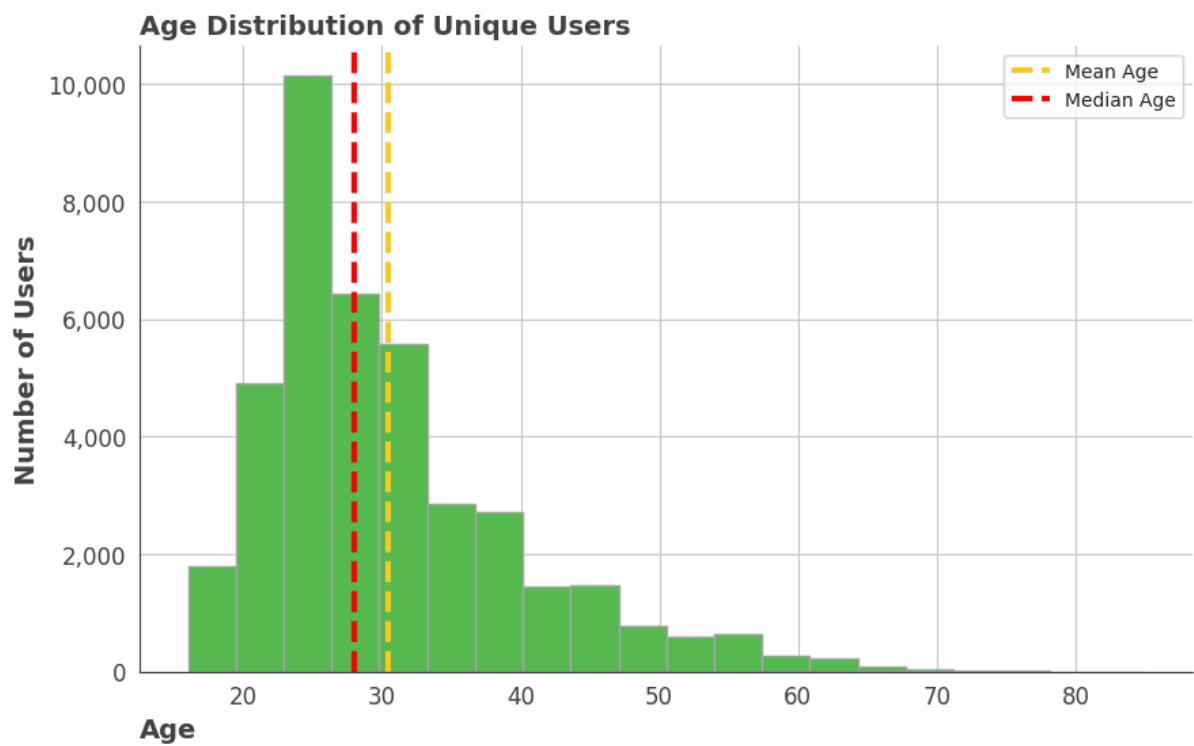
توزیع سن کاربران:

طبق شکل (۲-۲) بیشترین تعداد کاربران در بازه سنی ۲۵ تا ۲۹ سال قرار دارد، به طوری که بیش از ۱۰,۰۰۰ کاربر را شامل می‌شود. میانگین سنی کاربران حدود ۳۰ سال است که با خطچین زرد در نمودار مشخص شده و نشان‌دهنده تمکز بالای کاربران در رده‌های سنی جوان است.

از سن ۳۵ سال به بعد، تعداد کاربران به تدریج کاهش می‌یابد و این روند به‌ویژه پس از ۵۰ سالگی مشهودتر است. با این حال، همچنان کاربران بالای ۷۰ سال نیز در سیستم حضور دارند که نشان‌دهنده شمول سنی قابل قبول است.

الگوی توزیع به صورت چوله به راست (Right-Skewed) است، به این معنا که بخش عمده کاربران در رده‌های سنی پایین‌تر متتمرکز شده‌اند.

بطور کلی این تحلیل نشان می‌دهد که مخاطبان اصلی سامانه MiBici را جوانان، به‌ویژه افراد در دهه سوم زندگی، تشکیل می‌دهند. لذا پیشنهاد می‌شود جهت افزایش مشارکت سنی، برنامه‌های تشويقی و آگاهی‌رسانی برای گروه‌های سنی میانسال و سالمند طراحی و اجرا شود.



شکل (۲-۲) توزیع سن کاربران

نمودار شکل (۲-۳) نسبت (درصد) کاربران یکتا مربوط به هر گروه سنی را از کل کاربران یکتا در هر سال نشان می‌دهد. این نمودار با تفکیک روندهای هر گروه سنی، مشاهدهٔ واضح‌تر و دقیق‌تر تغییرات سهم هر گروه سنی در طول زمان را ممکن می‌سازد که به شرح زیر است:

گروه سنی 34-25

- این گروه در تمام سال‌ها بیشترین درصد از کاربران یکتا را به خود اختصاص داده است.
- سهم این گروه در طول زمان تقریباً ثابت باقی مانده و فقط با نوسانات جزئی همراه بوده است.
- این نشان می‌دهد که این گروه سنی، پایه‌ی اصلی کاربران را تشکیل می‌دهد.

گروه سنی 24-18

- این گروه در رتبه دوم قرار دارد و روندی افزایشی از سال 2015 تا 2018 داشته، سپس در سال 2019 کمی کاهش داشته است.
- این نشان می‌دهد که جذب کاربران جوان‌تر تا حدی موفق بوده، اما در سال پایانی رشد متوقف شده است.

گروه سنی 44-35

- این گروه سنی در کل سهم نسبتاً ثابتی در طول زمان داشته است.
- بعد از کاهش اولیه از 2015 تا 2018، در سال 2019 کمی افزایش داشته است.
- به طور کلی می‌توان گفت که تغییر قابل توجهی در رفتار این گروه سنی دیده نمی‌شود.

گروه سنی 54-45

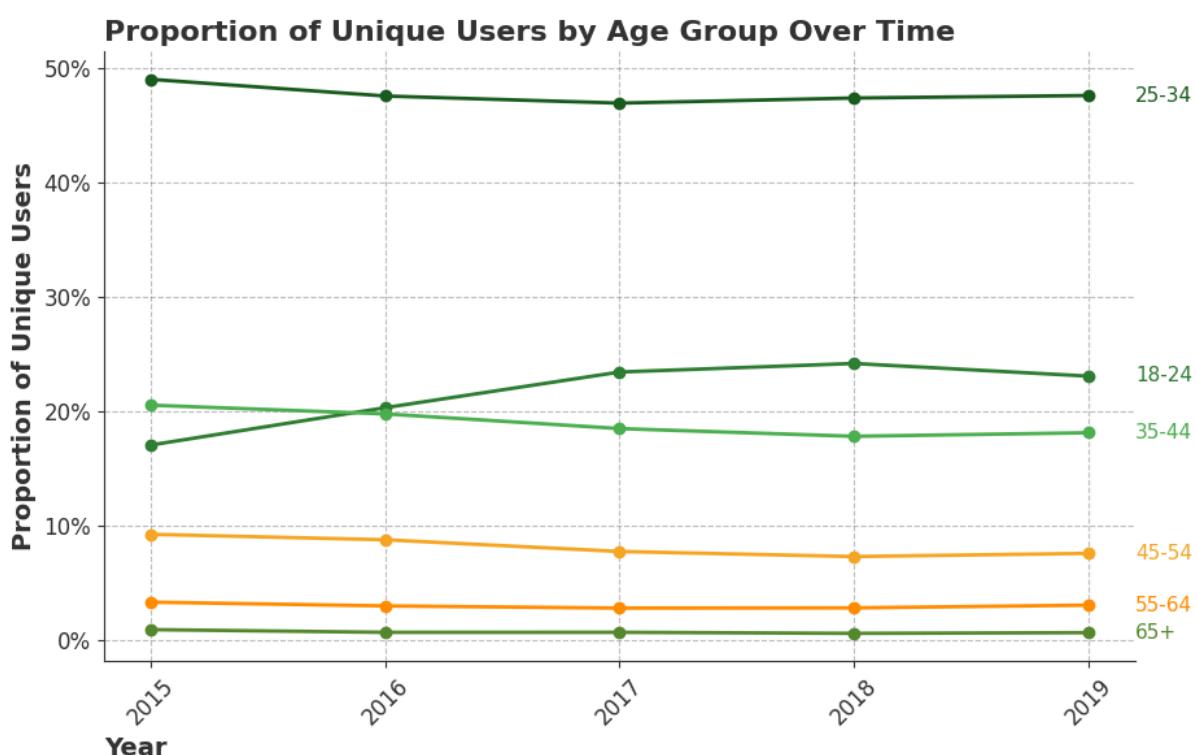
- این گروه در رتبه‌های پایین نمودار قرار دارد و روندی نزولی خفیف را در سهم کاربران یکتا طی کرده است.
- تغییرات این گروه اندک و یکنواخت بوده‌اند، که نشان می‌دهد میزان استفاده‌ی این گروه در حال کاهش یا ثبات است.

گروه سنی ۶۴-۵۵ و +۶۵

- هر دو گروه پایین‌ترین درصد از کاربران یکتا را داشته‌اند.
- سهم آن‌ها در طول زمان تقریباً ثابت و بسیار پایین باقی مانده است.
- به ویژه گروه ۶۵+، که هیچ تغییر محسوسی در مشارکت ندارد و به نظر می‌رسد کمترین مشارکت را دارد.

نتیجه‌گیری کلی:

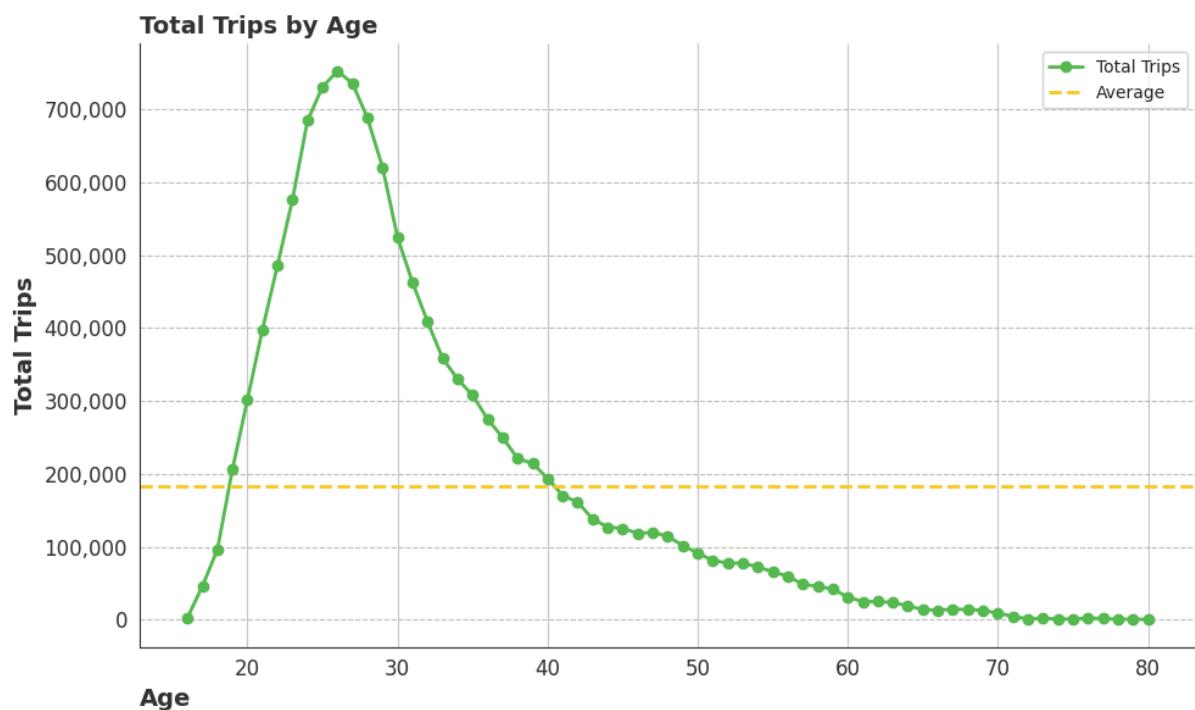
- گروه‌های سنی جوان‌تر (۱۸-۲۴ و ۲۵-۳۴) دارای بالاترین درصد از تعداد کاربران یونیک هستند و تمایل به رشد در این گروه‌ها وجود دارد.
- گروه‌های سنی بالاتر به ویژه ۵۴-۴۵، ۶۴-۵۵ و ۶۵+ روند ثابتی دارند و به نظر نمی‌رسد که این گروه‌ها به سرعت در حال جذب کاربران جدید باشند.



شکل (۳-۲) نسبت کاربران یکتا بر اساس گروه سنی در گذر زمان

نمودار شکل (۴-۲) تعداد کل سفرها براساس سن کاربران را نشان می‌دهد. بیشترین تعداد سفرها مربوط به سنین ۲۴ تا ۲۷ سالگی است که قله‌ی نمودار را تشکیل می‌دهد. پس از این بازه، نمودار روندی نزولی دارد و با افزایش سن کاربران، تعداد سفرها به طور پیوسته کاهش می‌یابد. کاربران بین ۱۸ تا ۳۵ سال بیشترین سهم از سفرها را دارند. میانگین تعداد

سفرها (خط زرد رنگ) نشان می‌دهد که بیشتر گروه‌های سنی زیر میانگین سنی ۴۵ سال بالاتر از حد متوسط سامانه عمل کرده‌اند. از سن ۵۰ به بالا، تعداد سفرها کاهش چشمگیرتری دارد که می‌تواند ناشی از محدودیت‌های جسمی، دسترسی، یا تمایل کمتر به استفاده از دوچرخه باشد.



شکل (۴-۲) تعداد کل سفرها براساس سن کاربران

روند رشد سالانه کاربران یکتا

نمودار شکل (۵-۲)، رشد سالانه کاربران یکتا شرکت MiBici رادر فاصله سال‌های ۲۰۱۵ تا ۲۰۱۹ را تحلیل می‌کند. نمودار بالایی (خطی) تعداد کل کاربران یکتا را در هر سال نشان می‌دهد، که بیانگر گسترش مطلق پایگاه کاربران است. مطابق این نمودار تعداد کاربران یکتا از ۵۱۸۶ نفر در سال ۲۰۱۵ به ۳۰,۷۳۶ نفر در سال ۲۰۱۹ رسیده است (حدود ۶ برابر افزایش در طول ۵ سال).

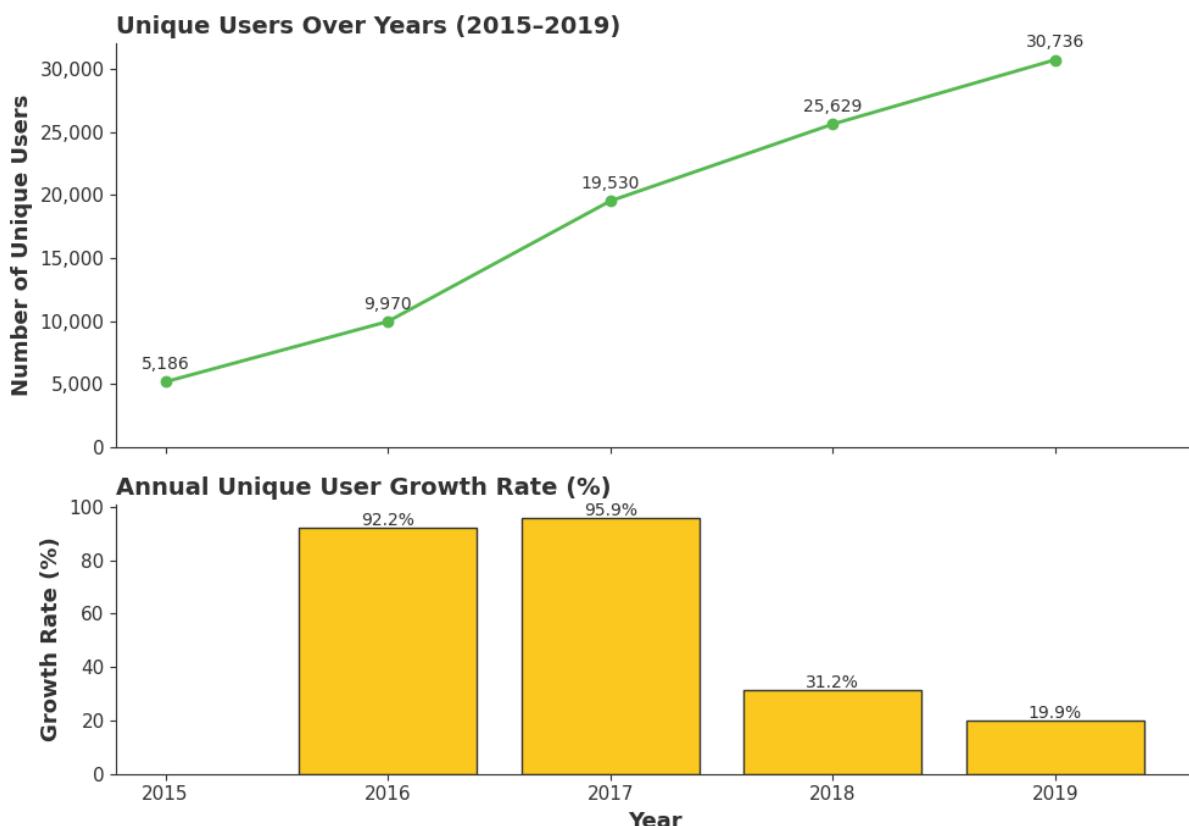
نمودار پایینی (میله‌ای) نرخ رشد درصدی این کاربران را به صورت سال به سال نمایش می‌دهد و به درک شتاب یا کندی روند جذب کاربر کمک می‌کند. بیشترین نرخ رشد مربوط به سال‌های ابتدایی است، بطوریکه:

- ۲۰۱۶: رشد ۹۲.۲٪
- ۲۰۱۷: رشد ۹۵.۹٪

اما از سال ۲۰۱۸ به بعد، روند رشد بهوضوح کاهش یافته:

- ۲۰۱۸: رشد ۳۱.۲٪
- ۲۰۱۹: رشد تنها ۱۹.۹٪

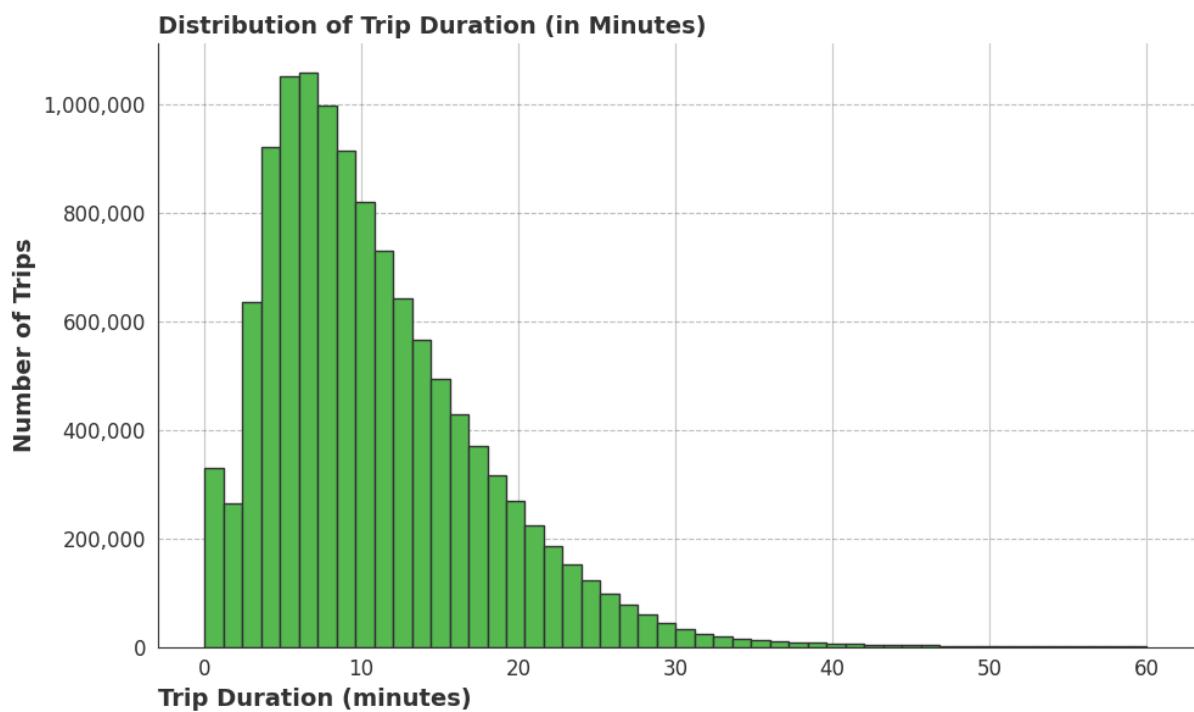
رشد کاربران در سال‌های ابتدایی بسیار چشمگیر بوده، که نشان‌دهنده استقبال اولیه قوی از سامانه MiBici است. اما در سال‌های بعد، هرچند تعداد کاربران همچنان افزایش یافته، اما نرخ رشد کاهش یافته است. این روند می‌تواند نشان‌دهنده رسیدن به مرحله بلوغ بازار یا نیاز به افزایش تلاش‌های بازاریابی و جذب کاربر جدید باشد.



شكل (۵-۲) روند رشد کاربران یکتا در گذر زمان

۳.۳.۲ الگوهای مدت زمان و مسافت سفر

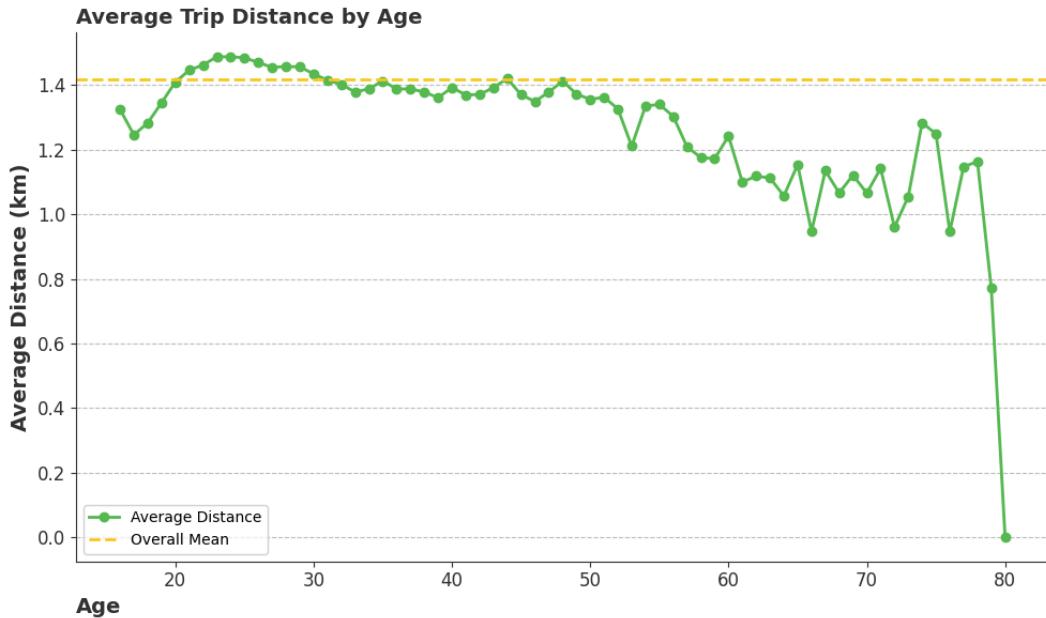
نمودار (۶-۲) توزیع مدت زمان سفرها بر حسب دقیقه برای کلیه کاربران MiBici را نشان می‌دهد. این توزیع فقط برای سفرهای کمتر از 60 دقیقه است. بخش عمده‌ای از سفرها در بازه‌ی کمتر از 30 دقیقه انجام شده‌اند. قله‌ی نمودار به وضوح در ابتدای محور X قرار دارد، که بیانگر تراکم بالای سفرهای کوتاه‌مدت است. با افزایش مدت زمان سفر، تعداد سفرها به صورت سریع کاهش می‌یابد. کاربران MiBici عمده‌ای از دوچرخه برای جابجایی‌های کوتاه و روزمره استفاده می‌کنند.



شکل (۶-۲) توزیع مدت زمان سفرها (کمتر از 60 دقیقه)

نمودار (۷-۲) میانگین فاصله سفر مستقیم (بر حسب کیلومتر) را بر اساس سن کاربران سامانه اشتراک دوچرخه MiBici نشان می‌دهد. لازم به ذکر است که برای محاسبه مسافت بین ایستگاهها از فرمول محاسبه فاصله مستقیم بین دو ایستگاه استفاده شده است.

- **کاربران جوان‌تر (تا حدود 25 سال)** میانگین مسافت بالاتری نسبت به دیگر گروه‌ها دارند، به‌ویژه در سنین حدود 23–25 که میانگین فاصله به بیشترین مقدار می‌رسد (بیش از 1.45 کیلومتر).
- از سن 30 سال به بالا، یک روند نزولی در میانگین فاصله سفرها مشاهده می‌شود. این کاهش با نوساناتی تا سن 70 ادامه دارد و بیانگر کاهش متوسط استفاده بلندمدت از دوچرخه با افزایش سن است.
- کاهش شدید پس از 78 سالگی به چشم می‌خورد؛ به‌ویژه در سن 80 سال که میانگین تقریباً به صفر می‌رسد.
- خط چین زرد، میانگین کلی تمام سفرها را نشان می‌دهد (حدود 1.41 کیلومتر). بخش بزرگی از کاربران تا حدود 50 سال، بالاتر یا نزدیک به این میانگین حرکت می‌کنند، در حالی که کاربران مسن‌تر، معمولاً کمتر از میانگین کل سفر می‌کنند.



شکل (۷-۲) میانگین مسافت طی شده با دوچرخه در گروه های سنی مختلف

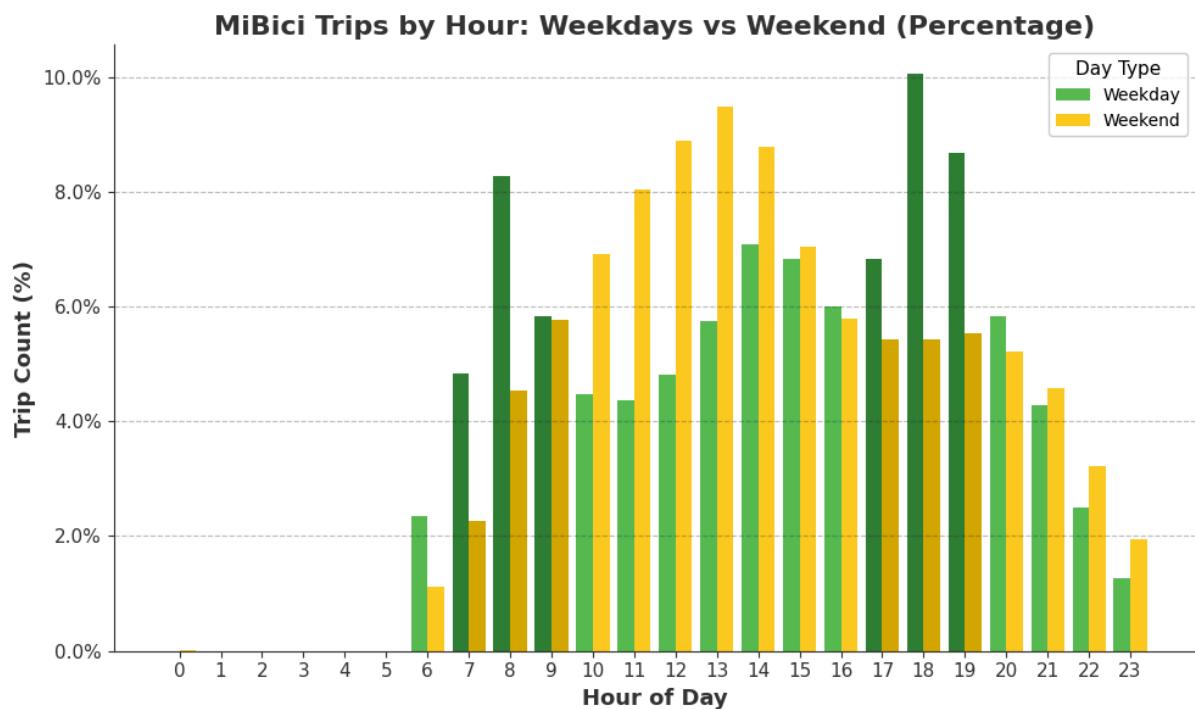
۴.۳.۲ روندهای زمانی (روزانه، ماهانه، فصلی و سالانه)

نمودار شکل (۸-۲) توزیع زمانی سفرهای MiBici را در طول شباهه روز به تفکیک روزهای کاری و تعطیل نشان می دهد. آنچه بهوضوح دیده می شود، تفاوت الگوی استفاده کاربران در این دو دسته از روزهاست.

در روزهای کاری، دو نقطه ای اوج مشخص دیده می شود که بهوضوح به ساعات شلوغ صبحگاهی (حدود ساعت ۷ تا ۹) و عصرگاهی (حدود ساعت ۱۷ تا ۱۹) مربوط هستند. این الگو بسیار شبیه به رفتار رفت و آمد شغلی یا تحصیلی است و نشان می دهد که MiBici نقش قابل توجهی در جابجایی روزمره کاربران دارد. این ساعات پیک، با رنگ سبز تیره تر در نمودار مشخص شده اند و شدت استفاده در آنها از سایر ساعات به مراتب بیشتر است.

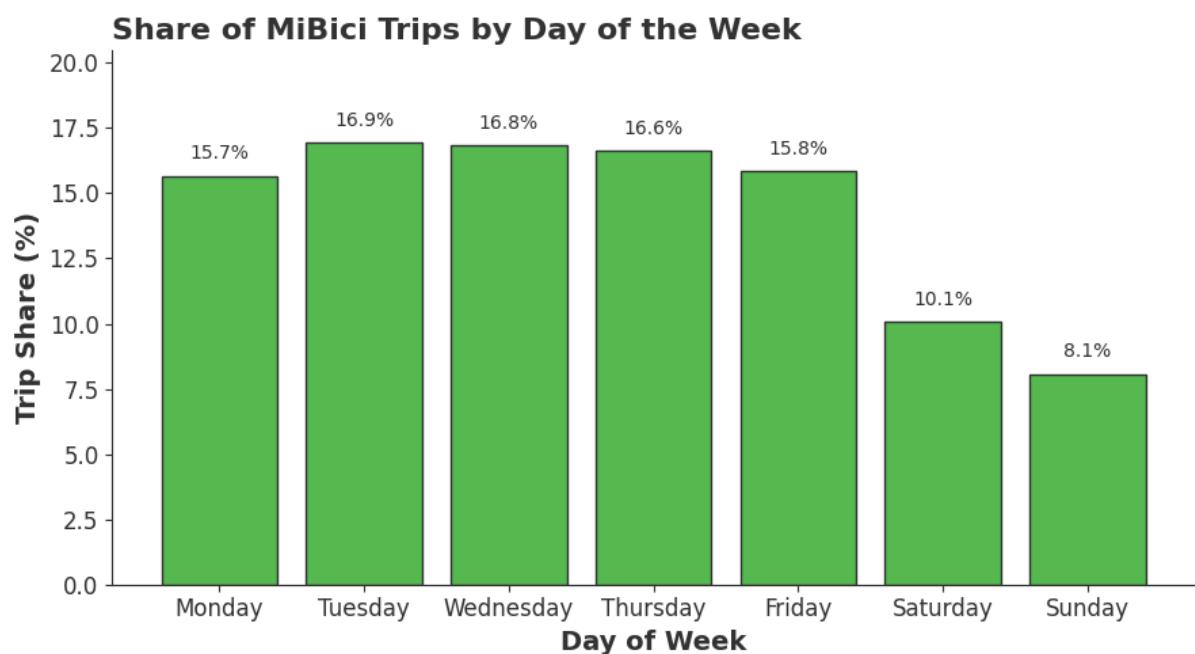
در مقابل، در روزهای تعطیل این تمرکز زمانی وجود ندارد. الگوی سفرها در روزهای تعطیل یکنواخت تر است و به جای دو قله مشخص، از حدود ساعت ۱۰ صبح به بعد تا حدود ساعت ۱۸، حجم سفرها تقریباً پایدار و بدون نوسان شدید باقی می ماند. چنین الگویی بیانگر استفاده تفریحی یا گردشگری از سامانه در روزهای تعطیل است، زمانی که الزام زمانی خاصی برای شروع یا پایان سفر وجود ندارد.

در مجموع، نمودار نشان می دهد MiBici نه تنها ابزار حمل و نقل شهری در روزهای کاری است بلکه در روزهای تعطیل نیز به عنوان وسیله ای تفریحی یا جایگزین وسایل نقلیه موتوری استفاده می شود. این تنوع کاربری، یکی از نقاط قوت سیستم محسوب می شود.



شکل (۸-۲) الگوی ساعتی سفرهای MiBici در روزهای هفته و تعطیل

نمودار شکل (۹-۲) سهم درصدی سفرهای MiBici در طول روزهای هفته را نمایش می‌دهد و نشان می‌دهد که تقاضای استفاده از دوچرخه در ایام کاری (Friday تا Monday) نسبت به تعطیلات آخر هفته (Sunday و Saturday) بسیار بیشتر است. MiBici عمدهً یک سرویس روزمره و کاربردی برای ترددهای کاری و تحصیلی است و پیشنهاد می‌شود در روزهای وسط هفته، ظرفیت ایستگاهها و تعمیرات دوچرخه‌ها تقویت شود.



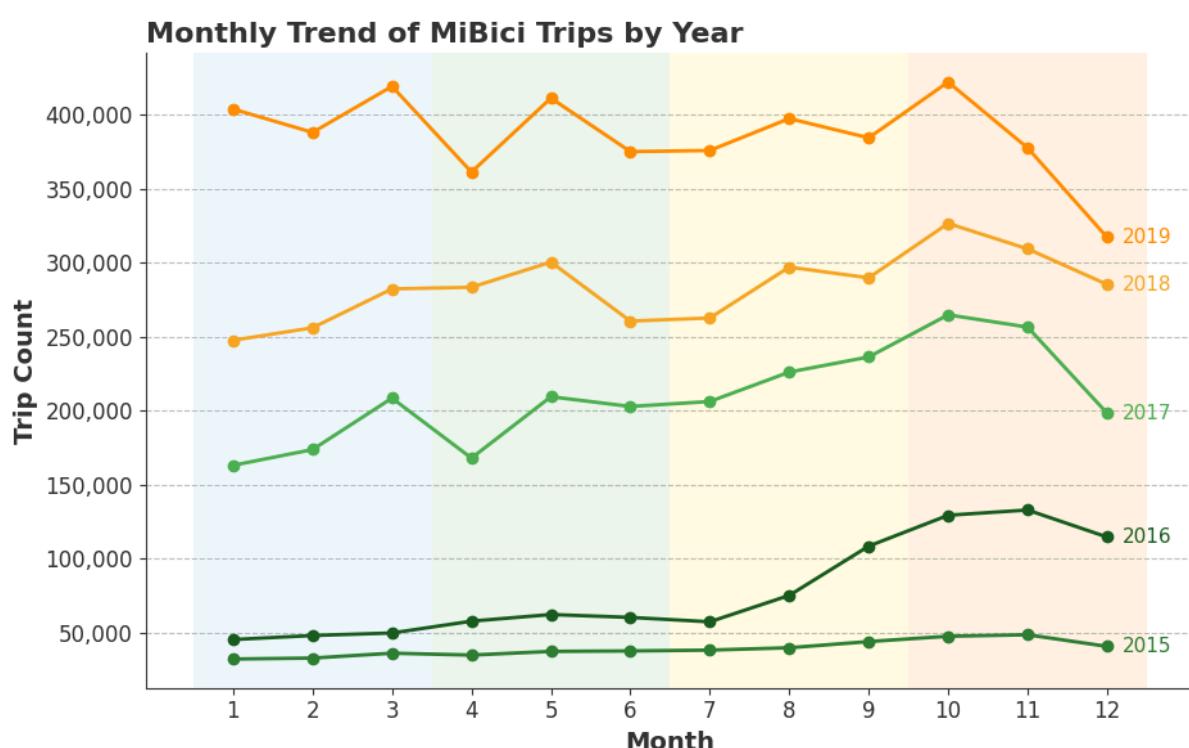
شکل (۹-۲) الگوی روزانه سفرهای MiBici

نمودار شکل (۱۰-۲) روند ماهانه‌ی تعداد سفرهای MiBici از سال ۲۰۱۵ تا ۲۰۱۹ را نشان می‌دهد و به خوبی تغییرات تقاضا و رشد سیستم را در طول سال‌ها و فضول مختلف منعکس می‌کند. حجم کلی سفرها از ۲۰۱۵ تا ۲۰۱۹ به صورت پیوسته افزایش یافته است. خطوط مربوط به سال‌های جدیدتر (نارنجی) در بالاترین سطوح نمودار قرار دارند، که نشان‌دهنده‌ی رشد پایدار و استقبال فزاینده از سامانه در شهر است. در سال ۲۰۱۹، MiBici به بالاترین سطح استفاده در تمام ماههای سال رسیده است.

در سال‌های اولیه مانند ۲۰۱۵ و ۲۰۱۶ (سبز تیره)، تعداد سفرها بسیار کمتر و دارای روند صعودی آرام است. این می‌تواند ناشی از محدود بودن پوشش جغرافیایی یا آشنایی مردم با سامانه در مراحل ابتدایی فعالیت آن باشد.

الگوی فصلی نیز در این نمودار بسیار واضح است. در اکثر سال‌ها، از ماههای ابتدایی سال (ژانویه تا مارس) روند نسبتاً پایینی وجود دارد و سپس در اواسط سال (بین مه تا اکتبر) اوج‌گیری سفرها دیده می‌شود. این الگو می‌تواند مرتبط با شرایط آب‌وهوازی بهتر در فصل بهار و تابستان باشد. در مقابل، در ماه دسامبر (ماه ۱۲)، در اغلب سال‌ها افت محسوسی دیده می‌شود که احتمالاً به دلیل تعطیلات، کاهش دمای هوا یا پایان تقویم آموزشی و کاری است.

نکته جالب اینکه با افزایش کاربران در سال‌های اخیر، **الگوی فصلی حفظ شده ولی سطح کل سفرها بالا رفته**، به ویژه در تابستان ۲۰۱۹ که بالاترین قله‌ها در طول این پنج سال دیده می‌شوند.

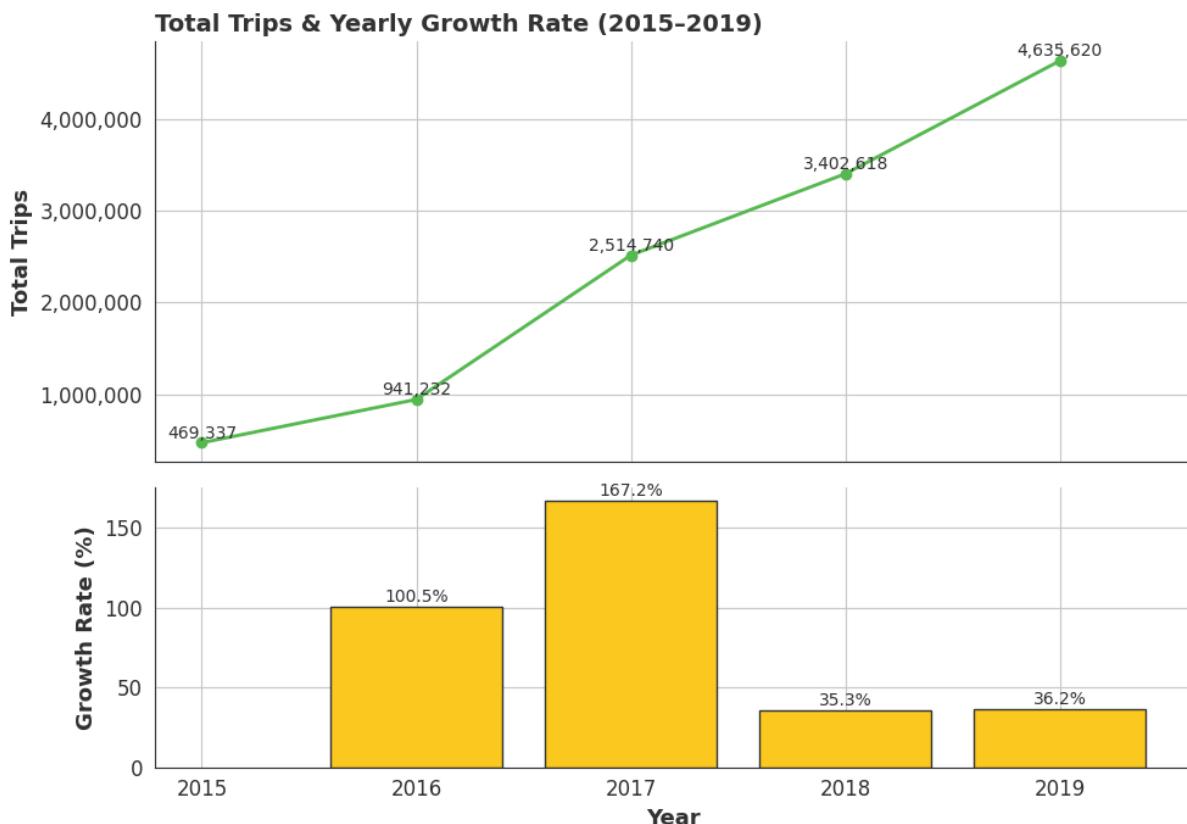


شکل (۱۰-۲) الگوی ماهانه و فصلی سفرهای MiBici در گذر زمان

شکل (۱۱-۲) دو نما از عملکرد سامانه اشتراک دوچرخه MiBici در بازه زمانی ۵ ساله را نمایش می‌دهد. نمودار خطی بالا مجموع تعداد سفرهای انجام‌شده در هر سال و نمودار ستونی پایین نرخ رشد سالانه (درصد افزایش نسبت به سال قبل) است.

تعداد سفرها از ۴۶۹,۳۳۷ در سال ۲۰۱۵ به ۴,۶۳۵,۶۲۰ در سال ۲۰۱۹ رسیده است. این یعنی طی ۵ سال، تعداد سفرها ۱۵ برابر شده است که نشان‌دهنده پذیرش گسترده این سرویس است.

نمودار نرخ رشد سالانه نشان می‌دهد که در سال ۲۰۱۶ تعداد سفرهای انجام شده نسبت به سال ۲۰۱۵ ۱۰۰.۵٪ و در سال ۲۰۱۷ نسبت به سال ۲۰۱۶ ۱۶۷.۲٪ رشد داشته است. اما از ۲۰۱۸ به بعد، نرخ رشد بهطور محسوسی کاهش می‌یابد. MiBici در دو سال ابتدایی رشد انواع تجربی تجربه کرده که احتمالاً ناشی از تبلیغات، توسعه زیرساخت، و ثبت نام گسترده اولیه بوده است. پس از آن، سامانه وارد مرحله تثبیت شده و رشد با شبیه ملائم‌تری ادامه یافته است. کاهش نرخ رشد از ۲۰۱۸ به بعد می‌تواند ناشی از اشباع نسبی بازار هدف یا نیاز به نوآوری و انگیزه‌بخشی مجدد باشد.

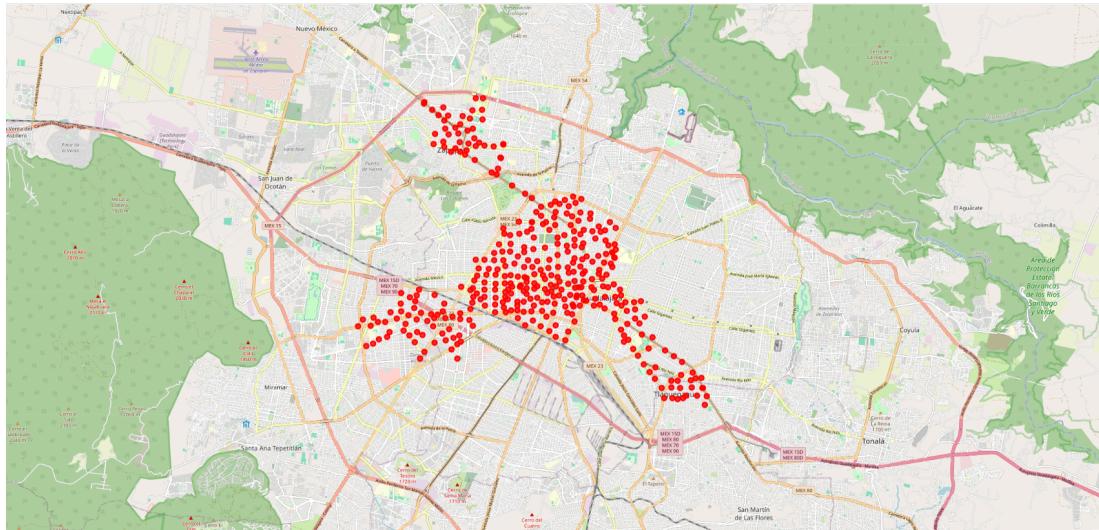


شکل (۱۱-۲) تعداد سفرهای انجام شده و رشد سالانه

۵.۳.۲ استفاده از ایستگاه و جریان ترافیک

شکل (۱۲-۲) موقعیت ایستگاه‌های سامانه اشتراک دوچرخه MiBici را روی نقشه شهر گوادالاخارای مکزیک نمایش می‌دهد.

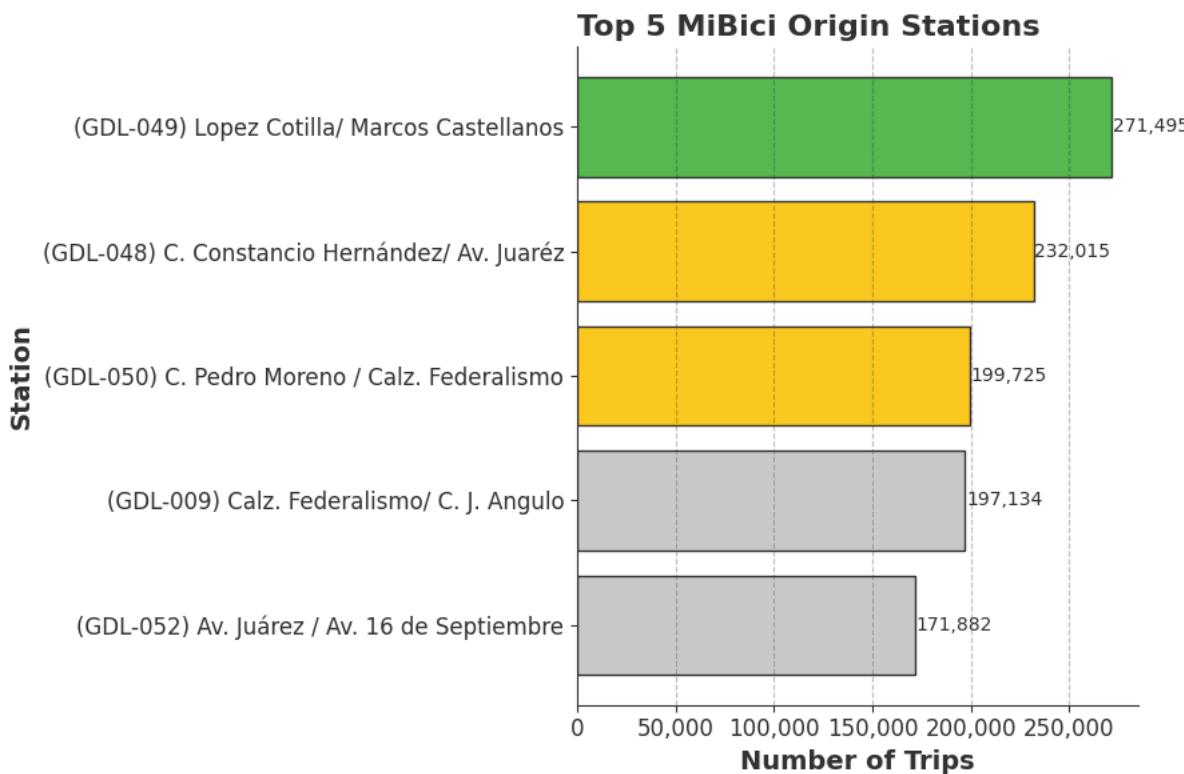
هر دایره نشان‌دهنده یک ایستگاه است و با رنگ مشخص قرمز مکان دقیق آن نمایش داده شده است. این نقشه برای بررسی پوشش جغرافیایی ایستگاه‌ها و شناسایی نقاط با تراکم بالا یا کمبود ایستگاه بسیار مفید است.



شکل (۱۲-۲) نقشه ایستگاه‌ها در شهر

پرترددترین ایستگاه‌های مبدأ:

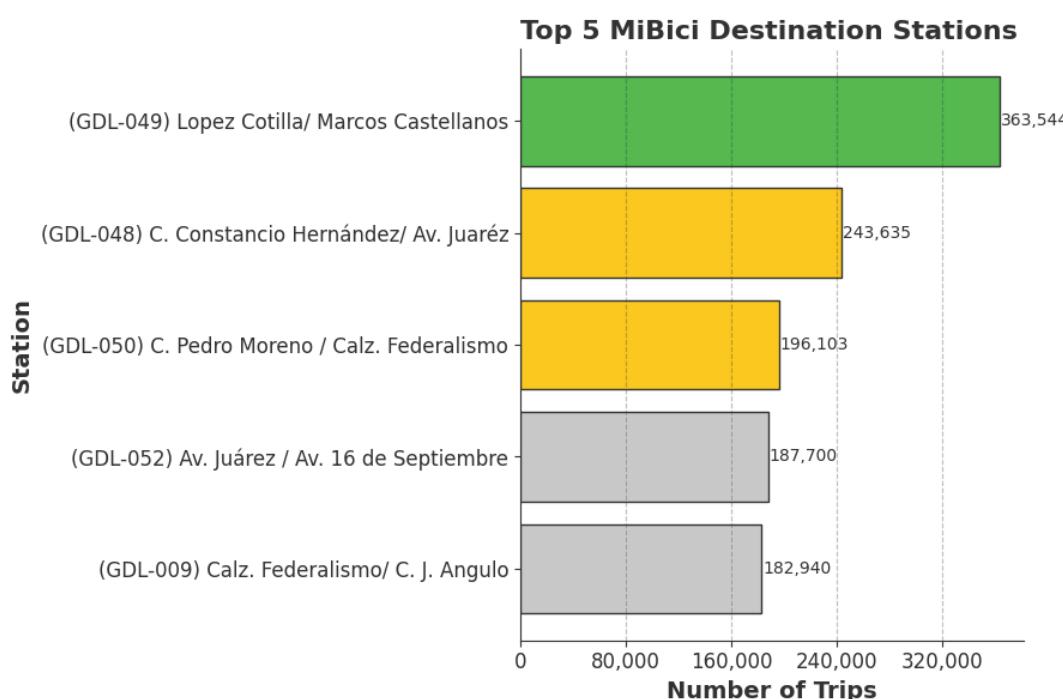
نمودار (۱۳-۲) ۵ ایستگاه پرتردد مبدأ سیستم MiBici را نمایش می‌دهد. ایستگاه (GDL-049) Lopez با بیش از ۲۷۱ هزار سفر، بیشترین تعداد آغاز سفر را در میان همه ایستگاه‌ها دارد. اختلاف بین رتبه اول و بقیه قابل توجه است و نشان می‌دهد که برخی ایستگاه‌ها هاب اصلی شروع سفر هستند.



شکل (۱۳-۲) پرترددترین ایستگاه‌های مبدأ

پرترددترین ایستگاههای مقصد:

نمودار (۱۴-۲) ۵ ایستگاه پرتردد مقصد در سیستم MiBici را نمایش می‌دهد. ایستگاه (GDL-049) Lopez Cotilla / Marcos Castellanos با بیش از ۳۶۳ هزار سفر پایان‌یافته، پرترددترین ایستگاه مقصد است که به احتمال زیاد در منطقه‌ای پرترافیک یا با دسترسی بالا واقع شده است. بیشتر ایستگاههای مقصد مشابه مبدأ هستند (مثل GDL-048 و GDL-050)، که نشان‌دهنده محورهای اصلی تردد روزانه است. توزیع سفرهای پایان‌یافته نیز نشان می‌دهد که مناطق مرکزی شهر مقصد اصلی کاربران است.

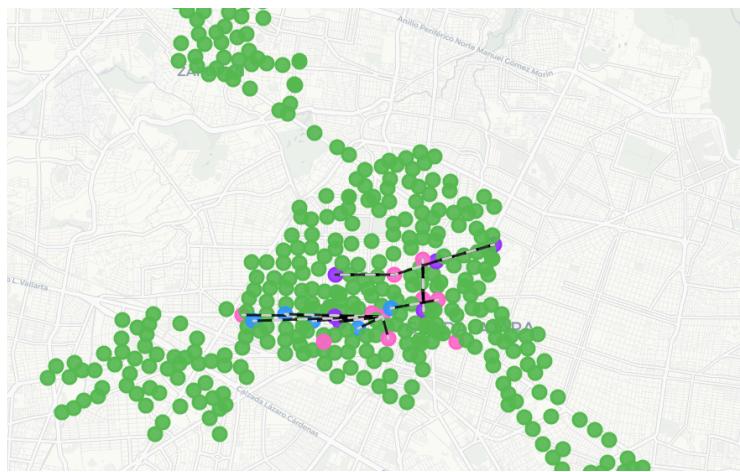


شکل (۱۴-۲) پرترددترین ایستگاههای مقصد

محبوب‌ترین مسیرها (جفت مبدأ-مقصد):

شکل (۱۵-۲) نقشه مسیرها و ایستگاههای پرکاربرد MiBici در سال ۲۰۱۹ را نشان می‌دهد که با رنگ‌بندی خاص، دسته‌بندی ایستگاهها و مسیرها را تفکیک کرده است:

- نقاط سبز: تمام ایستگاههای MiBici هستند که در سطح شهر گوادالاخارا پراکنده شده‌اند.
- نقاط آبی: ایستگاههایی که به عنوان مبدأ پرترکار (در بین ۲۰ مسیر پرکاربرد) شناخته شده‌اند.
- نقاط بنفس: ایستگاههایی که هم به عنوان مبدأ و هم مقصد پرترکار ظاهر شده‌اند.
- نقاط صورتی: ایستگاههایی هستند که در مسیرهای رفت‌وبرگشتی پرترکار (round-trip) قرار دارند.
- خطوط مشکی: مسیرهای پرترکار بین ایستگاههای مختلف (top 20 paths) را نشان می‌دهند که الگوی سفرهای شهری را بازنگاری می‌کنند.

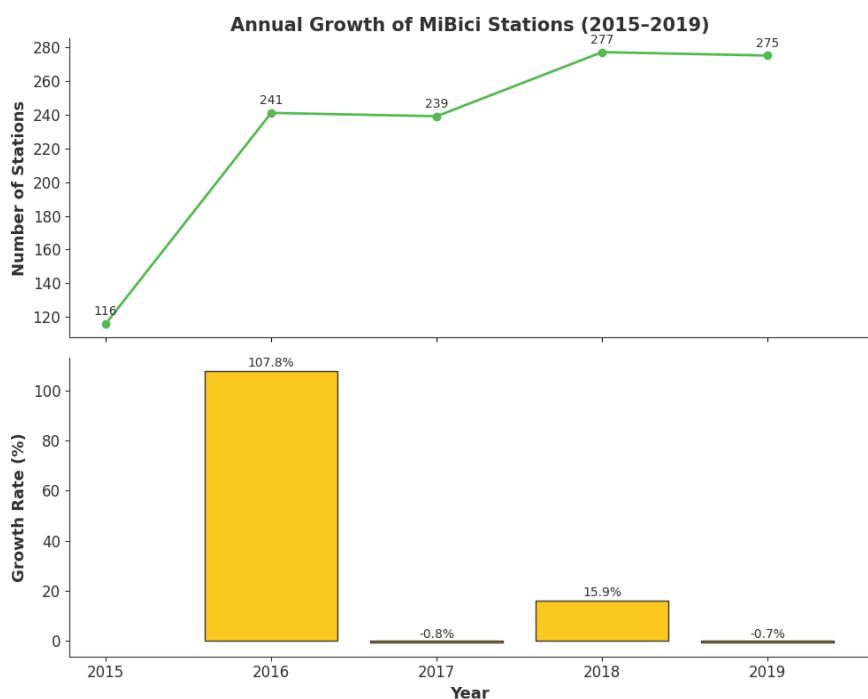


شکل (۱۵-۲) محبوب‌ترین مسیرها (جفت مبدأ و مقصد)

نمودار شکل (۱۶-۲) روند رشد سالانه ایستگاه‌های MiBici در بازه‌ی زمانی ۲۰۱۵ تا ۲۰۱۹ را نمایش می‌دهد. نمودار بالا تعداد کل ایستگاه‌ها در هر سال را نشان می‌دهد. در سال ۲۰۱۵ تنها ۱۱۶ ایستگاه فعال بوده است. در سال ۲۰۱۶ تعداد ایستگاه‌ها با جهشی چشمگیر به ۲۴۱ رسید (افزایش بیش از دو برابر). اما در سال‌های بعد، روند رشد متوقف شد و حتی در سال ۲۰۱۷ و ۲۰۱۹ اندکی کاهش یافت.

نمودار پایین نرخ رشد درصدی را نمایش می‌دهد. سال ۲۰۱۶ با رشد ۱۰۷٪، نقطه‌ی اوج توسعه‌ی فیزیکی شبکه ایستگاه‌ها است. از سال ۲۰۱۷ به بعد، نرخ رشد به صفر نزدیک یا حتی منفی شده، که نشان‌دهنده‌ی ثبات یا توقف گسترش زیرساخت ایستگاهی است.

پس از توسعه سریع در سال ۲۰۱۶ MiBici وارد مرحله‌ی تثبیت شد و تا پایان ۲۰۱۹ رشد جدیدی در تعداد ایستگاه‌ها اتفاق نیفتاد. این موضوع می‌تواند بیانگر تمرکز بر نگهداری، بهره‌برداری، یا اشباع پوشش منطقه‌ای در فاز فعلی باشد.



شکل (۱۶-۲) رشد ایستگاه‌ها

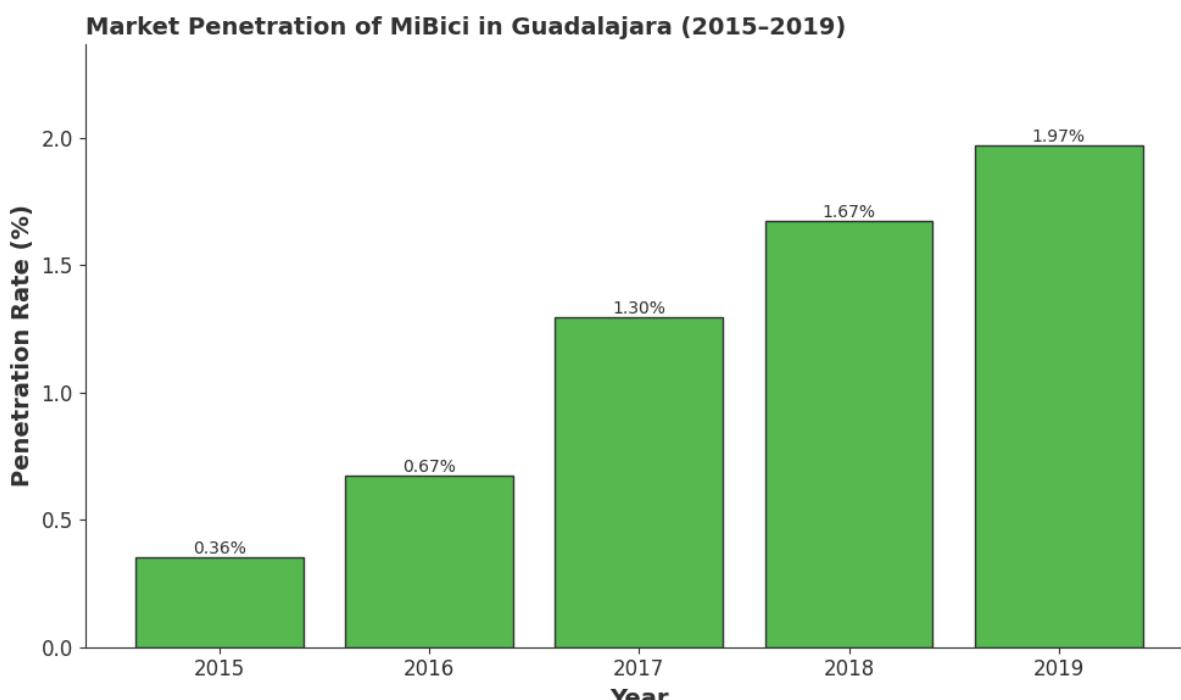
۶.۳.۲ سهم بازار

نمودار (۱۷-۲) نرخ نفوذ (Market Penetration) سرویس دوچرخه اشتراکی MiBici در شهر گوادالاخارا طی سال‌های ۲۰۱۵ تا ۲۰۱۹ را نشان می‌دهد. نرخ نفوذ یعنی چه درصدی از جمعیت شهر در هر سال حداقل یکبار از MiBici استفاده کرده‌اند و فرمول آن به صورت زیر است:

$$\text{نرخ نفوذ (\%)} = \frac{\text{تعداد کاربران منحصربهفرد}}{\text{جمعیت کل}} \times 100$$

مطابق نمودار (۱۷-۲) در سال ۲۰۱۵، تنها ۰.۳۶٪ از جمعیت گوادالاخارا از MiBici استفاده کردند. این نرخ به مرور افزایش یافته تا در سال ۲۰۱۹ به ۱.۹۷٪ نزدیک شده است. یعنی از هر ۵۰ نفر در شهر، ۱ نفر کاربر فعال MiBici بوده است.

در مدت کوتاهی توانسته است سهم قابل توجهی از بازار حمل و نقل شهری را به دست آورد، هرچند هنوز فاصله زیادی تا اشباع بازار وجود دارد. افزایش ایستگاه‌ها، تبلیغات هدفمند و بهبود دسترسی می‌توانند در ادامه این روند مؤثر باشند.



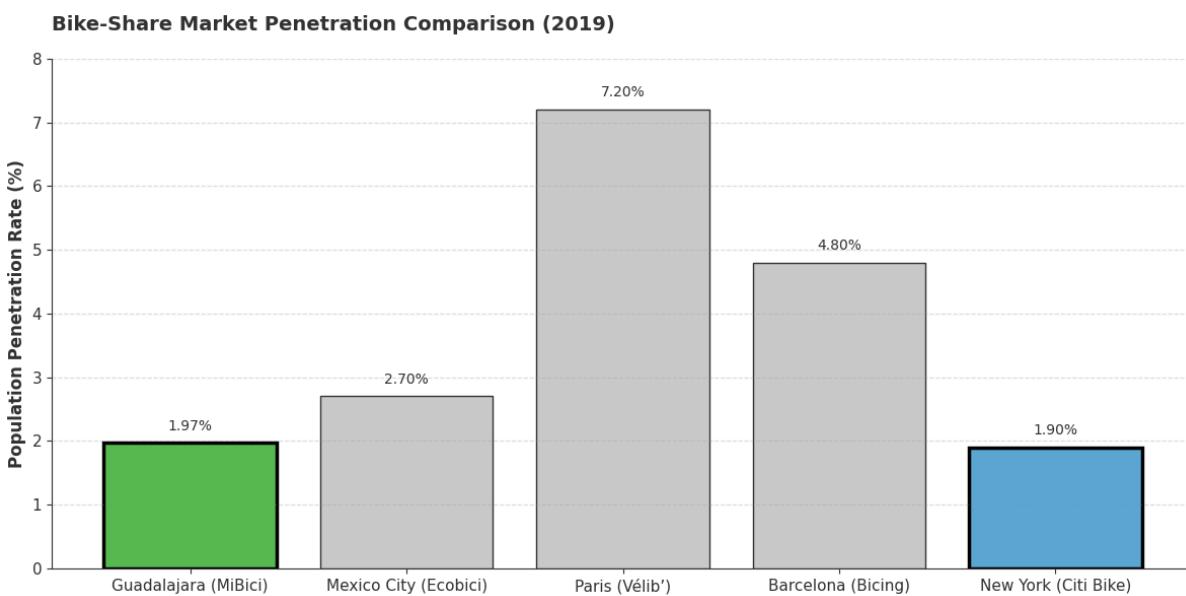
شکل (۱۷-۲) نرخ نفوذ MiBici

نمودار شکل (۱۸-۲) نرخ نفوذ (penetration rate) سیستم‌های اشتراک دوچرخه را در چند شهر بزرگ دنیا در سال ۲۰۱۹ مقایسه می‌کند.

- بالاترین نرخ نفوذ مربوط به شهر پاریس (Vélib') با ۷.۲٪ است. این نشان‌دهنده فراگیری گستره و موفقیت‌آمیز این سامانه در جذب کاربران است.
- بارسلونا (Bicing) با ۴.۸٪ و مکزیکو سیتی (Ecobici) با ۲.۷٪ در جایگاه‌های بعدی قرار دارند.

- گوادالاخارا (MiBici) با نرخ نفوذ ۱.۹۷٪ عملکردی قابل توجه اما پایین‌تر از میانگین اروپا دارد، با این حال نسبت به نیویورک (CitiBike) با ۱.۹٪ کمی بهتر عمل کرده است.

در گوادالاخارا هنوز ظرفیت بالایی برای رشد دارد. با توجه به تجربه موفق شهرهایی مانند پاریس و بارسلونا، می‌توان با توسعه ایستگاه‌ها، بهبود زیرساخت و بازاریابی مؤثر، نرخ نفوذ را افزایش داد.



شکل (۱۸-۲) مقایسه نرخ نفوذ در سامانه‌های دوچرخه اشتراکی متفاوت

بخش سوم:
پاسخ به سوالات کسب و کار

۱.۳ تبلیغات دوچرخه‌های مسافت بلند (سوال ۲)

۱.۱.۳ تشریح مسئله کسب و کار

بازخوردهای دریافتی از کاربران نشان می‌دهد که دوچرخه‌های مخصوص مسافت بلند به دلیل کاهش آسیب‌های جسمی و طراحی ارگونومیک، باید از دوچرخه‌های معمولی متمایز باشند. مدیر بازاریابی شرکت MiBici خواهان یک رویکرد داده‌محور جهت شناسایی و هدف‌گذاری دقیق کاربران است. سؤالات اصلی عبارت‌اند از:

- چه کسانی در حال حاضر بیشترین استفاده از دوچرخه‌های مسافت بلند را دارند؟
- چگونه با بهره‌گیری از ویژگی‌های رفتاری، دموگرافیک و مکانی می‌توان کاربران جدید یا کاربرانی با الگوی مشابه را شناسایی نمود؟

در این راستا، علاوه بر هدف‌گذاری بر روی کاربرانی که در سال گذشته عملکرد بالای ثبت سفرهای بلند داشته‌اند، از مدل پیش‌بینی برای ارزیابی کاربران جدید و همچنین از خوش‌بندی جهت تقسیم‌بندی دقیق کاربران بر اساس رفتار سفرهای بلند استفاده می‌شود.

۲.۱.۳ متداول‌وزی و معیارهای استفاده شده

الف) تعریف متريک «سفر بلند»

- استفاده از مسافت مستقیم:

به جای محاسبه مسافت واقعی طی‌شده (که نیازمند استفاده از API‌های مسیریابی شهری است و از نظر زمانی و هزینه‌ای چالش‌برانگیز می‌باشد)، از مسافت مستقیم بین ایستگاه‌های مبدأ و مقصد به عنوان شاخص (Proxy Metric) استفاده شده است.

- دلایل این انتخاب:

محاسبه مسافت واقعی که مسیرهای خیابانی و انحراف‌های شهری را در نظر می‌گیرد، مستلزم استفاده از API‌های مسیریابی مانند OpenRoute یا Google Maps است. این سرویس‌ها اغلب محدودیت درخواست دارند یا نیاز به پرداخت هزینه دارند که در این پروژه امکان‌پذیر نبوده است.

"مسافت مستقیم" (فاصله‌ی جغرافیایی بین مبدأ و مقصد) به‌راحتی و به صورت قابل‌تکرار (Reproducible) از داده‌ها استخراج می‌شود، و برای اهداف تحلیلی و مقایسه‌ای بین کاربران یک شاخص استاندارد فراهم می‌کند.

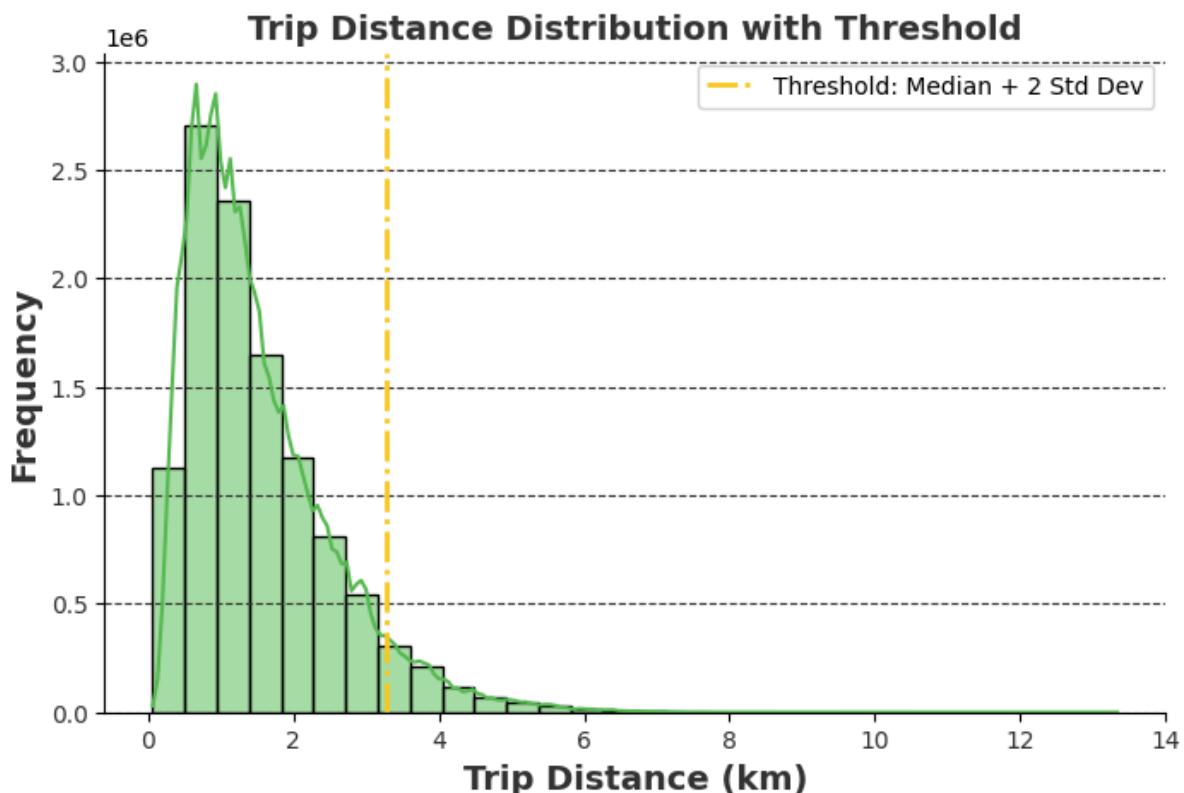
هرچند مسافت مستقیم معمولاً از مسیر واقعی کوتاه‌تر است، اما با استفاده از یک آستانه آماری بالا (میانه $+ 2 \times$ انحراف معیار)، سفرهایی که به‌طور معناداری طولانی (بیش از 3.27 کیلومتر) هستند به درستی شناسایی می‌شوند. این روش باعث می‌شود اثر خطای ناشی از کوتاه‌نمایی مسیر به حداقل برسد.

$$\text{Long Distance Threshold} = \text{Median}(\text{Direct Distance}) + 2 * \text{Standard Deviation}(\text{Direct Distance})$$

شکل (۱-۳) توزیع واقعی فاصله سفرها در قالب هیستوگرام + منحنی KDE برای نمایش روند را نشان می‌دهد. بیشتر سفرها بین ۰.۵ تا ۲.۵ کیلومتر انجام شده‌اند. نمودار بهشت چوله به راست است (Right-skewed)، یعنی بیشتر کاربران سفرهای کوتاه انجام می‌دهند.

آستانه‌ی زرد (حدود ۳.۵ کیلومتر) فاصله‌ای را نشان می‌دهد که از آن به بعد، سفرها به عنوان سفرهای طولانی (long-distance) طبقه‌بندی شده‌اند. این آستانه براساس ترکیب میانه و انحراف معیار تنظیم شده که در برابر داده‌های پرت مقاوم باشد.

تعداد سفرهای طولانی به‌طور چشمگیری کاهش یافته است و این نشان می‌دهد که فقط اقلیت کوچکی از کاربران سفرهای طولانی دارند.



شکل (۱-۳) توزیع مسافت طی شده در سفرها

ب) کنار گذاشتن «مدت‌زمان سفر»

- عدم قطعیت تفسیر:

توقف در مقصد یا پارک در محل کار می‌تواند باعث افزایش غیرواقعی مدت زمان سفر شود، در حالی که کاربر الزاماً مسافت طولانی را طی نکرده است.

- عدم تطابق مستقیم با مسافت:

زمان طولانی سفر همیشه معادل با طی کردن مسیر بلند نیست؛ بنابراین معیار مسافت مستقیم به عنوان شاخص دقیق‌تری انتخاب گردید.

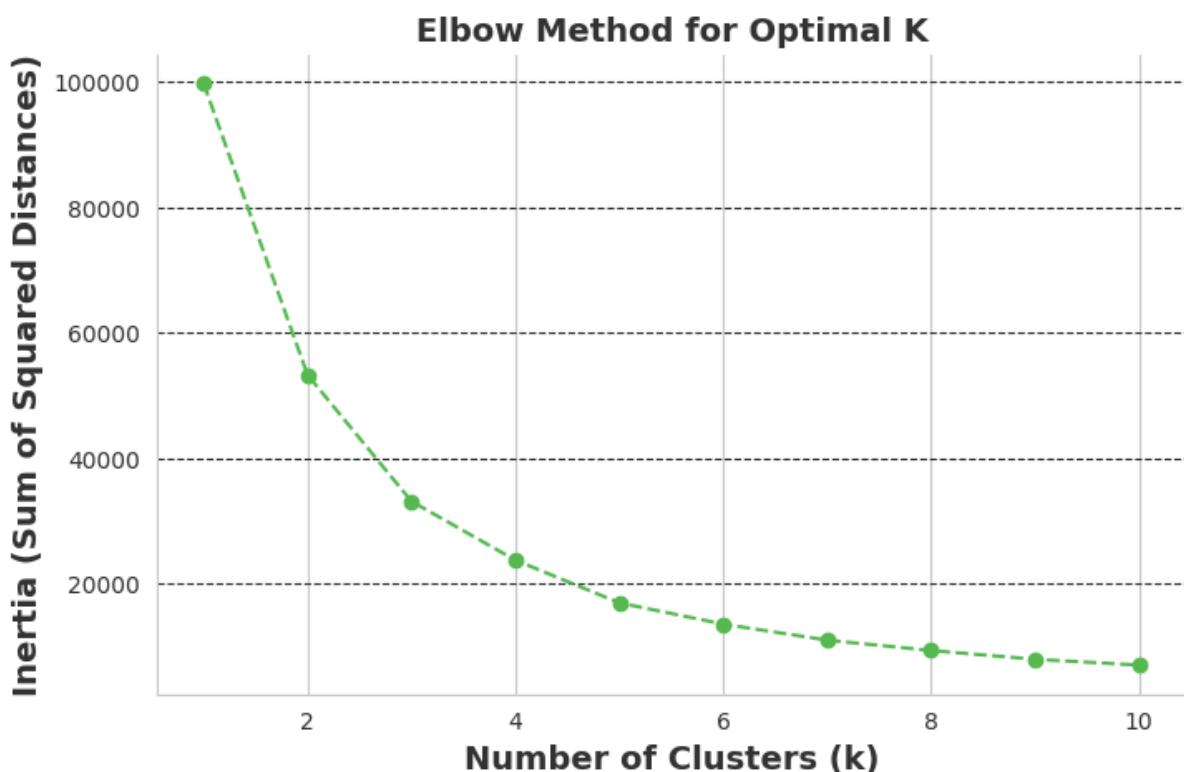
پ) شناسایی مشتریان هدف از طریق خوشبندی

به منظور هدفگذاری دقیق‌تر و داده‌محور، رویکرد خوشبندی به جای استفاده از یک تعداد ثابت کاربران به کار گرفته شده است. مراحل اصلی این بخش به شرح زیر است:

۳.۱.۳ مدل خوشبندی k_means

داده‌های کاربران براساس تعداد سفرهای بلند ثبت شده و نسبت سفرهای بلند به کل سفرها (long_trip_ratio) تجمعی شدند.

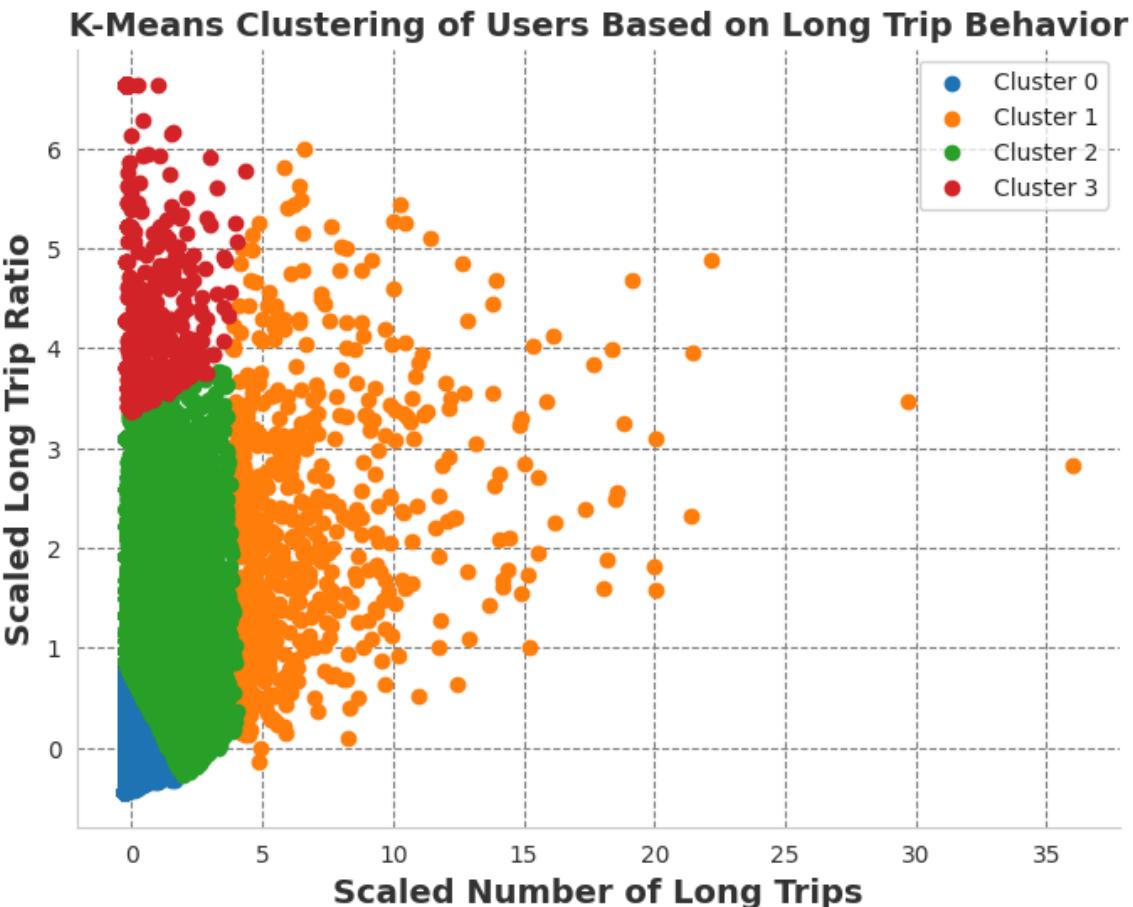
طبق نمودار Elbow (شکل ۲-۳) توصیه می‌شود که کاربران به چهار گروه (خوش) طبقه‌بندی شوند.



شکل (۲-۳) نمودار Elbow برای انتخاب بهترین تعداد خوش

خروجی خوشبندی:

تحلیل خوشبندی نشان می‌دهد خوشبندی که دارای میانگین بسیار بالای ثبت سفرهای بلند (حدود 391.25 سفر بلند به ازای هر کاربر) و نسبت متوسط 0.38 است، شامل 593 کاربر با شناسه‌های یکتا شده و به عنوان گروه هدف اصلی (Primary Target Group) انتخاب گردیده است.



شکل (۳-۳) خوشه‌بندی کاربران براساس تعداد و نسبت سفرهای طولانی با استفاده از الگوریتم K-Means

د) مدل پیش‌بینی

علاوه بر خوشه‌بندی، می‌توان برای کاربران جدید یا کاربران فعلی از یک مدل (با استفاده از الگوریتم رگرسیون لجستیک) بهره گرفت که با ورود ویژگی‌هایی همچون:

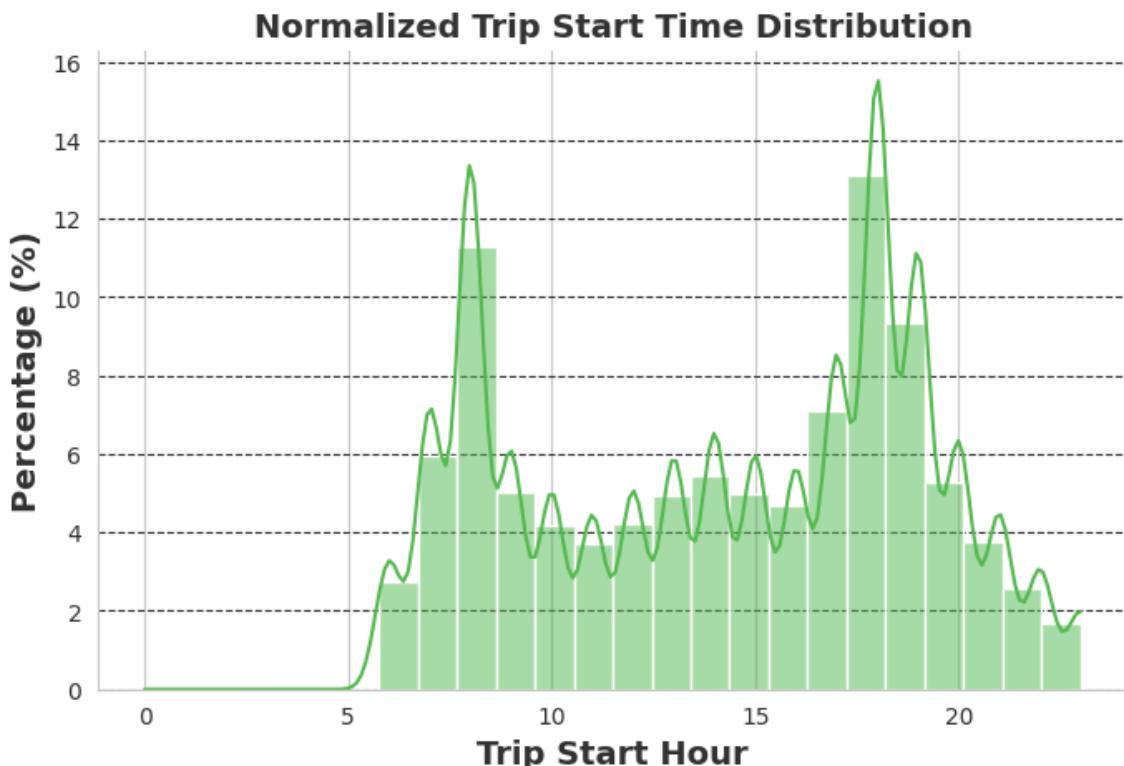
- سن (Age)
- جنسیت (Gender)
- زمان شروع سفر (Trip Start Hour)
- روز هفته (Day-of-Week)
- ماه سفر (Trip Month)
- ایستگاه مبدأ (Origin Station ID)

احتمال ثبت سفر بلند (long_prob) را برآورد می‌کند. این امکان به تیم بازاریابی کمک می‌کند تا در لحظه‌های ورود کاربر به سیستم، پیام‌های تبلیغاتی مناسب نمایش داده شود.

۴.۱.۳ یافته‌ها و بینش‌ها

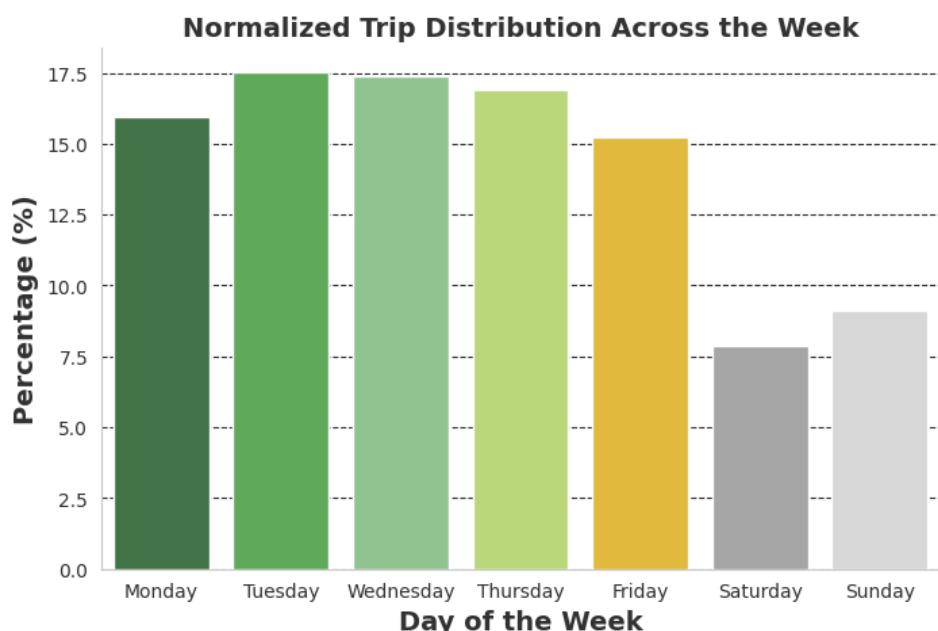
بینش‌های استخراج شده از نمودارهای تحلیلی

نمودار (۴-۳) توزیع زمان شروع سفر را نشان می‌دهد که سفرهای بلند عمدتاً در ساعت‌های اوج صبحگاهی و عصرگاهی آغاز می‌شوند. این الگو به تیم بازاریابی کمک می‌کند تا کمپین‌های تبلیغاتی در زمان‌های پرtraفیک اجرا شود.



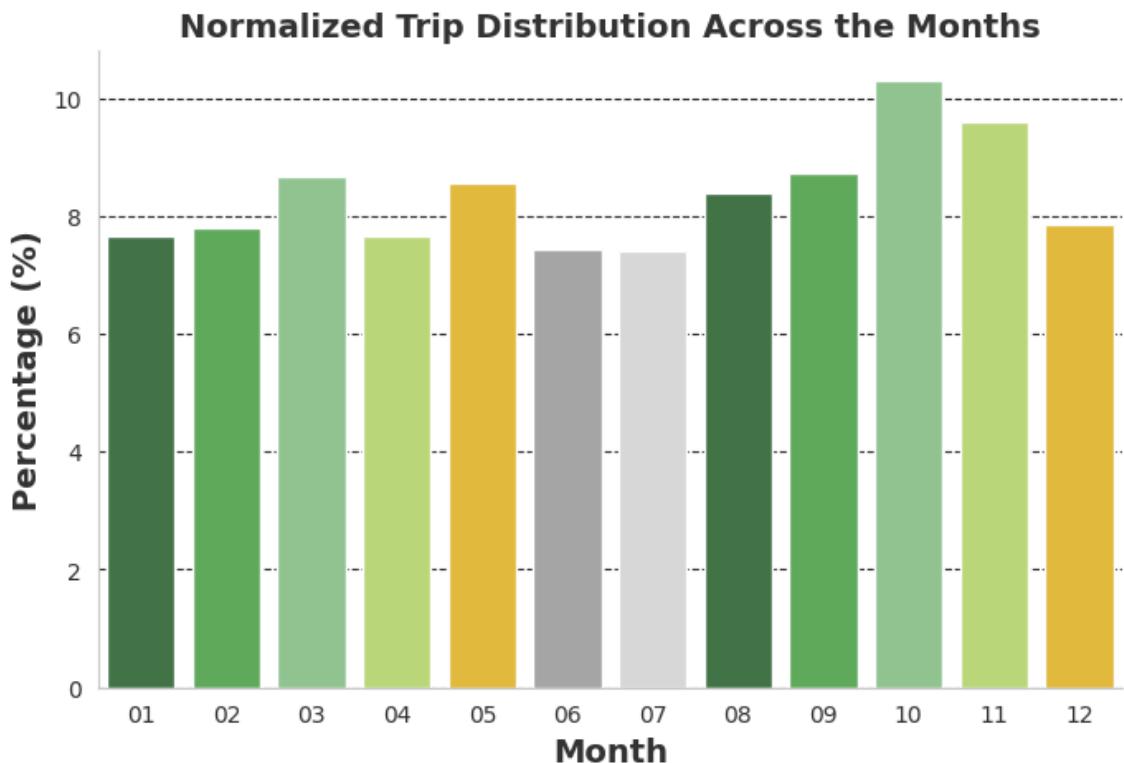
شکل (۴-۳) توزیع زمان (ساعت) شروع سفر در سفرهای طولانی

شکل (۵) نمودار توزیع نرمال روزهای هفته را نشان می‌دهد. الگوهای هفتگی ثبت سفرهای بلند بر اساس روزهای هفته مشخص شده است؛ به عنوان مثال، برخی روزها مانند آخر هفته‌ها فعالیت کمتری دارند.



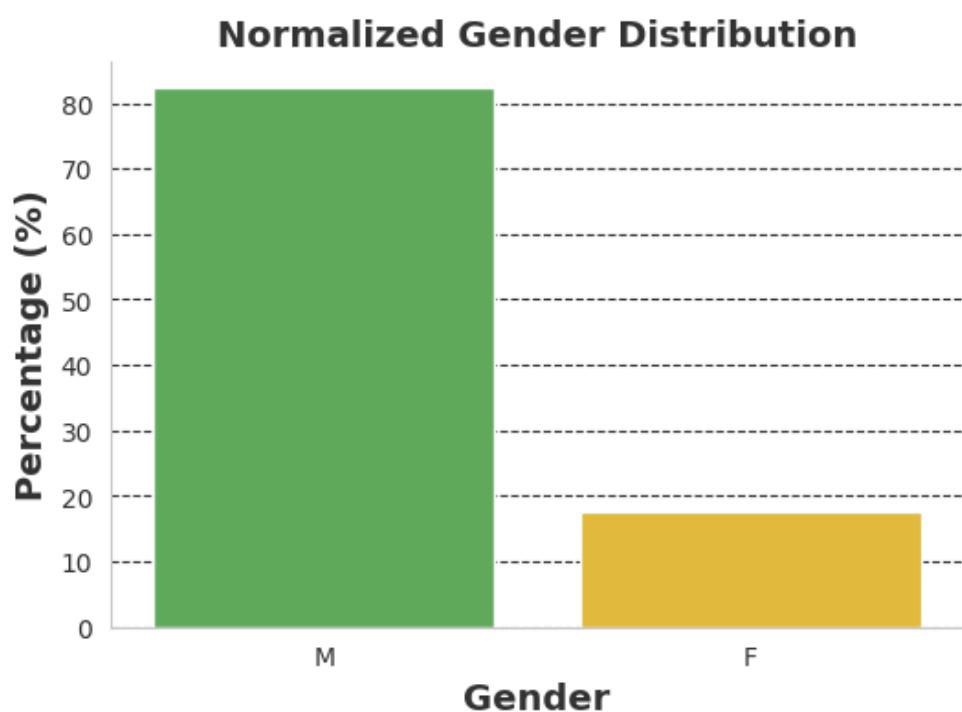
شکل (۵) توزیع روزهای هفته در سفرهای طولانی

شکل (۶-۳) نمودار توزیع ماهانه سفرهای طولانی به تفکیک ماه در طول سال را نشان می‌دهد. الگوهای فصلی (مثلاً افزایش فعالیت در ماههای پایانی سال) به برنامه‌ریزی کمپین‌های فصلی کمک کرده و امکان تنظیم پیام‌های متناسب با شرایط فصلی را فراهم می‌آورد.



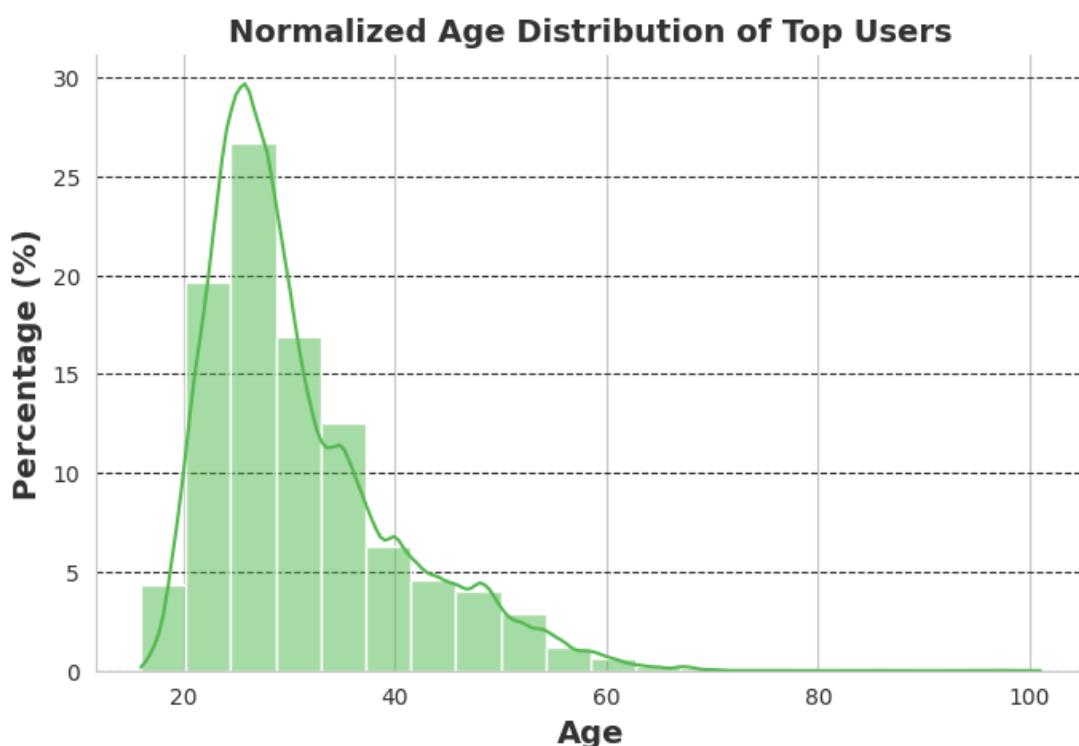
شکل (۶-۳) توزیع ماهانه سفرهای طولانی

شکل (۷-۳) نمودار توزیع جنسیتی کاربران در سفرهای طولانی را نشان می‌دهد. تقسیم‌بندی جنسیتی در میان کاربران ثبت‌کننده نشان می‌دهد که بالای ۸۰٪ از سفرهای بلند توسط مردان ثبت شده باشد. این بینش الزامی است تا پیام‌های تبلیغاتی به صورت ویژه و متناسب با نیازهای هر دو گروه جنسیتی تنظیم شوند.



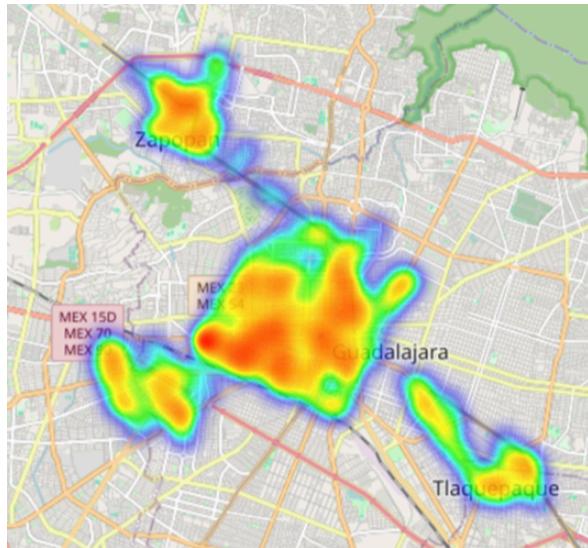
شکل (۷-۳) توزیع جنسیتی کاربران در سفرهای طولانی

شکل (۸-۳) درصد کاربران در بازه‌های سنی مختلف، به ویژه گروه سنی ۲۵ تا ۳۰ سال که بیشترین فعالیت را نشان می‌دهند، به تصویر کشیده شده است. بدین ترتیب، پیام‌های تبلیغاتی می‌توانند با تأکید بر ویژگی‌هایی مانند طراحی ارگونومیک و کاهش آسیب‌های جسمی، متناسب با نیازهای این گروه تنظیم شوند.



شکل (۸-۳) توزیع سنی کاربران در سفرهای طولانی

شکل (۹-۳) تحلیل استفاده تجمعی از ایستگاههای میدا را نشان می‌دهد. مطابق این تصویر برخی ایستگاهها به عنوان ایستگاههای اصلی در سطح شهر شناسایی شده‌اند که کاربران سفرهای بلند بیشترین استفاده را از آن‌ها دارند. این اطلاعات به تیم بازاریابی امکان می‌دهد تا فعالیت‌های محلی مانند نصب پوستر یا QR کدهای تبلیغاتی در این مناطقی را هدف قرار دهند.



شکل (۹-۳) استفاده تجمعی از ایستگاههای میدا

خوشبندی کاربران (K-Means Clustering)

خوشبندی کاربران بر اساس تعداد سفرهای بلند و نسبت آن (long_trip_ratio) نشان می‌دهد که چهار خوشبندی متفاوت به دست آمده است.

خوشبندی با 593 کاربر، که دارای میانگین سفرهای بلند بسیار بالا (حدود 391.25 سفر بلند) و نسبت متوسط 0.38 است، به عنوان گروه هدف اصلی انتخاب شده است.

نکات کلیدی استخراج شده

زمان‌های اوج ثبت سفر: ارائه بینش جهت اجرای کمپین‌های زمان‌بندی شده در ساعت پر ترافیک.

تقسیم‌بندی دقیق بر اساس رفتار و دموگرافیک: استفاده از نمودارهای توزیع (روز، ماه، جنسیت و سن) و نیز خوشبندی، امکان هدف‌گذاری دقیق‌تر را فراهم کرده است.

استفاده تجمعی از ایستگاههای میدا: شناسایی نقاط پرتردد جهت راه‌اندازی فعالیت‌های محلی تبلیغاتی.

گروه هدف اصلی خوشبندی شده: خروجی تحلیل خوشبندی (با 593 کاربر یکتا) به عنوان گروه اصلی هدف در نظر گرفته شده که از نظر کمی و کیفی متمایز است.

۵.۱.۳ توصیه‌ها و اقدامات عملی

جدول (۱-۳) توصیه‌ها و اقدامات عملی برای جذب و نگهداری کاربران از طریق ایمیل، کمپین‌های تبلیغاتی چندکاناله، پیام‌های تبلیغاتی هدفمند، و آزمایش‌های A/B برای بهینه‌سازی پیام‌ها را برای گروه‌های متفاوت شناسایی شده نشان می‌دهد.

جدول (۱-۳) توصیه‌ها و اقدامات عملی برای گروه‌های متفاوت مشتریان شناسایی شده

گام	مخاطب هدف	اقدام
۱	گروه هدف اصلی (خوشه ۵۹۳ کاربر)	ارسال ایمیل‌های هدفمند به همراه ارائه تخفیف اشتراکی ویژه برای سفرهای بلند
۲	کاربرانی با ویژگی‌های مشابه (Look-alike)	اجرای کمپین‌های تبلیغاتی چندکاناله (SMS, Push Notification) برای جذب کاربرانی با الگوی رفتاری مشابه
۳	کاربران جدید (Real-Time Scoring)	نمایش پیام‌های تبلیغاتی در لحظه هنگام رزرو با امتیازدهی فوری مدل پیش‌بینی احتمال ثبت سفر بلند
۴	تمام کاربران	اجرای آزمایش‌های A/B جهت بهینه‌سازی پیام‌های تبلیغاتی

۶.۱.۳ جمع‌بندی و نتیجه‌گیری

این گزارش جامع، با استفاده از رویکرد داده‌محور، مسئله تبلیغات دوچرخه‌های مسافت بلند را از منظر چندگانه بررسی می‌کند. از جمله نکات کلیدی این گزارش می‌توان به موارد زیر اشاره کرد:

استفاده از مسافت مستقیم: به منظور ارائه شاخصی قابل تکرار و کارآمد برای شناسایی سفرهای بلند، با در نظر گرفتن محدودیت‌های زمانی و هزینه‌ای تحلیل واقعی مسیر.

کنار گذاشتن مدت‌زمان سفر: به دلیل تأثیر توقف‌ها و عوامل نامرتبط، معیار دقیقت‌تری جهت شناسایی سفرهای بلند اتخاذ نشد.

خوبه‌بندی کاربران: استفاده از K-Means منجر به تقسیم‌بندی دقیق کاربران بر اساس رفتار سفر. خوبه‌ای با 593 کاربر به عنوان گروه هدف اصلی شناسایی شد که دارای شناسه‌های یکتا و ویژگی‌های رفتاری برجسته هستند.

بینش‌های حاصل از نمودارهای تحلیلی: نمودارهای توزیع زمان شروع سفر، روزهای هفته، ماه سفر، درصد جنسیت و توزیع سنی، در کنار تحلیل استفاده تجمعی از ایستگاه‌های مبدأ، اطلاعات بسیار مفیدی را جهت تنظیم پیام‌های تبلیغاتی متناسب با نیازهای کاربران ارائه می‌دهد.

این رویکرد جامع امکان تعریف دقیقت‌ترین گروه‌های هدف، اجرای کمپین‌های تبلیغاتی چند کاناله و شخصی‌سازی پیام‌ها را فراهم می‌آورد. از منظر داده و استراتژی بازاریابی، گزارش حاضر پاسخگوی نیازهای مدیر بازاریابی و معیارهای ارزیابی حرفه‌ای است.

۲.۳ شناسایی و پیش‌بینی دوچرخه‌های معیوب در ناوگان CitiBike (سوال ۳)

ناوگان دوچرخه‌ها اشتراکی CitiBike در نیویورک، سرویسی حیاتی برای حمل و نقل شهری است. کیفیت و در دسترس بودن دوچرخه‌ها تأثیر مستقیمی بر رضایت کاربران و کارایی عملیاتی شرکت دارد. دوچرخه‌های معیوب نه تنها تجربه کاربری را مختل می‌کنند، بلکه منجر به افزایش هزینه‌های نگهداری ناگهانی و کاهش بهره‌وری ناوگان می‌شوند. هدف از این گزارش، ارائه یک چارچوب تحلیلی جامع برای شناسایی فعالانه و پیش‌بینی دوچرخه‌های معیوب در ناوگان، با استفاده از داده‌های تردد روزانه از ژوئن تا دسامبر 2013 است. این تحلیل شامل رویکردهای تشخیصی (شناسایی مشکلات موجود) و پیش‌بینی‌کننده (شناسایی مشکلات در حال ظهر) خواهد بود.

۱.۲.۳ روش‌های تشخیصی (رویکرد اولیه)

این بخش به دو روش اصلی برای شناسایی دوچرخه‌هایی که در حال حاضر مشکل دارند یا از سرویس خارج شده‌اند، می‌پردازد.

1. تحلیل عدم فعالیت دوچرخه‌ها: شناسایی دوچرخه‌هایی که برای مدت طولانی یا به دفعات زیاد غیرفعال هستند.
2. تحلیل ناهنجاری در مدت زمان سفر: شناسایی دوچرخه‌هایی که درصد بالایی از سفرهای بسیار کوتاه را به خود اختصاص می‌دهند.

روش ۱: تحلیل عدم فعالیت دوچرخه‌ها

این روش بر این فرض استوار است که دوچرخه‌هایی که برای مدت زمان قابل توجهی (چه به صورت پیوسته و چه تکراری) غیرفعال هستند، ممکن است دچار خرابی و در انتظار تعمیر باشند، یا به درستی در چرخه عملیات قرار نگرفته باشند. تحلیل با ساخت یک جدول آغاز می‌شود که وضعیت فعالیت (تعداد سفرهای روزانه) هر دوچرخه را از اولین روز فعالیت آن در ناوگان تا آخرین روز سفرش ثبت می‌کند. این رویکرد تضمین می‌کند که عدم فعالیت تنها پس از ورود دوچرخه به ناوگان در نظر گرفته شود.

متداول‌تر و انتخاب آستانه‌ها:

• آستانه اولیه - ۷ روز عدم فعالیت متوالی:

- علت انتخاب: یک دوره ۷ روزه (یک هفته) عدم فعالیت مداوم، یک آستانه عملیاتی معنادار است.
- دوچرخه‌ای که یک هفتۀ کامل بدون استفاده می‌ماند، به احتمال زیاد دچار مشکلی است که فراتر از نوسانات عادی تقاضا یا جابجایی روتین است. این دوره زمانی به اندازه کافی بلند است تا شامل نگهداری‌های معمول نشود و نشانه‌ای از یک خرابی جدی‌تر یا "گم شدن" دوچرخه در چرخه عملیات باشد. از دست رفتن درآمد حاصل از یک هفتۀ عدم فعالیت نیز برای شرکت قابل توجه است.
- با انتخاب این آستانه ۴۹۸۷ دوچرخه شناسایی شدند که بیش از ۷ روز بدون فعالیت بودند.

• معیارهای تکمیلی

- تعداد دوره‌های سکون (inactivity_periods) تعداد دفعاتی که یک دوچرخه وارد دوره‌ی عدم فعالیت ۷ روزه یا بیشتر شده است.

- طولانی‌ترین دوره سکون (longest_inactivity): حداقل تعداد روزهای متوالی عدم فعالیت برای یک دوچرخه.
- مجموع روزهای سکون (total_inactive_days): مجموع تمام روزهای عدم فعالیت در دوره‌هایی که حداقل 7 روز طول کشیده‌اند.
- نسبت روزهای سکون به کل روزها (inactive_ratio): معیاری نرمال شده که درصد روزهای غیرفعال را از کل روزهای حضور دوچرخه در ناوگان نشان می‌دهد.

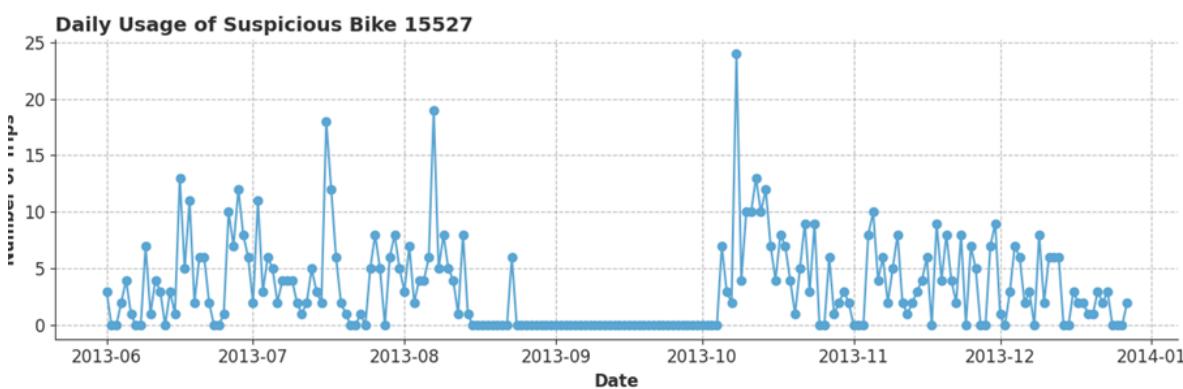
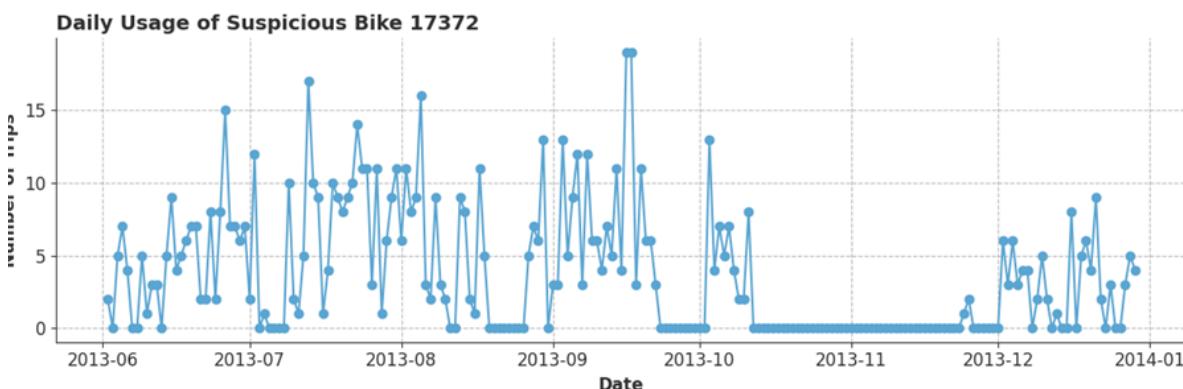
تحلیل نمودارهای روش اول:

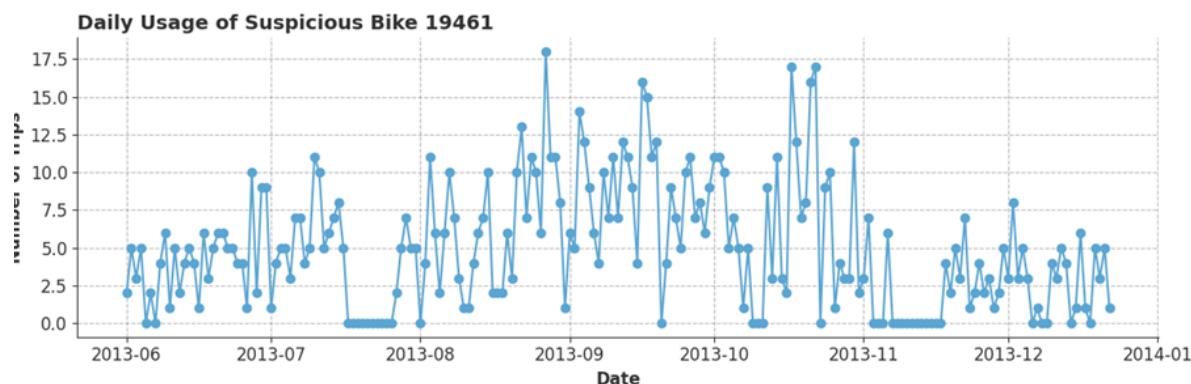
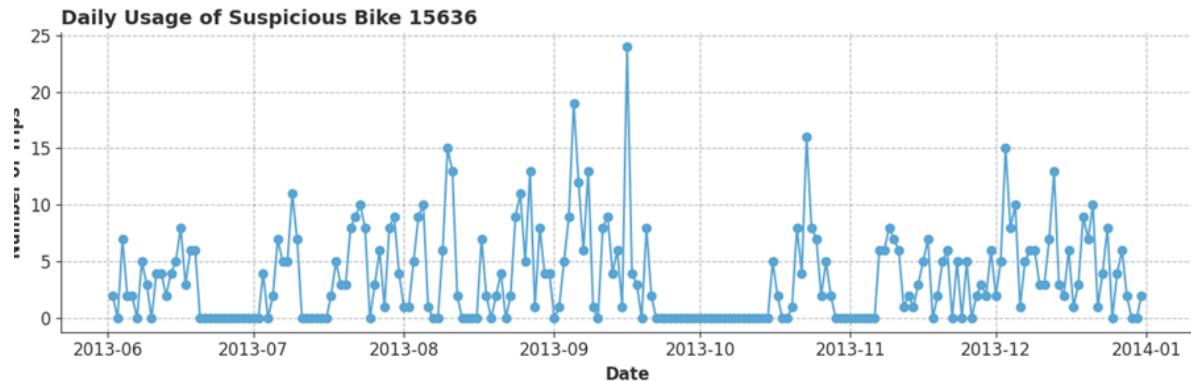
پلات‌های ارائه شده برای روش ۱، بینش‌های عمیقی در مورد ماهیت عدم فعالیت دوچرخه‌ها ارائه می‌دهند:

- نمودار الگوی مصرف چند نمونه دوچرخه مشکوک به خرابی

این نمودارها الگوی استفاده روزانه دوچرخه‌هایی را نشان می‌دهند که حداقل یک دوره 7 روزه عدم فعالیت داشته‌اند.

تحلیل: مشاهده می‌شود که دوچرخه‌های نمونه، دوره‌های طولانی (خطوط افقی در صفر) عدم فعالیت دارند. برخی از آن‌ها ممکن است پس از یک دوره عدم فعالیت طولانی دوباره فعال شده باشند، در حالی که برخی دیگر ممکن است پس از شروع عدم فعالیت، دیگر فعالیتی ثبت نکرده باشند (نشان‌دهنده خروج دائمی از سرویس یا خرابی کامل). این پلات‌ها تایید می‌کنند که این معیار به درستی دوچرخه‌های دارای سکون قابل توجه را شناسایی کرده است. این نمودارها برای بررسی موردي و تایید بصري «رفتار مشکوک» بسیار مفید هستند.

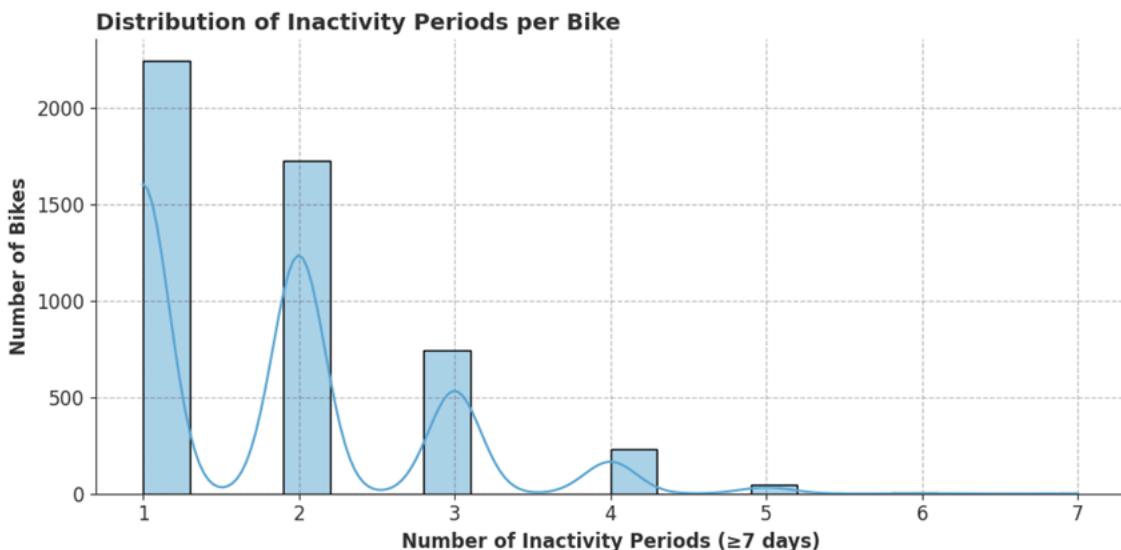




شکل (۱۰-۳) الگوی مصرف چند دوچرخه مشکوک به خرابی در روش اول

نمودار توزیع تعداد دوره‌های سکون در کل ناوگان

تحلیل: این هیستوگرام توزیع دوره‌های سکون را در بین ۴۹۸۷ دوچرخه مشکوک اولیه نشان می‌دهد. پیک نمودار به وضوح در عدد ۱ قرار دارد، به این معنی که اکثر دوچرخه‌های شناسایی شده، تنها یک دوره عدم فعالیت (با طول حداقل ۷ روز) داشته‌اند. این نشان می‌دهد که برای بسیاری از این دوچرخه‌ها، مشکل عدم فعالیت ممکن است یکبار و برای یک دوره طولانی رخ داده باشد. با این حال، انتهای نمودار به سمت راست کشیده شده است، که نشان‌دهنده وجود تعداد کمی دوچرخه با ۲، ۳، یا حتی بیش از ۱۰ دوره عدم فعالیت است. این دوچرخه‌ها (با دوره‌های متعدد) کاندیداهای خوبی برای شناسایی به عنوان «دوچرخه‌های مشکل‌ساز مزمن» هستند.



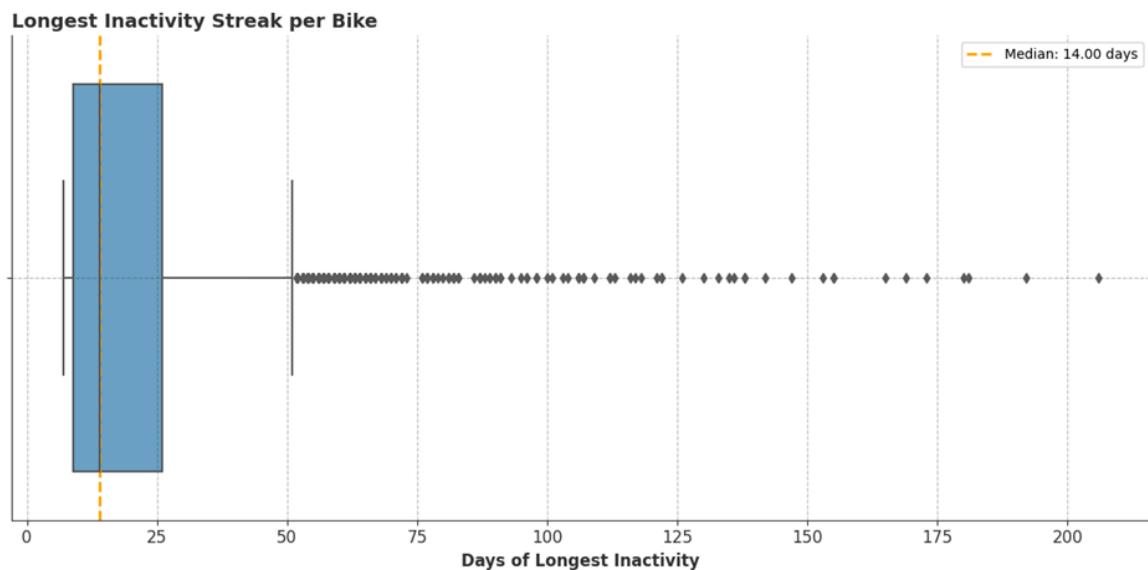
شکل (۱۱-۳) توزیع تعداد دوره‌های سکون در کل ناوگان

• نمودار طولانی‌ترین دوره سکون برای هر دوچرخه

این باکس‌پلات توزیع طولانی‌ترین دوره عدم فعالیت متوالی را برای دوچرخه‌های مشکوک را نشان می‌دهد. میانه این توزیع ۱۴ روز است، و بر اساس نتایج، مقادیری مانند ۱۳ و ۱۴ روز برای طولانی‌ترین دوره سکون در بین دوچرخه‌های مشکوک رایج است.

این یافته به ما کمک می‌کند تا آستانه اولیه ۷ روز را مورد ارزیابی مجدد قرار دهیم. از آنجایی که میانه توزیع در ۱۴ روز قرار دارد، به این معنی است که نیمی از دوچرخه‌هایی که ما مشکوک تشخیص داده‌ایم، حداقل ۱۴ روز یا بیشتر غیرفعال بوده‌اند. این نشان می‌دهد که آستانه اولیه ۷ روز، برای شناسایی "بسیاری" از دوچرخه‌های با مشکل عدم فعالیت کافی است، اما برای هدف قرار دادن دوچرخه‌هایی با "مشکلات واقعاً جدی‌تر و طولانی‌تر"، افزایش آستانه به ۱۴ روز یا بیشتر (مثلاً ۱۵ روز) می‌تواند منطقی‌تر باشد. این کار کمک می‌کند تا تمرکز بر دوچرخه‌هایی باشد که به طور معناداری برای مدت طولانی غیرفعال بوده‌اند و احتمالاً نیازمند تعمیرات اساسی‌تری هستند.

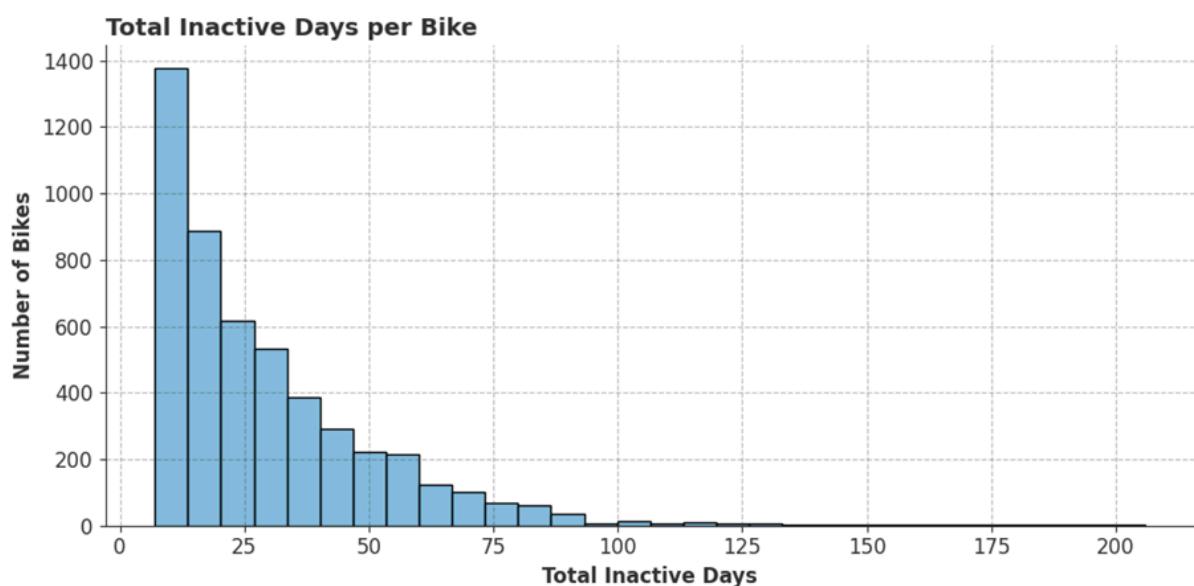
وجود مقادیر پرت بسیار دور از باکس همچنان نشان می‌دهد که برخی دوچرخه‌ها برای دوره‌های بسیار طولانی (چندین ماه) غیرفعال بوده‌اند. این دوچرخه‌ها، صرف نظر از میانه توزیع، همچنان از بالاترین اولویت برای بازرسی و تعمیر برخوردارند، زیرا به وضوح نشانه‌های خرابی یا فراموشی کامل را بروز می‌دهند و بیشترین ضرر مالی را به شرکت وارد می‌کنند.



شکل (۱۲-۳) طولانی ترین دوره سکون هر دوچرخه

نمودار مجموع روزهای سکون برای هر دوچرخه

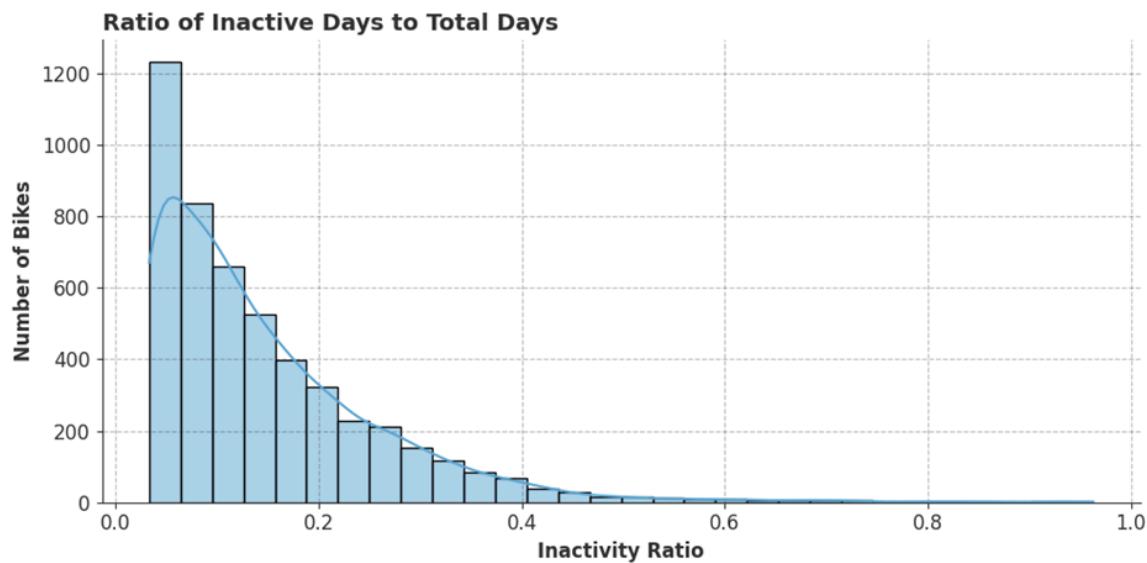
تحلیل: این هیستوگرام توزیع مجموع روزهای غیرفعال را برای هر دوچرخه نشان می‌دهد. نمودار به سمت مقادیر پایین کج شده است، به این معنی که بسیاری از دوچرخه‌ها مجموعاً روزهای غیرفعالی کمتری دارند. با این حال، انتهای طولانی به سمت راست نشان می‌دهد که تعداد قابل توجهی دوچرخه وجود دارد که مجموعاً روزهای بسیار زیادی را غیرفعال بوده‌اند (شاید دهها یا حتی صدها روز). این دوچرخه‌ها "سربار" ناوگان هستند و نشان می‌دهند که سرمایه‌گذاری روی آن‌ها بهینه نبوده است. شناسایی این دوچرخه‌ها برای تصمیم‌گیری در مورد تعمیرات اساسی، خروج از سرویس یا جایگزینی اهمیت دارد.



شکل (۱۳-۳) مجموع روزهای سکون برای هر دوچرخه

• نمودار نسبت روزهای سکون به کل روزها

تحلیل: این هیستوگرام نسبت روزهای غیرفعال به کل روزهای در ناوگان بوده است را نشان می‌دهد. این یک معیار قدرتمند برای درک کارایی عملیاتی دوچرخه است. پیک در سمت چپ (نسبت‌های پایین) قرار دارد، به این معنی که اکثر دوچرخه‌ها فعال بوده‌اند. با این حال، دمی به سمت راست وجود دارد که دوچرخه‌هایی با نسبت بالای عدم فعالیت را نشان می‌دهد. دوچرخه‌هایی با نسبت عدم فعالیت بالا، حتی اگر همیشه در یک دوره طولانی منفعل نباشند، نشان می‌دهند که به طور مزمن از دسترس خارج بوده‌اند. اینها کاندیداهای خوبی برای بررسی کامل یا حتی خارج کردن از ناوگان هستند. با تعریف یک آستانه می‌توان گروهی از دوچرخه‌های کمبازده را شناسایی کرد.



شکل (۱۴-۳) مجموع روزهای سکون برای هر دوچرخه

نتایج روش ۱:

روش ۱ با موفقیت ۴۹۸۷ دوچرخه را با دوره‌های عدم فعالیت قابل توجه شناسایی کرده است. این دوچرخه‌ها به گروه‌های زیر تقسیم می‌شوند که هر کدام نیاز به رویکرد خاصی دارند:

- **دوچرخه‌های با عدم فعالیت طولانی و یکباره:** ممکن است نیاز به تعمیرات اساسی یا بازبینی پس از یک خرابی جدی داشته باشند.
- **دوچرخه‌های با عدم فعالیت مکرر:** نشان‌دهنده مشکلات مزمن یا کیفیت پایین تعمیرات هستند و نیاز به بازبینی سیستماتیک‌تری دارند.
- **دوچرخه‌های با نسبت بالای عدم فعالیت:** نشان‌دهنده کارایی پایین کلی و نیاز به تصمیم‌گیری در مورد آینده آنها در ناوگان.

روش ۲: تحلیل ناهنجاری در مدت زمان سفر دوچرخه‌ها

این روش بر این فرض استوار است که دوچرخه‌هایی که درصد غیرعادی بالایی از سفرهای بسیار کوتاه دارند، ممکن است دچار مشکل باشند. سفرهای کوتاه غیرعادی می‌توانند ناشی از خرابی دوچرخه در حین سفر (کاربر مجبور به بازگشت سریع می‌شود)، مشکل در مکانیسم قفل/باز کردن، یا انجام «سفرهای آزمایشی» توسط کاربران یا تکنسین‌ها برای بررسی عملکرد دوچرخه باشد.

متدولوژی و انتخاب آستانه‌ها:

• آستانه ۵٪ برای مدت زمان سفر کوتاه

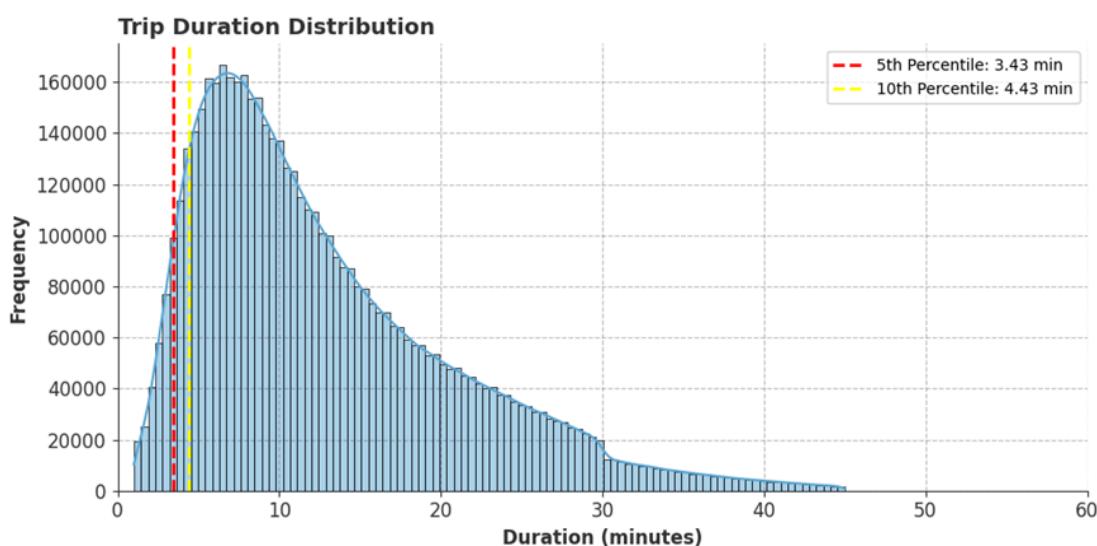
علت انتخاب: این آستانه‌ها برای شناسایی «سفرهای بسیار کوتاه» استفاده می‌شوند. ۵٪ پایین‌ترین مدت زمان سفر معمولاً شامل سفرهایی می‌شود که به دلایل غیرعادی کوتاه هستند (مثلاً کمتر از ۳-۲ دقیقه).

تحلیل نمودارهای روش دوم:

نمودار توزیع مدت زمان سفرها

این نمودار توزیع مدت زمان سفر را نشان می‌دهد. می‌بینیم که توزیع دارای یک دم بلند به سمت راست است (سفرهای طولانی‌تر)، اما یک پیک مشخص در سمت چپ (سفرهای کوتاه) نیز وجود دارد.

خط قرمز (آستانه ۵٪) نشان می‌دهند که چه مدت زمان‌هایی به عنوان «بسیار کوتاه» در نظر گرفته شده‌اند. این خطوط عمودی در ناحیه ابتدایی توزیع (احتمالاً در حدود 3.5 دقیقه) قرار می‌گیرند. مقایسه آستانه ۵٪ و ۱۰٪ به صورت تجربی و با بررسی توزیع سفرهای واقعی انجام شد تا مطمئن شویم که سفرهای «عادی کوتاه» (مثلاً جابجایی در داخل یک بلوک یا بازگشت فوری به دلیل مقصد اشتباه) را از سفرهای «غیرعادی کوتاه» (ناشی از مشکل) تفکیک کنیم. آستانه ۵٪ معمولاً محافظه‌کارانه‌تر و برای شناسایی موارد شدیدتر است.



شکل (۱۵-۳) توزیع مدت زمان سفرها و مقایسه آستانه ۵٪ و ۱۰٪ برای شناسایی سفرهای کوتاه

• آستانه پویا (90th percentile) برای تفکیک سفرهای کوتاه

علت انتخاب: به جای انتخاب یک آستانه ثابت برای نسبت سفرهای کوتاه، استفاده از صدک ۹۰ام یک رویکرد تطبیقی‌تر است. این رویکرد به طور خودکار ۱۰٪ از دوچرخه‌ها را که بالاترین نسبت سفرهای کوتاه را دارند، شناسایی می‌کند، فارغ از اینکه این نسبت دقیقاً چه عددی باشد.

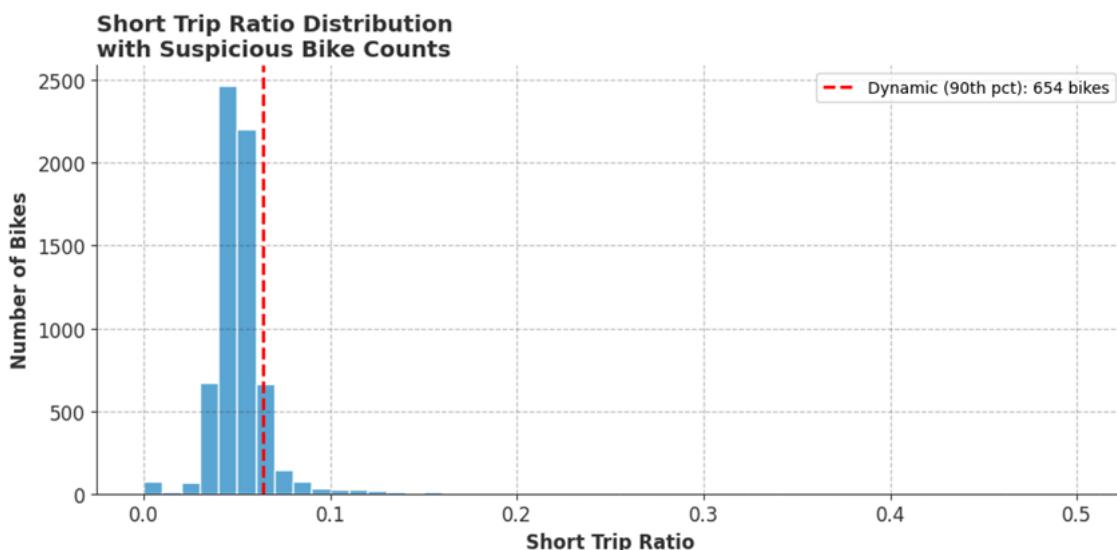
در این روش ۶۵۴ دوچرخه خراب شناسایی شده است.

تحلیل نمودارهای روش دوم:

• نمودار توزیع نسبت سفرهای کوتاه

این هیستوگرام توزیع نسبت سفرهای کوتاه را برای تمام دوچرخه‌ها نشان می‌دهد. پیک نمودار در سمت چپ (نسبت‌های پایین) قرار دارد، به این معنی که بیشتر دوچرخه‌ها نسبت کمی از سفرهای کوتاه دارند.

خط عمودی قرمز صدک ۹۰ام این توزیع را نشان می‌دهد. دوچرخه‌هایی که در سمت راست این خط قرار می‌گیرند، همان ۶۵۴ دوچرخه مشکوک هستند. این نمودار به وضوح نشان می‌دهد که این دوچرخه‌ها جزو ۱۰٪ با بالاترین نسبت سفرهای کوتاه در کل ناوگان هستند.



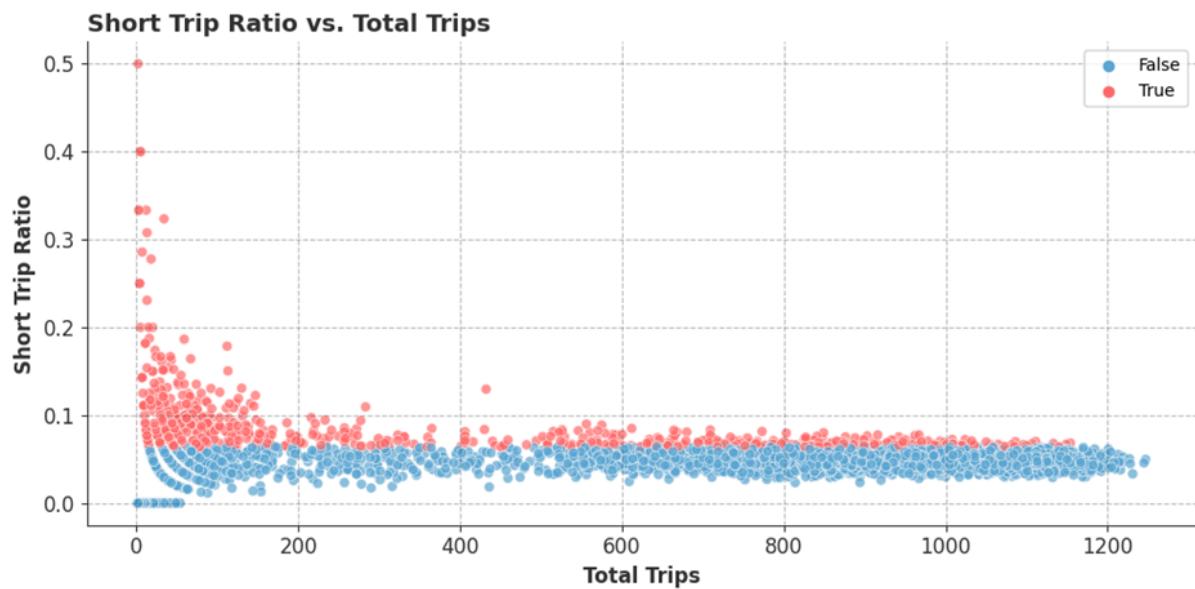
شکل (۱۶-۳) توزیع نسبت سفرهای کوتاه و آستانه پویای صدک ۹۰ام

• نمودار پراکندگی نسبت سفرهای کوتاه در مقابل تعداد کل سفرها

این نمودار رابطه بین تعداد کل سفرها و نسبت سفرهای کوتاه را نشان می‌دهد. نقاط قرمز (دوچرخه‌های مشکوک) عمدها در بالای نمودار (نسبت بالای سفرهای کوتاه) قرار دارند، که نشان‌دهنده عملکرد صحیح آستانه پویا است. این نمودار همچنین می‌تواند الگوهای جالبی را آشکار کند:

دوچرخه‌های با سفرهای کم و نسبت بالای سفرهای کوتاه: ممکن است دوچرخه‌هایی باشند که به تازگی خراب شده‌اند و هنوز سفرهای زیادی را ثبت نکرده‌اند.

دوچرخه‌های با سفرهای زیاد و نسبت بالای سفرهای کوتاه: اینها دوچرخه‌هایی هستند که به طور مکرر مورد استفاده قرار می‌گیرند، اما به نظر می‌رسد به طور مداوم دچار مشکل می‌شوند و نیاز به بررسی فوری دارند.

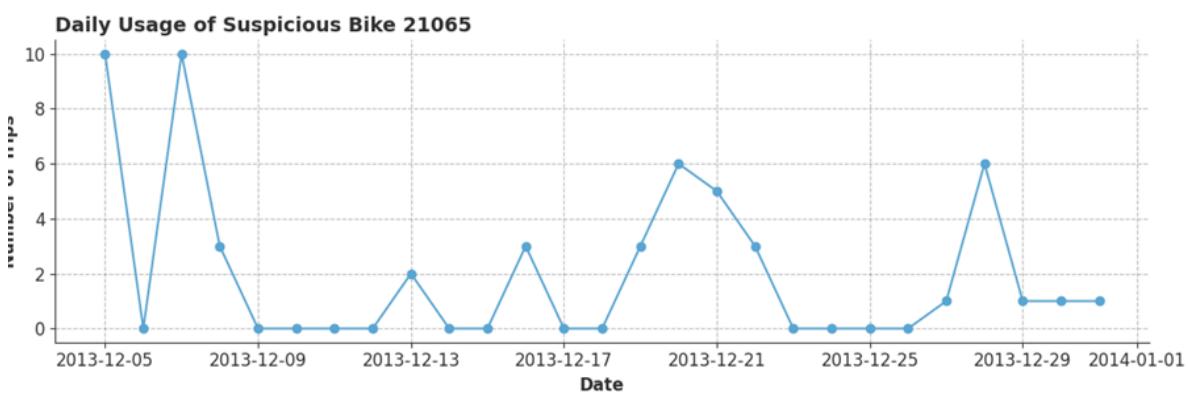


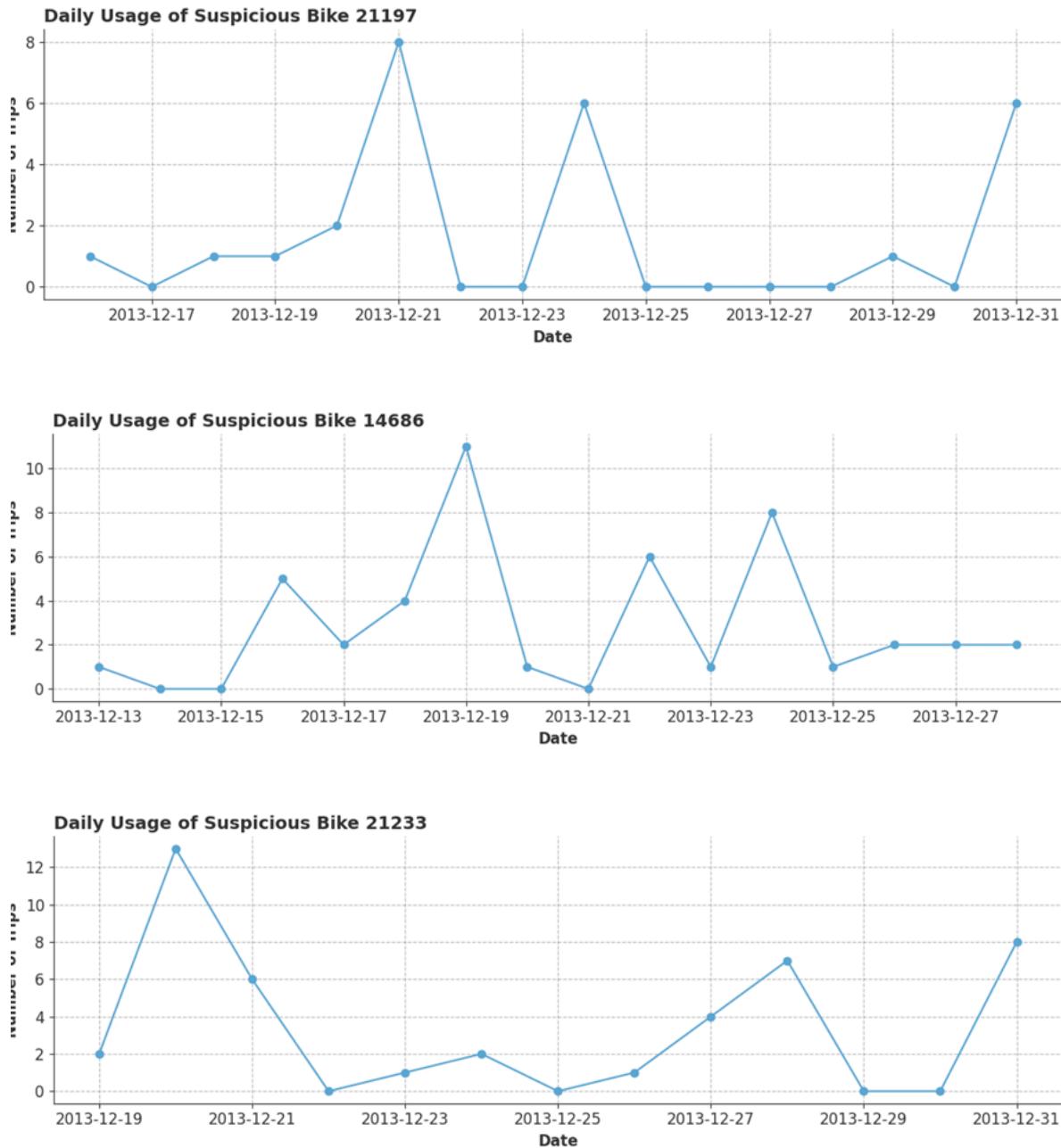
شکل (۱۷-۳) پراکندگی سفرهای کوتاه در مقابل کل سفرها

نمودار الگوی مصرف چند نمونه دوچرخه مشکوک به خرابی در روش دوم

تحلیل: این نمودارها نشان‌دهنده دوره‌هایی از فعالیت با تعداد سفرهای روزانه متفاوت و همچنان دوره‌هایی از عدم فعالیت (خطوط افقی در صفر) است.

این دوچرخه توسط روش ۲ شناسایی شده و به این معنی است که حتی در روزهایی که فعال بوده و سفر ثبت کرده، درصد بالایی از این سفرها بسیار کوتاه بوده‌اند. این نمودار نشان می‌دهد که دوچرخه، حتی با وجود مشکل سفرهای کوتاه، همچنان در دوره‌هایی از زمان در حال فعالیت بوده است. دوره‌های طولانی عدم فعالیت در این نمودار، نشان‌دهنده این است که این دوچرخه دارای مشکل ترکیبی است. هم مشکل فنی (که منجر به سفرهای کوتاه می‌شود) و هم مشکل عدم در دسترس بودن/خرابی طولانی‌مدت.





شکل (۱۸-۳) الگوی مصرف چند دوچرخه مشکوک به خرابی در روش دوم

نتایج روش ۲:

روش ۲ با شناسایی ۶۵۴ دوچرخه که درصد بالایی از سفرهای کوتاه را دارند، رویکرد مکمل مهمی را ارائه می‌دهد. این دوچرخه‌ها ممکن است به دلیل موارد زیر دچار مشکل باشند:

- خرابی‌های جزئی که منجر به لغو سفر یا بازگشت فوری می‌شود.
- مشکلات در قفل/باز کردن که باعث شروع و پایان سریع سفرهای سفرهای می‌شود.
- نیاز به نگهداری پیشگیرانه برای جلوگیری از خرابی‌های بزرگتر.

۲.۲.۳ روش‌های پیش‌بینی‌کننده (رویکرد مکمل)

این بخش به رویکردهای پیشرفت‌تری می‌پردازد که با استفاده از تحلیل روند و هموارسازی داده‌ها، قادر به شناسایی دوچرخه‌هایی هستند که هنوز به طور کامل خراب نشده‌اند اما نشانه‌هایی از افت عملکرد یا فرسودگی را بروز می‌دهند. این رویکرد به ویژه برای نگهداری پیشگیرانه حیاتی است.

افزودن ویژگی‌های پیشرفت‌هه:

پیش از اعمال این روش‌ها، چند ویژگی کلیدی به داده‌ها اضافه شد:

- شناسایی روزهای کاری و تعطیلات: با استفاده از کتابخانه `holidays`, روزهای غیرکاری (آخر هفته و تعطیلات رسمی) مشخص شدند. این امر از شناسایی نادرست دوچرخه‌ها به عنوان "غیرفعال" صرفاً به دلیل نبود تقاضا در روزهای غیرکاری جلوگیری می‌کند.
- هموارسازی با فیلتر کالمن (Kalman Filter)
 - فیلتر کالمن بر روی `this_day_avg_speed`, `this_day_trips`, `this_day_avg_time_usage` اعمال شد تا نویز و نوسانات کوتاه‌مدت کاهش یابد و روندهای اصلی پدیدار شوند (ایجاد ستون‌های `(trips_kfilter, time_usage_kfilter, speed_kfilter)`)
 - مزیت: این کار به شناسایی مشکلات بلندمدت و در حال توسعه کمک می‌کند.
 - محاسبه شیب (Slope) روندها:
 - شیب (نرخ تغییر) برای هر یک از ستون‌های فیلتر شده کالمن محاسبه شد (`(time_usage_slope, speed_slope)`)
 - مزیت: شیب منفی در این شاخص‌ها می‌تواند نشان‌دهنده شروع فرسودگی یا خرابی باشد، حتی قبل از اینکه دوچرخه کاملاً از کار بیفتد. این یک رویکرد پیشگیرانه (Proactive) است.

روش 3: شناسایی غیرمستقیم پیشرفت‌هه (Enhanced Passive Detection)

این روش، نسخه بهبود یافته‌ای از تحلیل عدم فعالیت است که با بهره‌گیری از ویژگی‌های پیشرفت‌هه، دقت شناسایی را افزایش می‌دهد.

متدولوژی

معیار اصلی: `this_day_trips == 0` (عدم فعالیت در روز مورد بررسی).

شرطی هوشمندانه برای تایید عدم فعالیت مشکوک:

- روز مورد بررسی باید یک روز کاری باشد (`is_working_day == True`).
- دوچرخه باید سابقه فعالیت قبلي داشته باشد (`total_previous_trips > 1`).
- مقدار هموار شده سفرهای دوچرخه (`trips_kfilter`) باید کمتر یا مساوی میانه `trips_kfilter` برای دوچرخه‌هایی باشد که در روزهای کاری غیرفعال بوده‌اند. این نشان می‌دهد که دوچرخه در یک روند کلی کمکاری قرار دارد، نه صرفاً یک افت لحظه‌ای.

- نتیجه: این روش به شناسایی روزهای خاصی که یک دوچرخه با رعایت شرایط هوشمندانه، مشکوک به غیرفعال بودن (یا کمکاری) در نظر گرفته می‌شود، کمک می‌کند و پایه‌ای برای تعریف شروع دوره‌های عدم فعالیت (inactive_start_day) فراهم می‌آورد.

نتایج روش ۳:

نتایج بدست آمده از تحلیل انجام شده بر ۷ ماه ثانویه سال ۲۰۱۳، منجر به شناسایی تعداد قابل توجهی از دوچرخه‌های غیرفعال شد که در جدول (۲-۳) قابل مشاهده است. این مقادیر نشان می‌دهد که با این روش می‌توان موارد غیرفعالی را به صورت ساده ولی مؤثر تشخیص داد.

جدول (۲-۳) تعداد دوچرخه‌های معیوب شناسایی شده در روش سوم (رویکرد پیش‌بینی‌کننده)

تعداد دوچرخه‌های غیرفعال	ماه
1373	Jun-13
1074	Jul-13
2941	Aug-13
1299	Sep-13
1224	Oct-13
1304	Nov-13
1791	Dec-13

روش ۴: شناسایی فعال (Active Detection)

این روش، قلب رویکرد پیش‌بینی‌کننده است و دوچرخه‌هایی را شناسایی می‌کند که علائم فرسودگی یا مشکلات در حال ظهور را نشان می‌دهند، حتی اگر هنوز کاملاً از کار نیفتاده باشند و ممکن است سفر نیز ثبت کنند.

متداول‌وزی

- معیارهای ترکیبی: این روش ترکیبی از چندین شاخص عملکردی روزانه (this_day_trips, trips_slope, time_usage_slope,) و روندهای آنها (this_day_avg_time_usage, this_day_avg_speed) را بررسی می‌کند.

- آستانه‌های پویا و معنادار: برای هر یک از 6 شاخص فوق، آستانه از میانه همان شاخص‌ها در روزهایی که دوچرخه‌ها توسط "روش غیرمستقیم پیشرفته" به عنوان غیرفعال/مشکوک شناسایی شده‌اند، تعیین می‌شود.

مثال: اگر `this_day_trips` یک دوچرخه فعال، کمتر یا مساوی میانه شیب دوچرخه‌های غیرفعال باشد، مشکوک تلقی می‌شود.

همچنین اگر شیب عملکرد (`trips_slope`) یک دوچرخه فعال، کمتر یا مساوی میانه شیب دوچرخه‌های غیرفعال باشد (یعنی روند نزولی مشابه)، مشکوک است.

شرابط اضافی:

- دوچرخه باید سابقه فعالیت قبلی داشته باشد (`total_previous_trips > 1`)
- هم روز فعال و هم روز بعدی باید روز کاری باشند (`next_is_working_day == 1` و `is_working_day == True`)
- این شرط، به خصوص `next_is_working_day`، تضمین می‌کند که افت عملکرد در یک زمینه عملیاتی (`True`) عادی رخ داده باشد.

هدف: شناسایی دوچرخه‌هایی که عملکرد ضعیف یا روند نزولی دارند (`suspicious_bicycles`) و نیاز به بازررسی پیشگیرانه دارند.

نتایج روش ۴:

در این تحلیل، هدف آن بود که بتوان دوچرخه‌هایی را که احتمال خرابی یا غیرفعال شدن در آینده نزدیک دارند، شناسایی کرد. این پیش‌بینی‌ها بر اساس روند کاهشی تدریجی در متغیرهای کلیدی (تعداد سفر، سرعت و زمان استفاده) و مشتقات آنها صورت گرفت. مدل طراحی‌شده توانست برای هر ماه، لیستی از دوچرخه‌های پربریسک را تولید کند. نتایج در جدول (۳-۳) قابل ملاحظه است.

جدول (۳-۳) تعداد دفعات خرابی شناسایی شده در روش سوم (رویکرد پیش‌بینی‌کننده)

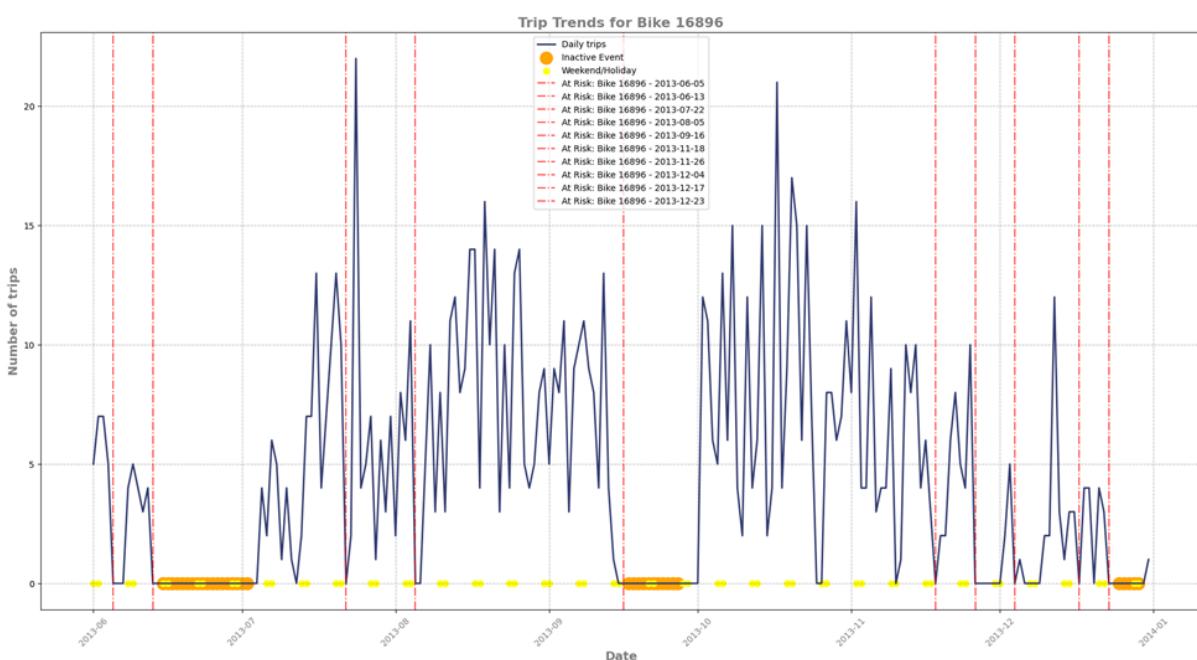
تعداد دفعات خرابی شناسایی شده	ماه
7229	Jun-13
5256	Jul-13
12598	Aug-13
8995	Sep-13

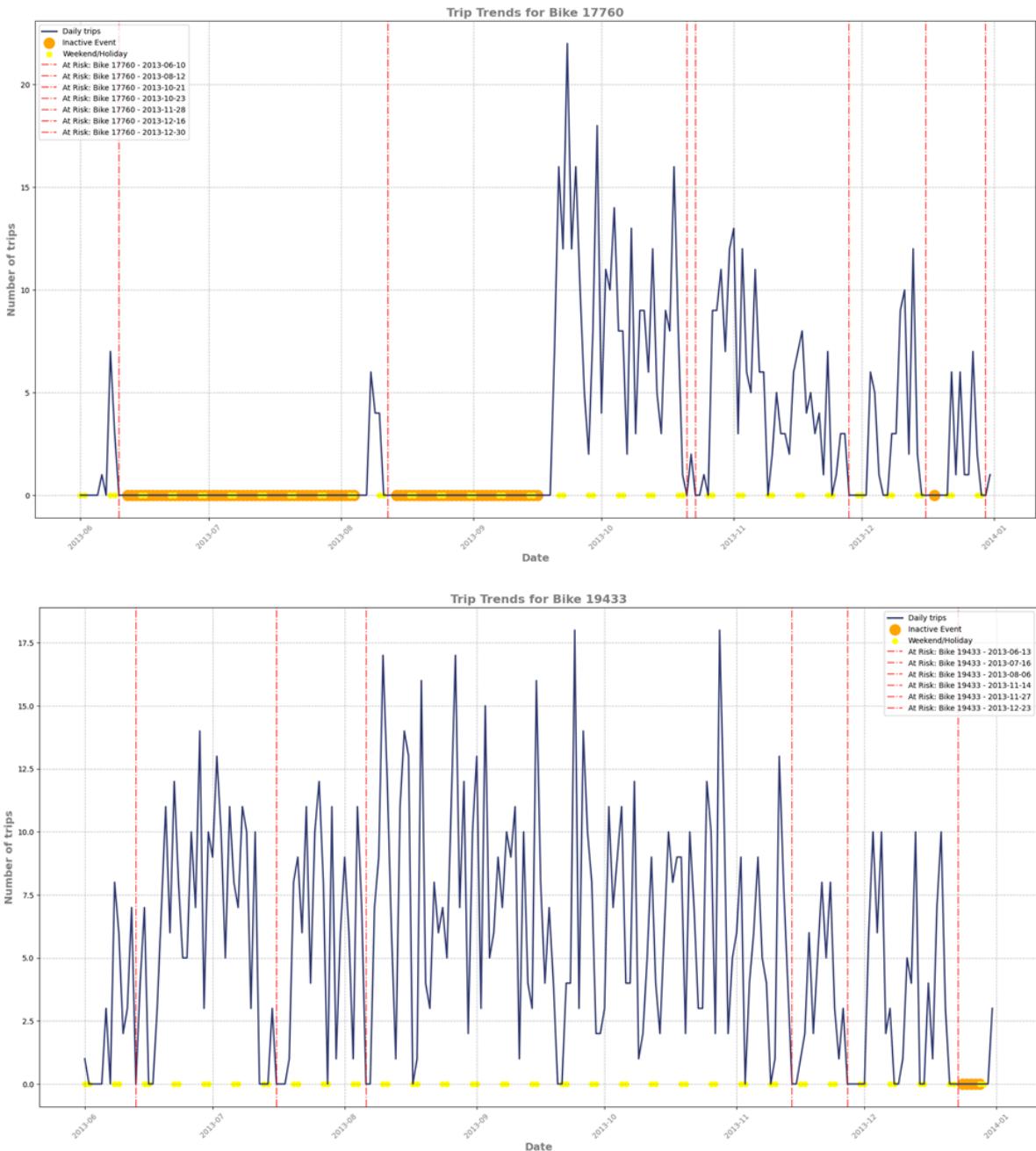
6272	Oct-13
7104	Nov-13
10050	Dec-13

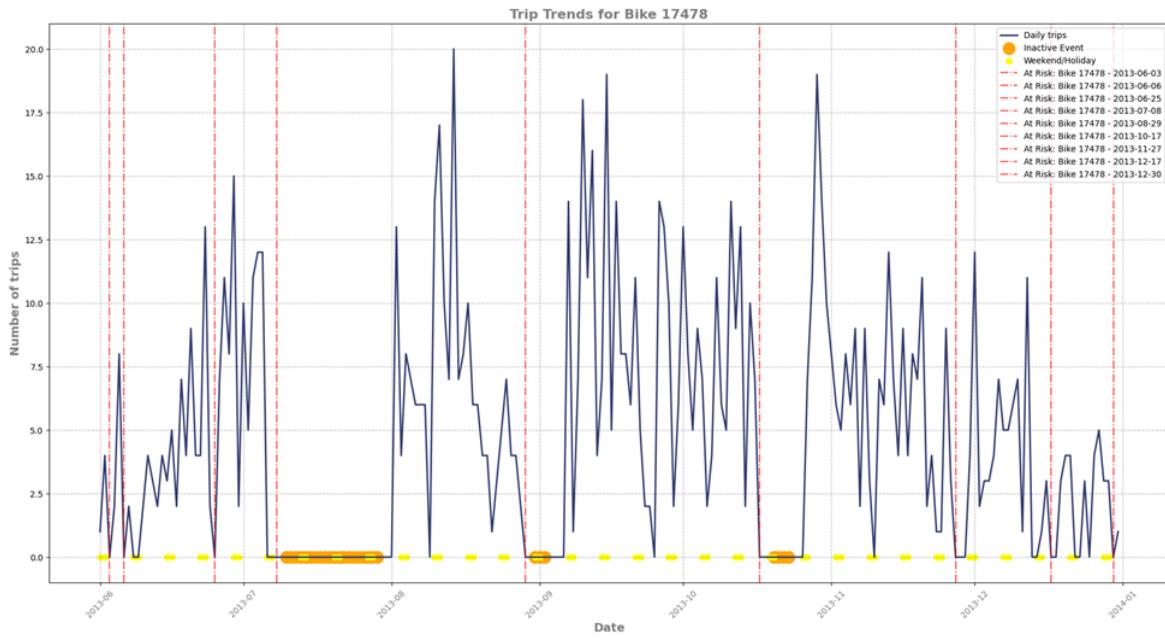
تحلیل بصری روش‌های پیشرفتی (نمودار ترکیبی)

نمودار روند سفر چند دوچرخه مشکوک در شکل (۳-؟؟) مشاهده می‌شود. این نمودار به وضوح قدرت روش‌های پیشرفتی را نشان می‌دهد:

- **خط آبی پرنگ:** سفرهای روزانه واقعی دوچرخه (Daily trips) (Daily trips)
- **دایره‌های نارنجی:** نشان‌دهنده رویدادهای عدم فعالیت (Inactive Event) که توسط روش غیرمستقیم پیشرفتی شناسایی شده‌اند. این نقاط روزهای بدون فعالیت در روزهای کاری هفته هستند که فیلتر کالم‌من نیز روند کمکاری را تأیید کرده است.
- **دایره‌های زرد:** نشان‌دهنده آخر هفته/تعطیلات (Weekend/Holiday). این به وضوح نشان می‌دهد که چرا فیلتر روزهای کاری مهم است؛ از علامت‌گذاری نادرست دوچرخه‌ها به عنوان "غیرفعال" در این روزها جلوگیری می‌کند.
- **خطوط عمودی قرمز نقطه‌چین:** نشان‌دهنده روزهایی که دوچرخه توسط روش فعلی به عنوان "در معرض خطر" (At Risk) شناسایی شده است. نکته کلیدی اینجاست که این خطوط قرمز ممکن است در روزهایی ظاهر شوند که خط آبی (سفرهای واقعی) هنوز بالای صفر است، اما شاخص‌های عملکردی یا شیب کمتر از آستانه قرار گرفته‌اند. این همان قابلیت پیش‌بینی‌کننده است که به شرکت اجازه می‌دهد قبل از خرابی کامل دوچرخه وارد عمل شود.







شکل (۱۹-۳) الگوی مصرف چند دوچرخه مشکوک به خرابی با استفاده از رویکرد پیش‌بینی‌کننده

محدودیت‌ها:

با اینکه مدل در شناسایی روندهای مشکوک موفق عمل کرده و تعداد قابل توجهی از دوچرخه‌ها را به عنوان پررسی‌سک علامت‌گذاری کرده است، اما با محدودیت در دسترس نبودن داده‌ی دقیق از وضعیت تعمیر یا خرابی واقعی دوچرخه‌ها در آن بازه‌ی زمانی مواجه است.

در نتیجه، امکان محاسبه‌ی دقیق معیارهای ارزیابی مدل وجود ندارد. با این حال، نمودارها و مطالعه‌های موردي نشان می‌دهند که:

- مدل می‌تواند الگوهای غیرفعال شدن را حتی پیش از شروع کامل بی‌تحرکی تشخیص دهد.
- مدل قادر است نوسانات و نویز ناشی از تعطیلات، شرایط خاص یا خطاهای ثبت داده را تا حد زیادی نادیده بگیرد.
- نمودار مشتق روندها (slope) برای دوچرخه‌های پررسیک معمولاً سیری نزولی پیوسته دارد که نشانه‌ای از افت عملکرد مستمر است.

با توجه به این شواهد، می‌توان نتیجه گرفت که مدل از لحاظ منطقی و بصری عملکرد رضایت‌بخشی دارد. در صورت پیاده‌سازی عملی این سیستم و مقایسه پیش‌بینی‌ها با داده‌های نگهداری واقعی، می‌توان دقت نهایی آن را اندازه‌گیری و بهبود داد.

۳.۲.۳ نتیجه‌گیری نهایی

با ادغام تحلیل‌های عمیق از رویکردهای مختلف، می‌توان یک سیستم شناسایی جامع برای دوچرخه‌های معیوب در ناوگان CitiBike ارائه داد.

روش‌های تشخیصی:

- **معیار عدم فعالیت:** در شناسایی دوچرخه‌های کاملاً غیرفعال یا "فراموش شده" بسیار مؤثر است. این روش به مشکلات جدی در دسترس نبودن فیزیکی یا لجستیکی دوچرخه اشاره می‌کند.
- **معیار سفرهای کوتاه:** در شناسایی دوچرخه‌هایی که فعال هستند اما کیفیت خدمات آنها پایین است (احتمالاً به دلیل مشکلات فنی جزئی) بسیار کارآمد است.

روش‌های پیشرفتی/پیش‌بینی‌کننده (مکمل):

- **شناسایی غیرمستقیم پیشرفتی:** نسخه بهبود یافته‌ای از تحلیل عدم فعالیت است که با لحاظ کردن روزهای کاری و روندهای هموار شده، دقت تشخیص را بالا می‌برد.
- **شناسایی فعال:** حیاتی‌ترین ابزار برای نگهداری پیشگیرانه است. این روش، دوچرخه‌هایی را شناسایی می‌کند که علائم اولیه فرسودگی یا مشکلات را نشان می‌دهند، حتی قبل از اینکه به طور کامل از کار بیفتد. این امکان مداخله زودهنگام و جلوگیری از خرابی‌های بزرگتر را فراهم می‌کند.

۴.۲.۳ پیشنهادات برای اجرا در سیستم واقعی

1. پیاده‌سازی یک سیستم هشدار ترکیبی: ایجاد یک داشبورد یا سیستم خودکار که نتایج هر چهار روش را ادغام کرده و یک "امتیاز خطر" برای هر دوچرخه ارائه دهد. این امتیاز می‌تواند بر اساس همپوشانی شناسایی‌ها (مثلًاً اگر یک دوچرخه توسط چند روش شناسایی شود، امتیاز بالاتری بگیرد).
2. اولویت‌بندی تعمیرات:
 - اولویت بالا: دوچرخه‌هایی که توسط هر سه روش شناسایی شده‌اند (هم غیرفعال، هم دارای سفرهای کوتاه، و هم دارای افت عملکرد فعال).
 - اولویت متوسط: دوچرخه‌هایی که توسط یک یا دو روش تشخیص (روش 1 یا 2) شناسایی شده‌اند.
 - اولویت پایین (پیشگیرانه): دوچرخه‌هایی که تنها توسط روش 4 (شناسایی فعال) شناسایی شده‌اند و نیاز به بازرسی زودهنگام دارند.
3. بهره‌گیری از داده‌های تعطیلات و روزهای کاری: حفظ منطق `is_working_day` در تمام تحلیل‌ها برای افزایش دقت و جلوگیری از شناسایی نادرست.
4. کالیبراسیون مداوم: آستانه‌ها می‌توانند با جمع‌آوری داده‌های بیشتر و بازخورد از تیم نگهداری (مثلًاً تأیید اینکه آیا دوچرخه‌های شناسایی شده واقعاً معیوب بوده‌اند یا خیر) به صورت مداوم کالیبره و بهینه شوند.
5. بررسی محدودیت‌ها:
 - سوابق تعمیرات: دسترسی به سوابق تعمیرات انجام شده می‌تواند به اعتبارسنجی مدل و درک بهتر انواع خرابی‌ها کمک کند.
 - متغیرهای محیطی: در نظر گرفتن عواملی مانند آب و هوا (بارندگی، دما) می‌تواند بینش‌های عمیق‌تری در مورد الگوهای استفاده و خرابی‌ها ارائه دهد.
 - اطلاعات موقعیتی: دانستن دقیق مکان دوچرخه در لحظه عدم فعالیت می‌تواند به تشخیص مشکلات لجستیکی کمک کند.

این رویکرد جامع، شرکت CitiBike را قادر می‌سازد تا مدیریت ناوگان خود را از یک مدل واکنشی به یک مدل پیشگیرانه و هوشمندانه ارتقا دهد، که در نهایت منجر به بهبود تجربه کاربری و بهینه‌سازی عملیات خواهد شد.

۳.۳ گسترش و نوسانات شرکت (سوال ۴)

۱.۳.۳ تشریح مسئله کسبوکار

با گذشت چندین سال از راهاندازی سامانه دوچرخه اشتراکی MiBici در شهر گوادالاخارا (مکزیک)، شرکت بهرهبردار تصمیم گرفته است وضعيت توسعه اين سامانه را از زمان تأسيس آن مورد بررسی قرار دهد. در همين راستا، مدیرعامل شرکت تيمی را مأمور کرده است که چگونگی بزرگ شدن سامانه از سال ۲۰۱۵ تا ۲۰۱۹ را بررسی کرده و گزارش جامعی در اين باره ارائه دهد. مدیر تيم از اعضای تيم تحليل داده خواسته است که با استفاده از داده‌هاي موجود، به دو پرسش اصلی پاسخ دهیم:

۱. آيا سفرهای انجام شده با این سامانه در ماههای تحت بررسی، بهنسبت یکنواخت در سطح شهر گسترش یافته‌اند؟ چگونه می‌توان از این اطلاعات برای کمک به تیم بازاریابی استفاده کرد؟
۲. نوساناتی که در تعداد سفرها مشاهده می‌شود به چه عواملی نسبت داده می‌شود و سهم هر یک از عوامل چقدر است؟

۲.۳.۳ ابزارها و فناوری‌های مورد استفاده

برای تحلیل این سامانه، از ترکیبی از زبان برنامه‌نویسی پایتون، پایگاه داده تحلیلی DuckDB و کتابخانه‌های تصویری و مکانی زیر استفاده شده است:

- duckdb: پایگاه داده درونحافظه‌ای برای اجرای سریع کوئری‌های SQL بر روی داده‌های CSV •
- pandas, numpy: برای پردازش و تحلیل داده‌های جدولی •
- matplotlib, seaborn: برای رسم نمودارهای آماری و تصویری •
- folium, pydeck: برای مصورسازی مکانی و تولید نقشه‌های تعاملی با لایه‌های سه‌بعدی •
- requests: برای دریافت داده‌های محیطی و مکانی از سرویس‌های خارجی •

۳.۳.۳ آماده‌سازی داده‌ها

در مراحل ابتدایی، سه مجموعه داده پیش‌پردازش شده به شرح زیر باگذاري شد:

- اطلاعات ایستگاهها: شامل شناسه، نام ایستگاه، مختصات جغرافیایی و ارتفاع از سطح دریا •
- اطلاعات سفرها: شامل زمان شروع و پایان سفر، شناسه کاربر، ایستگاه مبدا و مقصد، سال تولد و جنسیت •
- اطلاعات آب‌وهوایی: شامل میانگین دما، میزان بارندگی، نوع روز (عادی، تعطیل، آخر هفته) •

این داده‌ها از منابع داخلی سامانه و همچنین داده‌های خارجی تهیه شده‌اند. برای تکمیل تحلیل، از API‌های خارجی زیر نیز استفاده شد:

- Visual Crossing Weather API: برای دریافت داده‌های روزانه دما، بارندگی و شرایط جوی شهر گوادالاخارا در ۲۰۱۹ تا ۲۰۱۵ بازه •
- Open Elevation API: برای استخراج ارتفاع ایستگاه‌ها بر اساس مختصات آنها •

- برای جمعآوری مکان‌های عمومی مانند دانشگاه‌ها، مدارس، بیمارستان‌ها و ادارات در محدوده شهری
- برای تولید تقویم رسمی مکزیک و شناسایی تعطیلات ملی و محلی (مثل روز بنیاد گوادالاخارا و جمعه نیک)

سپس با استفاده از DuckDB، داده‌ها در قالب جداول تحلیل پذیر بارگذاری شده و عملیات‌هایی از جمله:

- استانداردسازی نام ستون‌ها
- فیلتر کردن سفرهای معتر (سفرهایی با ساعت و تاریخ صحیح)
- ایجاد جدول کاربران یکتا (unique_users) شامل تعداد سفر، سن کاربر، تاریخ اولین و آخرین سفر

انجام گرفت. این جدول‌ها پایه اصلی تحلیل‌های بعدی در این پروژه را تشکیل می‌دهند.

۴.۳.۳ سوال اول: گسترش یکسان سرویس در شهر

۱.۴.۳.۳ متداول‌وزی و معیارهای استفاده شده

برای ارزیابی توزیع فضایی سفرهای انجام‌شده با MiBici، از چهار محور تحلیلی استفاده شده است:

۱- تحلیل آماری نابرابری استفاده از ایستگاه‌ها

یکی از روش‌های کلاسیک برای سنجش تمرکز در توزیع، استفاده از منحنی لورنز (Lorenz Curve) و ضریب جینی (Gini Coefficient) است.

- منحنی لورنز میزان سهم تجمعی سفرها را نسبت به سهم تجمعی ایستگاه‌ها نمایش می‌دهد. اگر تمامی ایستگاه‌ها به یک اندازه استفاده می‌شوند، این منحنی به شکل یک خط ۴۵ درجه درمی‌آمد.
- ضریب جینی که از منحنی لورنز مشتق می‌شود، معیاری عددی برای سنجش نابرابری است. این ضریب بین ۰ (توزیع کاملاً یکنواخت) تا ۱ (تمرکز کامل در یک ایستگاه) قرار می‌گیرد.

در این مطالعه، ضریب جینی محاسبه شده برابر با ۰.۴۹۱ بوده است که نشان‌دهنده نابرابری متوسط به بالا در توزیع استفاده از ایستگاه‌هاست.

۲- تحلیل مکانی (Spatial Distribution Mapping)

با استفاده از کتابخانه‌های folium و pydeck، نقشه‌هایی تولید شده‌اند که ایستگاه‌ها را بر اساس تعداد سفر ثبت شده (مبدأ یا مقصد) نمایش می‌دهند. این نقشه‌ها دارای ویژگی‌های زیر هستند:

- سایز یا رنگ ایستگاه‌ها متناسب با تعداد سفرهای آن‌ها است.
- با استفاده از folium.HeatMap و pydeck.Layer، نواحی پرترکم شناسایی شدند.

این نقشه‌ها به‌وضوح نشان می‌دهند که اکثر سفرهای در نواحی مرکزی شهر، به‌ویژه اطراف دانشگاه‌ها، ایستگاه‌های حمل و نقل و مراکز اداری متمرکز شده‌اند.

۳- بررسی خوشه‌های مکانی و ایستگاه‌های پرتردد

ایستگاه‌ها بر اساس میزان استفاده به سه دسته تقسیم شده‌اند:

- ایستگاه‌های پرتردد (Top 20%)
- ایستگاه‌های میان‌رده
- ایستگاه‌های کم‌استفاده یا غیرفعال

تحلیل آماری نشان داد که کمتر از ۲۰٪ ایستگاه‌ها مسئول بیش از ۵۰٪ سفرها هستند، در حالی‌که تعداد قابل‌توجهی از ایستگاه‌ها کمتر از ۱۰ سفر در ماه داشته‌اند.

۴- تحلیل مجاورت ایستگاه‌ها با مکان‌های عمومی (POI)

برای بررسی عوامل مؤثر بر تمرکز سفرها در برخی نواحی، فاصله ایستگاه‌ها تا مکان‌های عمومی مانند مدارس، دانشگاه‌ها، دفاتر دولتی محاسبه شد.

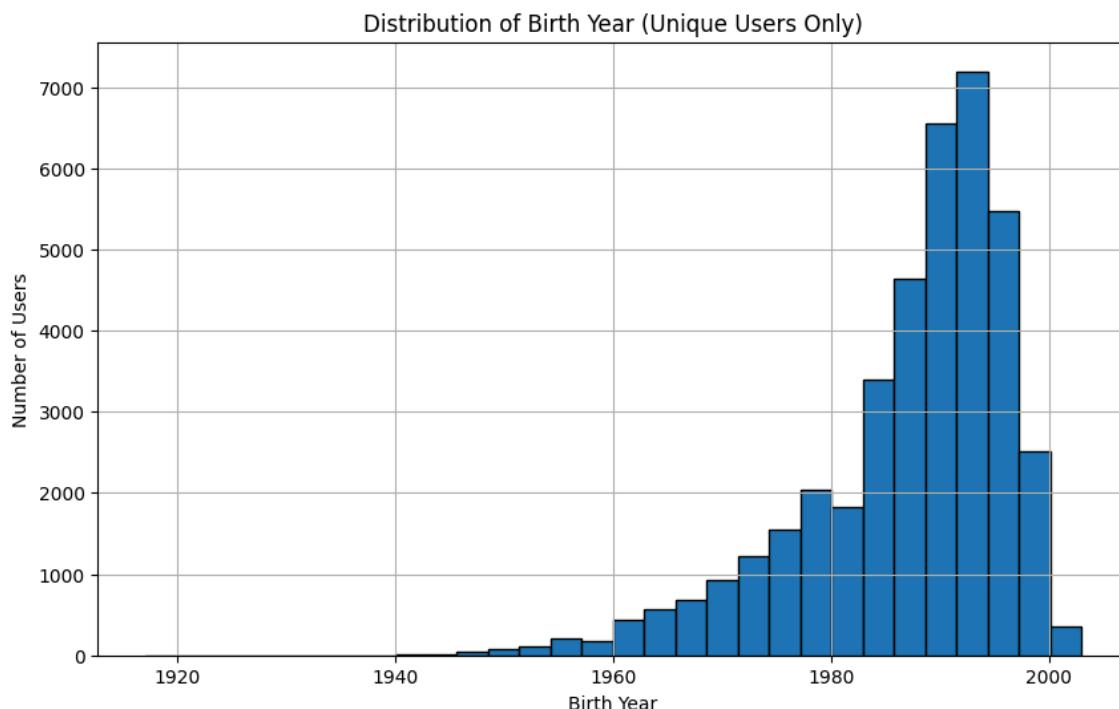
- ایستگاه‌هایی که در شعاع ۳۰۰ متر از یک دانشگاه یا مرکز مهم قرار داشتند، به‌طور میانگین دو برابر بیشتر از سایر ایستگاه‌ها استفاده شده‌اند.
- داده‌های مکانی با استفاده از API OpenStreetMap Overpass به صورت مستقیم استخراج شدند.

ابزار / معیار	هدف	محور
منحنی لورنز، ضریب جینی	سنجدش تمرکز استفاده از ایستگاه‌ها	نابرابری آماری
Folium، Pydeck	نمایش توزیع فضایی سفرها	تصویرسازی مکانی
دسته‌بندی بر حسب تعداد سفر	شناسایی خوشه‌ها	طبقه‌بندی ایستگاه‌ها
OpenStreetMap API، Factual مکانی	بررسی عوامل بیرونی مؤثر	تحلیل مجاورت با POI

جدول(۴-۳) متداول‌ترین و معیارهای استفاده شده

۲.۴.۳.۳ یافته‌ها و بینش‌ها

۱- ترکیب سنی کاربران MiBici



شکل (۲۰-۳) توزیع تعداد سفر ها بر حسب سال تولد کاربران

نمودار توزیع سال تولد کاربران (بر اساس داده های کاربران یکتا) نشان می دهد که:

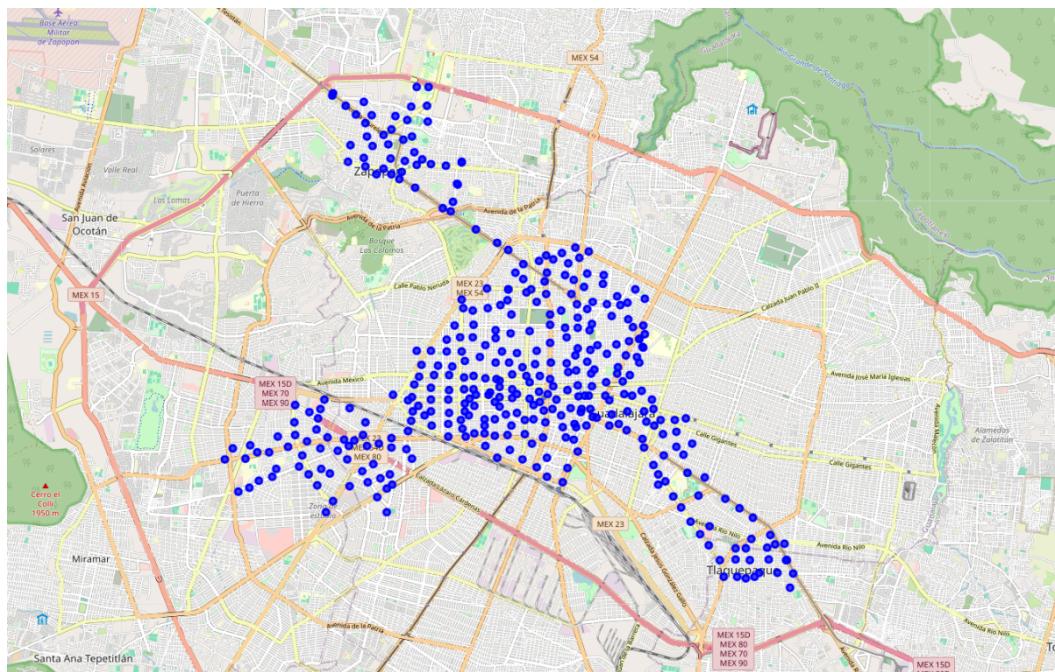
- بیشترین تراکم کاربران متولدین بین سال های ۱۹۸۰ تا ۲۰۰۰ هستند، به عبارتی بخش بزرگی از کاربران در محدوده سنی ۲۰ تا ۴۰ سال قرار دارند.
- یک پیک قابل توجه بین سال های ۱۹۹۵ تا ۱۹۹۰ مشاهده می شود که به احتمال زیاد نشان دهنده مشارکت بالای دانشجویان دانشگاه ها و جوانان شاغل است.
- کاربران متولد قبل از ۱۹۶۰ کمتر از ۲٪ جامعه کاربران را تشکیل می دهند، که نشان دهنده استقبال بسیار پایین از سوی افراد مسن تر است.

تحلیل بازاریابی:

- کمپین های تبلیغاتی و اطلاع رسانی باید بر گروه سنی ۱۸ تا ۳۵ سال متمرکز باشد.
- دانشگاه ها، خوابگاه ها و مراکز آموزشی باید به عنوان نقاط کلیدی تبلیغاتی هدف گذاری شوند.
- می توان از همکاری با مؤسسات آموزشی، کارت های اشتراک ویژه برای دانشجویان، و تخفیف برای ثبت نام گروهی بهره گرفت.

۲- تحلیل مکانی ایستگاه های فعال

۱-۲ نقشه ایستگاه های فعال (تمام ایستگاه ها)



شکل (۲۱-۳) توزیع ایستگاه ها در شهر

شرح نقشه:

نقاط آبی روی نقشه نشان دهنده تمام ایستگاه های سامانه MiBici هستند

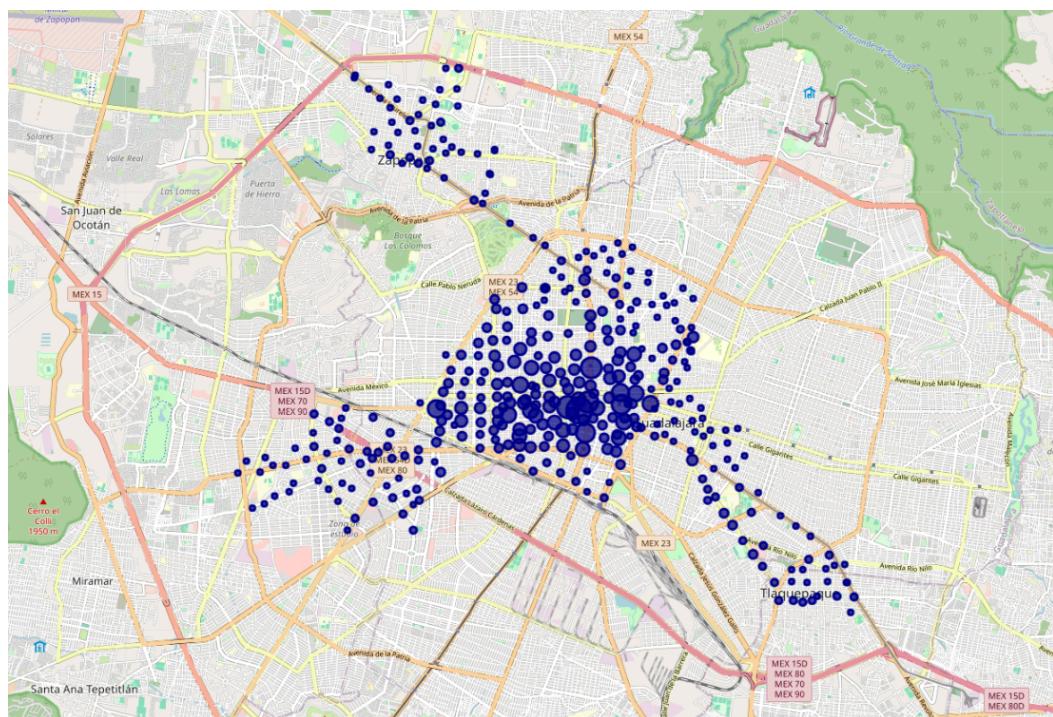
مشاهدات کلیدی:

- تمرکز بالای ایستگاه ها در مرکز شهر (به ویژه نواحی تاریخی و پرتردد)
- کاهش تراکم در نواحی شرقی و جنوبی شهر
- مناطق غربی و شمال غربی هنوز از تراکم کافی برخوردار نیستند

کاربرد مدیریتی:

این نقشه پایه تحلیل های پوشش مکانی است و نشان می دهد که توسعه زیرساختی تاکنون نامتقاض است.

۲-۲ نقشه ایستگاههای پرتردد به عنوان مبدا



شکل (۲۲-۳) توزیع ایستگاههای پرتردد به عنوان مبدا در شهر

شرح نقشه:

ایستگاهها با دایره‌هایی با اندازه متناسب با تعداد سفرهایی که از آن‌ها شروع شده‌اند نمایش داده شده‌اند.

مشاهدات کلیدی:

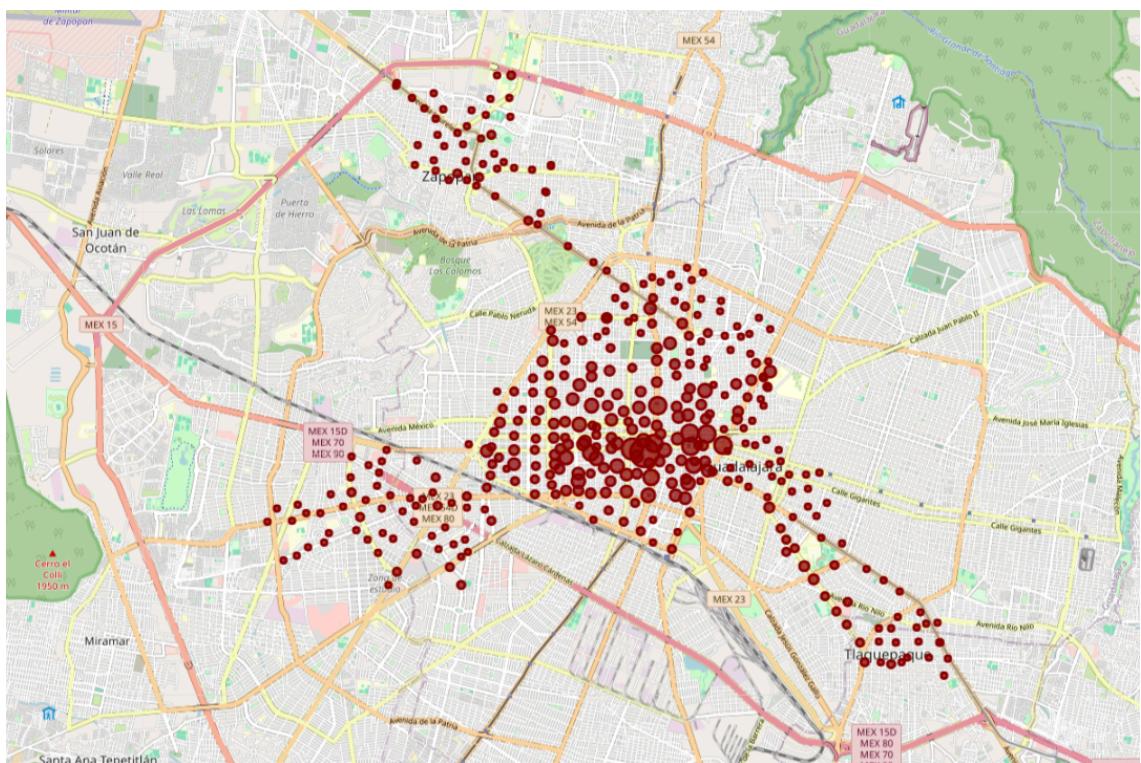
- بیشترین مبدأها در نواحی مرکزی و نزدیک به مراکز حمل و نقل عمومی و دانشگاهی
برخی ایستگاهها در مناطق حاشیه‌ای سفرهای بسیار کمی به عنوان مبدأ ثبت کرده‌اند

کاربرد مدیریتی:

این نقشه برای بهینه‌سازی استراتژی‌های rebalancing (جابجایی دوچرخه‌ها) بسیار مؤثر است، چون نقاط پُررفت‌وآمد قابل شناسایی‌اند.

۳-۲ نقشه ایستگاه‌های پرتردد به عنوان مقصد

شکل (۲۳-۳) توزیع ایستگاه‌های پرتردد به عنوان مقصد در شهر



شرح نقشه:

دایره‌ها نشان‌دهنده حجم سفرهایی هستند که در هر ایستگاه خاتمه یافته‌اند.

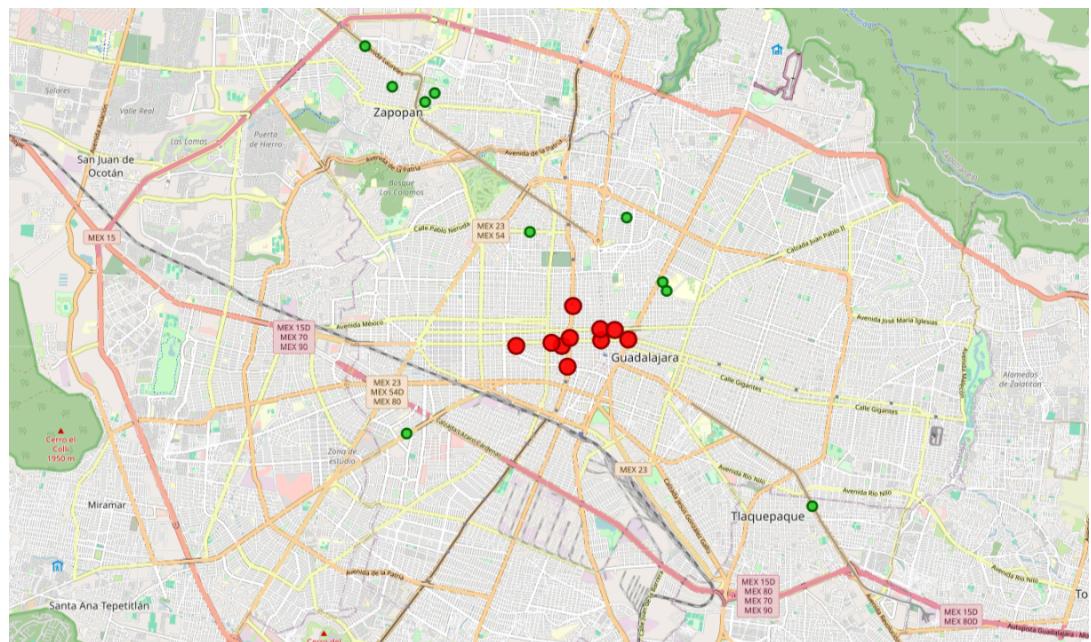
مشاهدات کلیدی:

- الگوی کلی مشابه نقشه مبدا است، اما برخی ایستگاه‌ها مقصد پرترکارتری نسبت به مبدأ دارند.
- نشان می‌دهد که برخی ایستگاه‌ها بیشتر برای دسترسی نهایی به مقصد هایی خاص (مانند ادارات یا مراکز تجاری) استفاده می‌شوند.

کاربرد مدیریتی:

درک تفاوت بین مبدأ و مقصد می‌تواند به طراحی مسیرها، جانمایی امکانات خدماتی و تبلیغاتی کمک کند.

۴-۲ نقشه ۱۰ ایستگاه برتر و ۱۰ ایستگاه کم استفاده



شکل (۲۴-۳) توزیع ۱۰ ایستگاه‌های پرتردد و ۱۰ ایستگاه کم استفاده

شرح نقشه:

- دایره‌های قرمز بزرگ: ۱۰ ایستگاه با بیشترین استفاده
- دایره‌های سبز کوچک: ۱۰ ایستگاه با کمترین استفاده

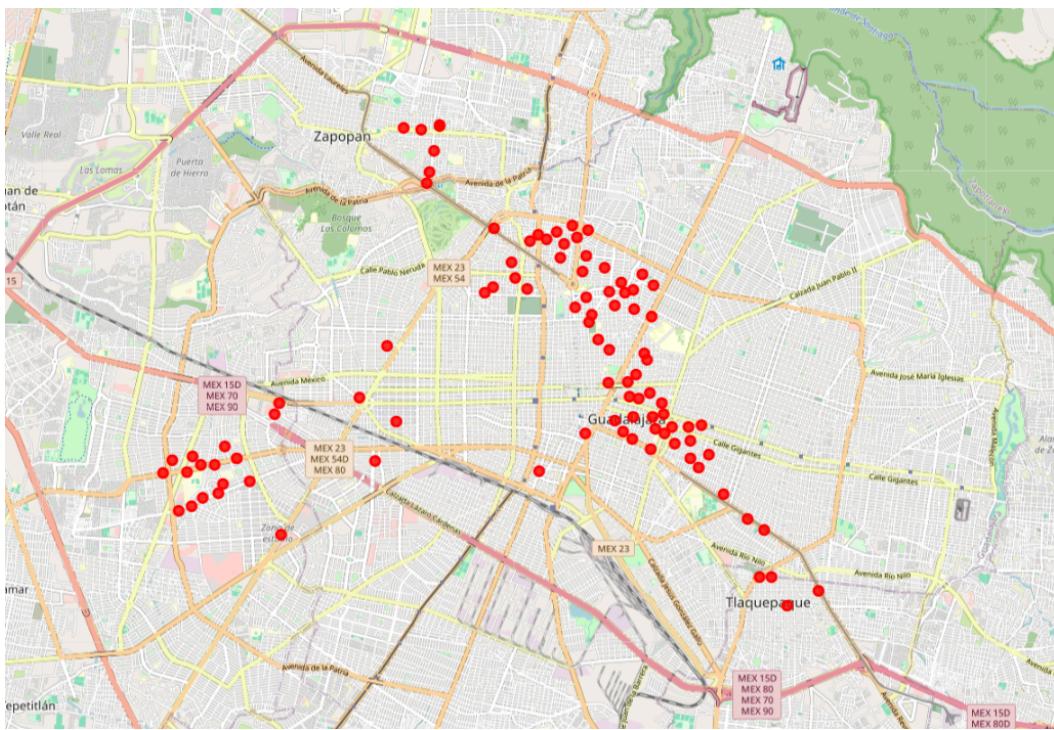
مشاهدات کلیدی:

- ایستگاه‌های پرتردد در مرکز شهر و نزدیکی به حمل و نقل عمومی و دانشگاه‌ها قرار دارند.
- ایستگاه‌های کم استفاده معمولاً در حاشیه یا نواحی کم جمعیت قرار گرفته‌اند.

کاربرد مدیریتی:

تحلیل اختلاف استفاده بین این دو گروه، بینشی حیاتی برای بازنگری در جانمایی ایستگاه‌ها و طراحی کمپین‌های بازاریابی محلی فراهم می‌آورد.

۵-۲ نقشه ایستگاههای ساخته شده بعد از بازه داده‌ها (بدون استفاده)



شکل (۲۶-۳) توزیع ۱۰ ایستگاههای شاخته شده بعد از بازه‌ی بررسی دیتا ها

شرح نقشه:

ایستگاههایی که در داده‌ها هیچ سفری (نه مبدأ، نه مقصد) ثبت نکرده‌اند.

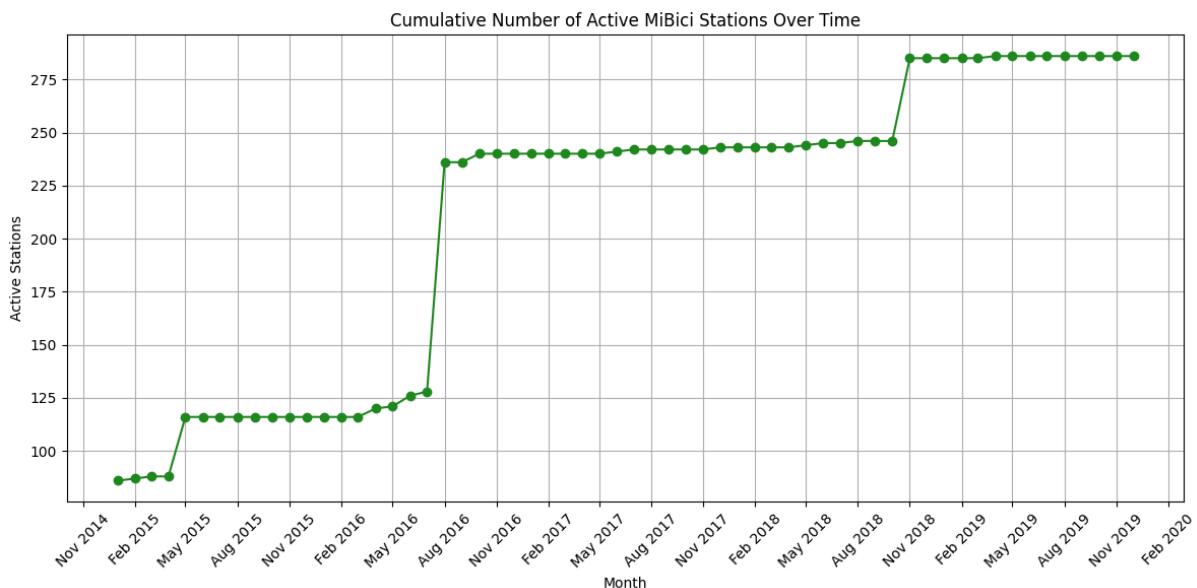
مشاهدات کلیدی:

- تمامی ایستگاههای بدون استفاده مربوط به زیرساختهای ساخته شده پس از سال ۲۰۱۹ هستند.
- این ایستگاهها به درستی در تحلیل زمانی وارد نشده‌اند اما باید در برنامه توسعه دیده شوند.

کاربرد مدیریتی:

ضروری است که داده‌های تازه‌تر نیز وارد تحلیل شوند تا وضعیت این ایستگاهها نیز پایش شود. همچنین، این نقشه نشان می‌دهد لزوم هماهنگی بین داده و اجرا تا چه اندازه مهم است.

۳- رشد تجمعی ایستگاههای MiBici در گذر زمان



شکل (۳-۲۷) نمودار رشد تعداد ایستگاه ها

این نمودار نشان‌دهنده‌ی رشد تجمعی تعداد ایستگاههای فعال سامانه MiBici در شهر گوادالاخارا از ابتدای سال ۲۰۱۵ تا پایان سال ۲۰۱۹ است. هر نقطه، تعداد کل ایستگاههایی را که تا آن زمان راهاندازی شده‌اند نشان می‌دهد.

نکات کلیدی نمودار:

۱. آغاز به کار (اواخر ۲۰۱۴):

سامانه با حدود ۸۶ ایستگاه اولیه آغاز به کار کرده و در ماه‌های ابتدایی سال ۲۰۱۵ ۲۰۱۵ روند نسبتاً ثابتی دارد.

۲. موج اول توسعه (میانه ۲۰۱۵):

در این مقطع، سیستم به بیش از ۱۱۵ ایستگاه گسترش می‌یابد؛ این مرحله به عنوان تثبیت هسته اولیه شبکه قابل تفسیر است.

۳. موج دوم توسعه (تابستان ۲۰۱۶):

بزرگ‌ترین جهش در این نمودار بین ژوئن تا آگوست ۲۰۱۶ دیده می‌شود. تعداد ایستگاهها در این دوره از حدود ۱۲۵ به بیش از ۲۳۵ ایستگاه می‌رسد.

این توسعه احتمالاً همراه با افزایش ظرفیت عملیاتی، ورود مناطق جدید به شبکه و پاسخ به تقاضای فزاینده بوده است.

۴. موج سوم توسعه (پاییز ۲۰۱۸):

آخرین جهش عمده در پاییز ۲۰۱۸ رخ داده که تعداد ایستگاهها را به بیش از ۲۸۰ عدد می‌رساند. پس از آن، شبکه نسبتاً ثابت باقی مانده است.

دلایل رشد و سیاست‌های توسعه‌ای

مطابق با مستندات عمومی، گسترش MiBici با حمایت نهادهای دولتی و محلی انجام شده است:

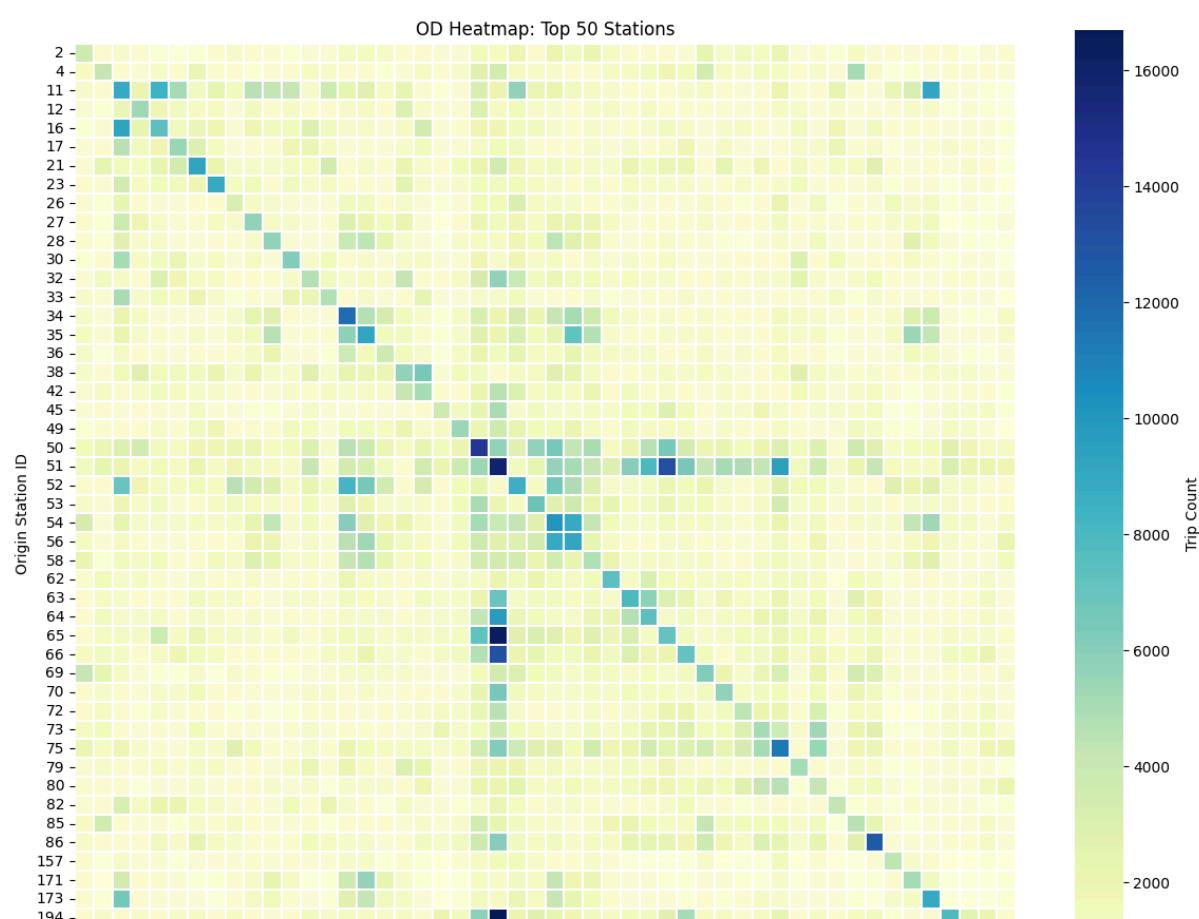
- نهادهایی نظیر دولت ایالت خالیسکو و مؤسسه برنامه‌ریزی منطقه کلان‌شهری گوادالاخارا (IMEPLAN) نقش کلیدی در تأمین بودجه و سیاست‌گذاری ایفا کردند.
- این توسعه بخشی از استراتژی حمل و نقل غیرمоторی شهری بوده که اهداف آن شامل:
 - کاهش ازدحام ترافیکی
 - کاهش انتشار گازهای گلخانه‌ای
 - کاهش واپسیگی به خودروهای شخصی
 - افزایش دسترسی شهروندان به حمل و نقل پاک و ارزان
- توسعه سال ۲۰۱۸ به ویژه با هدف اتصال مناطق حاشیه‌ای به شبکه مرکزی و ایجاد پیوستگی با خطوط حمل و نقل عمومی همزمان شده است.

منابع استنادی:

[در ویکی‌پدیا \(اسپانیایی\)](#)

[مکانیکوستی](#)

۴- نقشه حرارتی (Heatmap) مبدا-مقصد (OD) برای ۵۰ ایستگاه برتر



شکل (۳-۲۸) نمودار Heatmap مبدا-مقصد (OD) برای ۵۰ ایستگاه برتر

نمودار حرارتی فوق، حجم سفرها بین ۵۰ ایستگاه با بیشترین تعداد سفر آغاز شده (origin) را نمایش می‌دهد:

- ردیف‌ها نشان‌دهنده ایستگاه‌های مبدا
- ستون‌ها نشان‌دهنده ایستگاه‌های مقصد
- رنگ تیره‌تر نشان‌دهنده تعداد بیشتر سفرها بین دو ایستگاه مشخص است

مشاهدات کلیدی:

1. تمرکز شدید در محور قطري:

بسیاری از خانه‌های روی قطر (مبدأ = مقصد) نسبتاً تیره هستند که نشان‌دهنده‌ی بازگشت کاربران به همان ایستگاه مبدا یا وجود استفاده‌های محلی (مثلاً اجاره دوچرخه برای چرخ زدن در اطراف) است.

همچنین، برخی ایستگاه‌ها با چند ایستگاه خاص ارتباط قوی دارند که بیانگر محورهای رفت‌وآمد پرتکرار است.

2. غلبه تعداد کمی ایستگاه بر کل شبکه:

ایستگاه‌هایی با شناسه‌هایی مانند ۵۱، ۵۳، ۶۵ (و چند مورد دیگر) هم به عنوان مبدا و هم مقصد، بیشترین بار ترافیکی را دارند.

این ایستگاه‌ها به احتمال زیاد در مکان‌هایی قرار دارند که نقش مرکزی یا اتصال به سایر سامانه‌های حمل و نقل دارند؛ مانند دانشگاه‌ها، پایانه‌های اتوبوس، ایستگاه‌های مترو یا نواحی تجاري پر جمعیت.

3. عدم تقارن در برخی ایستگاه‌ها:

برخی ایستگاه‌ها تنها در یکی از دو محور (ردیف یا ستون) فعال هستند، یعنی فقط مبدا یا فقط مقصد اصلی هستند.

این ایستگاه‌ها می‌توانند اهداف بالقوه‌ای برای بازتعریف کاربری یا ایجاد کمپین‌های جذب سفر برگشتی باشند.

بینش‌های راهبردی:

این heatmap اجازه می‌دهد تمرکز منابع و برنامه‌ریزی‌های عملیاتی بر محورهای پرتردد و نقاط اتصال پرتکرار صورت گیرد.

ایستگاه‌هایی که هم مبدا و هم مقصد پرتکرار هستند باید:

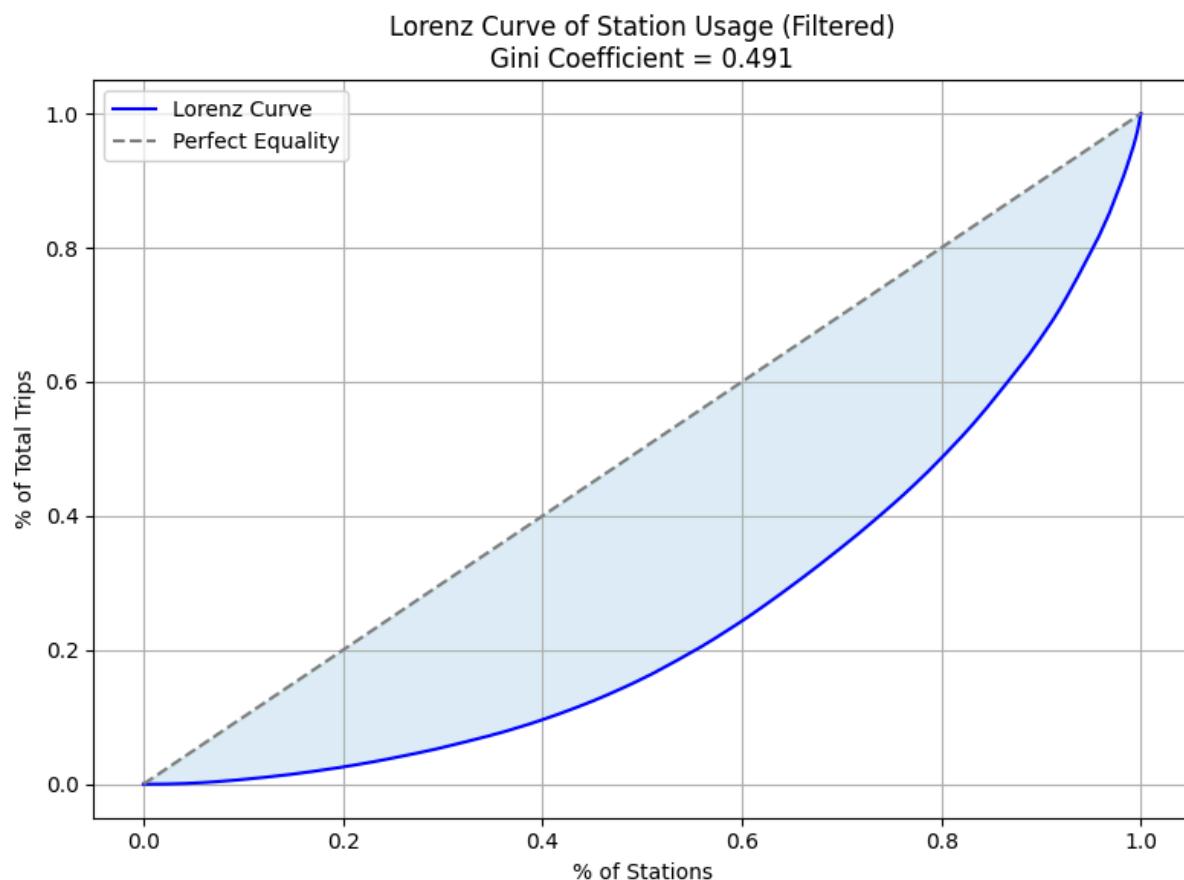
- به صورت اولویت‌دار نگهداری، نوسازی و تقویت شوند
- به عنوان هسته‌های ترافیکی در طراحی مسیرها و پیشنهادات سیاستی در نظر گرفته شوند

در مقابل، ایستگاه‌هایی با تراکم یک‌طرفه باید:

- مورد بررسی عملکردی قرار گیرند (چرا فقط خروجی یا ورودی دارند؟)
- در بازاریابی محلی یا ایجاد انگیزه‌های بازگشتی وارد شوند

این heatmap یکی از مهم‌ترین ابزارها برای درک تمرکز جریان حرکتی در شبکه دوچرخه‌سواری شهر است. تحلیل مبدأ-مقصد نه تنها می‌تواند باعث افزایش بهره‌وری شود، بلکه به بهینه‌سازی جانمایی ایستگاه‌ها، تبلیغات هدفمند، و طراحی سیستم متعادل‌تر کمک می‌کند.

۵- نابرابری استفاده از ایستگاه‌ها – تحلیل منحنی لورنز و ضریب چینی



شکل (۲۹-۳) نمودار منحنی لورنز

مفهوم منحنی لورنز:

منحنی لورنز ابزاری است برای اندازه‌گیری میزان نابرابری در توزیع یک متغیر—در اینجا، تعداد سفرهای ثبت‌شده برای ایستگاه‌های دوچرخه.

- محور افقی (x-axis): درصد تجمعی ایستگاه‌ها، مرتب شده از کم‌استفاده‌ترین تا پرtraفیک‌ترین.
- محور عمودی (y-axis): درصد تجمعی سفرهایی که توسط آن ایستگاه‌ها انجام شده‌اند.

جزئیات نمودار:

- خط آبی (Lorenz Curve): نشان‌دهنده‌ی توزیع واقعی سفرها بین ایستگاه‌هاست.
- خط خاکستری قطعی (Perfect Equality): سناریویی فرضی که در آن همه ایستگاه‌ها دقیقاً به یک اندازه استفاده شده‌اند.

• سطح زیر خط خاکستری و بالای منحنی آبی: هرچه این سطح بیشتر باشد، نابرابری بیشتر است.

در نمودار حاضر، انحنای واضح خط آبی نسبت به خط برابری نشان می‌دهد که توزیع بهوضوح ناهمگن است.

مقدار ضریب جینی (Gini Coefficient)

مقدار محاسبه شده در این تحلیل برابر با ۴۹۱/۰ است.

تفسیر:

- اگر مقدار Gini صفر باشد: یعنی برابری کامل (همه ایستگاهها دقیقاً به یک میزان استفاده شده‌اند).
- اگر مقدار Gini یک باشد: یعنی نابرابری کامل (همه سفرها فقط از طریق یک ایستگاه انجام شده‌اند).
- مقدار ۴۹۱/۰ در این زمینه به معنای نابرابری متوسط به بالا است.

یافته‌های کلیدی:

- تقریباً ۲۰٪ از ایستگاهها، بیش از ۵۰٪ سفرها را پوشش می‌دهند.
- در حالی‌که بخش بزرگی از ایستگاهها (در نیمه‌ی چپ منحنی) کمتر از ۱۰٪ سفرها را به خود اختصاص داده‌اند.
- این تمرکز می‌تواند ناشی از:
 - مکان‌یابی ضعیف یا غیر بهینه برخی ایستگاهها
 - نبود زیرساخت‌های مناسب (مثل مسیر دوچرخه‌سواری یا امنیت محلی)
 - عدم آشنایی کاربران با ایستگاه‌های حاشیه‌ای

پیامدهای مدیریتی و بازاریابی:

تیم بازاریابی می‌تواند:

- برای ایستگاه‌های کم‌استفاده کمپین‌های تشویقی برگزار کند.
- با نصب بنرها، تابلوها یا QR کدهای تبلیغاتی، آگاهی نسبت به این ایستگاهها را افزایش دهد.

برنامه‌ریزان شهری و حمل و نقل می‌توانند:

- برای ایستگاه‌های کم‌استفاده، تحلیل محیطی و مکانی انجام دهند (مثلاً بررسی دسترسی پیاده، روشنایی شبانه، نزدیکی به مکان‌های عمومی).
- مسیرهای دوچرخه‌سواری ایمن و پیوسته به آن ایستگاهها طراحی کنند.

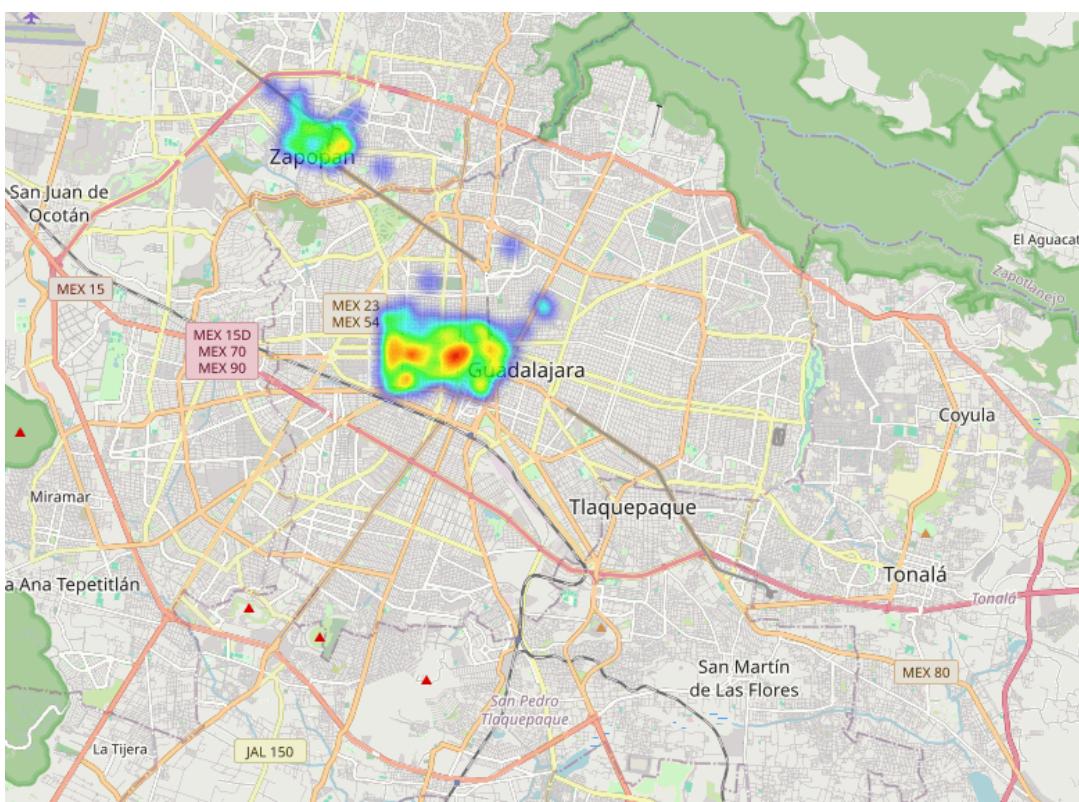
مدیران عملیاتی MiBici می‌توانند:

- از این شاخص برای تصمیم‌گیری در جایجایی ایستگاهها، ادغام ایستگاه‌های غیرفعال، یا توسعه در نواحی جدید استفاده کنند.

منحنی لورنزو و ضریب جینی ابزارهایی حیاتی برای پایش عدالت در دسترسی و بهره‌برداری از سامانه حمل و نقل عمومی مانند MiBici هستند. این تحلیل نشان می‌دهد که برای دستیابی به شبکه‌ای پایدارتر و عادلانه‌تر، لازم است تعادل در استفاده ایستگاه‌ها هدف‌گذاری شود.

۶- نقشه حرارتی تراکم سفرهای (۲۰۱۵-۲۰۱۹) MiBici

این نقشه حرارتی (Heatmap) نشان‌دهنده‌ی شدت فضایی استفاده از سامانه MiBici در شهر گوادالاخارا طی بازه زمانی ۵ ساله است. داده‌ها به صورت ماهانه تجمعی شده و شدت استفاده از طریق رنگ‌ها نمایش داده شده است:



شکل (۳۰-۳) نقشه حرارتی تراکم سفرهای

- رنگ‌های قرمز و زرد نشان‌دهنده مناطق با بیشترین تراکم سفر هستند
- رنگ‌های آبی و سبز نمایانگر تراکم پایین‌ترند
- نواحی بدون رنگ، قادر داده‌های سفر یا ایستگاه فعال در آن بازه‌اند

نکات کلیدی مشاهدات:

- ۱- تمرکز اولیه بر مرکز گوادالاخارا:
 - از سال ۲۰۱۵، بیشترین تراکم سفرها در مرکز شهر دیده می‌شود، بهویژه در نواحی مجاور مراکز اداری، دانشگاه‌ها، ایستگاه‌های حمل و نقل عمومی و بازارها.

۲- گسترش تدریجی به نواحی اطراف:

- از اواسط ۲۰۱۶، با توسعه ایستگاهها (مطابق با نمودار رشد تجمعی)، لکه‌های قرمز به نواحی غربی، شرقی و شمالی گسترش می‌یابند.

۳- ظهور نقاط پرتردد در زاپوپان (Zapopan) و تلاکوپاکه (Tlaquepaque):

- با نصب ایستگاههای جدید در این مناطق، فعالیت سفر به صورت محلی افزایش یافته و نقاط جدیدی به نقشه حرارتی افزوده شده‌اند.
- این پدیده تأیید می‌کند که توسعه زیرساخت مستقیماً با گسترش تقاضا و مشارکت فضایی در ارتباط است.

پیامدهای تحلیلی

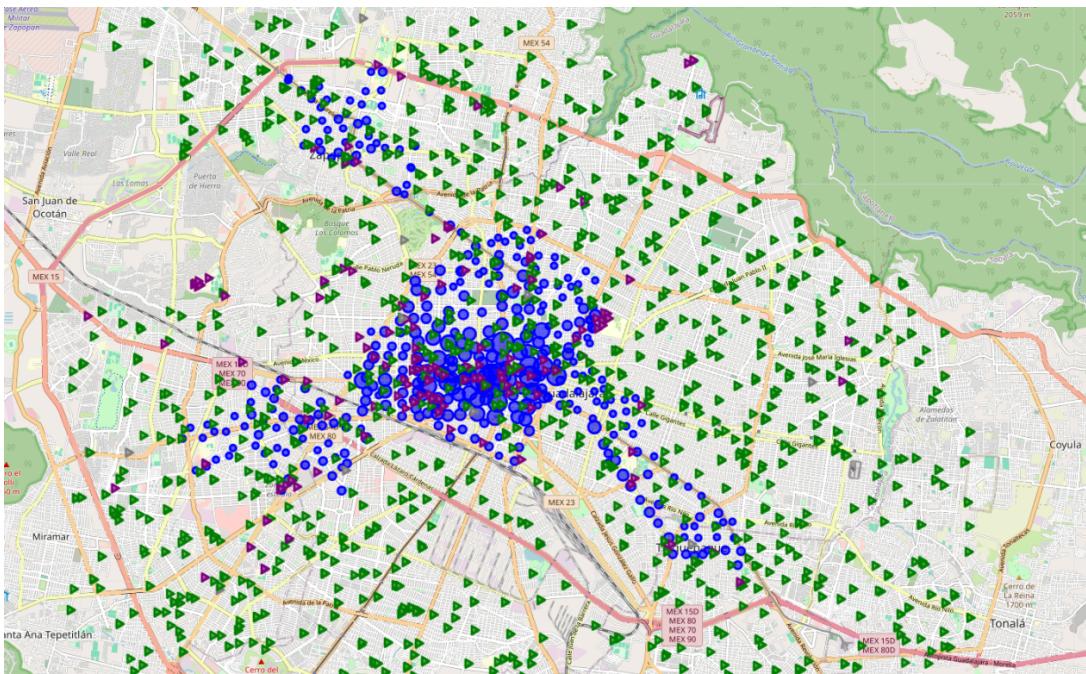
کاربردهای مدیریتی	مشاهدات	بعد تحلیلی
تمرکز عملیات و تبلیغات در هسته‌ها	تمرکز مرکزی، گسترش شعاعی	الگوی فضایی استفاده
ارزش سرمایه‌گذاری در گسترش فیزیکی	ایجاد hotspot در مناطق تازه توسعه یافته	واکنش به توسعه ایستگاهها
بهبود دسترسی برای ساکنان حاشیه‌ای	پراکندگی در سال‌های آخر بیشتر شده	توزیع عادلانه خدمات
همسویی استراتژی توسعه با نیازهای محلی	افزایش استفاده بلا فاصله بعد از نصب ایستگاه جدید	تأثیر زیرساخت بر رفتار کاربران

جدول(۵-۳) متداول‌ترین و معیارهای استفاده شده

این نقشه حرارتی اثبات می‌کند که رشد سامانه MiBici نه تنها از نظر تعداد سفر، بلکه از نظر پوشش جغرافیایی نیز گسترش یافته است. این گسترش در هماهنگی کامل با روند توسعه ایستگاهها بوده و نشان‌دهنده‌ی آن است که افزایش زیرساخت، تقاضا را تحریک می‌کند.

از این تحلیل می‌توان برای برنامه‌ریزی استراتژیک توسعه آینده، بهینه‌سازی محل ایستگاههای جدید و پیش‌بینی میزان مشارکت کاربران در مناطق جدید استفاده نمود.

۷- استفاده از ایستگاهها در ارتباط با مکان‌های عمومی



شکل (۳۱-۳) وضعیت توزیع ایستگاه‌ها و مدارس و دانشگاه‌ها در شهر

شرح اجزای نقشه:

- دایره‌های آبی: ایستگاه‌های سامانه MiBici، با اندازه متناسب با میزان استفاده (مجموع سفرهای ورودی و خروجی).
- مثلث‌های سبز: مدارس (Schools)
- مثلث‌های بنفش: دانشگاه‌ها (Universities)

این نقشه با تلفیق داده‌های مکانی ایستگاه‌ها و مکان‌های آموزشی، بستر مناسبی برای تحلیل ارتباط فضایی بین زیرساخت حمل و نقل و کاربری‌های شهری فراهم کرده است.

مشاهدات کلیدی:

- ۱- تراکم بالای استفاده در مجاورت نهادهای آموزشی:
 - در مناطقی که دانشگاه‌ها یا مدارس پر تعداد هستند (بهویژه در مرکز گوادالاخارا)، ایستگاه‌های با استفاده بالا نیز غالب‌اند.
 - این همبستگی مکانی، نقش دانشجویان، دانشآموزان و کارکنان آموزشی را در افزایش استفاده از دوچرخه تأیید می‌کند.

- ۲- ایستگاه‌های کم استفاده در نواحی فاقد زیرساخت عمومی:

- در مناطق حاشیه‌ای یا جنوب شرقی شهر که تراکم مدارس یا دانشگاه‌ها کم است، ایستگاه‌ها اغلب کم‌کاربرد یا بی‌کاربرد باقی مانده‌اند.
- این یافته از تحلیل‌های قبلی (منحنی لورنز و Gini) نیز پشتیبانی می‌کند.

۳- فقدان ایستگاه در مناطقی با تراکم مراکز آموزشی:

- در مناطقی از شمال شرق و غرب گوادالاخارا، نهادهای آموزشی متعددی دیده می‌شوند، اما ایستگاه MiBici در نزدیکی آن‌ها موجود نیست یا پوشش مناسبی ندارد.
- این نواحی می‌توانند اهداف بالقوه برای گسترش شبکه باشند.

پیامدهای راهبردی و سیاستی:

برای بازاریابی:

- برگزاری کمپین‌های تبلیغاتی در دانشگاه‌ها و مدارس مجاور ایستگاه‌های فعال (پیک استفاده) برای تشویق به اشتراک بلندمدت یا استفاده روزانه.
- ارائه تخفیف یا اعتبار ویژه برای اساتید، دانشجویان، یا والدین در مناطقی با ایستگاه‌های کم‌استفاده.
- همکاری با مؤسسات آموزشی برای تبلیغ مشترک و افزایش آگاهی عمومی.

برای توسعه شهری و برنامه‌ریزی:

- توسعه ایستگاه‌های جدید در شعاع ۳۰۰ تا ۵۰۰ متری از مدارس یا دانشگاه‌های بدون پوشش.
- اولویت‌بندی سرمایه‌گذاری بر اساس هم‌پوشانی ایستگاه‌ها و نهادهای آموزشی.
- استفاده از داده‌های API‌های شهری برای بهروزرسانی منظم نقشه‌های پوشش خدمات.

نوع توسعه	دلیل انتخاب	منطقه پیشنهادی
نصب ایستگاه جدید	تراکم بالای مدارس، نبود ایستگاه	شمال‌شرقی Zapopan
بازطراحی موقعیت ایستگاه یا تبلیغات هدفمند	کم‌کاری ایستگاه‌ها، ضعف زیرساخت	جنوب شرقی Guadalajara
افزودن ایستگاه جدید با مشارکت دانشگاه	خوشه دانشگاهی فعال بدون پوشش ایستگاه	امتداد Avenida Revolución

جدول(۳-۶) روش‌ها بهبود و توسعه

این نقشه تصویری روشی از تأثیر بافت شهری و نهادهای عمومی بر رفتار استفاده کنندگان سامانه دوچرخه اشتراکی ارائه می‌دهد. نتایج تحلیل نشان می‌دهد که موفقیت MiBici نه فقط به مکان‌یابی ایستگاه‌ها، بلکه به پیوند آن‌ها با زیرساخت‌های عمومی موجود وابسته است.

۳.۴.۳ خلاصه تحلیل‌های انجام شده

۱. تحلیل توزیع سنی کاربران:

- اکثریت کاربران بین ۲۰ تا ۴۰ ساله هستند، با تمرکز در دهه ۹۰ میلادی.
- فرصت بازاریابی هدفمند برای جوانان، دانشجویان و کارمندان شهری فراهم است.

۲. تحلیل فضایی ایستگاه‌ها:

- تمرکز بالا در مرکز شهر و کمبود پوشش در نواحی حاشیه‌ای.
- توزیع ایستگاه‌ها به صورت شعاعی، اما نامتوازن.
- ایستگاه‌های بدون استفاده یا ایستگاه‌های ساخته شده پس از سال ۲۰۱۹ باید مجزا بررسی شوند.

۳. نقشه‌های مبدأ، مقصد، استفاده کلی:

- برخی ایستگاه‌ها به طور خاص یا فقط به عنوان مبدأ یا مقصد غالب هستند.
- الگوهای رفت‌وآمد نابرابر، نیازمند بازنگری در جانمایی، و استراتژی‌های rebalancing هستند.

۴. منحنی لورنزو و ضریب جینی:

- $Gini = 0.491 \rightarrow$ نابرابری متوسط به بالا در استفاده از ایستگاه‌ها.
- حدود ۲۰٪ ایستگاه‌ها بیش از ۵٪ سفرها را به خود اختصاص داده‌اند.

۵. نقشه‌های حرارتی تراکم استفاده در زمان:

- همگام با رشد ایستگاه‌ها، استفاده نیز از مرکز به سمت حومه گسترش یافته است.
- توسعه زیرساخت باعث افزایش طبیعی مشارکت شده است.

۶. تحلیل همپوشانی ایستگاه‌ها با نهادهای عمومی:

- ایستگاه‌های نزدیک به دانشگاه‌ها و مدارس بسیار فعال هستند.
- مناطقی با تراکم نهاد آموزشی بدون ایستگاه، کاندیدای توسعه هستند.

۴.۴.۳ جمع‌بندی سوال اول

استفاده از سامانه MiBici در سطح شهر به صورت یکنواخت توزیع نشده است. تمرکز سفرها و استفاده از ایستگاه‌ها در نواحی مرکزی شهر و در مجاورت مراکز آموزشی، تجاری و حمل‌ونقل عمومی بسیار بالاست، در حالی‌که نواحی حاشیه‌ای و کمترکم چهار ضعف در جذب تقاضا هستند.

پیشنهادات کاربردی برای تیم بازاریابی:

توصیه	محور اقدام
تمرکز بر جوانان، دانشجویان، و مناطق با تراکم دانشگاهی	تبلیغات هدفمند
کمپین‌های انگیزشی، تخفیف، اعتبار رایگان	تشویق استفاده از ایستگاه‌های کم‌کاربرد
استفاده از داده‌های نهادهای عمومی برای توسعه مبتنی بر تقاضای بالقوه	انتخاب محل ایستگاه‌های آینده
بهروزرسانی استراتژی‌های rebalancing با توجه به مناطق با تفاوت مبدأ-مقصد	مدیریت تعادل شبکه
استفاده از ابزارهایی مانند heatmap و Gini برای تصمیم‌سازی داده‌محور	پایش مداوم توزیع مکانی

جدول(۷-۳) راهکار ها و توصیه ها

۵.۳.۳ عوامل موثر بر نوسانات سفرها

۱.۵.۳.۳ متداول‌وزی و معیارهای استفاده شده

شناسایی و تبیین عواملی که باعث افزایش یا کاهش تعداد سفرهای روزانه/ماهانه با دوچرخه‌های اشتراکی MiBici شده‌اند و برآورد سهم نسبی هر عامل در این نوسانات.

۱. ساخت دیتا فریم روزانه جامع (Daily Master Table)

برای انجام این تحلیل، یک جدول اصلی با سطح تجمعی روزانه ساخته شد که شامل موارد زیر است:

توضیح	ستون‌ها
تاریخ	date

تعداد سفرهای انجام شده در آن روز	ride_count
نوع روز (کاری، تعطیل رسمی، آخر هفته)	day_type
دماه میانگین روزانه ($^{\circ}\text{C}$)	mean_temp
بارش روزانه (میلیمتر)	rain_mm
سرعت باد	wind_kph
رطوبت نسبی (%)	humidity
روز هفته (شنبه تا جمعه)	weekday

جدول(۸-۳) توضیحات جدول Daily Master

۲. منابع داده خارجی (API‌ها و داده‌های برون‌سازمانی):

برای تبیین عوامل محیطی و اجتماعی مؤثر، از داده‌های غیرمستقیم و برون‌سازمانی استفاده شده است:

کاربرد	نوع داده	منبع
دما، بارش، رطوبت، باد	آب و هوای تاریخی گوادالاخارا	Open-Meteo API
طبقه‌بندی روزهای رسمی	نوع روز	جدول روزهای تعطیل مکزیک

بررسی الگوهای روزانه	استخراج نام روز هفته	تقویم میلادی و هفتگانگاری
----------------------	----------------------	---------------------------

جدول(۹-۳) API

این داده‌ها در جدول weather_with_day_type.csv ادغام و پیش‌پردازش شده‌اند.

۳. شاخص‌های کلیدی بررسی شده:

هدف تحلیلی	شاخص
متغیر وابسته (پاسخ)	تعداد سفر در روز (ride_count)
اثر منفی بر استفاده از دوچرخه	بارش روزانه (rain_mm)
تعیین نقش تعطیلی رسمی یا آخر هفته	نوع روز (day_type)
بررسی اثر چرخه‌های کاری/تفریحی	روز هفته (weekday)
سنجش قدرت رابطه هر متغیر با ride_count	همبستگی پارامترها
بررسی روندهای بلندمدت	تحلیل گرایش‌های ماهانه و فصلی

جدول(۱۰-۳) شاخص‌های کلیدی

۴. روش‌های آماری و مصوروسازی:

در تحلیل این بخش، از روش‌های زیر استفاده شده است:

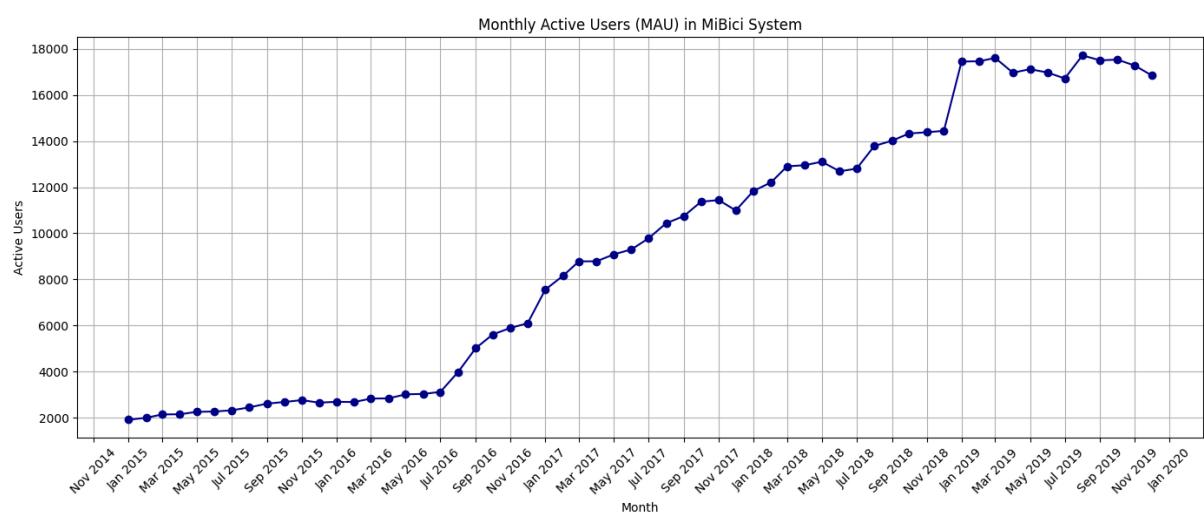
- تحلیل روند زمانی: با نمودارهای خطی برای بررسی نوسانات ماهانه
- نمودار بارش/سفر: تحلیل رابطه معکوس

در این تحلیل، متغیر پاسخ ما تعداد سفرهای روزانه است، و متغیرهای مستقل شامل مجموعه‌ای از ویژگی‌های تقویمی، جوی و اجتماعی هستند. هدف، درک کامل از این است که:

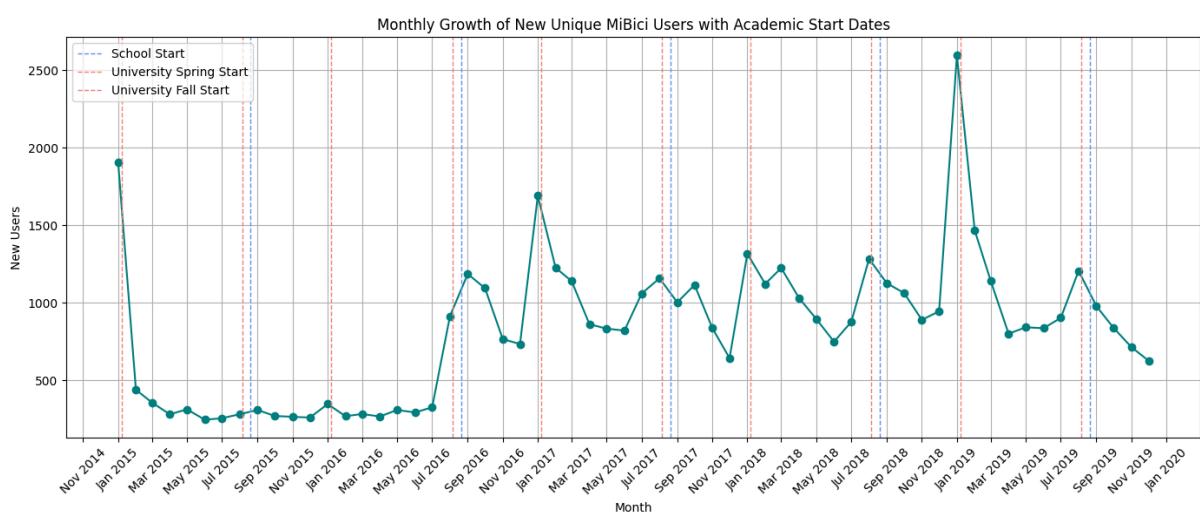
- چه عواملی باعث اوج یا افت سفرها می‌شوند؟
- کدامیک از این عوامل سهم بیشتری در نوسانات دارند؟
- چه بینش‌هایی می‌توان برای بهینه‌سازی تقاضا، اطلاع‌رسانی آب‌وهوايی، یا سیاست‌گذاري زمان‌بندی توسعه از این تحلیل گرفت؟

۲.۵.۳.۳ یافته‌ها و بینش‌ها

الف) تحلیل روند رشد کاربران جدید و کاربران فعال ماهانه



شکل (۳۲-۳) نمودار کاربران فعال ماهانه



شکل (۳۳-۳) نمودار رشد کاربران

۱. رشد ماهانه کاربران جدید (MiBici) New Unique Users

نمودار اول تعداد کاربران منحصر به فردی را که در هر ماه به سامانه MiBici پیوسته‌اند نشان می‌دهد. نقاط عطف کلیدی در این نمودار به شرح زیر هستند:

مشاهدات کلیدی:

- در اکثر ماه‌ها بین ۵۰۰ تا ۱۳۰۰ کاربر جدید جذب شده‌اند که نشان‌دهنده رشد پایدار اما تدریجی است.
- دو جهش بسیار بزرگ در جذب کاربران مشاهده می‌شود:
 - ژانویه ۲۰۱۷ با بیش از ۱۶۰۰ کاربر جدید
 - ژانویه ۲۰۱۹ با بیش از ۲۵۰۰ کاربر جدید (بزرگ‌ترین رشد ماهانه ثبت شده)

تحلیل ژانویه ۲۰۱۹:

- همزمانی این جهش با شروع ترم بهار دانشگاه‌ها و بحران کمبود بنزین در گوادالاخارا.
- منابع خبری تأیید کرده‌اند که در ده روز نخست ژانویه ۲۰۱۹، بیش از ۶۰۰ کاربر جدید به سامانه پیوستند و استفاده از دوچرخه به‌طور چشمگیری افزایش یافت.
- نتیجه: MiBici به عنوان یک گزینه حمل و نقل جایگزین در زمان بحران عمل کرده است.

نتیجه‌گیری:

رشد کاربران جدید نه تنها به مناسبت‌های دوره‌ای مانند شروع ترم‌های دانشگاهی وابسته است، بلکه به رویدادهای ناگهانی و بحران‌های شهری نیز حساس است.

۲. همبستگی بین شروع ترم‌های تحصیلی و افزایش کاربران جدید

نمودار دوم، داده‌های رشد کاربران را با تاریخ‌های تقریبی شروع مدارس و دانشگاه‌ها در گوادالاخارا همپوشانی کرده است.

عناصر کلیدی:

- خطوط عمودی آبی → شروع سال تحصیلی مدارس (اواسط تا اواخر آگوست)
- خطوط قرمزگ → شروع ترم پاییز دانشگاه‌ها (اوایل آگوست)
- خطوط قرمز پرنگ → شروع ترم بهار دانشگاه‌ها (اوایل ژانویه)

مشاهدات تحلیلی:

- همپوشانی واضح بین پیک‌های کاربران جدید و تقویم دانشگاهی:
 - ژانویه‌ها و آگوست‌ها شاهد افزایش ثبت‌نام چشمگیر هستند.
 - نشان‌دهنده وابستگی بالا به جمعیت دانشجویی و تحصیل کرده شهر.
- قوی‌ترین اوج، ژانویه ۲۰۱۹، هم‌زمان با هر دو عامل: شروع ترم + بحران شهری (کمبود بنزین)

منابع:

- [تقویم رسمی آموزش و بورش مکزیک \(SEP\)](#)
- [تقویم دانشگاه گوادالاخارا \(UdeG\)](#)
- [مقاله Emerald Insight درباره بحران حمل و نقل سال ۲۰۱۹](#)

۳. کاربران فعال ماهانه (MAU) در سیستم MiBici

نمودار اول نشان دهنده تعداد کاربرانی است که در هر ماه دستکم یک سفر با سامانه انجام داده اند.

بینش‌ها

روندی صعودی و پایدار از سال ۲۰۱۵ تا ۲۰۱۹ مشاهده می‌شود.

رشد قابل توجه بعد از میانه سال ۲۰۱۶ همزمان با توسعه ایستگاه‌ها.

اوج فعالیت در اوایل ۲۰۱۹، به بیش از ۱۷,۵۰۰ کاربر فعال ماهانه می‌رسد.
پس از این اوج، سطح کاربران فعال در سطح بالایی ثبت می‌شود، با نوسانات فصلی جزئی.

نتیجه‌گیری

شاخص MAU نشان می‌دهد که نه تنها جذب کاربران موفق بوده، بلکه نگهداشت و تداوم استفاده نیز در سطح بالایی حفظ شده است.

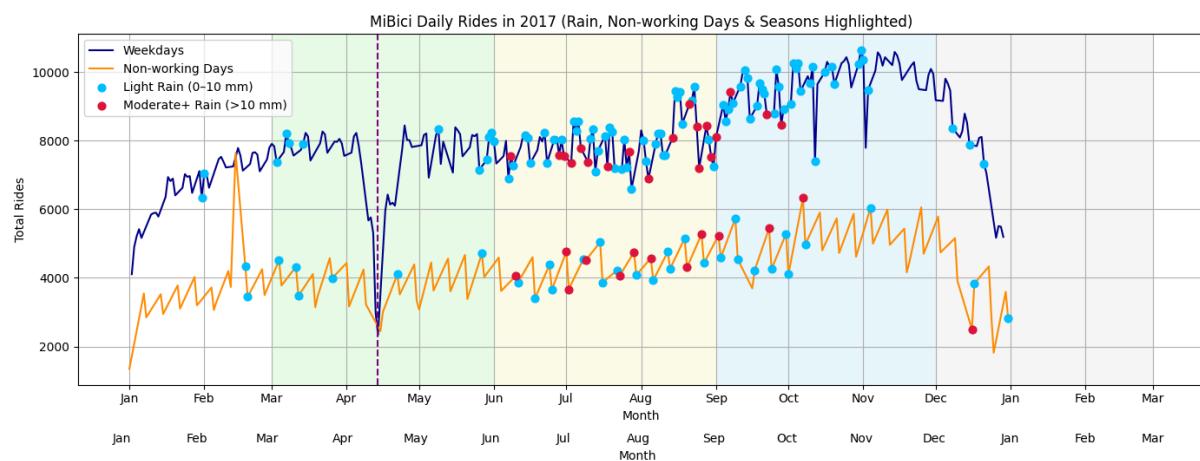
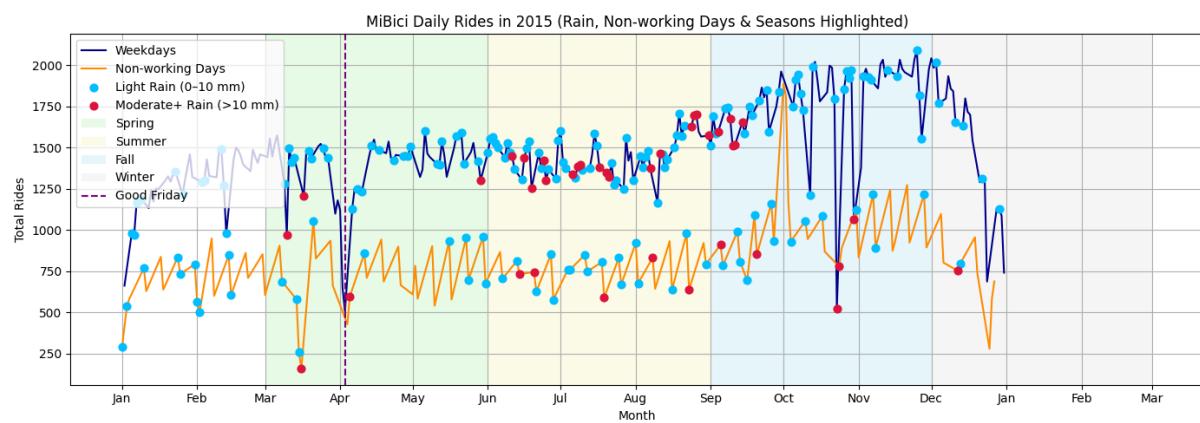
جمع‌بندی یافته‌ها درباره رشد کاربران

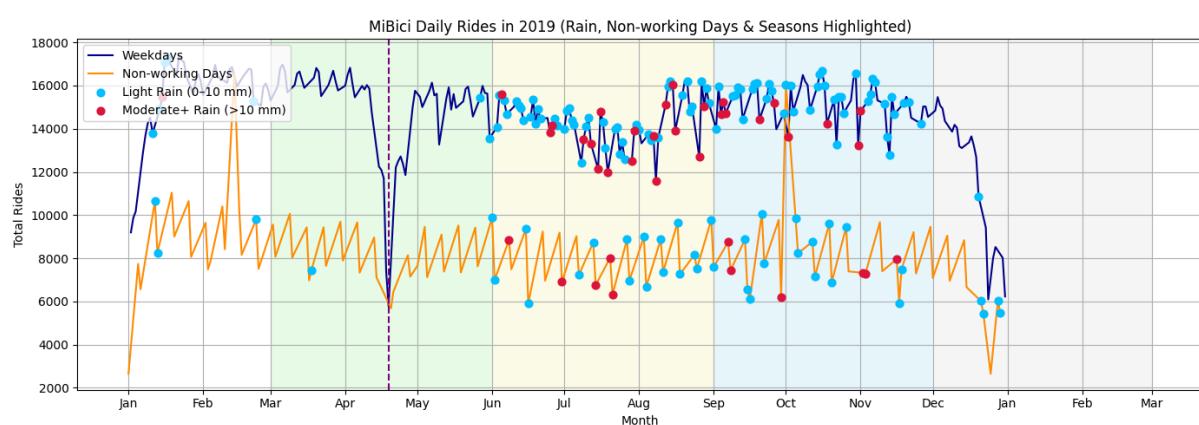
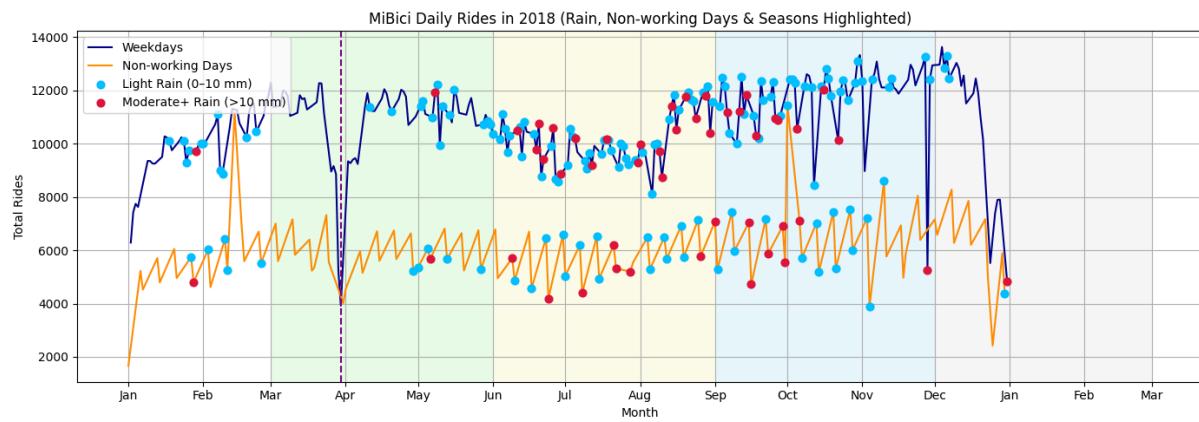
نقش در رشد	مؤلفه
عامل تکرارشونده و قابل پیش‌بینی برای جذب کاربران جدید	شروع ترم‌های دانشگاهی
عامل غیرقابل پیش‌بینی با تأثیر بسیار شدید و مثبت بر استفاده	بحران شهری (کمبود بنزین)
پشتیبان افزایش کاربران فعال بلندمدت	توسعه ایستگاه‌ها
تأیید تمرکز سامانه بر جامعه دانشجویی و حرفه‌ای جوان	گروه سنی غالب (۱۹۸۵-۲۰۰۰)

پیامدهای استراتژیک:

- کمپین‌های جذب کاربران جدید باید با تقویم دانشگاهی تنظیم شوند (زانویه، آگوست).
- آمادگی برای شرایط بحرانی شهری (مانند اعتصاب یا کمبود سوخت) می‌تواند باعث افزایش سریع مشارکت شود.
- پایش MAU و تحلیل فصلی برای شناخت زمان‌های اوج و افول استفاده جهت بهینه‌سازی عملیات حیاتی است.

ب) تحلیل روزانه استفاده از دوچرخه در طول سال (۲۰۱۹-۲۰۱۵)





شکل (۳۴-۳) نمودار تعداد سفر ها در بازه مورد بررسی

نمودارهای پنجگانه ارائه شده، تعداد سفرهای روزانه MiBici را به تفکیک سال نشان می دهند، در حالی که پس زمینه هر نمودار با رنگ های مختلف نشان دهنده فصل هاست:

- بهار (مارس-مای) – سبز روشن
- تابستان (ژوئن-آگوست) – زرد کمرنگ
- پاییز (سپتامبر-نوامبر) – آبی کمرنگ
- زمستان (دسامبر-فوریه) – خاکستری روشن

بینش های کلیدی:

۱- چرخه هفتگی کاملاً مشخص است:

- در همه سال ها، استفاده در روزهای هفته به مرتب بالاتر از روزهای تعطیل است.
- این الگو نشان دهنده ماهیت کاری و رفت و آمدی سیستم است.

۲- اثر فصل‌ها:

- در بیشتر سال‌ها، پاییز و بهار بیشترین استفاده را به خود اختصاص داده‌اند.
- تابستان با افت نسبی همراه است؛ احتمالاً به دلیل گرمای هوا یا تعطیلی مدارس و دانشگاه‌ها.
- زمستان، بهویژه در تعطیلات دسامبر، شاهد کاهش چشمگیر در میزان سفرهای است.

۳- افزایش روند کلی در طول سال‌ها:

- از ۲۰۱۵ تا ۲۰۱۹، حجم کل سفرهای روزانه افزایش یافته که این موضوع بازتاب‌دهنده گسترش شبکه ایستگاه‌ها و پذیرش عمومی بیشتر است.

ج) تحلیل تفصیلی اثرات محیطی و فرهنگی

برای درک دقیق‌تر نوسانات، نمودارها علاوه‌بر فصل‌ها، اطلاعاتی درباره موارد زیر نیز نمایش می‌دهند:

۱- نوع روز (کاری/تعطیل):

- در تمامی سال‌ها، روزهای کاری (خط آبی)، بیشترین سفرها را ثبت کرده‌اند.
- روزهای تعطیل (خط نارنجی)، مانند آخر هفته‌ها و تعطیلات رسمی، کاهش قابل‌توجهی را نشان می‌دهند.
- با این حال، در فصل‌های پاییز (اکتبر و نوامبر)، این شکاف بین روزهای کاری و تعطیل کاهش می‌یابد، که ممکن است به دلیل هوای معتدل‌تر و رویدادهای شهری باشد.

۲- بارندگی:

- نقاط آبی (بارش سبک ۰-۱۰ میلی‌متر) و قرمز (بارش متوسط یا زیاد >۱۰ میلی‌متر) نشان‌دهنده اثرات آب‌وهوایی هستند:
- باران متوسط و شدید تقریباً همیشه با کاهش چشمگیر تعداد سفرها همراه است.
- باران سبک نیز اثر منفی دارد، اما در مقایسه با بارش شدید، کمتر بازدارنده است.

۳- جمعه نیک (خط عمودی بنفسن):

- در تمام سال‌ها، در روز جمعه نیک کاهش چشمگیر در تعداد سفرها مشاهده می‌شود.
- این موضوع نشان‌دهنده تأثیر مناسبت‌های مذهبی و فرهنگی خاص بر رفتار کاربران است و لزوم در نظر گرفتن این ایام در پیش‌بینی تقاضا را تأیید می‌کند.

جمع‌بندی تحلیلی

- استفاده روزانه از MiBici به شدت تحت تأثیر متغیرهای محیطی و فرهنگی است.
- فصل، بارندگی، تعطیلات و حتی مناسبت‌های خاص، همگی نقش مهمی در شکل‌گیری نوسانات سفرها دارند.
- با افزایش مقیاس سیستم، اثرات این متغیرها پرزنگ‌تر شده و نیاز به سیاست‌گذاری دقیق‌تری برای پشتیبانی از تقاضا، مدیریت دوچرخه‌ها و اطلاع‌رسانی مناسب وجود دارد.

- تحلیل‌های انجام‌شده نشان می‌دهد که باید کاربران را براساس نوع استفاده (رفت‌وآمدی، تفریحی، فصلی) طبقه‌بندی کرد و برنامه‌ریزی حمل‌ونقل شهری را متناسب با الگوهای استفاده و شرایط محیطی تنظیم نمود.

۳.۵.۳. جمع‌بندی سوال دوم

در این بخش، تمرکز اصلی بر پاسخ به این سؤال کلیدی بود:

چه عواملی باعث نوسانات تعداد سفرها در سیستم MiBici می‌شوند و سهم نسبی هر عامل چقدر است؟

تحلیل‌های انجام‌شده بر اساس داده‌های روزانه و ماهانه از سال ۲۰۱۵ تا ۲۰۱۹، به همراه داده‌های تقویمی، بارندگی و مناسبت‌های شهری، ما را به یافته‌های مهم زیر رساند:

۱- فصل‌های سال نقش کلیدی دارند

- پاییز و بهار پر تقاضاترین فصل‌ها برای استفاده از دوچرخه هستند.
- زمستان (به‌ویژه دسامبر) و تابستان (به‌خصوص جولای و آگوست) با کاهش محسوسی در تعداد سفرها همراه‌اند.
- این الگوها با شرایط آب‌وهواهی، تعطیلی دانشگاه‌ها و کاهش رفت‌وآمدگاه‌های روزانه هماهنگ هستند.

۲- آب‌وهوا، به‌ویژه بارندگی، اثر مستقیم دارد

- بارش‌های متوسط یا سنگین (بیش از ۱۰ میلی‌متر) موجب افت شدید در استفاده می‌شوند، حتی در روزهای کاری.
- باران سبک نیز باعث کاهش نسبی می‌شود، اما کاربران وفادار در روزهای خشک‌تر فعال‌تر هستند.
- این موضوع اهمیت پیش‌بینی هواشناسی در عملیات روزانه سیستم را نشان می‌دهد.

۳- نوع روز (کاری/تعطیل) عامل تعیین‌کننده است

- استفاده در روزهای کاری به‌طور واضح بالاتر از تعطیلات و آخر هفته‌هاست.
- این اختلاف، نشان می‌دهد که MiBici بیشتر برای سفرهای کاری، تحصیلی و روزمره استفاده می‌شود.
- البته در ماه‌های اوچ (اکتبر، نوامبر)، حتی در روزهای تعطیل نیز تقاضا افزایش می‌یابد.

۴- مناسبت‌های فرهنگی و مذهبی (مانند جمیعه نیک) تأثیرگذار هستند

- به صورت مکرر دیده شد که در روز جمیعه نیک، که یک مناسبت مذهبی ملی در مکزیک است، افت مشخصی در سفرها رخ می‌دهد.
- این موضوع نشان می‌دهد که پیش‌بینی تقاضا باید به الگوهای تقویمی محلی نیز حساس باشد.

۵- رخدادهای غیرمنتظره شهری می‌توانند جهش ایجاد کنند

- در زانویه ۲۰۱۹، به‌دلیل بحران کمبود سوخت در شهر گوادالاخارا، استفاده از MiBici به‌طور بی‌سابقه‌ای افزایش یافت.

- این نشان می‌دهد که سیستم در مواجهه با بحران‌ها نقش پشتیبان برای حمل و نقل پایدار ایفا می‌کند.

۶.۳.۳ نتیجه‌گیری‌نهایی

نوسانات تعداد سفرهای MiBici کاملاً تصادفی یا فصلی نیستند، بلکه به شدت تحت تأثیر مجموعه‌ای از عوامل قابل پیش‌بینی (فصل، باران، تقویم آموزشی) و عوامل غیرقابل پیش‌بینی (بحران‌ها، مناسبات‌ها) قرار دارند.

این یافته‌ها می‌توانند راهگشای اقدامات زیر باشند:

- بهینه‌سازی لجستیک و توزیع دوچرخه‌ها در فصل‌ها و روزهای مختلف
- افزایش آمادگی عملیاتی در مواجهه با شرایط جوی یا رویدادهای خاص
- هدف‌گذاری تبلیغات و برنامه‌های تشویقی بر اساس تقویم آموزشی و شهری
- طراحی سیاست‌های بلندمدت برای تقویت نقش MiBici به عنوان ابزار مقاوم حمل و نقل در شرایط بحران

در مجموع، شناخت عمیق از این عوامل زمینه‌ساز تصمیم‌گیری دقیق‌تر برای توسعه پایدار، افزایش بهره‌وری و بهبود تجربه کاربری در سیستم MiBici خواهد بود.

۴.۳ بخش‌بندی کاربران وفادار و موردي (سوال ۵)

۱.۴.۳ تشریح مسئله

مدیر قیمتگذاری و مدیر بازاریابی شرکت MiBici، با هدف بهینه‌سازی درآمد و افزایش وفاداری کاربران، تصمیم دارند کاربران را بر اساس الگوهای رفتاری به دو دسته‌ی «وفادار» (Loyal) و «موردی» (Occasional) تقسیم کنند. این دسته‌بندی مبنای طراحی دو استراتژی متفاوت خواهد بود:

□ مدیر قیمتگذاری قصد دارد مدل‌های قیمتی متمایزی برای کاربران وفادار و موردی در نظر بگیرد، به‌گونه‌ای که این استراتژی‌ها منجر به حفظ کاربرهای وفادار و تشویق کاربرهای موردی شود.

□ همچنین مدیر بازاریابی به دنبال این است که کمپین‌های تبلیغاتی هدفمندی برای کاربران موردی یا غیرفعال اجرا کند و از طریق اطلاع‌رسانی مزایای قیمتگذاری جدید، آن‌ها را به استفاده بیشتر از سرویس تشویق کند.

برای تحقق این اهداف، لازم است ابتدا با استفاده از داده‌های تاریخی سفرها، کاربران را به‌طور دقیق و قابل اتکا در این دو دسته طبقه‌بندی کنیم و سپس رفتار این دو گروه را در ویژگی‌هایی مثل سن، جنسیت، الگوی استفاده و ایستگاه‌های ترجیحی تحلیل کنیم تا استراتژی‌های طراحی شده به شکلی اثربخش و داده‌محور اجرا شوند.

۲.۴.۳ دسته‌بندی کاربران

برای پاسخ به این سؤال که چه تفاوتی میان کاربران وفادار و کاربران موردی وجود دارد، نخست لازم بود تا کاربران را به دو دسته‌ی معنادار و مبتنی بر رفتار استفاده از سرویس تقسیم کنیم. این دسته‌بندی باید به گونه‌ای باشد که بتواند تفاوت‌های معناداری را در رفتار، ویژگی‌ها و ترجیحات دو گروه آشکار کند.

با توجه به تحقیقات، معمولاً برای بررسی میزان وفاداری کاربران، دو فاکتور بسیار مهم وجود دارد. این دو فاکتور که میزان یا حجم استفاده و پیوستگی استفاده هستند و در بازه‌های زمانی ۶ ماه تا ۱ ساله مورد بررسی قرار می‌گیرند. البته بازه‌های طولانی‌تر برای محصولاتی که نوسان فصلی دارند رایج‌تر است.

• **معیار بررسی حجم استفاده:** برای بررسی این موضوع فاکتور تعداد سفر در ماه برای کاربرها مورد بررسی قرار داده شد.

• **معیار بررسی پیوستگی استفاده:** برای بررسی این موضوع فاکتور تعداد روز فعال در ماه برای کاربرها مورد بررسی قرار داده شد.

دسته‌بندی نهایی با توجه به بررسی‌ها

• **تعريف کاربر وفادار (Loyal):** کاربرهای وفادار رو به صورت زیر تعریف کردیم: «کاربرانی که در حداقل ۵ ماه از ۶ ماه دوم سال ۱۴۰۹، در هر ماه دستکم ۱۴ سفر و در حداقل ۸ روز متفاوت سفر داشته‌اند.»

• **تعريف کاربران موردی (Occasional):** تمامی کاربران دیگر که معیارهای بالا را احراز نکردند، به عنوان کاربران موردی (Occasional Users) طبقه‌بندی شدند. این گروه شامل طیف وسیعی از کاربران است، از کسانی که تنها چند بار در ماه از سرویس استفاده کرده‌اند تا کسانی که استفاده‌ی منظمی نداشته‌اند یا استفاده‌شان محدود به یک یا دو ماه بوده است.

منطق انتخاب معیارها و آستانه‌ها

تعریف دقیق و داده محور از وفاداری کاربران در سامانه اشتراک دوچرخه، هم از تحلیل آماری داده‌های MiBici در بازه ۶ ماهه (دوم سال ۲۰۱۹) و هم از بررسی پژوهش‌های معتبر بین‌المللی در این حوزه الهام گرفته شد. نتیجه‌ی این دو مسیر، ما را به سمت انتخاب آستانه‌ها و تعاریفی رساند که در ادامه تشریح می‌شود.

۱. چرایی انتخاب بازه ۶ ماه دوم سال ۲۰۱۹

طبق بررسی‌هایی که روی داده‌ها انجام دادیم، نوسان فصلی قابل توجهی در داده‌ها نداشتیم و از این جهت برای بررسی کاربرهای وفادار بازه ۶ ماهه آخر ۲۰۱۹ را در نظر گرفتیم. همچنین این بازه بازتاب خوبی از رفتار اخیر کاربران پیش از ورود به سال جدید (۲۰۲۰) را به ما می‌دهد و برای تصمیمات بازاریابی بسیار مفید است.

۲. انتخاب آستانه «۱۴ سفر» و «۸ روز فعال» در هر ماه:

در انتخاب این اعداد، از بررسی میانه و میانگین این شاخص‌ها در بازه مورد بررسی (۶ ماه دوم ۲۰۱۹) استفاده شد و با توجه به اینکه میانه دید بهتری از کلیت کاربرها به ما می‌دهد، از این شاخص برای آستانه استفاده کردم.

تعداد سفرهای کاربرهای هر ماه:

- میانگین تعداد سفر در ماه برای کاربرهای همان ماه: ۲۱.۹۷ سفر
- میانه تعداد سفر در ماه برای کاربرهای همان ماه: ۱۴ سفر

تعداد روزهای فعال کاربرها در هر ماه:

- میانگین تعداد روز فعال کاربرها در هر ماه: ۸ روز
- میانه تعداد روز فعال کاربرها در هر ماه: ۸ روز

۳. شرط پیوستگی: حداقل در ۵ ماه از ۶ ماه

یکی از ویژگی‌های وفاداری واقعی، «پایداری رفتار» در طول زمان است، نه صرفاً حجم مصرف. به همین دلیل، صرفاً داشتن یک یا دو ماه پرتردد برای ورود به گروه وفادار کافی نیست. شرط اینکه کاربر باید در حداقل ۵ ماه از ۶ ماه تعریف شده، هر دو شرط (۱۴ سفر و ۸ روز فعال) را داشته باشد، تضمین می‌کند که کاربر رفتار مصرف پیوسته و مستمر دارد.

نتیجه نهایی دسته‌بندی:

با انجام هر دو شرط و اشتراک گرفتن از آن‌ها خروجی به شرح زیر است:

- تعداد کاربران وفادار: ۵۲۰۳ نفر
- درصد کاربران وفادار از کل کاربران ۶ ماه آخر: %۲۱
- درصد کاربران وفادار از کل کاربران بیزنس در سال‌های ۲۰۱۹ تا ۲۰۱۵: %۱۰.۴۴

۱.۳.۴.۳ بررسی و تحلیل توزیع ویژگی‌های کاربران وفادار در مقایسه با کاربران موردي و آزمون‌های آماري

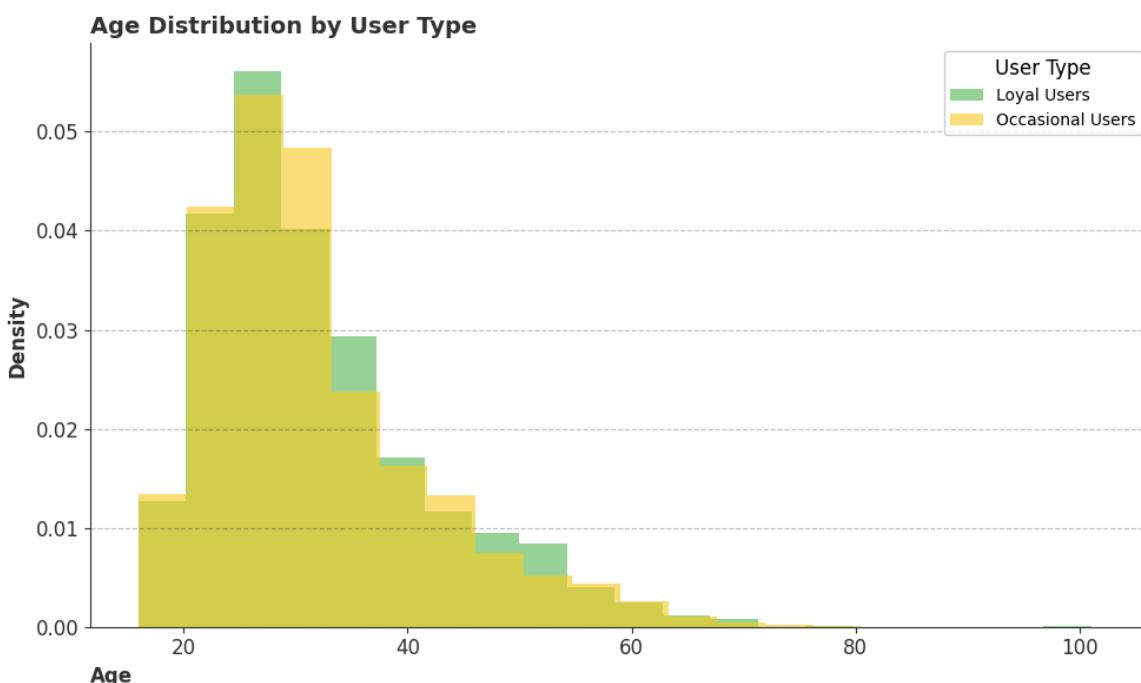
برای بررسی تفاوت کاربران وفادار و موردی از نظر ویژگی‌های جمعیتی و رفتاری، و به منظور شناسایی فرصت‌های طراحی کمپین‌های بازاریابی هدفمند برای هر گروه از کاربران، سه ویژگی سن، جنسیت و میانگین مدت زمان استفاده از دوچرخه‌ها در کاربران مورد بررسی قرار گرفت.

الف) سن کاربران

شکل (۳۵-۳) توزیع سن کاربران وفادار و موردی را نشان می‌دهد. توزیع سن کاربران وفادار و موردی شباهت زیادی به هم دارد و تفاوت معناداری بین آن‌ها وجود ندارد. با بررسی توزیع سنی کاربران وفادار و موردی مشخص شد که:

- هر دو گروه بیشترین تراکم سنی را در بازه ۲۰ تا ۳۵ سال دارند. با این حال، توزیع کاربران وفادار به‌طور نسبی به سمت افراد جوان‌تر تمایل دارد.
- نمودار چگالی سنی نشان می‌دهد که قله توزیع کاربران وفادار کمی زودتر (در سنین پایین‌تر) نسبت به کاربران موردی اتفاق می‌افتد.
- آزمون آماری **Mann–Whitney U** برای بررسی معناداری این تفاوت انجام شد. مقدار **p-value** بالاتر از سطح معناداری نشان داد که تفاوت مشاهده شده از نظر آماری معنادار نیست ($p\text{-value} > 0.05$). این یعنی اگرچه تفاوت‌های ظاهری کوچکی وجود دارد، اما نمی‌توان با اطمینان گفت که سن عاملی در تعیین وفاداری است.

نتیجه: تفاوت‌های جزئی در الگوی سنی دیده می‌شود (تمایل کاربران وفادار به سنین پایین‌تر)، اما این تفاوت‌ها از نظر آماری معنادار نیستند و نمی‌توانند مبنای تصمیم‌گیری بازاریابی قرار گیرند.



شکل (۳۵-۳) مقایسه توزیع سنی کاربران موردی و وفادار

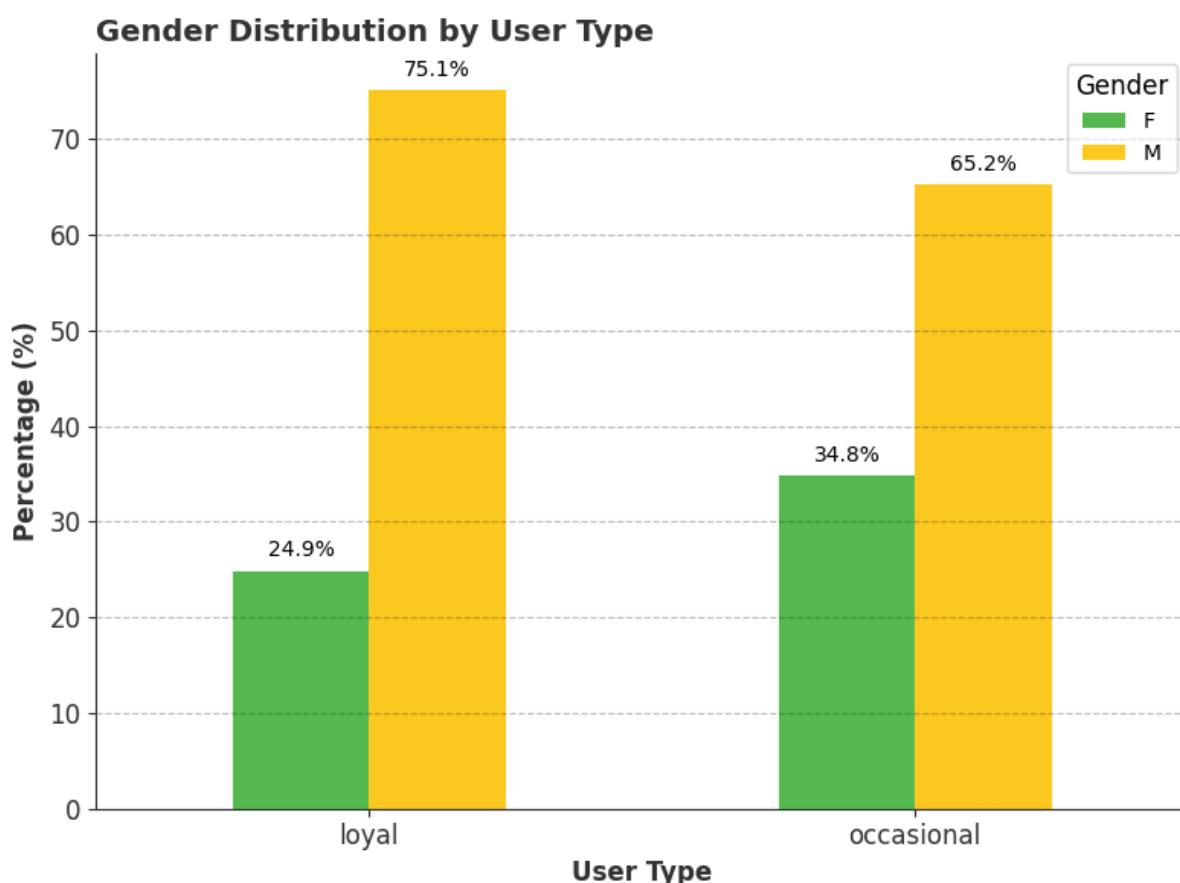
ب) جنسیت کاربران

نمودار (۳۶-۳) توزیع حنсیت کاربران وفادار و موردی را نشان می‌دهد.

- کاربران وفادار حدود ۷۵٪ مرد و ۲۵٪ زن هستند.
- کاربران موردی حدود ۶۵٪ مرد و ۳۵٪ زن هستند.

با توجه به آزمون کای دو و مقدار $p\text{-value} < 0.05$, این تفاوت از نظر آماری معنادار است.

نتیجه: با توجه به اینکه سهم خانم‌ها در کاربرهای وفادار کمتر از موردی است، به عنوان یک استراتژی می‌توان روی وفادارسازی خانم‌ها تمرکز بالاتری کرد.



شكل (۳۶-۳) مقایسه جنسیت کاربران موردی و وفادار

ج) میانگین مدت زمان استفاده از دوچرخه

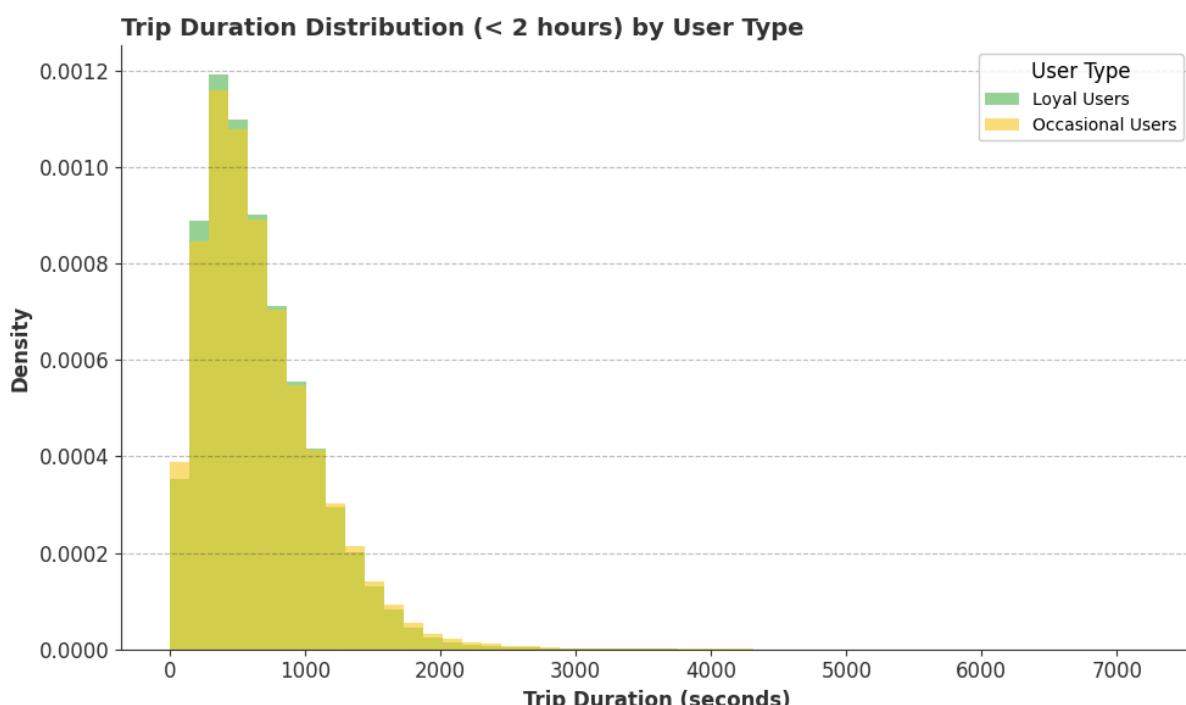
با استفاده از آزمون Mann-Whitney U تفاوت معناداری بین دو گروه مشاهده شد ($p\text{-value} < 0.001$). کاربران وفادار به طور میانگین کمتر از کاربران موردی از دوچرخه‌ها استفاده می‌کنند (در زمان هر سفر). میانگین و میانه هر دو تفاوت مشابهی را نشان می‌دهند. کاربران وفادار احتمالاً از دوچرخه برای رفت‌وآمدهای کوتاه و روزمره (مثل رفتن به محل کار یا خرید) استفاده می‌کنند که سفرهای کوتاه‌تری دارد و کاربران موردی معمولاً سفرهای طولانی‌تری دارند، احتمالاً در تعطیلات، گردش یا شرایط خاص از سیستم استفاده می‌کنند. و نشان‌دهنده عادت‌های استفاده سریع‌تر و کوتاه‌تر کاربران وفادار است که ممکن است از سرویس دوچرخه اشتراکی برای مسیرهای ثابت کاری یا روزمره استفاده نمایند.

میانگین مدت استفاده (ثانیه):

- کاربران وفادار: 654.8
- کاربران موردي: 673.2

میانه مدت استفاده (ثانیه):

- کاربران وفادار: 568.0
- کاربران موردي: 577.0



شکل (۳۷-۳) مقایسه مدت زمان استفاده از دوچرخه در کاربران موردی و وفادار

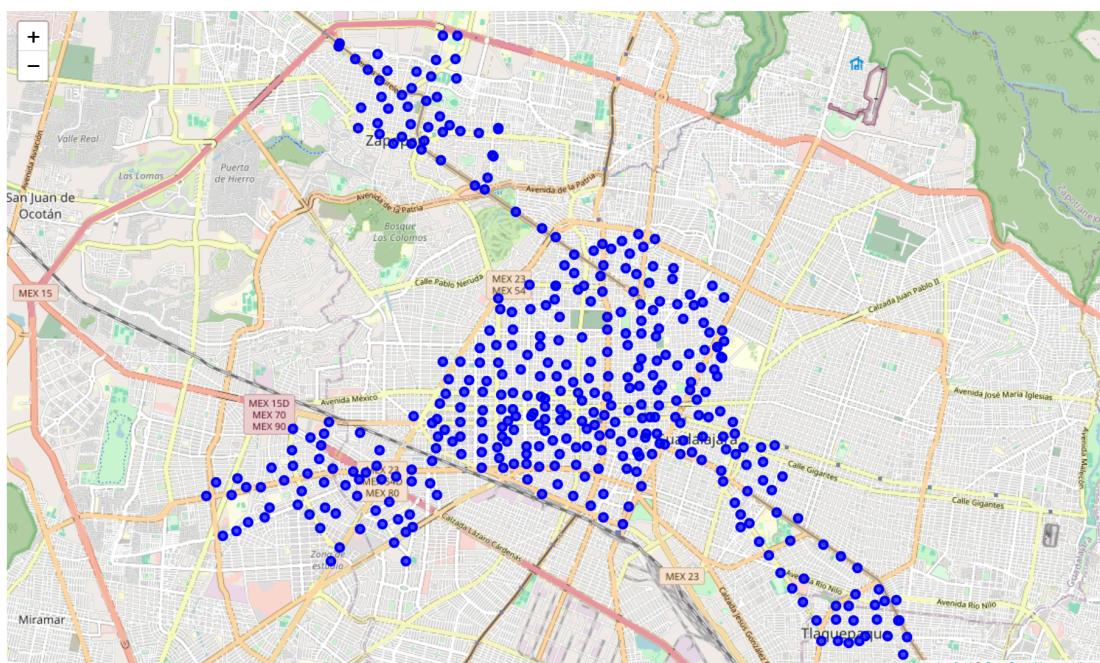
۲.۳.۴.۳ یافته‌ها و توصیه‌هایی برای تیم بازاریابی

- **کمپین‌های وفادارسازی برای خانم‌ها:** با توجه به نرخ پایین‌تر وفاداری در خانم‌ها می‌توان کمپین‌های انگیزشی ویژه برای آن‌ها طراحی کرد (مثلاً تخفیف برای استفاده‌های پیاپی یا در مسیرهای مشخص).
- **هدف‌گیری کاربران با سفرهای بلندمدت (از نظر زمان):** سفرهای طولانی‌تر بیشتر به کاربران موردی تعلق داشته دارد. می‌توان بسته‌هایی برای کاهش هزینه سفرهای طولانی در نظر گرفت تا آن‌ها را به استفاده منظم‌تر تشویق کند.
- **تمرکز بر سنین جوان برای وفادارسازی در بازاریابی:** با توجه به اینکه توزیع کاربران به طور نسبی، به سمت افراد جوان‌تر تمرکز دارد، می‌توان از کمپین‌های وفادارسازی مخصوص کاربرهای رنج سنی جوان‌تر بهره برد.

۳.۱۴.۳ الگوی استفاده کاربران وفادار از ایستگاهها

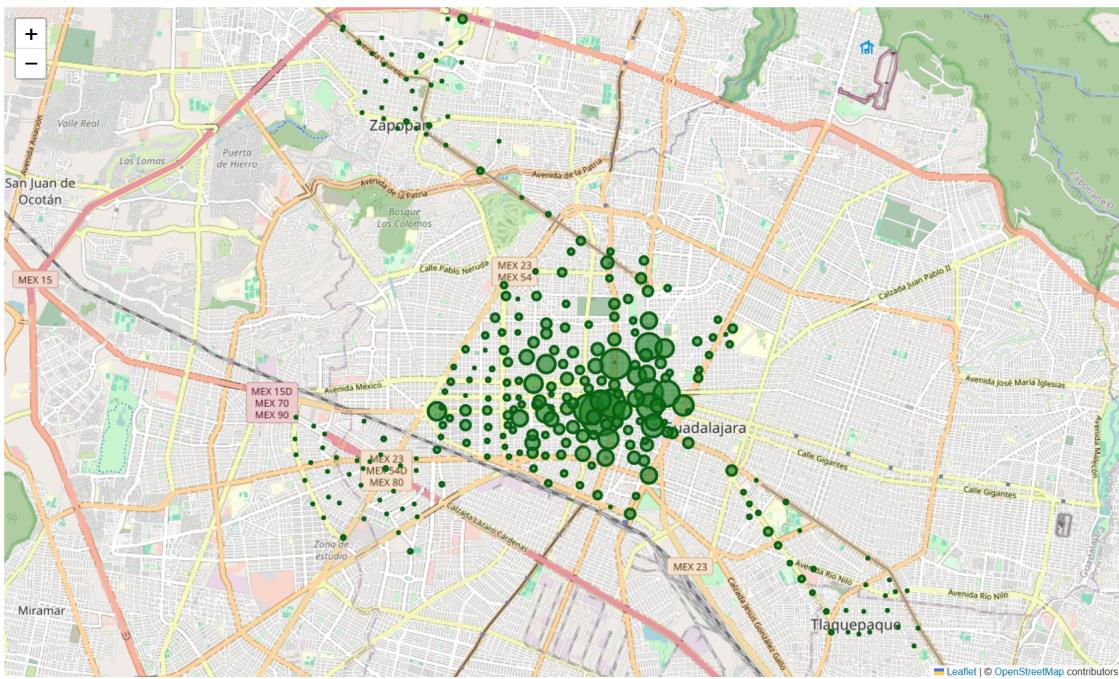
برای بررسی اینکه آیا کاربران وفادار تمایل دارند فقط از تعداد محدودی ایستگاه استفاده کنند یا رفتارشان پراکنده و متنوع است، مجموع استفاده آن‌ها از ایستگاه‌های مختلف (مجموع دفعاتی که هر ایستگاه به عنوان مبدأ یا مقصد مورد استفاده قرار گرفته) مورد بررسی گرفت. سپس به این مسئله پرداخته شد که آیا ایستگاه‌های خاصی سهم بیشتری در سفرهای کاربران وفادار دارند یا این سهم در ایستگاه‌های مختلف مشابه است. در شکل (۳۸-۳)، ابتدا توزیع ایستگاه‌ها در سطح شهر و سپس توزیع ایستگاه‌های مورد استفاده توسط کاربران وفادار قابل مشاهده است.

نمایش کل ایستگاه‌ها در سطح شهر:



شکل (۳۸-۳) نمایش همه ایستگاه‌ها در نقشه شهر

نمایش ایستگاه‌های مورد استفاده کاربران وفادار بر حسب میزان استفاده:



شکل (۳۹-۳) نمایش ایستگاههای پر استفاده توسط کاربران وفادار در نقشه شهر

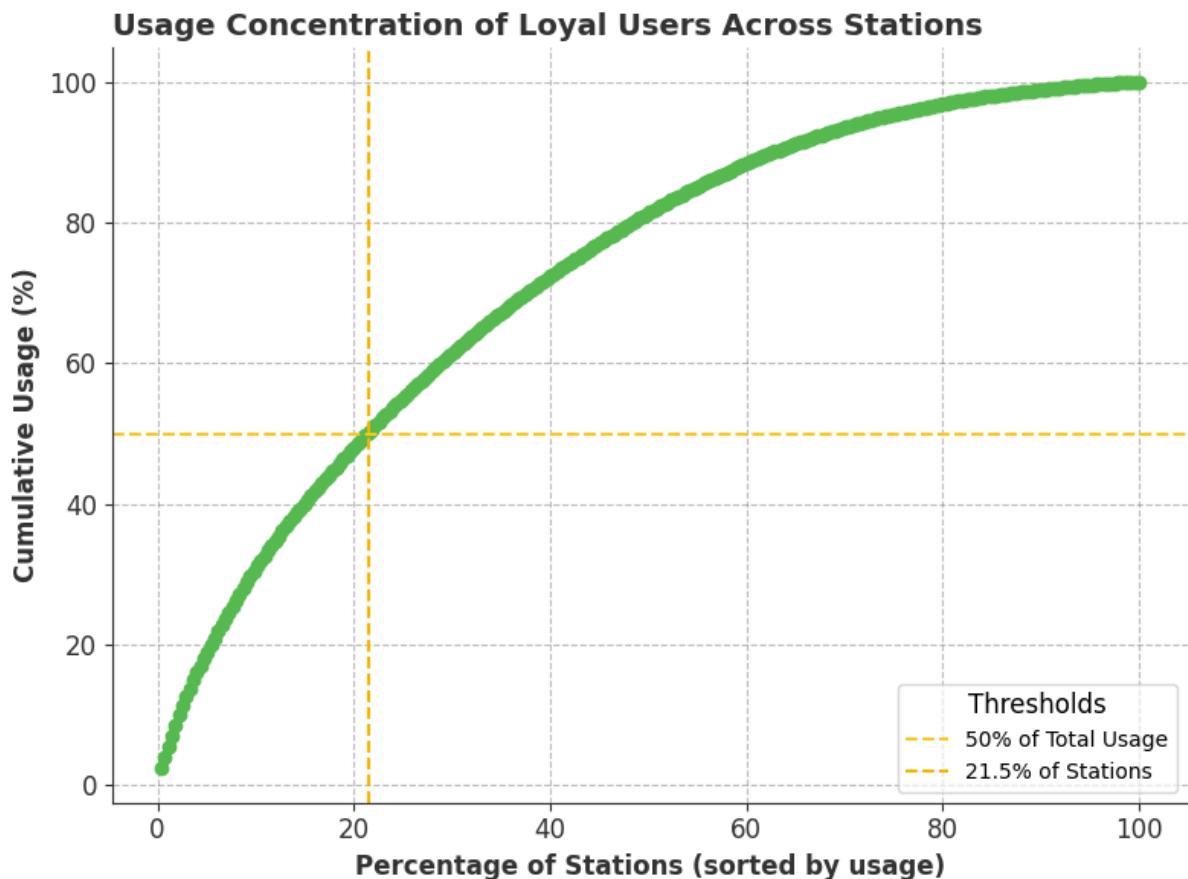
با مقایسه دو نقشه ارائه شده، تفاوت قابل توجهی در الگوی پراکندگی مکانی کل ایستگاهها و ایستگاههای مورد استفاده کاربران وفادار مشاهده می شود. در واقع، تراکم دایره های بزرگتر (نمایانگر دفعات استفاده بیشتر) در نقشه دوم، در نواحی مرکزی شهر متتمرکز است، در حالی که ایستگاههای موجود در نواحی حاشیه ای یا شمال غربی و جنوب شرقی، یا اصلأً توسط کاربران وفادار استفاده نشده اند یا استفاده بسیار محدودی داشته اند. این موضوع نشان می دهد که بخش اعظم استفاده کاربران وفادار از ایستگاهها متتمرکز بر مرکز شهر Guadalajara است.

۱.۳.۴.۳ تحلیل مرکز بر اساس ضریب جینی

برای تکمیل تحلیل مرکز استفاده کاربران وفادار از ایستگاهها، از شاخص استاندارد ضریب جینی (Gini Coefficient) استفاده شد. این شاخص نشان می دهد که آیا استفاده کاربران از ایستگاههای مختلف به طور یکنواخت پراکنده شده یا بیشتر بر روی ایستگاههای خاصی متتمرکز است.

نتایج نشان داد که مقدار ضریب جینی برای توزیع مجموع استفاده از ایستگاهها توسط کاربران وفادار برابر با 0.4515 است. این عدد نشان دهنده تمرکز نسبتاً بالای استفاده است؛ به عبارت دیگر، تعداد محدودی ایستگاه بیشترین حجم سفرها را به خود اختصاص داده اند. در ادامه بررسی می کنیم که این مقادیر دقیقاً چقدر هستند.

نمودار جینی نشان داد که فقط **۵۹ ایستگاه** (معادل حدود **۲۱.۵%** از کل ایستگاههای مورد استفاده کاربران وفادار) توانسته اند **۵۰٪ کل سفرهای این کاربران را پوشش دهند**. این نتیجه بهوضوح نشان دهنده تمرکز بالای استفاده در بین کاربران وفادار است.



شکل (۴۰-۳) تحلیل تمرکز براساس ضریب جینی

۱. ۲.۳.۴.۳ یافته‌ها و بینش‌ها

۱. تمرکز استفاده:

تنها ۲۱.۵٪ از کل سفرهای کاربران وفادار را پوشش می‌دهند. این نشان دهنده‌ی تمرکز بالا در رفتار استفاده‌ی این دسته از کاربران است.

۲. ضریب گینی:

مقدار Gini نشان می‌دهد که توزیع استفاده بین ایستگاه‌ها نابرابر و متمایل به چند نقطه پرتکرار است.

۳. تمرکز جغرافیایی:

کاربران وفادار بیشتر از ایستگاه‌های واقع در مرکز شهر استفاده می‌کنند. این تمرکز مکانی در نقشه به‌وضوح قابل مشاهده است. این تمرکز می‌تواند از نزدیکی محل سکونت یا کار کاربران به این ایستگاه‌ها، مسیرهای پرتردد روزمره، یا دسترسی بهتر این ایستگاه‌ها به خدمات حمل و نقل عمومی نشأت گرفته باشد.

۴. پیشنهاد عملیاتی:

استراتژی‌های نرخ‌گذاری، تبلیغات و بهینه‌سازی خدمات می‌توانند با تمرکز بر این ایستگاه‌های پرتکرار، اثربخش‌تر و مقرن‌به‌صرفه‌تر باشند.

۳.۵ شناسایی کاربران با کاهش یا توقف استفاده (سوال ۶)

۱.۵.۳ تشریح مسئله

در چارچوب سیستم اشتراک دوچرخه‌ی MiBici که در شهر گوادالاخارا، مکزیک فعالیت می‌کند، تیم بازاریابی با چالشی مهم روبروست: شناسایی کاربرانی که در گذشته فعال بوده‌اند اما با گذر زمان دفعات استفاده‌ی آن‌ها کاهش یافته‌ی به‌طور کامل متوقف شده است. هدف اصلی، احیای مجدد مشارکت این کاربران از طریق ارائه‌ی مشوق‌های مالی است.

این مشوق‌ها قرار است در قالب سه بسته‌ی مختلف طراحی شوند، به‌گونه‌ای که با توجه به میزان فعالیت گذشته‌ی کاربران، به هر دسته از آن‌ها بسته‌ای مناسب اختصاص داده شود. از ما خواسته شده است تا با تحلیل داده‌های مربوط به استفاده‌ی کاربران، آن‌ها را به سه گروه مشخص دسته‌بندی کنیم:

در این پژوهه، با استفاده از داده‌های خام مربوط به سفرهای کاربران سیستم اشتراک دوچرخه‌ی MiBici در شهر گوادالاخارا، هدف این بوده است که کاربران بر اساس تاریخچه‌ی استفاده و روند کاهش فعالیت آن‌ها دسته‌بندی شوند. ابتدا اطلاعات استفاده‌ی کاربران به صورت تجمعی و ویژگی‌محور استخراج شد. سپس با استفاده از تحلیل همبستگی و کاهش ویژگی‌های تکراری، پنج شاخص رفتاری کلیدی انتخاب شد تا از آن‌ها برای خوشبندی کاربران استفاده شود. در نهایت، سه دسته کاربر با عنوان‌ی "فعال"، "در معرض ریش"، و "ریش‌کرده" شناسایی شدند

۲.۵.۳ معرفی داده‌ها و آماده‌سازی اولیه

در این پژوهه از دو مجموعه داده‌ی اصلی استفاده شده است:

۱. داده‌ی ایستگاه‌ها (`stations_with_elevation.csv`)
شامل اطلاعات مربوط به کد، نام، موقعیت جغرافیایی، و ارتفاع ۳۸۲ ایستگاه مختلف MiBici.

۲. داده‌ی سفرها (`Mexico_data.csv`)
شامل ۱۱,۹۶۸,۴۶۰ سفر ثبت‌شده توسط کاربران در بازه‌ی زمانی ۲۰۱۵ تا ۲۰۱۹.

۱.۲.۵.۳ مراحل پردازش و پاکسازی داده‌ها

۱. استانداردسازی و تغییر نام ستون‌ها

در طی مراحل آماده‌سازی داده‌ها، به‌منظور افزایش خوانایی و سهولت در تحلیل، نام ستون‌های اسپانیایی موجود در جدول سفرها به معادله‌ای انگلیسی استاندارد تغییر یافت. به عنوان نمونه، ستون `User_Id` به `Usuario_Id`، ستون `start_time` به `Inicio_del_viaje`، و ستون‌های `Origen_Id` و `Destino_Id` به ترتیب به `origin_id` و `destination_id` تغییر نام یافته‌اند. این اقدام باعث شد ساختار داده‌ها برای تحلیل در محیط‌های پایتون و SQL شفافتر و سازگارتر شود.

۲. بررسی اعتبار زمانی رکوردها

در راستای سیاست عملیاتی MiBici، فقط سفرهایی معتبر تلقی می‌شوند که:

در بازه‌ی زمانی مجاز (بین ۵ صبح تا ۱۲ شب) انجام شده باشند. تاریخ شروع و پایان آن‌ها در یک روز باشد، یا در صورت عبور از نیمه‌شب، تا ساعت ۰۰:۰۰ روز بعد به پایان برسند.

با اعمال این شرایط، از مجموع ۱۱,۹۶۸,۴۶۰ سفر، تعداد ۱۱,۹۶۳,۵۴۷ رکورد به عنوان رکوردهای معتبر شناخته شد و ۴,۹۱۳ سفر (۰٪ کل داده‌ها) به دلیل ناسازگاری زمانی حذف شدند.

نتیجه‌ی این فیلترینگ در قالب جدولی جدید با نام `rides_valid` ذخیره شد تا فقط سفرهای واقعی، قابل اعتماد و تحلیل‌پذیر در ادامه استفاده شوند.

۲.۲.۵.۳ هدف از تحلیل داده‌ها

در گام‌های بعدی، از این داده‌های پاکسازی‌شده استفاده شد تا شاخص‌هایی رفتاری برای هر کاربر استخراج گردد، از جمله:

روند تغییر دفعات سفر در ماههای متوالی
فاصله زمانی بین سفرها
میزان استفاده در ماههای اخیر در مقایسه با گذشته
و سایر فاکتورهای زمانی و آماری مرتبط

در نهایت این شاخص‌ها در قالب یک جدول به نام `unique_users` خلاصه شدند تا امكان خوشبندی کاربران و پاسخ به سؤال تحقیق فراهم شود.

۳.۵.۳ متدولوژی و معیارهای استفاده شده

برای شناسایی کاربران دوچرخه اشتراکی MiBici که در طول زمان استفاده‌ی خود را کاهش داده‌اند یا به‌طور کامل متوقف کرده‌اند، ما نیاز به تحلیل دقیق رفتار گذشته‌ی کاربران داشتیم. به این منظور، ابتدا با استخراج و مهندسی ویژگی‌های رفتاری از جدول تراکنش‌های کاربران، یک جدول پروفایل کاربری با عنوان `unique_users` ساختیم که در آن هر سطر نماینده‌ی یک کاربر و ستون‌ها شامل شاخص‌های متنوعی از الگوی استفاده‌ی او در طول زمان بودند.

۱.۳.۵.۳ مهندسی ویژگی‌ها

در مجموع ۲۰ ویژگی عددی و زمانی برای هر کاربر استخراج شد که طیف وسیعی از ابعاد رفتاری را پوشش می‌دادند. این ویژگی‌ها شامل موارد زیر بودند:

۱.۱.۳.۵.۳ ویژگی‌های پایه‌ای:

۱. `user_id`: شناسه یکتا برای هر کاربر که به عنوان کلید اصلی در تمامی جداول تحلیلی به کار رفته است.

۲. `birth_year`: سال تولد کاربر که از آخرین ثبت موجود استخراج شده است. این مقدار برای برخی کاربران ممکن است تهی (NULL) باشد و بیشتر برای تحلیل‌های جمعیتی کاربرد دارد.

۳. **ride_count**: تعداد کل سفرهای معتبر که هر کاربر انجام داده است. این ویژگی معیاری کلی از مشارکت بلندمدت کاربر با سامانه MiBici را ارائه می‌دهد.

۴.۱.۳.۵.۳ شاخص‌های فعالیت زمانی:

۴. **first_ride**: زمان نخستین سفر ثبت‌شده کاربر. این شاخص شروع تعامل کاربر با سیستم را نشان می‌دهد.

۵. **last_ride**: زمان آخرین سفر انجام‌شده کاربر که نمایانگر پایان (یا آخرین نشانه) از مشارکت اوست.

۶. **active_days**: تعداد روزهای منحصربه‌فردی که کاربر حداقل یک سفر انجام داده است. شاخصی از میزان پراکندگی فعالیت در طول سال‌ها.

۷. **active_months**: تعداد ماههای متمایزی که کاربر در آن‌ها فعال بوده است. این ویژگی پایداری رفتار در بازه زمانی بلند را مشخص می‌کند.

۸. **days_since_last_ride**: تعداد روزهایی که از آخرین سفر کاربر تا تاریخ ۰۱/۰۱/۲۰۲۰ گذشته است. هر چه این عدد بیشتر باشد، احتمال ترک سرویس بیشتر است.

۴.۱.۳.۵.۴ شدت و فراوانی استفاده:

۹. **avg_rides_per_month**: میانگین سفرهای کاربر در ماههای فعالش. شاخصی از شدت استفاده در طول دوره فعالیت.

۱۰. **ride_in_last_3m**: تعداد سفرهایی که کاربر در سه ماه پایانی سال ۲۰۱۹ (اکتبر تا دسامبر) انجام داده است. این شاخص به عنوان معیاری از فعالیت اخیر استفاده شده است.

۱۱. **ride_in_last_6m**: تعداد سفرها در شش ماه پایانی سال ۲۰۱۹. نشان‌دهنده روند استفاده در نیمه دوم سال است.

۱۲. **recent_monthly_avg**: میانگین ماهانه سفرهای کاربر در سه ماه پایانی سال ۲۰۱۹. این مقدار نسبت به **ride_in_last_3m** نرمال‌شده‌تر بوده و قابلیت مقایسه میان کاربران را بهبود می‌بخشد.

۱۳. **lifetime_monthly_avg**: میانگین تعداد سفر ماهانه در کل ماههای فعال کاربر. شاخصی از سطح کلی تعامل کاربر در طول عمر عضویتش در سامانه.

۴.۱.۳.۵.۵ شاخص‌های روند و افت استفاده:

۱۴. **peak_monthly_rides**: بیشترین تعداد سفرهایی که کاربر در یک ماه خاص انجام داده است. نشان‌دهنده قله فعالیت کاربر است.

۱۵. **peak_month**: ماهی که بیشترین استفاده توسط کاربر در آن ثبت شده است. به عنوان نقطه اوج الگوی استفاده تلقی می‌شود.

۱۶. به نسبت میانگین فعالیت اخیر (سه ماه پایانی ۲۰۱۹) بیشترین فعالیت ماهانه. این ویژگی برای تشخیص میزان افت استفاده نسبت به دوران اوج بسیار کلیدی است.

۱۷. درصد افت استفاده نسبت به ماه اوج. اگر برابر با ۱۰۰٪ باشد، کاربر استفاده خود را به صفر رسانده است.

۵.۱.۳.۵.۳ شاخص‌های نوسان و پایداری رفتاری:

۱۸. **slope_monthly_usage**: شب خط رگرسیون بین ماههای فعالیت و تعداد سفرها. این مقدار مثبت باشد نشان‌دهنده روند افزایشی، منفی باشد روند کاهشی، و نزدیک به صفر روند پایدار را نشان می‌دهد.

۱۹. انحراف معیار تعداد سفرهای ماهانه. شاخصی از نوسان و غیرپایداری در رفتار کاربر است.

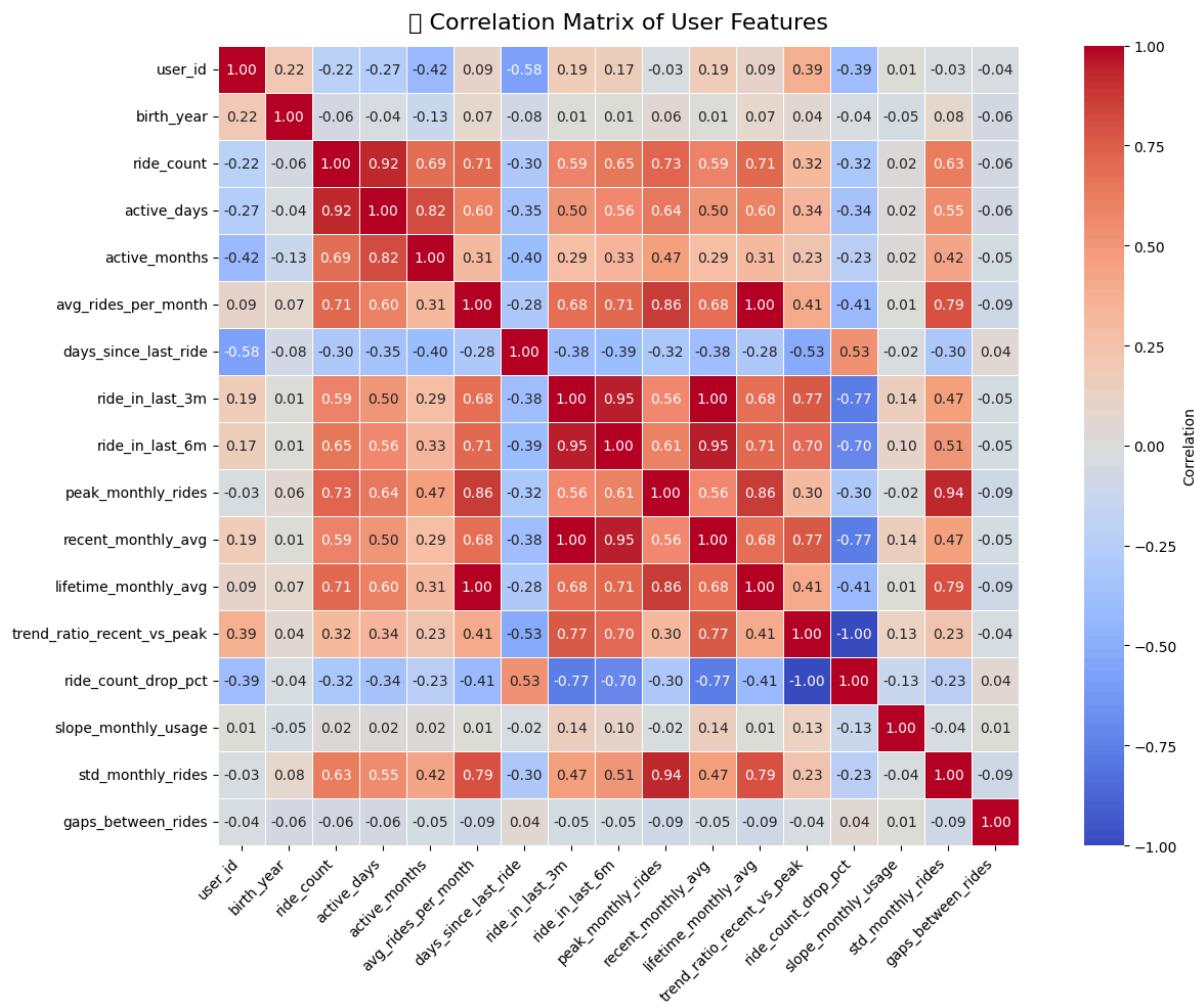
۲۰. **gaps_between_rides**: میانه فاصله روزانه بین دو سفر متوالی کاربر. مقدار بالاتر نشان‌دهنده وقفه‌های طولانی‌تر بین استفاده‌هاست.

۴.۵.۳ تحلیل همبستگی و انتخاب ویژگی‌های رفتاری برای خوشبندی کاربران

پس از استخراج ۲۰ ویژگی متنوع برای هر کاربر در جدول نهایی `unique_users`، گام بعدی در مسیر تحلیل، بررسی روابط آماری بین این متغیرها و کاهش ابعاد ویژگی‌ها برای مرحله خوشبندی بود. هدف از این گام، شناسایی ویژگی‌های تکراری (Redundant)، هم خط (Collinear) یا غیر مؤثر، و در مقابل، انتخاب ویژگی‌هایی متمایز و معنادار در راستای هدف بازاریابی (تشخیص افت یا ترک استفاده توسط کاربران) بود.

۱.۴.۵.۳ استفاده از ماتریس همبستگی

برای این منظور، از ماتریس همبستگی پیرسون (Pearson Correlation Matrix) بین تمامی ویژگی‌های عددی استفاده شد. این ماتریس روابط خطی بین هر دو ویژگی را به صورت عددی بین -۱ تا +۱ نمایش می‌دهد. در نمودار حرارتی حاصل، رنگ‌های قرمز نشان‌دهنده همبستگی مثبت بالا، رنگ‌های آبی نشانه همبستگی منفی، و رنگ‌های سفید/خنثی بیانگر همبستگی ناچیز هستند.



شکل (۴۱-۳) ماتریس همبستگی پیرسون برای همه ویژگی‌های عددی

۱.۱.۴.۵.۱۳ یافته‌های کلیدی از ماتریس همبستگی:

۱. ویژگی‌های شدیداً همبسته (تکراری)

تحلیل ماتریس همبستگی نشان داد که برخی از ویژگی‌ها به شدت با یکدیگر همبستگی دارند و اطلاعات تقریباً مشابهی را منتقل می‌کنند. در راستای جلوگیری از افزونگی و بهبود تفسیر مدل، فقط یکی از این ویژگی‌ها حفظ شد:

نتیجه‌گیری	ضریب همبستگی	زوج ویژگی
تنها یکی از آن‌ها حفظ شود	0.92	ride_count و active_days

کاملاً مشابه – یکی حذف شد	1.00	avg_rides_per_month و lifetime_monthly_avg
یکی برای نمایندگی فعالیت اخیر حفظ شد	0.95	ride_in_last_3m و recent_monthly_avg
وارونگی کامل – تنها یکی کافی است	-1.00	trend_ratio_recent_vs_peak و ride_count_drop_pct

جدول(۱۱-۳) ویژگی های وابسته

۲. روابط منفی معنادار

برخی ویژگی ها دارای همبستگی منفی قابل توجه با شاخص های فعالیت بودند که نشانه ای از کاهش استفاده یا احتمال ترک کاربر تلقی شدند:

تفسیر	ضریب همبستگی	زوج ویژگی
کاربرانی که اخیراً سوار شده اند، فاصله کمتری با آخرین استفاده دارند	-0.77	days_since_last_ride و ride_in_last_3m
کاربران دارای افت شدید، مدت طولانی تری از آخرین استفاده شان می گذرد	-0.53	days_since_last_ride و trend_ratio_recent_vs_peak

جدول(۳۲-۳) روابط منفی معنادار

۳. ویژگی های غیرهمبسته (مستقل و ارزشمند)

برخی ویژگی ها دارای همبستگی پایین با سایر متغیرها بودند و اطلاعات منحصر به فردی درباره روند یا رفتار کاربر ارائه می دهند. این ویژگی ها برای خوشه بندی بسیار ارزشمند تلقی شدند:

(slope_monthly_usage: روند بلند مدت استفاده (مثبت، منفی یا ثابت)

ride_count: سابقه کلی تعامل کاربر

۲.۴.۵.۳ ویژگی‌های انتخاب شده برای خوشبندی کاربران

با تلفیق نتایج آماری فوق و در نظر گرفتن اهداف تجاری مسئله، پنج ویژگی زیر به عنوان نماینده‌گان مؤثر برای خوشبندی انتخاب شدند:

ویژگی	توضیح
ride_count	حجم کلی استفاده کاربر از سامانه در طول زمان
days_since_last_ride	فاصله زمانی تا آخرین استفاده (معیار «جدید بودن» تعامل)
ride_in_last_3m	شدت استفاده در سه ماه پایانی سال ۲۰۱۹ (فعالیت اخیر)
trend_ratio_recent_vs_peak	نسبت استفاده اخیر به اوج تاریخی – شاخص افت استفاده
slope_monthly_usage	روند ماهانه سفرها – افزایشی، کاهشی یا پایدار

جدول(۱۲-۳) ویژگی‌های انتخاب شده

این ویژگی‌ها به شکلی متعادل نمایانگر خصوصیات زیر می‌باشند:

- شدت و حجم گذشته‌ی استفاده
- میزان فعالیت اخیر یا ترک تعامل
- جهت‌گیری رفتاری در طول زمان
- پایداری یا افت استفاده

این فرآیند تحلیل همبستگی و انتخاب ویژگی، نقش کلیدی در موفقیت خوشبندی ایفا کرد. ویژگی‌های انتخاب شده نه تنها همپوشانی آماری ندارند بلکه از نظر تفسیری نیز معنادار و در راستای هدف اصلی پروژه (تشخیص کاربران ترک‌کرده یا در معرض ترک) هستند. این مجموعه ویژگی‌ها، مبنای قدرتمندی برای اعمال الگوریتم KMeans و تقسیم‌بندی کاربران MiBici به سه گروه رفتاری کاربردی فراهم نمودند.

۳.۴.۵.۳ روش خوشه‌بندی کاربران بر اساس الگوهای رفتاری

پس از انتخاب نهایی پنج ویژگی رفتاری کلیدی و انجام تحلیل‌های آماری، گام بعدی اجرای یک الگوریتم خوشه‌بندی بدون نظارت برای گروه‌بندی کاربران سامانه MiBici بر اساس الگوهای استفاده آن‌ها در بازه زمانی ۲۰۱۵ تا ۲۰۱۹ بود. هدف این بخش، شناسایی خوشه‌هایی از کاربران با رفتار مشابه از نظر میزان و روند استفاده از دوچرخه‌هاست؛ به‌گونه‌ای که امکان هدف‌گذاری مجزا برای هریک از گروه‌ها فراهم شود.

۱.۳.۴.۵.۳ مراحل پیاده‌سازی خوشه‌بندی

۱. استخراج و آماده‌سازی داده‌ها

ابتدا جدول unique_users از پایگاه داده بارگذاری و پنج ویژگی منتخب برای خوشه‌بندی استخراج شد

۲. نرم‌السازی ویژگی‌ها

از آنجا که الگوریتم‌های مبتنی بر فاصله (از جمله KMeans) نسبت به مقیاس داده‌ها حساس هستند، ویژگی‌ها با استفاده از روش Z-Score Normalization استاندارد شدند. این کار با استفاده از StandardScaler در کتابخانه Scikit-learn انجام شد و باعث شد میانگین داده‌ها صفر و انحراف معیار آن‌ها یک شود.

۳. اعمال الگوریتم KMeans

برای خوشه‌بندی، از الگوریتم معروف KMeans استفاده شد. در این روش، فرض شد که کاربران را می‌توان به سه خوشه (k=3) تقسیم کرد، که در راستای درخواست اولیه تیم بازاریابی برای تقسیم کاربران به سه دسته رفتاری است.

۴. کاهش ابعاد برای مصورسازی

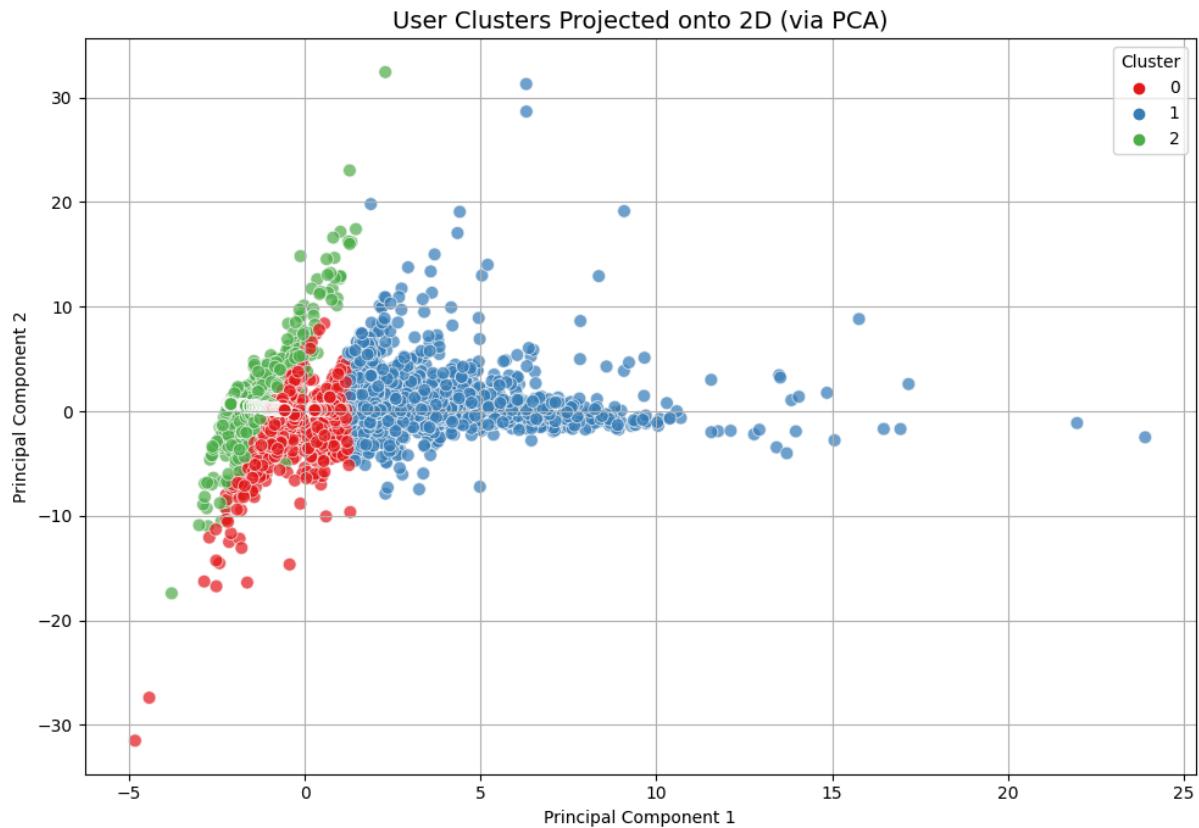
برای نمایش دو بعدی ساختار خوشه‌ها، از تکنیک تحلیل مؤلفه‌های اصلی (PCA) استفاده شد. این روش داده‌های پنج بعدی را به دو مؤلفه اصلی (PC1 و PC2) کاهش داد که بیشترین واریانس داده‌ها را حفظ می‌کنند. سپس نمودار پراکندگی کاربران بر اساس این دو مؤلفه و بر حسب رنگ خوشه رسم شد.

فرآیند خوشه‌بندی با استفاده از الگوریتم KMeans و داده‌های مهندسی شده، کاربران را به سه گروه مجزا تقسیم کرد که در مراحل بعدی، هر خوشه با تحلیل‌های آماری و تصویری مورد بررسی قرار گرفت. این رویکرد امکان تعریف بسته‌های تشویقی متفاوت برای گروه‌های رفتاری مشخص (کاربران فعال، در حال افت، و غیرفعال) را فراهم می‌سازد.

۵.۵.۳ یافته‌ها و بینش‌ها:

پس از ساخت جدول ویژگی‌های رفتاری کاربران و خوشه‌بندی آن‌ها بر اساس الگوهای استفاده از دوچرخه در سال‌های ۲۰۱۵ تا ۲۰۱۹، هدف این بخش تحلیل خروجی‌های حاصل از الگوریتم خوشه‌بندی و ارائه بینش‌های عملیاتی برای تیم بازاریابی است. این تحلیل با تفسیر گرافیکی نتایج شروع شده و در ادامه به بررسی آماری خوشه‌ها و پیشنهاد اقدامات برای هر گروه از کاربران می‌پردازد.

۱.۵.۵.۳ تحلیل نمودار PCA (خوشه‌بندی)



شکل (۴۲-۳) ماتریس همبستگی پیرسون برای همه ویژگی‌های عددی

در شکل بالا، سه خوشی به دست آمده از الگوریتم KMeans با استفاده از تکنیک تحلیل مؤلفه‌های اصلی (PCA) به فضای دوبعدی نگاشته شده‌اند. این نگاشت، نمایی ساده‌شده از فضای پنج بعدی ویژگی‌ها را فراهم می‌کند و امکان تفسیر بصری ساختار خوشی‌ها را می‌دهد.

۱.۱.۵.۵.۳ مشاهدات کلیدی:

در این نمودار، سه خوشی مجزا به خوبی از یکدیگر تفکیک شده‌اند. این جداسازی نشان‌دهنده‌ی کفايت و دقت ویژگی‌های انتخاب شده برای تفکیک کاربران است.

خوشی قرمز (Cluster 0) در سمت چپ و به صورت فشرده قرار دارد که نشان‌دهنده کاربرانی با فعالیت کم یا کاهش‌یافته است.

خوشی آبی (Cluster 1) گسترده‌ی زیادی در امتداد مؤلفه اول دارد و بیشتر در بخش پایینی محور PC2 دیده می‌شود. این الگو ممکن است بیانگر کاربران با رفتارهای متغیر یا در حال افت باشد.

خوشی سبز (Cluster 2) بیشترین پراکندگی را در امتداد مؤلفه اول دارد و می‌تواند نمایانگر کاربران با میزان استفاده بالا یا رفتارهای نامنظم باشد.

به طور کلی، این نگاشت دوبعدی، اعتبار و کیفیت خوشی‌بندی را از نظر ساختار داخلی تأیید می‌کند و پایه‌ای مناسب برای تحلیل آماری و سیاست‌گذاری هدفمند برای هر خوشی فراهم می‌سازد.

۲.۱.۵.۵.۳ تحلیل آماری خوشها و تعریف رفتار هر گروه

برای درک بهتر ویژگی‌های رفتاری هر خوش، میانگین مقادیر پنج ویژگی کلیدی در هر گروه محاسبه شد. این ویژگی‌ها شامل حجم کلی استفاده، زمان آخرین استفاده، استفاده اخیر، نسبت استفاده فعلی به بیشینه، و روند تغییرات در طول زمان هستند.

خوشه	تعداد کاربران	میانگین تعداد سفر	میانگین روز از آخرین سفر	سفرهای سه ماهه اخیر	نسبت استفاده فعلی به بیشینه	شیب روند استفاده
0	26,782 نفر	196.81	169.8 روز	7.82 سفر	0.10	-1.22
1	8,417 نفر	675.20	8.61 روز	107.15 سفر	0.60	+0.91
2	14,631 نفر	68.99	982.84 روز	0.00 سفر	0.00	-0.45

جدول(۳-۱۳) آمار هر خوش رفتار هر گروه

۳.۱.۵.۵.۳ تعریف خوشها و تفسیر رفتاری

بر اساس تحلیل آماری بالا، کاربران به سه گروه رفتاری اصلی تقسیم شدند:

خوشه ۱ – کاربران فعال (Active Users)

تعداد سفر بسیار بالا و استفاده منظم

فاصله‌ی بسیار کم تا آخرین استفاده (میانگین ۸ روز)

رونده رو به رشد در استفاده (شیب مثبت)

این کاربران به وضوح مشتریان وفادار هستند و می‌توانند با خدمات ویژه، اشتراک‌های بلندمدت یا پاداش‌های وفاداری حفظ شوند.

خوشه ۰ – کاربران در معرض ریزش (At-Risk Users)

تعداد سفر متوسط ولی کاهش شدید در استفاده اخیر

فاصله‌ی نسبتاً زیاد از آخرین استفاده (حدود ۵ ماه)

روند منفی واضح (شیب -1.22)

این کاربران هنوز ارتباط خود را با سرویس قطع نکرده‌اند ولی در حال فاصله گرفتن‌اند. ارسال یادآور، پیشنهادات تخفیف یا ایمیل‌های بازگشتی برای این گروه مؤثر خواهد بود.

خوشه ۲ – کاربران ریزش‌کرده (Churned Users)

تعداد سفر پایین

بیش از ۹۸۰ روز فاصله از آخرین سفر

هیچ استفاده‌ای در سه ماه پایانی سال ۲۰۱۹

روند استفاده به شدت کاهشی

این کاربران عملأً از سیستم خارج شده‌اند. برای جذب مجدد آن‌ها نیاز به مشوق‌های قوی مانند سفر رایگان، عضویت آزمایشی یا کمپین‌های بازگشت وجود دارد.

۴.۱.۵.۵.۳ نتیجه‌گیری از خوشه‌بندی

با استفاده از پنج ویژگی منتخب و الگوریتم KMeans، کاربران MiBici به سه دسته‌ی مشخص و قابل اقدام تقسیم شدند. این خوشه‌بندی پایه‌ای برای اجرای استراتژی‌های بازاریابی تفکیک‌شده، بهینه‌سازی هزینه‌های تبلیغات و ارتقاء نرخ بازگشت کاربران خواهد بود.

۲.۵.۵.۳ تحلیل تصویری دوبه‌دوی ویژگی‌ها بین خوشه‌ها

برای درک بهتر تفکیک خوشه‌ها، از بین پنج ویژگی کلیدی انتخاب‌شده برای خوشه‌بندی، تمامی ترکیب‌های دوبه‌دو ترسیم شدند. این ۱۰ نمودار پراکندگی، نحوه‌ی توزیع داده‌ها و مرزبندی طبیعی بین سه گروه کاربری را به خوبی نشان می‌دهد.

مشاهدات کلیدی از نمودارهای دوبه‌دو

`ride_count vs days_since_last_ride .1`

○ کاربران فعال (قرمز) در ناحیه‌ای با تعداد سفر بالا و فاصله‌ی زمانی کم از آخرین سفر متمرکز شده‌اند.

○ کاربران ریزش‌کرده (آبی) در نواحی با تعداد کم سفر و وقفه‌های طولانی قرار دارند.

○ کاربران در معرض ریزش (سبز) در ناحیه‌ی بینایینی با حجم استفاده متوسط و تأخیر متوسط قرار گرفته‌اند.

`ride_in_last_3m vs .2`

○ در بیشتر نمودارها، کاربران فعال به‌وضوح دارای استفاده‌ی اخیر بالا هستند، که آن‌ها را از سایر گروه‌ها متمایز می‌کند.

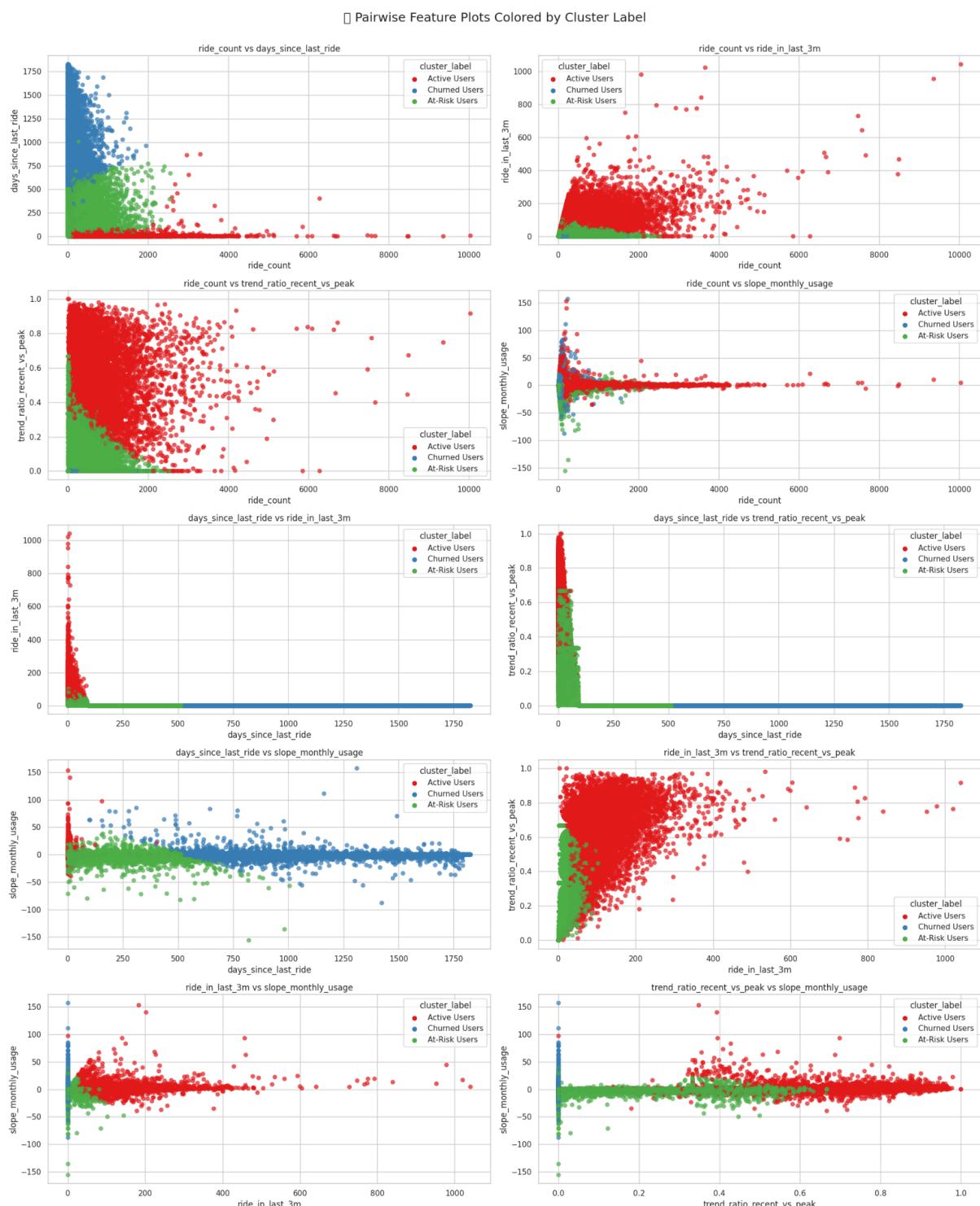
کاربران ریزش کرده تقریباً در تمامی این نمودارها به صورت خوش‌های در ناحیه‌ی صفر قرار گرفته‌اند.
○
(بدون فعالیت اخیر).

trend_ratio_recent_vs_peak vs slope_monthly_usage .۳

این نمودار به‌وضوح کاربران فعال با روند رو به رشد و نسبت استفاده‌ی بالا را از دیگران جدا می‌کند.
○
کاربران در معرض ریزش دارای روند منفی و نسبت استفاده‌ی بسیار کم هستند.
○

slope_monthly_usage vs .۴ سایر متغیرها

خوشبندی در این محور اغلب خطی یا قطری دیده می‌شود. کاربران فعال شیب مثبت و سایر گروه‌ها اغلب شیب منفی دارند.
○



شکل (۳-۱۴) نمودار های دوبه‌دوی ویژگی‌ها بین خوش‌ها

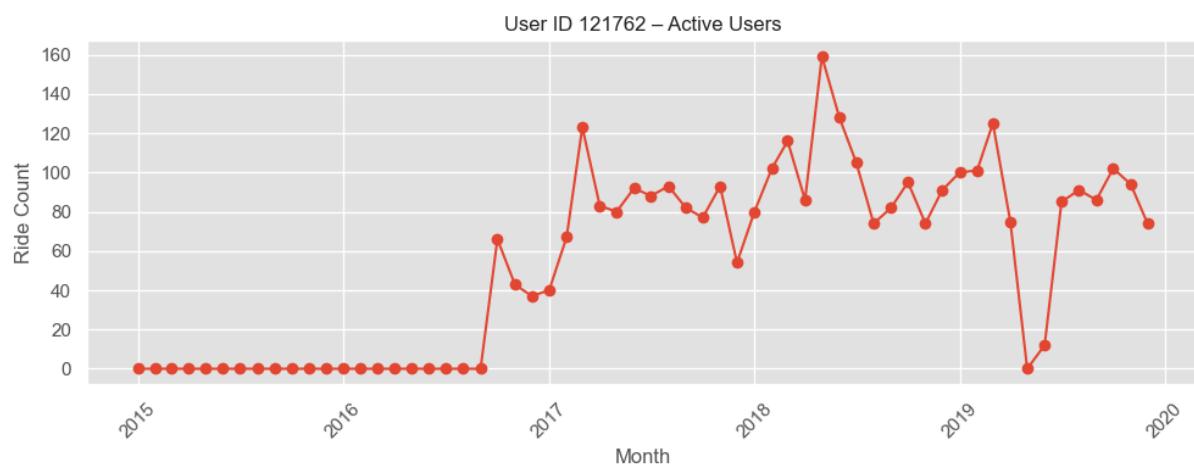
این نمودارها صحت خوش‌بندی را از لحاظ بصری تأیید می‌کنند. کاربران در سه دسته‌ی رفتاری کاملاً متمایز قرار دارند:

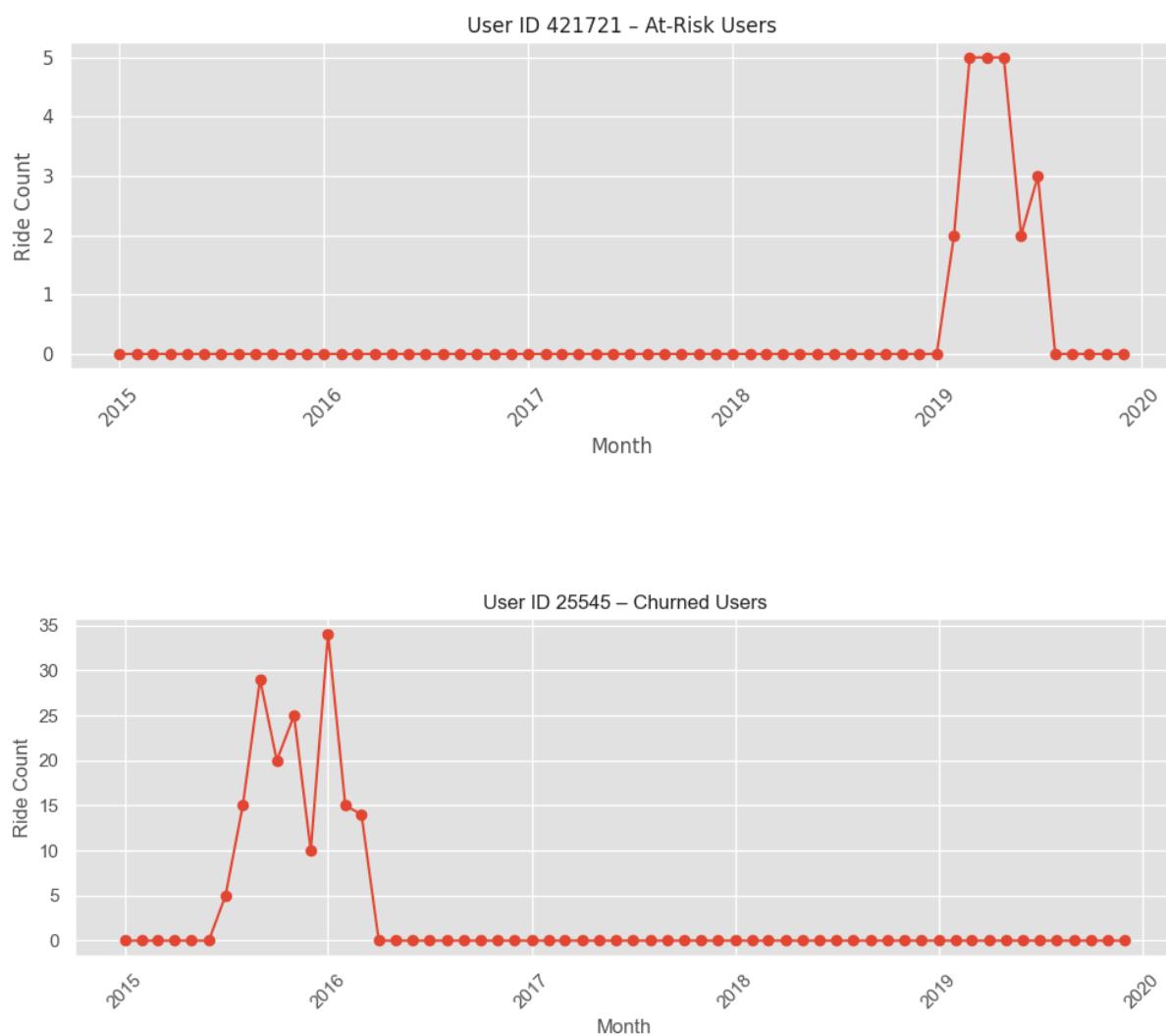
- کاربران فعال: رفتار صعودی، استفاده‌ی مکرر، بدون وقفه‌های طولانی
- کاربران در معرض ریزش: حجم استفاده‌ی متوسط ولی با روند کاهشی و وقفه‌های روبه‌افزایش
- کاربران ریزش‌کرده: فعالیت بسیار کم، شبیب منفی، بدون استفاده‌ی اخیر

با توجه به جداسازی واضح در این نمودارها، می‌توان از این ویژگی‌ها برای طراحی دقیق کمپین‌های بازاریابی و بازگشت کاربران بهره گرفت.

۳.۵.۵.۳ بررسی رفتار زمانی کاربران نمونه در هر خوش

برای درک بهتر ماهیت هر خوش و اعتبارسنجی نتایج خوش‌بندی، رفتار زمانی سه کاربر نمونه از هر خوش انتخاب شده و نمودار تعداد سفرهای ماهانه‌ی آن‌ها طی سال‌های ۲۰۱۵ تا ۲۰۱۹ رسم گردید. این نمودارها نمایانگر الگوهای زیر هستند:





شکل (۴۴-۳) نمودار رفتار زمانی کاربران نمونه در هر خوش

کاربر خوش Active Users نشان‌دهنده استفاده‌ی منظم و پُرچم در اکثر ماه‌های است، با نوسانات طبیعی ولی بدون کاهش شدید.

کاربر خوش At-Risk Users الگوی کاهش تدریجی در میزان استفاده را دارد؛ بهویژه در ماه‌های پایانی فعالیت، روند نزولی مشهود است.

کاربر خوش Churned Users تنها در بازه‌ی کوتاهی فعال بوده و سپس برای مدت طولانی غیرفعال باقی مانده است که نشانه‌ی ترک کامل سیستم است.

این نمودارها بهخوبی با تعاریف نظری هر خوش همخوانی دارند و صحت مدل خوشبندی را تأیید می‌کنند.

۶.۵.۳ راهبردهای بازاریابی برای هر خوش

– کاربران فعال – Active Users

ویژگی‌ها: استفاده‌ی مداوم، حجم بالای سفر، روند افزایشی یا پایدار در طول زمان

هدف: حفظ رضایت و افزایش وفاداری

پیشنهاد بازاریابی:

ارائه‌ی جوايز وفاداري مانند اشتراك رايگان يكماهه يا دسترسی به ايستگاههای ویژه اطلاع‌رسانی زودهنگام درباره‌ی تغییرات سرویس یا تخفیف‌های ویژه

– کاربران در معرض ریزش – At-Risk Users

ویژگی‌ها: سابقه‌ی متوسط یا خوب، اما کاهش قابل توجه در ماههای اخیر

هدف: جلوگیری از ریزش کامل

پیشنهاد بازاریابی:

ارسال پیام‌های شخصی‌سازی‌شده با تحلیل رفتار گذشته ("مدت زیادی گذشته از آخرين بارتون!")

پیشنهاد تخفیف در سفرهای آینده یا بازگشت با بسته‌ی تشویقی (مثلًا ۳ سفر رايگان)

فعال‌سازی کمپین‌های اعتمادسازی مجدد (مانند معرفی تغییرات جدید در ايستگاه‌ها یا اپليکيشن)

– کاربران ریزشی – Churned Users

ویژگی‌ها: استفاده‌ی موقتی در گذشته، بدون هیچ فعالیت در سال ۲۰۱۹

هدف: بازیابی کاربران از دست‌رفته

پیشنهاد بازاریابی:

ارائه‌ی پیشنهاد بازگشت ویژه با عنوان‌هایی مانند «MiBici دلتنگ شماست!»

عرضه‌ی کدهای اعتباری برای اولین سفرهای بازگشت

دعوت به شرکت در رویدادهای رايگان دوچرخه‌سواری یا چالش‌های اجتماعی

این پیشنهادها با توجه به ویژگی‌های هر خوش به گونه‌ای طراحی شده‌اند که حداکثر تطابق با انگیزه‌ها و الگوی رفتاری کاربران داشته باشند و بتوانند نرخ بازگشت، وفاداری و استفاده مجدد از سیستم اشتراك دوچرخه را به‌طور مؤثری افزایش دهند.

۷.۵.۴ جمع‌بندی

در این مطالعه، با تحلیل دقیق داده‌های کاربران سامانه اشتراك دوچرخه MiBici و بررسی الگوهای زمان، رفتاری و آماری آن‌ها، توانستیم یک چارچوب داده‌محور برای شناسایی کاربران با کاهش یا توقف استفاده توسعه دهیم. از طریق ساخت جدول جامع unique_users شامل ۲۰ ویژگی کلیدی، تحلیل همبستگی و انتخاب متغیرهای مناسب، فرآیند خوشبندی کاربران به‌صورت هدفمند انجام شد.

نتایج خوشبندی با الگوریتم KMeans بهوضوح نشان داد که کاربران MiBici به سه گروه اصلی تقسیم می‌شوند:

۱. کاربران فعال که رفتار پایدار و صعودی دارند
۲. کاربران در معرض ریزش که سابقه استفاده دارند ولی روندشان نزولی است
۳. کاربران ریزشی که در ماههای اخیر کاملاً غیرفعال شده‌اند

این تقسیم‌بندی نه تنها از نظر آماری معتبر است (براساس شاخص‌های روند، زمان، و نوسان)، بلکه از نظر رفتار واقعی کاربران نیز از طریق نمودارهای زمانی تأیید شد.
در نهایت، برای هر خوش، یک پیشنهاد بازاریابی هدفمند جهت افزایش مشارکت، حفظ کاربران فعلی، یا بازگرداندن کاربران ریزشی ارائه گردید. این پیشنهادها، در صورت اجرای درست، می‌توانند به شکل معناداری باعث افزایش نرخ بازگشت، رضایت و بهره‌وری اقتصادی سامانه MiBici شوند.

این تحلیل نشان داد که استفاده‌های مؤثر از داده‌های رفتاری کاربران می‌تواند نقش کلیدی در طراحی کمپین‌های بازاریابی هوشمند و مبتنی بر شواهد ایفا کند.

۶.۳ افزایش سود و بهینه‌سازی ناوگان (سوال ۷)

با توجه به گزارش‌های ماهانه سیستم دوچرخه‌رانی شرکت، مشاور مدیرعامل پیشنهاد کرده است که برای افزایش سود، علاوه بر توسعه شبکه و افزایش تعداد ایستگاه‌ها، باید به دو روش زیر توجه نمود:

- (الف) سرمایه‌گذاری بر روی عملیات باز توزیع (جابجایی دوچرخه‌ها بین ایستگاه‌ها)
- (ب) سرمایه‌گذاری بر روی خرید دوچرخه‌های جدید و توسعه زیرساخت‌های مربوط (ایجاد داک و ایستگاه‌های جدید)

با این حال، مدیرعامل نسبت به بهینه بودن استفاده از دوچرخه‌های موجود ابهام دارد؛ به عبارت دیگر، او نمی‌داند که آیا سیستم فعلی از دوچرخه‌های موجود به طور مطلوب استفاده می‌شود یا خیر. از این رو، هدف از این پژوهه، بررسی هر دو رویکرد به همراه یک استراتژی ترکیبی است تا عملکرد، سودآوری و حساسیت نتایج نسبت به فرضیه مورد استفاده به تفصیل تحلیل شود.

۱.۶.۳ ارائه و تحلیل داده‌های استخراج شده از گزارش‌های ماهانه

داده‌های گزارش ماهانه

بر اساس گزارش‌های شرکت در سال ۲۰۱۳ (ژوئن تا دسامبر) موارد زیر استخراج شده است:

جدول (۱۴-۳) اطلاعات به دست آمده از گزارش‌های ماهانه

Lost Revenue	Revenue	Neither Empty Nor (%) Full			Available Bikes	Rides of Month	Rebalance d per Day	Available Stations	Full/Empty Instances	Total Rebalance Bikes	Day	Month
		Avg	Peak	Non Peak								
\$175,143	\$4,203,421	96.0%	96%	96%	5,130	618,572	815	322	17,776	24,448	30	Jun
\$126,242	\$2,398,607	95.0%	95%	95%	5,531	953,872	2,303	328	22,125	71,383	31	JUL
\$190,903	\$2,195,380	92.0%	92%	92%	5,681	1,109,428	1968	332	48,784	60,999	31	Aug
\$96,833	\$1,663,763	94.5%	96%	93%	5,513	1,104,527	2,770	332	77,698	83,085	30	Sep
\$37,419	\$1,031,697	96.5%	97%	96%	5,623	1,087,170	1940	332	91,629	60,126	31	Oct

\$11,584	\$451,762	97. 5%	97%	98%	5,900	701,61 9	1,640	332	42,859	49,211	30	Nov
\$4,788	\$234,600	98. 0%	98%	98%	6,499	460,87 1	1429	332	21,200	44,301	31	Dec
\$642,91 1	\$12,179,23 0	-	-	-	-	-	12,865		321,095	393,553	-	TTL

اطلاعات بدست آمده از بنچمارک‌ها

جداول زیر اطلاعات بدست آمده از بنچمارک‌ها هستند:

جدول (۱۵-۳) ویژگی‌های انواع وسایل نقلیه جابجایی دوچرخه‌ها

سه‌چرخه یا تریلر بار (Cargo-Trike/Trailer)	کامیون بزرگ (Box Truck)	ون حمل و نقل (Van)	ویژگی
۶~ دوچرخه	۴۰-۲۰~ دوچرخه	۲۰-۱۰~ دوچرخه	ظرفیت بار هر بار
۱۶~ بار (تقريباً ۹۶ دوچرخه در روز)	۱۰~ بار (۱۵۰-۳۰۰ دوچرخه در روز)	۱۲-۱۵~ بار (۲۰۰-۱۰۰ دوچرخه در روز)	تعداد بارهای هر شیفت
۲۰۰۰-۱۵۰۰~ دوچرخه	۱۰,۰۰۰-۵,۰۰۰~ دوچرخه	۶۰۰۰-۳,۰۰۰~ دوچرخه	تعداد دوچرخه‌های ماهانه
۷-۵ دلار (عمر ۱۰,۰۰۰-۵,۰۰۰ سال)	۶۰,۰۰۰-۴۰,۰۰۰ دلار (عمر ۷ سال)	۵۰,۰۰۰-۳۰,۰۰۰ دلار (عمر ۷ سال)	هزینه سرمایه‌ای
۳۰~ دلار (شامل تمامي هزينه‌ها)	۵۰-۴۰~ دلار (شامل تمامي هزينه‌ها)	۴۵-۳۰~ دلار (شامل تمامي هزينه‌ها)	هزينه نيروي کار (ساعتني)
۲.۵۰~ دلار	۱.۲۰-۰.۹۰~ دلار	۱.۸۰-۱.۲۰~ دلار	هزينه نيروي کار به ازاي هر دوچرخه
۰.۱۰-۰.۰۵~ دلار به ازاي هر دوچرخه	۰.۱۵-۰.۰۸~ دلار به ازاي هر دوچرخه	۰.۱۵-۰.۰۵~ دلار به ازاي هر دوچرخه	هزينه سوخت و نگهداري
۲.۶۰-۲.۵۵~ دلار	۱.۳۵-۱.۰۰~ دلار	۲.۰۰-۱.۲۵~ دلار	هزينه کل به ازاي هر دوچرخه
داده‌های آزمایش اپراتور سه‌چرخه بار (نيروي کار و استهلاک حداقل)	تحليل هزينه‌های نيروي کار، سوخت و استهلاک برای ناوگان با توزيع در نيوپوريک	بنچمارک هزينه‌های عملياتي (دستمزد راننده و نرخ‌های جابجایي) و محاسبات سوخت/استهلاک	منبع داده

جدول (۱۶-۳) تعداد سفرهای به دست آمده به ازای هر دوچرخه جابجا شده در سیستم‌های مختلف

منبع	تعداد متوسط سفرهای به ازای هر دوچرخه جابجا شده	منبع	ویژگی
مطالعه UNC درباره تأثیرات جابجایی دستی دوچرخه‌ها با ون	۱.۲ – ۱.۱	Shaheen et al. (2011), Transportation Research Record	Capital BikeShare (واشنگتن)
تخمين مبتنی بر مدل، بسته به چگالی ایستگاه‌ها متفاوت است	۱.۴ – ۱.۰	Midgley (2012), Intl. Journal of Sustainable Transportation	'Vélib' (پاریس)
بیشتر (≈ 1.3) در ساعت اوج تقاضا؛ کمتر (≈ 0.8) در ساعت غیر اوج	۱.۳ – ۰.۸	Fishman, Washington & Haworth . (2017), Int. J. Transp. Sci. & Tech	Chicago Divvy (شیکاگو)

جدول (۱۷-۳) ویژگی‌های دوچرخه‌ها و هزینه‌های مرتبط

دوچرخه الکتریکی نقره‌ای (Silver e-bike, 2022)	دوچرخه الکتریکی اورجینال (Original e-bike, 2018)	دوچرخه پدالی استاندارد (Standard pedal bike)	ویژگی
\$۱۴,۰۰۰ (میانه)	\$۲,۵۰۰ (میانگین)	\$1200	هزینه واحد
۵ سال	۷ سال	۱۰ سال	عمر مفید
\$۸۰۰	\$۳۵۷	\$۱۲۰	استهلاک سالانه
\$۴۰۰	\$۲۵۰	\$۱۲۰	هزینه نگهداری سالانه
۴۵۰~ سفر (≈ 15 سفر/روز)	۱۵۰~ سفر (≈ 5 سفر/روز)	۵۰~ سفر (≈ 0.9 سفر/روز)	میانگین تعداد سفر در ماه
هزینه به ازای هر دوچرخه شامل زبرساخت: \$۳,۰۰۰–۵,۰۰۰ (itskrs.its.dot.gov)	محدوده قیمت برای دوچرخه‌های الکتریکی با کیفیت بالا 5,000–2,000\$ (finmodelslab.com)	هزینه واحد: گزارش‌های ناوگان Citi Bike ($\approx \$1,200$) ((en.wikipedia.org))	منبع داده
استفاده و عمر مفید (۱۵ سفر/روز → عمر ۵ سال) (curbed.com)	استفاده و عمر مفید (۵ سفر/روز → عمر ۵ سال) (curbed.com)	قانون نگهداری معمول (%) (finmodelslab.com)	قوانين نگهداری

جدول (۱۸-۳) تعداد سفرهای به ازای هر دوچرخه جدید در سرویس‌های دوچرخه اشتراکی متفاوت

منبع	تعداد سفرها به ازای هر دوچرخه در ماه	تعداد سفرهای اوج به ازای هر دوچرخه در روز	سیستم
۳۰۰,۰۰۰ سفر در ماه $1,100 \div 273 \approx 4\text{ سفر در اوج، اما در تابستان } 4\text{ سفر/روز} \approx 120 \text{ سفر در ماه (malkaram.com)}$	۱۲۰ ≈ سفر	۴ ≈ سفر/روز	Capital Bikeshare (واشنگتن)
۲,۶۹۳ سفر به ازای هر دوچرخه در ماه ژوئن $2024 \Rightarrow 2.69 \times 80 \approx 30 \times 80 \approx 80 \text{ سفر در ماه (en.wikipedia.org)}$	۸۰ ≈ سفر	۲.۷~ سفر/روز (ژوئن) (۲۰۲۴)	Citi Bike (نیویورک)
"پاریس بیشترین تعداد سفر در روز را دارد، با رسیدن به ۸ سفر در روز" (journalistsresource.org)	۲۴۰ ≈ سفر	۸ سفر/روز (سپتامبر)	Vélib' (پاریس)
"هر دوچرخه به طور متوسط ۶ سفر در روز انجام می‌دهد" (toolsforchange.com)	۱۸۰ ≈ سفر	۶ سفر/روز	BIXI (مونترال)

همچنین، درآمد هر سفر دوچرخه‌ای بر اساس استانداردهای صنعتی برابر با 3.50 دلار محاسبه شده (با در اختیار داشتن تعداد کل سفرهای انجام شده در هر ماه و درآمد ماهانه شرکت (منبع گزارشات ماهانه) و میانگین‌گیری ساده عدد ۳.۵ دلار به عنوان درآمد هر سفر محاسبه شد) و ضریب تبدیل به ازای هر جابجایی (حاصلضرب دوچرخه شده به ازای 1.3) منجر به درآمد تقریبی 4.55 دلار در هر جابجایی گردیده است.

همچنین برای سیستم خرید دوچرخه جدید، تولید سفر دوچرخه با دوچرخه‌های "Mix" با ظرفیت 160 سفر در ماه در نظر گرفته شده و محدودیت سطح اشباح سیستم به عنوان حداقل تعداد دوچرخه‌های فعال (۹۰۰۰ دوچرخه) تعیین شده است.

بررسی سطح دسترسی (Availability) و مقایسه با بنچمارک‌ها

سطح دسترسی سیستم به عنوان شاخص کل کیفیت سرویس و عملکرد دوچرخه‌رانی مورد توجه قرار گرفته است. طبق گزارش‌های ماهانه، سطوح اعلام شده به شرح زیر است:

- ژوئن: %96
- ژوئیه: %95
- آگوست: %92
- سپتامبر: %94.5
- اکتبر: %96.5
- نوامبر: %97.5
- دسامبر: %98

از آنجایی که مطالعات بین‌المللی (مانند تحقیق Van Hoeve et al. 2010) و دستورالعمل‌های صنعتی (NACTO/ITDP) استاندارد دسترسی بین 90 تا 95 درصد را به عنوان حداقل سطح قابل قبول و هدف اصلی در نظر گرفته‌اند، ارقام به دست آمده برای سیستم شرکت (که در برخی ماه‌ها حتی به ۹۸٪ می‌رسد) نشان‌دهنده وضعیت

بسیار مطلوبی است.

بنابراین، می‌توان نتیجه گرفت که از نظر سطح دسترسی (Availability) و تضمین کیفیت سرویس، عملیات ری‌بالیانسینگ شرکت با بنچمارک‌های صنعتی همخوانی داشته و وضعیت عملکرد سیستم از لحاظ دسترسی بسیار رضایت‌بخش است.

۲.۶.۳ بهبود بازتوزیع دوچرخه‌ها بین ایستگاه‌ها با سرمایه‌گذاری بر روی وسائل نقلیه

این بخش به بررسی این موضوع می‌پردازد که آیا سرمایه‌گذاری بیشتر بر روی وسائل نقلیه بازتوزیع‌کننده دوچرخه، می‌تواند مشکل عدم تعادل عرضه و تقاضا در ایستگاه‌های مختلف را بهبود بخشد و در نتیجه منجر به افزایش درآمد شود. عدم توازن در تعداد دوچرخه‌ها بین ایستگاه‌های مختلف، یکی از چالش‌های اصلی سیستم‌های دوچرخه اشتراکی است که می‌تواند منجر به از دست رفتن فرصت‌های درآمدی شود. با انتقال دوچرخه‌ها از ایستگاه‌های پر به ایستگاه‌های خالی، می‌توان اطمینان حاصل کرد که دوچرخه‌ها در دسترس کاربرانی که به آن‌ها نیاز دارند، قرار می‌گیرند.

۱.۲.۶.۳ پارامترها و فرضیات کلیدی

جایه‌جایی‌های نامتوازن ماهانه (ژوئن تا دسامبر ۲۰۱۳): این داده‌ها که از گزارشات ماهانه شرکت citibike استخراج شده‌اند، نمایانگر تعداد نیاز به جایه‌جایی دوچرخه در هر ماه بین ایستگاه‌ها هستند. این مقادیر به عنوان محدودیت‌های حداکثری برای جایه‌جایی‌های ماهانه در نظر گرفته شده‌اند؛ به این معنی که هیچ سیستمی نمی‌تواند بیش از این تعداد در هر ماه جایه‌جایی انجام دهد.

درآمد حاصل از هر جایه‌جایی: براساس اطلاعات بدست آمده از بنچمارک فرض بر این است که هر عملیات بازتوزیع (جایه‌جایی یک دوچرخه) منجر به تولید ۱.۳ سفر جدید می‌شود و هر سفر ۳.۵ دلار درآمدزایی دارد. با در اختیار داشتن تعداد کل سفرهای انجام شده در هر ماه و درآمد ماهانه شرکت (منبع گزارشات ماهانه) و میانگین‌گیری ساده عدد ۳.۵ دلار به عنوان درآمد هر سفر محاسبه شد. بنابراین، درآمد ناشی از هر جایه‌جایی به صورت $1.3 \times 3.5 = 4.55$ دلار محاسبه می‌شود. این فرض نشان‌دهنده ارزش استراتژیک بازتوزیع در افزایش بهره‌وری دوچرخه‌ها و درآمد کلی است.

ضریب سالانه سازی (Annualization factor): از آنجا که داده‌های عدم توازن برای ۷ ماه در دسترس هستند، برای سالانه کردن نتایج و مقایسه آن‌ها در مقیاس سالانه، مجموع جایه‌جایی‌های مؤثر طی این ۷ ماه در ضریب $12/7$ ضرب می‌شود.

انواع سیستم‌های بازتوزیع (System Types): چهار نوع وسیله نقلیه برای بازتوزیع در نظر گرفته شده‌اند که هر کدام دارای مشخصات خاص خود هستند:

- **ون باری (Cargo Van):** ظرفیت ۴۵۰۰ جایه‌جایی در ماه در هر واحد، هزینه خرید ۴۰۰۰۰ دلار، طول عمر ۷ سال، و هزینه عملیاتی ۱.۶۰ دلار به ازای هر جایه‌جایی.
- **کامیون باری (Box Truck):** ظرفیت ۷۵۰۰ جایه‌جایی در ماه در هر واحد، هزینه خرید ۵۰۰۰۰ دلار، طول عمر ۷ سال، و هزینه عملیاتی ۱.۱۷ دلار به ازای هر جایه‌جایی.
- **سه چرخه‌باری (Cargo-Trike):** ظرفیت ۱۷۵۰ جایه‌جایی در ماه در هر واحد، هزینه خرید ۷۵۰۰ دلار، طول عمر ۶ سال، و هزینه عملیاتی ۲.۵۸ دلار به ازای هر جایه‌جایی.

- سیستم ترکیبی (Mix): این گزینه نشان دهنده ترکیبی از سیستم‌ها (۰.۵ ون باری، ۰.۳ کامیون باری، ۰.۲ سه‌چرخه باری) با ظرفیت، هزینه، طول عمر و هزینه عملیاتی میانگین وزن دار است. این گزینه برای شبیه‌سازی یک رویکرد ترکیبی در سرمایه‌گذاری گنجانده شده است.

محدوده واحدها (Units range): تحلیل حساسیت برای تعداد واحدهای مختلف، از ۱ تا ۲۴۹ واحد از هر سیستم بازتوزیع، انجام می‌شود تا تأثیر مقیاس سرمایه‌گذاری بر نتایج بررسی شود.

۲.۲.۶.۳ محاسبات انجام شده

برای هر نوع سیستم بازتوزیع و برای هر تعداد واحد در محدوده تعیین شده، محاسبات زیر انجام می‌شود:

ظرفیت کل ماهانه نامحدود (Total Capacity (moves/month)): این مقدار برابر با تعداد واحدها ضربدر ظرفیت هر واحد در ماه است. این نشان دهنده حداکثر جابه‌جایی‌هایی است که ناوگان می‌تواند در یک ماه انجام دهد، در صورتی که هیچ محدودیتی از سوی تقاضای بازتوزیع وجود نداشته باشد.

جابه‌جایی‌های مؤثر سالانه (Annual Effective Moves): برای هر ماه، تعداد جابه‌جایی‌های واقعی (مؤثر) برابر با حداقل (ظرفیت کل ماهانه، نیاز به جابه‌جایی ماهانه) است. این مرحله نشان می‌دهد که ظرفیت اضافی سیستم بازتوزیع، در صورت عدم نیاز، بدون استفاده می‌ماند. سپس، مجموع این جابه‌جایی‌های مؤثر ماهانه برای ۷ ماه محاسبه و با ضریب سالانه سازی ضرب می‌شود تا به یک رقم سالانه تبدیل شود.

درآمد سالانه (Annual Revenue): جابه‌جایی‌های مؤثر سالانه ضربدر درآمد به ازای هر جابه‌جایی (\$ ۴.۵۵) این مقدار نشان دهنده کل درآمدی است که شرکت می‌تواند از طریق بازتوزیع کارآمد دوچرخه‌ها به دست آورد.

هزینه عملیاتی (Operating Cost): جابه‌جایی‌های مؤثر سالانه ضربدر هزینه عملیاتی به ازای هر جابه‌جایی. این هزینه شامل نیروی کار و سوخت/نگهداری مورد نیاز برای انجام عملیات بازتوزیع است.

سرمایه گذاری اولیه (Capital Investment): تعداد واحدها ضربدر (هزینه خرید سیستم / طول عمر سیستم). این هزینه سرمایه اولیه مورد نیاز برای خرید وسایل نقلیه را در طول عمر مفید آن‌ها به صورت سالانه محاسبه می‌کند.

سرمایه‌گذاری کل سالانه (Investment): سرمایه گذاری اولیه + هزینه عملیاتی. این مقدار مجموع تمام هزینه‌های سالانه (هم سرمایه‌ای و هم عملیاتی) مرتبط با عملیات بازتوزیع را نشان می‌دهد.

سود خالص (Net Profit): درآمد سالانه - سرمایه‌گذاری کل سالانه این معیار کلیدی، سودآوری خالص عملیات بازتوزیع را نشان می‌دهد و می‌تواند منفی باشد.

بازگشت سرمایه (%)ROI: (سود خالص / سرمایه‌گذاری کل سالانه) ضربدر ۱۰۰. ROI نشان دهنده کارایی سرمایه‌گذاری است و به درصد بیان می‌شود. این معیار به مدیر عامل کمک می‌کند تا میزان سودآوری هر دلار سرمایه‌گذاری شده را درک کند.

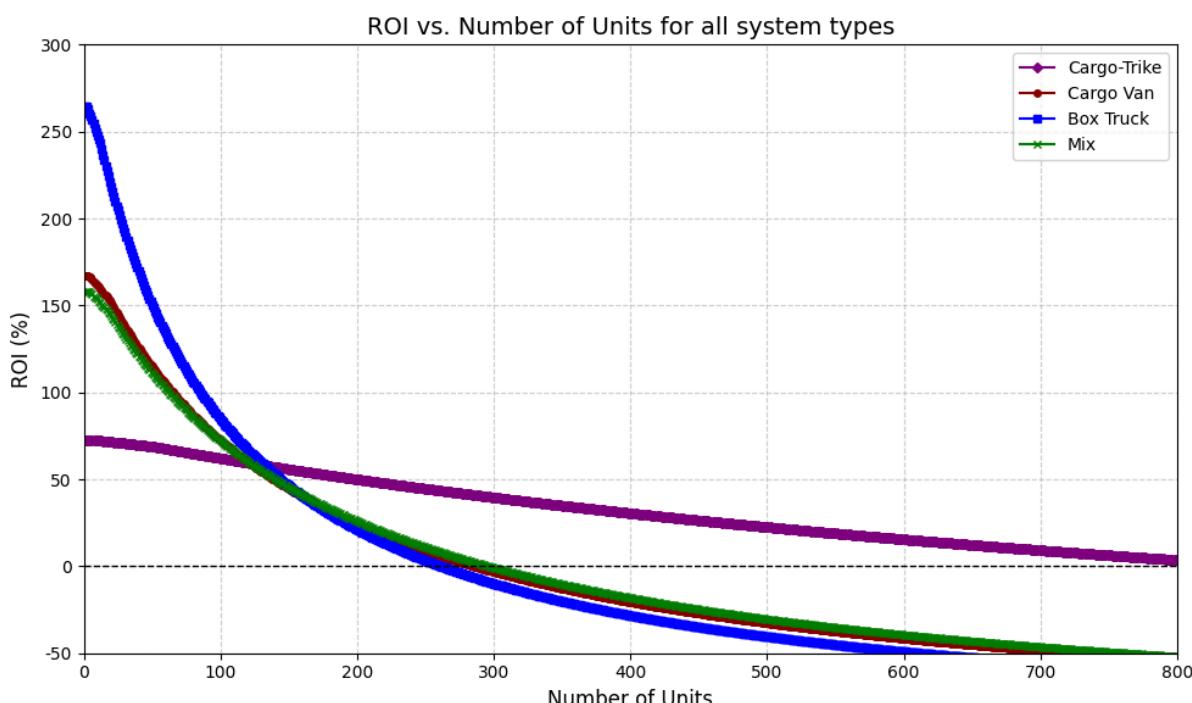
سفرهای جدید سالانه (Annual New Rides): جابه‌جایی‌های مؤثر سالانه * ۱.۳ این مقدار تعداد کل سفرهای جدیدی را که به دلیل بازتوزیع مؤثر دوچرخه‌ها ایجاد شده‌اند، نشان می‌دهد.

۳.۲.۶.۳ تحلیل حساسیت و نمودارها

شش نمودار حساسیت که هر کدام جنبه‌های مختلفی از تأثیر سرمایه‌گذاری بر سیستم‌های بازتوزیع را نشان می‌دهند در بخش زیر ارائه شده است. در این نمودارها، هر خط رنگی نمایانگر یک نوع سیستم بازتوزیع است و نقاط روی خط نشان‌دهنده نتایج برای تعداد واحدهای مختلف هستند.

نمودار ۱: نرخ بازگشت سرمایه در مقابل تعداد واحدهای وسایل نقلیه باز توزیع (ROI vs. Number of Units)

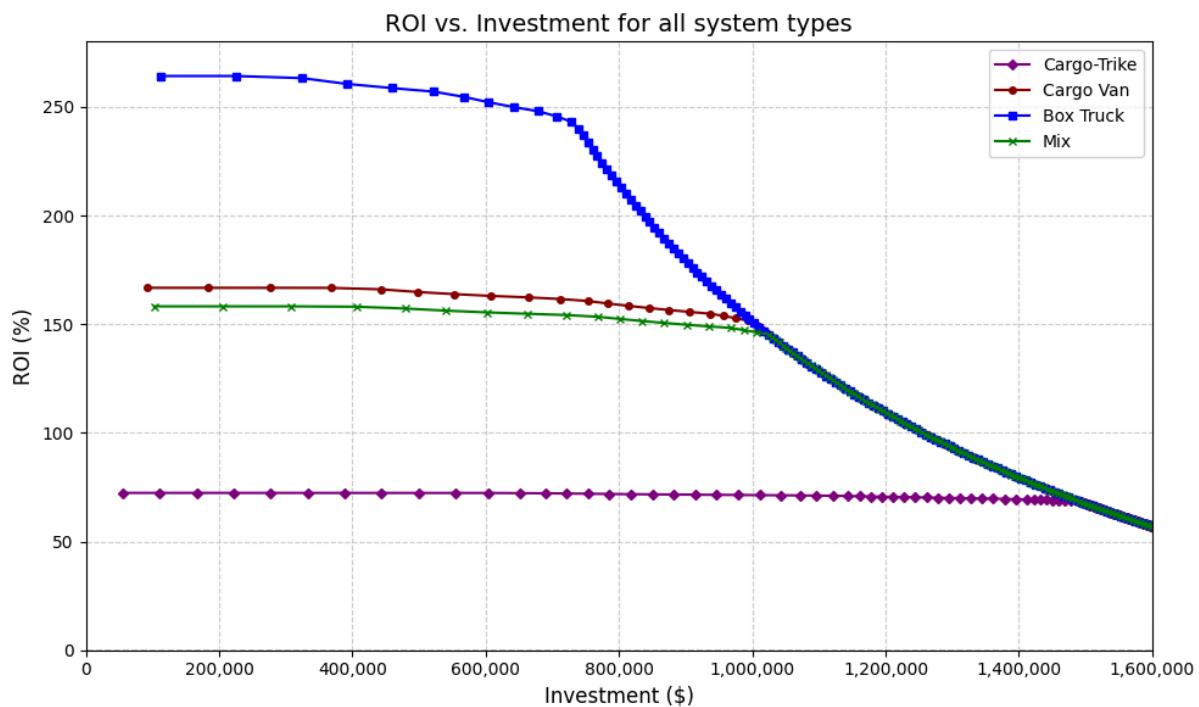
این نمودار نشان می‌دهد که چگونه بازگشت سرمایه برای هر نوع سیستم بازتوزیع با افزایش تعداد واحدهای خریداری شده تغییر می‌کند. به طور معمول، ROI در ابتدا با افزایش بهره‌وری افزایش می‌یابد، اما پس از رسیدن به نقطه بهینه، ممکن است به دلیل رسیدن به ظرفیت اشبع (نیازهای بازتوزیع) و افزایش هزینه‌های عملیاتی و سرمایه‌ای، کاهش یابد یا ثابت بماند. خط افقی صفر درصد نشان‌دهنده نقطه سربه سر است.



شکل (۴۵-۳) نرخ بازگشت سرمایه در مقابل تعداد واحدهای وسایل نقلیه باز توزیع

نمودار ۲: نرخ بازگشت سرمایه در مقابل سرمایه‌گذاری (ROI vs. Investment)

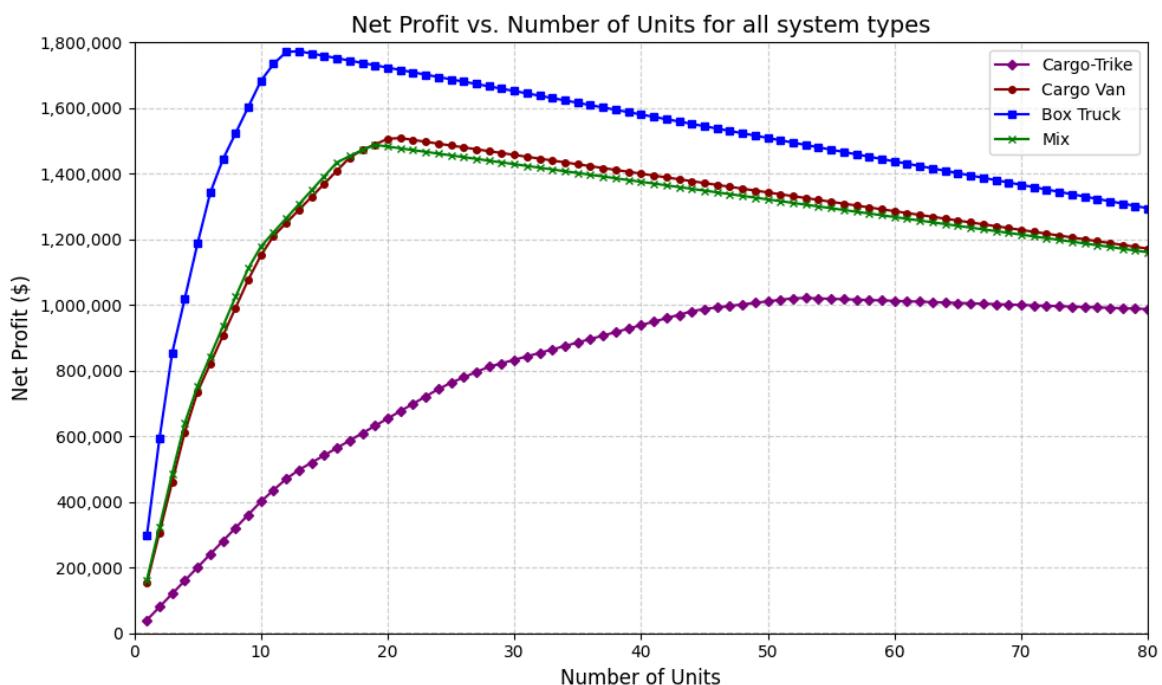
این نمودار نیز ROI را نشان می‌دهد، اما این بار در مقابل کل سرمایه‌گذاری سالانه. این به مدیر عامل کمک می‌کند تا متوجه شود که چه میزان سرمایه‌گذاری، بیشترین بازده را به همراه دارد. ممکن است مشاهده شود که پس از یک نقطه مشخص، افزایش سرمایه‌گذاری دیگر منجر به افزایش متناسب در ROI نمی‌شود و حتی ممکن است آن را کاهش دهد.



شکل (۴۶-۳) نرخ بازگشت سرمایه در مقابل تعداد واحدهای وسایل نقلیه باز توزیع

نمودار ۳: سود خالص در مقابل تعداد واحدها (Net Profit vs. Number of Units)

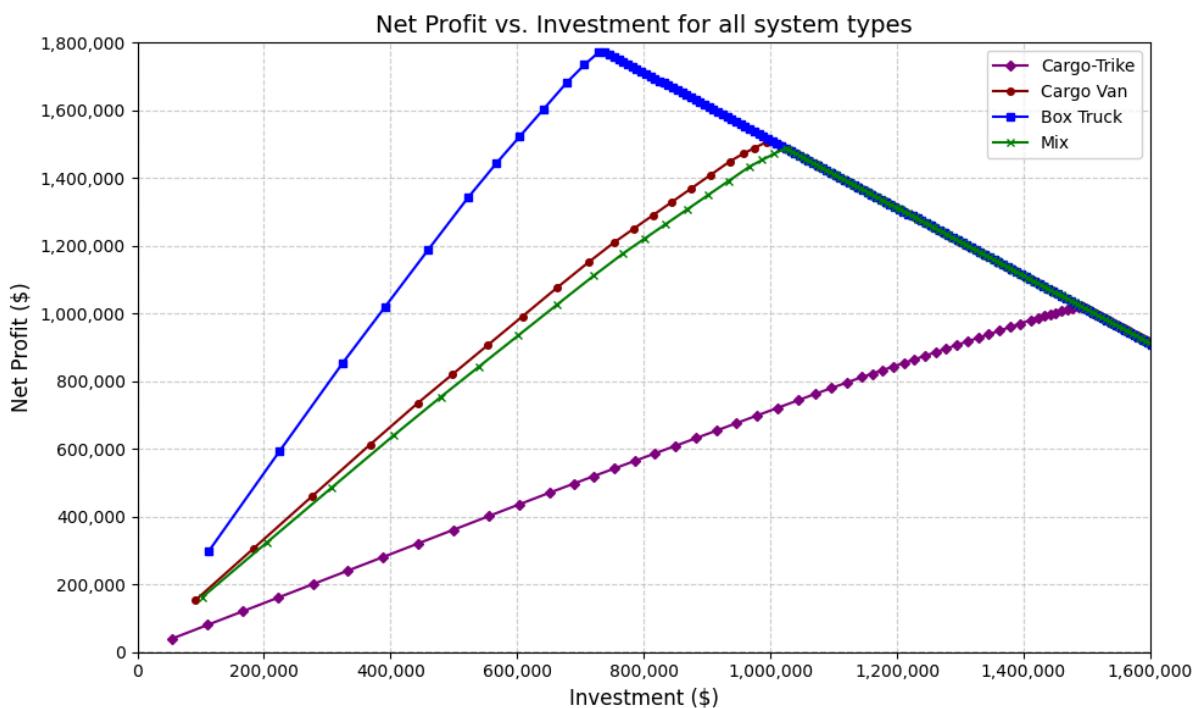
این نمودار مستقیماً سود خالص (درآمد منهای هزینه‌ها) را در برابر تعداد واحدهای سیستم بازتوزیع ترسیم می‌کند. این نمودار به شناسایی تعداد بهینه واحدها برای حداکثر کردن سود کمک می‌کند. نقطه اوج هر منحنی نشان‌دهنده حداکثر سود خالص است. خط افقی صفر دلار، نقطه سربه سر را مشخص می‌کند.



شکل (۴۷-۳) سود خالص در مقابل تعداد واحدها

نمودار ۴: سود خالص در مقابل سرمایه‌گذاری (Net Profit vs. Investment) (Net Profit vs. Investment for all system types)

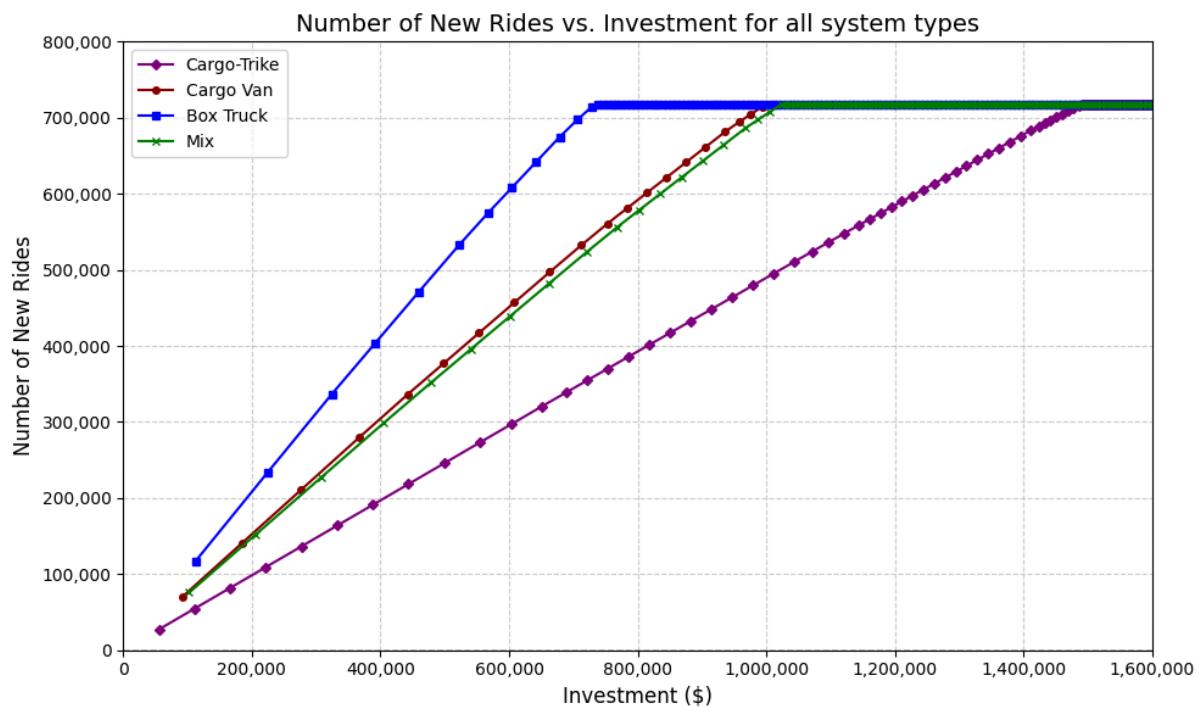
این نمودار سود خالص را براساس میزان سرمایه‌گذاری سالانه نمایش می‌دهد. این نمودار مکمل نمودار ROI است و به مدیر عامل اجاره می‌دهد تا ببیند چه مقدار سرمایه‌گذاری منجر به چه میزان سود دلاری می‌شود. معمولاً، ابتدا سود افزایش می‌یابد و سپس با رسیدن به نقطه اشباع بازار (یعنی دیگر نیاز به جابه‌جایی دوچرخه بیشتر از یک حد مشخص وجود ندارد)، شیب آن کاهش یافته یا به دلیل هزینه‌های اضافی، ممکن است حتی کاهش یابد.



شکل (۴۸-۳) سود خالص در مقابل تعداد واحدها

نمودار ۵: تعداد سفرهای جدید در مقابل سرمایه‌گذاری (Number of New Rides vs. Investment) (Number of New Rides vs. Investment for all system types)

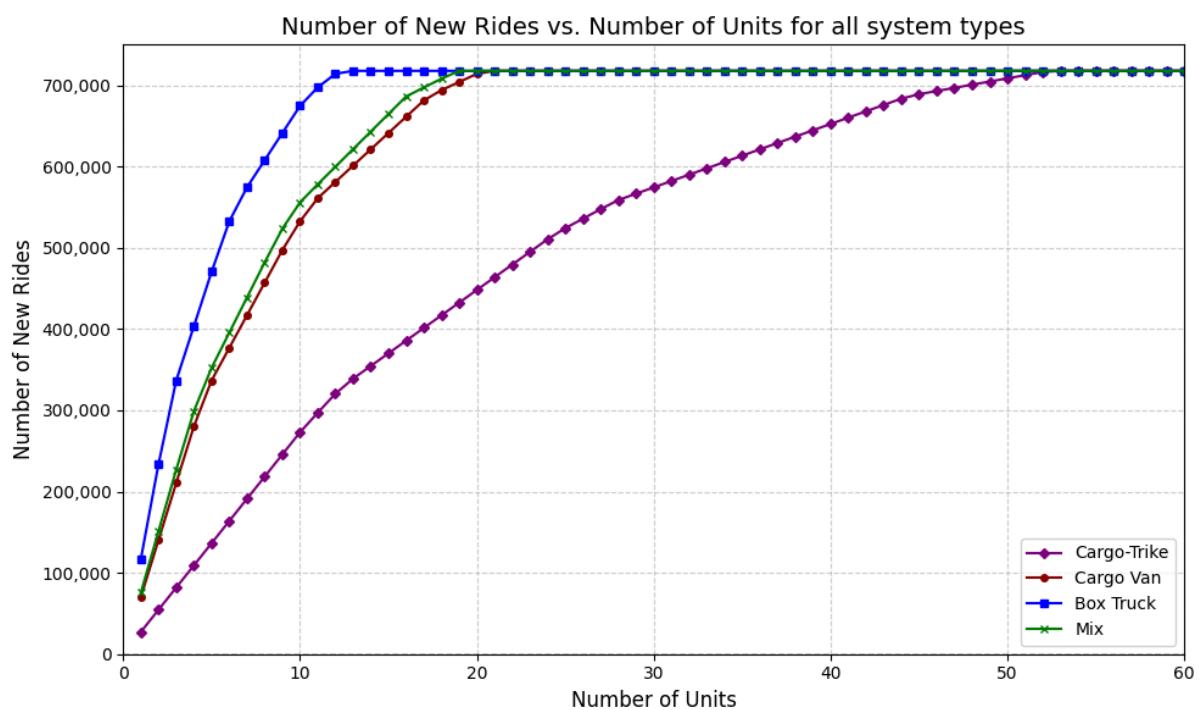
این نمودار رابطه بین تعداد سفرهای جدید ایجاد شده (نتیجه بازتوزیع مؤثر) و میزان سرمایه‌گذاری را نشان می‌دهد. این نمودار به مدیر عامل کمک می‌کند تا ارزش افزوده‌ای که از هر دلار سرمایه‌گذاری به دست می‌آید (به شکل سفرهای جدید) را ارزیابی کند. انتظار می‌رود که تعداد سفرهای جدید با افزایش سرمایه‌گذاری افزایش یابد، اما این افزایش ممکن است پس از یک نقطه به حالت اشباع برسد.



شکل (۴۹-۳) تعداد سفرهای جدید در مقابل سرمایه‌گذاری

نمودار 6: تعداد سفرهای جدید در مقابل تعداد واحدها (Number of New Rides vs. Number of Units)

این نمودار تعداد سفرهای جدید تولید شده را در برابر تعداد واحدهای سیستم بازتوزیع نشان می‌دهد. این نمودار به وضوح نشان می‌دهد که چه تعداد واحد از هر سیستم برای پوشش حداکثر نیازهای بازتوزیع و در نتیجه تولید حداکثر سفرهای جدید، مورد نیاز است. پس از رسیدن به حداکثر نیاز (اشباع افزایش تعداد واحدها دیگر منجر به افزایش سفرهای جدید نمی‌شود).



شکل (۵۰-۳) تعداد سفرهای جدید در مقابل تعداد واحدها

۴.۲.۶.۳ نتایج تحلیل روش بازتوزیع (الف)

- افزایش تعداد واحدهای بازتوزیع منجر به افزایش ظرفیت جابجایی می‌شود، ولی به دلیل وجود محدودیت‌های واقعی (مانند تعداد جابجایی‌های گزارش شده در ماه) افزایش بیش از حد تعداد واحدها تأثیر چندانی بر افزایش ROI ایجاد نمی‌کند.
- نمودارهای تهیه شده (مانند ROI vs. تعداد واحدها و ROI vs. سرمایه‌گذاری) نشان می‌دهند که با افزایش تعداد واحدها، مقدار بازگشت سرمایه به تدریج به نقطه‌ای می‌رسد که افزایش بیشتر، تغییر چندانی در نتایج ندارد.

۳.۶.۳ گسترش ناوگان دوچرخه‌ها و ساخت ایستگاه‌های جدید

این بخش به بررسی این موضوع می‌پردازد که آیا گسترش ناوگان دوچرخه‌ها و ساخت یا توسعه ایستگاه‌های جدید، راه حل مناسبی برای افزایش سودآوری شرکت است و چگونه می‌توان در این زمینه تصمیم‌گیری کرد. افزایش تعداد دوچرخه‌ها و ایستگاه‌ها می‌تواند دسترسی کاربران را افزایش داده و در نتیجه تعداد سفرها و درآمد را بالا ببرد، اما این کار با هزینه‌های سرمایه‌ای قابل توجهی همراه است.

۱.۳.۶.۳ پارامترها و فرضیات کلیدی

ناوگان موجود: ۶۰۰۰ دستگاه دوچرخه

جمعیت منطقه تحت پوشش: ۳۰۰۰۰۰۰ نفر

آستانه اشباع: ۳ دوچرخه به ازای هر ۱۰۰۰ نفر جمعیت. این به معنای یک ناوگان اشباع $(3,000,000/1000) \times 3 = 9000$ دستگاه دوچرخه است. این آستانه نشان‌دهنده حداکثر تعداد دوچرخه‌هایی است که بازار می‌تواند جذب کند.

درآمد متوسط هر سفر: ۳.۵ دلار (یکسان برای همه انواع دوچرخه).

معیارهای هزینه سرمایه (Capital Cost Benchmarks) :

- داک‌های جدید به ازای هر دوچرخه: ۰.۸ داک جدید به ازای هر دوچرخه جدید. این نشان می‌دهد که برای هر دوچرخه جدید، ۸۰٪ یک داک جدید نیز لازم است.
- هزینه هر داک: ۱۰۰۰ دلار.
- داک در هر ایستگاه: ۲۰ داک در هر ایستگاه.
- هزینه هر ایستگاه: ۱۵۰۰۰ دلار.
- هزینه نیروی کار نصب هر ایستگاه: ۳۰۰۰ دلار. این هزینه‌ها برای نصب داک‌ها و ایستگاه‌های جدید در نظر گرفته شده‌اند.

انواع دوچرخه

- دوچرخه استاندارد (Standard): هزینه ۱۲۰۰ دلار، ۵۰ سفر در ماه به ازای هر دوچرخه.
- دوچرخه برقی (E-Bike): هزینه ۲۵۰۰ دلار، ۱۵۰ سفر در ماه به ازای هر دوچرخه.

- دوچرخه برقی نقره‌ای (**Silver E-Bike**) : هزینه 4000 دلار، 450 سفر در ماه به ازای هر دوچرخه.
- حالت ترکیبی : ۵۰٪ دوچرخه استاندارد، ۳۰٪ دوچرخه برقی و ۲۰٪ دوچرخه برقی نقره‌ای با هزینه و تعداد سفر ماهانه میانگین وزن‌دار.

محدوده دوچرخه‌های جدید (New Bikes Range): تحلیل برای افزودن ۱ تا ۴۹۹۹ دوچرخه جدید از هر نوع انجام می‌شود.

۲.۳.۶.۳ محاسبات انجام شده

برای هر نوع دوچرخه و برای هر تعداد دوچرخه جدید، محاسبات زیر انجام می‌شود:

کل ناوگان (**Total Fleet**) : ناوگان موجود + دوچرخه‌های جدید.

سفرهای ماهانه جدید (Monthly New Rides)

- اگر ناوگان کل کوچکتر و یا مساوی آستانه اشباع باشد: تعداد دوچرخه‌های جدید ضربدر سفرهای ماهانه هر دوچرخه.
- در غیر این صورت (اشباع): (آستانه اشباع - ناوگان موجود) ضربدر سفرهای ماهانه هر دوچرخه. این شرط نشان می‌دهد که پس از رسیدن به نقطه اشباع بازار، افزودن دوچرخه‌های بیشتر منجر به تولید سفر اضافی نخواهد شد.

درآمد سالانه جدید (Annual New Revenue): سفرهای ماهانه جدید * ۱۲ * درآمد هر سفر (\$3.50). این مقدار درآمد اضافی سالانه‌ای است که از دوچرخه‌های جدید به دست می‌آید.

محاسبات هزینه سرمایه (Capital Cost Computations)

- دک‌های جدید (**New Docks**) : تعداد دوچرخه‌های جدید * ۰.۸.
- ایستگاه‌های جدید (**New Stations**) : تعداد دک‌های جدید / ۲۰.
- هزینه دوچرخه (**Cost Bikes**) : تعداد دوچرخه‌های جدید * هزینه خرید هر دوچرخه.
- هزینه دک‌ها (**Cost Docks**) : تعداد دک‌های جدید * ۱۰۰۰ دلار.
- هزینه ایستگاه‌ها (**Cost Stations**) : تعداد ایستگاه‌های جدید * ۱۵۰۰۰ دلار.
- هزینه نصب نیروی کار (**Installation Labor**) : تعداد ایستگاه‌های جدید * ۳۰۰۰ دلار.

کل سرمایه‌گذاری اولیه (TotalCapex): مجموع تمام هزینه‌های سرمایه‌ای (دوچرخه‌ها، دک‌ها، ایستگاه‌ها، و نصب). این مقدار به عنوان Investment (\$) برای نمایش در نمودارها نامگذاری شده است.

سود خالص (Net Profit): درآمد سالانه جدید - کل سرمایه‌گذاری اولیه. این معیار نشان‌دهنده سود خالص (یا زیان) حاصل از سرمایه‌گذاری در گسترش ناوگان و زیرساخت است.

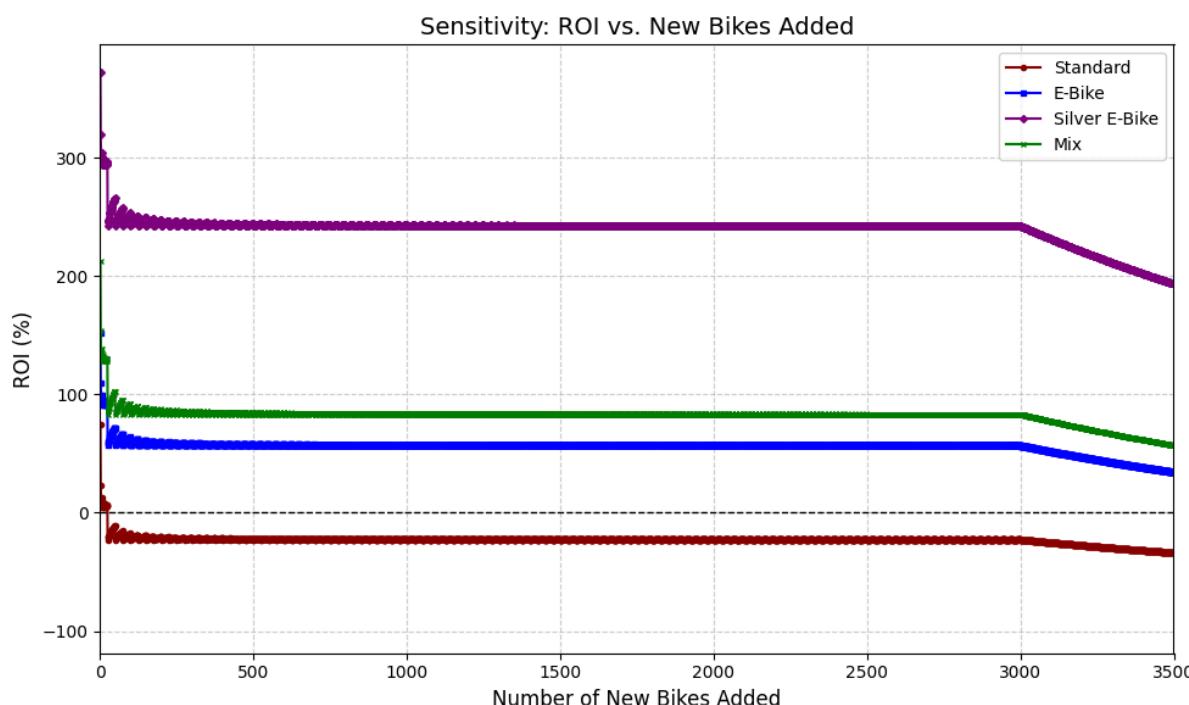
بازگشت سرمایه (%)ROI (سود خالص / کل سرمایه‌گذاری اولیه) * ۱۰۰. این معیار کارایی سرمایه‌گذاری را به درصد نشان می‌دهد.

۳.۳.۶.۳ تحلیل حساسیت و نمودارها

چهار نمودار حساسیت برای بخش گسترش ناوگان و ایستگاه‌ها ارائه شده است که به مدیر عامل کمک می‌کند تا گزینه‌های مختلف سرمایه‌گذاری را ارزیابی کند:

نمودار ۱: نرخ بازگشت سرمایه در مقابل دوچرخه‌های جدید اضافه شده (ROI vs. New Bikes Added)

این نمودار نشان می‌دهد که چگونه ROI با افزودن تعداد فزاینده‌ای از دوچرخه‌های جدید از هر نوع تغییر می‌کند. معمولاً، ROI ابتدا افزایش می‌یابد زیرا درآمدها سریع‌تر از هزینه‌ها رشد می‌کنند، اما پس از رسیدن به نقطه اشباع بازار (۹۰۰۰ دوچرخه کل) یا افزایش هزینه‌های زیرساخت، ممکن است کاهش یابد.

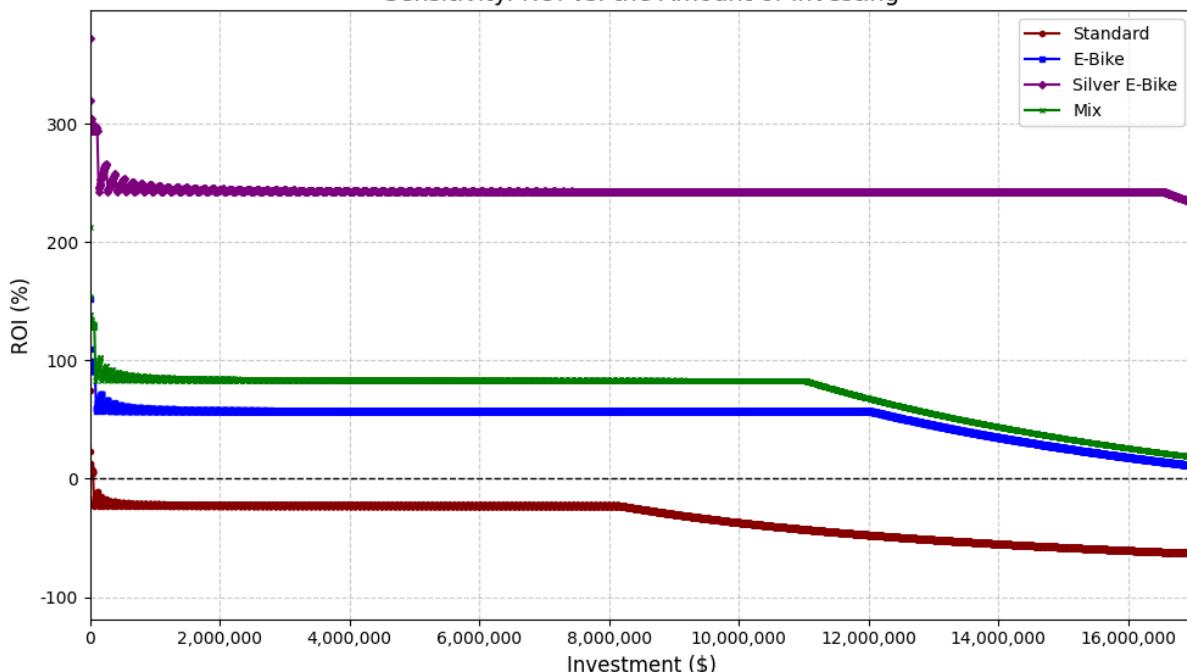


شکل (۵۱-۳) نرخ بازگشت سرمایه در مقابل دوچرخه‌های جدید اضافه شده

نمودار ۲: نرخ بازگشت سرمایه در مقابل میزان سرمایه‌گذاری (ROI vs. the Amount of Investing)

این نمودار ROI را در برابر کل سرمایه‌گذاری اولیه (CAPEX) نشان می‌دهد. این به مدیر عامل کمک می‌کند تا رابطه بین میزان سرمایه‌گذاری و بازده آن را درک کند. برای تصمیم‌گیری‌های کلان مالی، این نمودار بسیار کاربردی است. محورهای نمودار برای خوانایی بهتر با فرمت کاما نمایش داده می‌شوند.

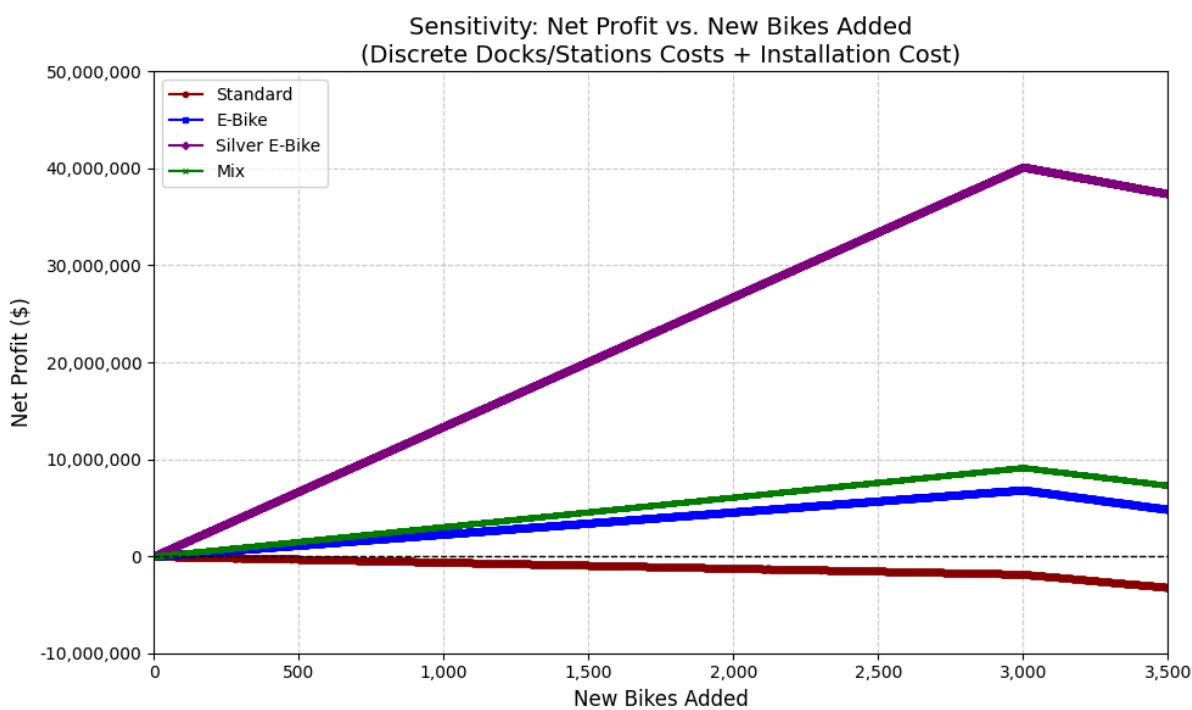
Sensitivity: ROI vs. the Amount of Investing



شکل (۵۲-۳) نرخ بازگشت سرمایه در مقابل میزان سرمایه‌گذاری

نمودار ۳: سود خالص در مقابل دوچرخه‌های جدید اضافه شده (Net Profit vs. New Bikes Added)

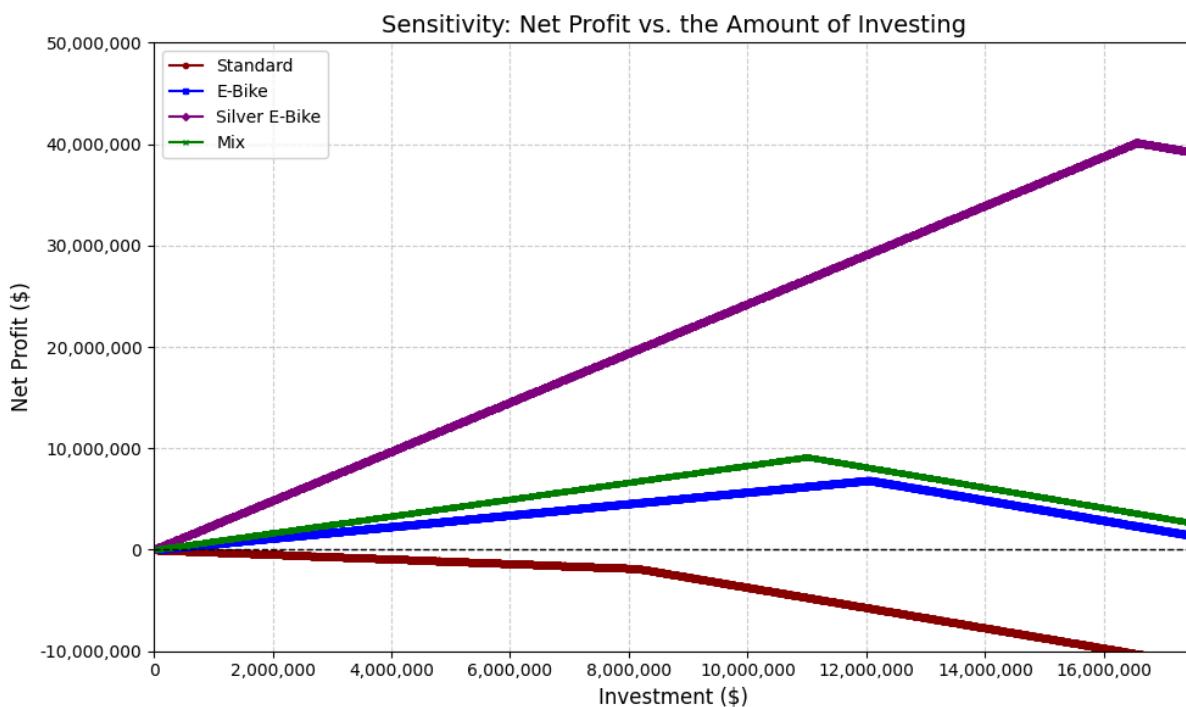
این نمودار مستقیماً سود خالص (با زیان) را بر اساس تعداد دوچرخه‌های جدید اضافه شده نشان می‌دهد. این نمودار به شناسایی تعداد بهینه دوچرخه‌های جدید برای حداکثر کردن سود کمک می‌کند. نقطه اوج منحنی (در صورت وجود) نشان‌دهنده حداکثر سود ممکن است.



شکل (۵۳-۳) سود خالص در مقابل دوچرخه‌های جدید اضافه شده

نمودار ۴: سود خالص در مقابل میزان سرمایه‌گذاری (Net Profit vs. the Amount of Investing)

این نمودار سود خالص را بر اساس میزان کل سرمایه‌گذاری (CAPEX) ترسیم می‌کند. این نمودار دیدگاه مالی کاملی از سودآوری پروژه‌های گسترش ناوگان و ایستگاهها ارائه می‌دهد. این نمودار نیز به شناسایی میزان سرمایه‌گذاری مورد نیاز برای دستیابی به اهداف سودآوری کمک می‌کند و نقاط سربه سر را مشخص می‌نماید.



شکل (۵۴-۳) سود خالص در مقابل میزان سرمایه‌گذاری

۴.۳.۶.۳ ایستگاه‌هایی با بیشترین عدم توازن

۵.۳.۶.۳ نتایج تحلیل روش گسترش ناوگان و ایستگاه‌ها (ب)

- با افزودن دوچرخه‌های جدید، درآمد سالیانه افزایش می‌یابد؛ با این حال، پس از رسیدن ناوگان به سطح اشباع (۹۰۰۰ دوچرخه)، سودآوری به دلیل افزایش هزینه‌های سرمایه‌ای کاهش می‌یابد.
- نمودارهای ROI و سود خالص نسبت به تعداد دوچرخه‌های اضافه شده و نیز نسبت به سرمایه‌گذاری، حساسیت نتایج را نسبت به تغییرات فروض زیرساختمان و درآمد نشان می‌دهند.

۴.۶.۳ استراتژی ترکیبی (بهبود بازتوزیع و گسترش ناوگان)

این استراتژی سعی در ترکیب سرمایه‌گذاری در دوچرخه‌های جدید و سیستم‌های بازتوزیع به صورت همزمان دارد. فرض کلیدی این استراتژی، تقسیم 50-50 سرمایه‌گذاری بین دو بخش است. همچنین، فرض می‌شود که با وجود سیستم‌های بازتوزیع، محدودیت اشباع برای دوچرخه‌های جدید تا حدی برطرف می‌شود و دوچرخه‌ها بیشتر می‌توانند مولد باشند.

بخش خرید دوچرخه‌های جدید:

- محدودیت سطح اشباع (تا ۹۰۰۰ دوچرخه فعال) لاحاظ می‌شود؛ یعنی اگر تعداد دوچرخه‌های جدید باعث افزایش ناوگان از این مقدار شود، تنها تعداد مؤثر دوچرخه‌های جدید برابر با $(9000 - 6000)$ محاسبه می‌شود.
- بهمنظور جبران کاهش تعداد مؤثر، ضریب تبدیل تعداد سفر به ازای هر دوچرخه از مقدار پایه ($1.1^{(3)}$) به عددی بالاتر (مثلًاً 1.5) افزایش می‌باید.

بخش بازتوزیع:

- واحدهای بازتوزیع بدون محدودیت در ظرفیت (با فرض عملکرد کامل) در محاسبات درآمد و هزینه لاحاظ می‌شوند.

۲.۴.۶.۳ محاسبات نهایی

- نسبت r_b نسبت به nb از طریق یک رابطه تعیین می‌شود ($r_b = nb \times r_ratio$) به طوری که هزینه مؤثر هر دوچرخه (تقرباً 4150 دلار) برابر هزینه مؤثر هر واحد ری‌باليانسینگ شود.
- درآمد کل (از دو بخش دوچرخه و بازتوزیع) و سرمایه کل محاسبه شده و در نهایت ROI و سود خالص استخراج می‌شود.

۳.۴.۶.۳ تحلیل حساسیت و نمودارها

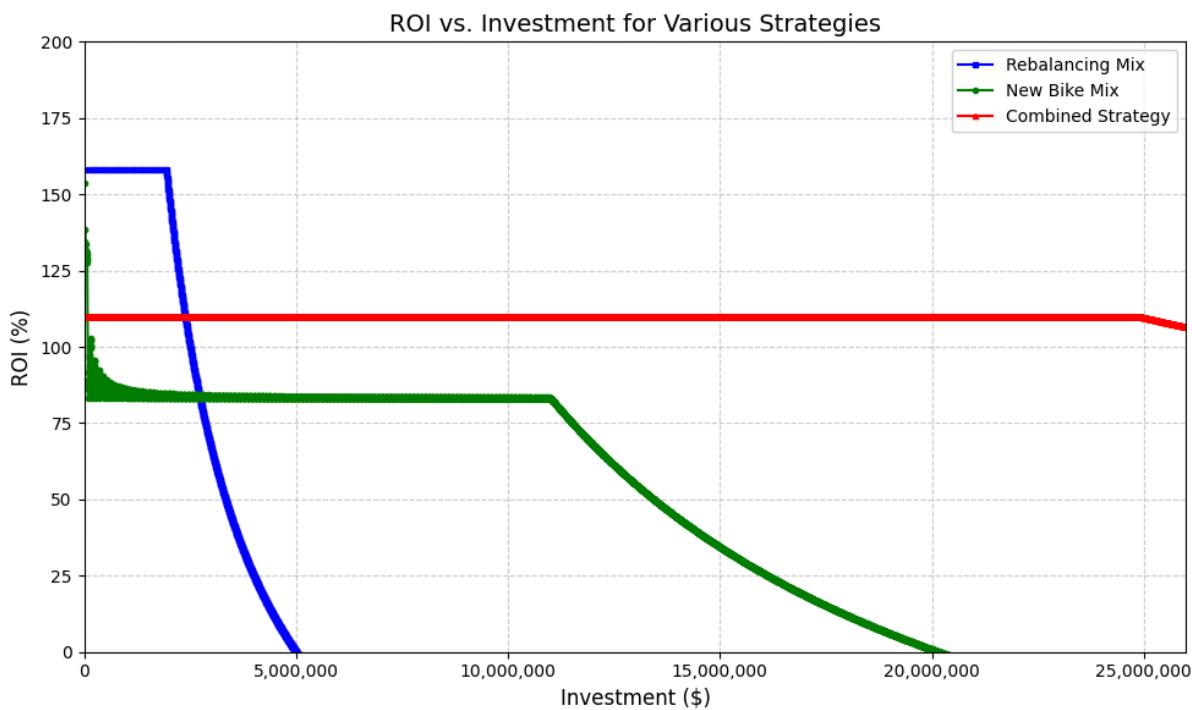
نمودار ۱: نرخ بازگشت سرمایه در مقابل سرمایه‌گذاری برای استراتژی‌های مختلف (Strategies)

این نمودار نشان می‌دهد که چگونه بازگشت سرمایه (ROI) برای هر یک از سه استراتژی (فقط دوچرخه جدید، فقط بازتوزیع، و ترکیبی) با افزایش میزان سرمایه‌گذاری تغییر می‌کند. هر خط رنگی نمایانگر یک استراتژی است.

استراتژی اول (Rebalancing Mix): این استراتژی در مقادیر پایین سرمایه‌گذاری ممکن است ROI بالاتری داشته باشد، زیرا بهبود بازتوزیع می‌تواند بهره‌وری دوچرخه‌های موجود را به سرعت افزایش دهد. اما با افزایش سرمایه‌گذاری، به دلیل محدودیت `max_moves` (حداکثر نیاز به بازتوزیع)، ممکن است ROI آن به تدریج کاهش یابد یا ثابت بماند.

استراتژی دوم (New Bike Mix): نرخ بازگشت سرمایه این استراتژی ممکن است در ابتدا کنتر رشد کند، اما با افزایش تعداد دوچرخه‌ها، درآمدها افزایش می‌یابند. با این حال، با رسیدن به نقطه اشباع (۹۰۰۰ دوچرخه کل)، ROI آن نیز به دلیل عدم تولید سفرهای بیشتر توسط دوچرخه‌های اضافی، کاهش خواهد یافت.

استراتژی سوم (Combined Strategy): نرخ بازگشت سرمایه این استراتژی به دلیل بهره‌برداری از هر دو منبع درآمد (افزایش ناوگان و بهبود بهره‌وری) و فرض برطرف شدن محدودیت بازتوزیع، احتمالاً در بلندمدت بالاترین ROI و سود خالص را ارائه می‌دهد. این نمودار می‌تواند نقاط بھینه سرمایه‌گذاری و استراتژی‌ای که بهترین تعادل بین ریسک و بازده را ارائه می‌دهد، نشان دهد.



شکل (۵۵-۳) نرخ بازگشت سرمایه در مقابل سرمایه‌گذاری برای استراتژی‌های مختلف

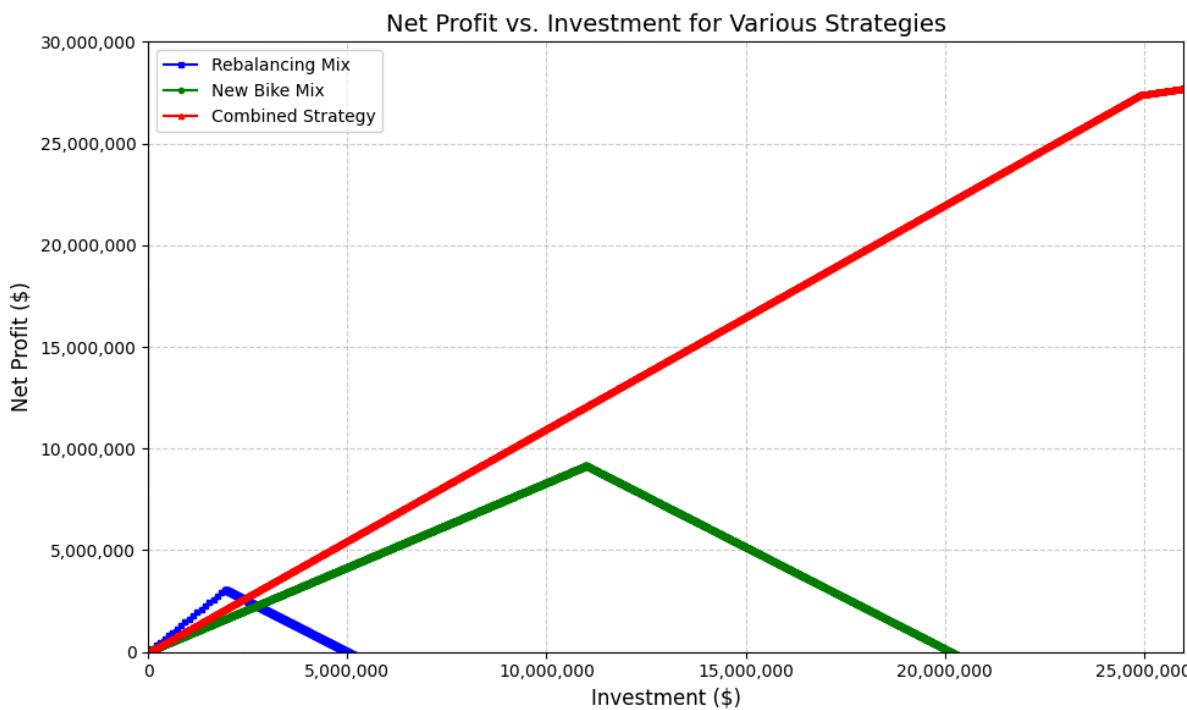
نمودار ۲: سود خالص در مقابل سرمایه‌گذاری برای استراتژی‌های مختلف (Strategies)

شرح نمودار: این نمودار سود خالص (درآمد منهای هزینه‌ها) را در برابر میزان سرمایه‌گذاری برای هر سه استراتژی نشان می‌دهد. این نمودار نیز از همان رنگ‌بندی و نشانه‌گذاری استراتژی‌ها استفاده می‌کند.

استراتژی اول (Rebalancing Mix): در حالی که ممکن است ROI خوبی در سرمایه‌گذاری‌های کوچک داشته باشد، پتانسیل سود خالص آن ممکن است محدودتر باشد زیرا تنها بهینه‌سازی عملیات موجود را هدف قرار می‌دهد و به افزایش اندازه بازار نمی‌پردازد.

استراتژی دوم (New Bike Mix): پتانسیل سود خالص بالاتری دارد، به خصوص تا رسیدن به نقطه اشباع. پس از آن، شیب افزایش سود خالص کاهش می‌یابد یا حتی ممکن است سود به دلیل هزینه‌های اضافی، کاهش یابد.

استراتژی سوم (Combined Strategy): انتظار می‌رود که این استراتژی پتانسیل بالاتری برای حداکثر کردن سود خالص داشته باشد، زیرا هم بازار را گسترش می‌دهد (دوچرخه‌های جدید) و هم بهره‌وری عملیاتی را افزایش می‌دهد (بازتوزیع). این نمودار به مدیر عامل کمک می‌کند تا میزان سرمایه‌گذاری لازم برای دستیابی به اهداف سود خالص مشخص را درک کند.



شکل (۳-۵۶) سود خالص در مقابل سرمایه‌گذاری برای استراتژی‌های مختلف

۴.۴.۶.۳ نتایج تحلیل روش ترکیبی (ج)

- استراتژی ترکیبی، با ترکیب سرمایه‌گذاری در خرید دوچرخه‌های جدید (با در نظر گرفتن محدودیت اشباع) و عملیات ری‌باليانسینگ (با افزایش ضریب تبدیل به منظور جبران محدودیت) به نتایج مطلوب‌تری دست می‌یابد.
- نمودارهای استخراج شده نشان می‌دهند که این استراتژی نسبت به تغییرات در فروض (هم در هزینه‌های عملیاتی و هم درآمد هر سفر) پایداری نسبی بیشتری از خود نشان می‌دهد.

۵.۶.۳ نتیجه‌گیری و توصیه‌ها

با استناد به تحلیل‌های حساسیت انجام شده بر پایه داده‌های واقعی گزارش‌های ماهانه شرکت (جدول ۱ در پیوست)، از جمله معیارهای کلیدی نظیر تعداد جابجایی‌های ماهانه (322,071 حرکت در ۷ ماه) و سطح دسترسی سیستم (بین ۹۲٪ تا ۹۸٪) و همچنین بررسی بنچمارک‌های صنعتی (سطح دسترسی مطلوب بین ۹۰–۹۵٪ و شرایط سیستم‌های مشابه مانند Hubway)، موارد زیر توصیه می‌شود:

- استراتژی بازتوزیع:** به طور محدود، افزایش تعداد واحدهای ری‌باليانسینگ می‌تواند در بهبود دسترسی (Availability) و کاهش ناهمانگی بین ایستگاه‌ها موثر باشد؛ اما به خاطر محدودیت‌های گزارش شده، افزایش بیش از حد تعداد واحدها نتایج قابل توجهی به همراه نخواهد داشت.
- استراتژی خرید دوچرخه‌های جدید:** افزایش ناوگان دوچرخه از طریق خرید دوچرخه‌های جدید و توسعه زیرساخت‌ها می‌تواند به افزایش درآمد

منجر شود؛ اما نباید حد اشباع (۹۰۰۰ دوچرخه فعال) را طی کرد، چرا که افزایش بیش از حد ناوگان موجب کاهش بهرهوری سرمایه‌گذاری خواهد شد.

- استراتژی ترکیبی:

ترکیب استراتژی‌های فوق (با تخصیص ۵۰-۵۰ سرمایه بین خرید دوچرخه‌ها و عملیات ری‌بالیانسینگ) بهبود قابل توجهی در هر دو جنبه درآمد و ROI ایجاد می‌کند. در این مدل، با وجود محدودیت سطح اشباع، افزایش ضریب تبدیل (از ۱.۳ به ۱.۵) برای دوچرخه‌ها، اثر مثبت ری‌بالیانسینگ در افزایش سفرها را به خوبی منعکس می‌کند.

توصیه نهایی:

با توجه به داده‌های گزارش‌های ماهانه (همچون میزان دسترسی بین ۹۲ تا ۹۸ درصد) و بررسی بنچمارک‌های صنعتی، استراتژی ترکیبی به عنوان بهترین گزینه توصیه می‌شود. در این راستا:

- ابتدا باید عملیات ری‌بالیانسینگ به‌گونه‌ای بهینه گردد که سطح دسترسی ایستگاه‌ها (بین ۹۰-۹۵٪ به عنوان حداقل قابل قبول و در برخی مواقع هم نزدیک به معیارهای جهانی) حفظ شود؛
- سپس با افزودن دوچرخه‌های جدید تا سقف ۹۰۰۰ دوچرخه فعال (با بهره‌برداری کامل، به کمک افزایش ضریب تبدیل) درآمد بهبود یافته و هزینه‌های اضافی کاهش می‌یابد.

۶.۶.۳ جمع‌بندی نهایی

این گزارش کل مسیر تحلیل را از تعریف مسئله و انتخاب فروض (بر اساس داده‌های گزارش‌های ماهانه شرکت و بنچمارک‌های بین‌المللی) تا استخراج متغیرهای کلیدی (ROI، سود خالص، تعداد سفرهای جدید) به‌طور کامل شرح می‌دهد. همچنین جداول گزارش‌های ماهانه (مانند "جدول ۱") به همراه نمودارهای حساسیت منتشر خواهند شد تا خواننده بتواند تمامی مراحل تحلیل را به‌طور شفاف بازتولید کند.

با اتخاذ استراتژی ترکیبی و با توجه به سطح دسترسی گزارش شده (که با بنچمارک‌های صنعتی همخوانی دارد) و نتایج تحلیل حساسیت، توصیه می‌شود که شرکت به صورت همزمان در حوزه‌های افزایش بهره‌برداری از ناوگان (از طریق ری‌بالیانسینگ) و افزایش دوچرخه‌های فعال (با رعایت محدودیت اشباع) سرمایه‌گذاری نماید تا ضمن بهبود کیفیت سرویس، سودآوری کلی نیز افزایش یابد.

جمع‌بندی نهایی: مدیر عامل باید هر دو راهکار را به صورت همزمان یا مرحله‌ای مورد توجه قرار دهد. بهبود بازتوزیع می‌تواند بهره‌وری ناوگان موجود را افزایش داده و سودآوری را در کوتاه‌مدت بهبود بخشد. در حالی که گسترش ناوگان و ایستگاه‌ها، پتانسیل رشد بلندمدت شرکت را فراهم می‌کند. تصمیم‌گیری نهایی باید بر اساس اهداف استراتژیک شرکت (رشد سریع در مقابل سودآوری پایدار)، میزان ریسک‌پذیری و دسترسی به منابع مالی صورت گیرد. تحلیل حساسیت ارائه شده در این گزارش ابزاری قدرتمند برای حمایت از این تصمیمات است.