### پروژه تحلیل اثرگذاری رجیستری بر بازار آیفون ایران

### مقدمه:

در دنیای امروز، بازار گوشی‌های هوشمند به یکی از بخش‌های مهم و تاثیرگذار در اقتصاد جهانی تبدیل شده است. در ایران، با اعمال سیاست‌های جدید، از جمله طرح رجیستری گوشی‌های موبایل، بازار این محصولات با تغییرات عمده‌ای مواجه شده است. اجرای طرح رجیستری آیفون در تاریخ 9 آبان 1403 به‌عنوان یک اقدام عمده در راستای مقابله با قاچاق گوشی‌های همراه و ساماندهی بازار، تاثیری قابل توجه بر قیمت‌ها، عرضه و تقاضا، و واکنش‌های اجتماعی و اقتصادی گذاشته است.

این پروژه به تحلیل و بررسی تاثیرات این طرح بر بازار گوشی‌های آیفون در ایران می‌پردازد. پروژه به چهار زیرمسئله تقسیم شده است:

**مسائل اقتصادی** :

* **بررسی تاثیر رجیستری بر قیمت گوشی ها**
* **بررسی تاثیر قیمت ارز بر قیمت گوشی های آیفون**
* **تحلیل روند تغییرات بازار موبایل قبل و بعد از اجرای رجیستری**

**مسائل اجتماعی** :

* **واکنش کاربران به طرح رجیستری و دسته بندی موضوعات**

**مسائل فنی و اجرایی مسائل حقوقی** :

* **مشکلات رجیستری از نظر کاربران و کارشناسان**
* **پاسخ‌گویی به نیاز بازار (آیا طرح رجیستری به‌طور کامل نیاز بازار به گوشی‌های آیفون را تأمین می‌کند؟ آیا بازار به اندازه کافی تأمین شده است یا نیاز به تنظیمات و اصلاحات بیشتر دارد؟)**

داده‌های مختلف از منابع متنوع همچون داده‌های مربوط به قیمت ارز، تغییرات بازار، توئیت‌ها و نظرات کاربران جمع‌آوری شده‌اند و مورد تحلیل قرار گرفته‌اند. در این گزارش، ابتدا اقدامات انجام‌شده و روش تحلیل توضیح داده خواهد شد و سپس نتایج به‌دست‌آمده در قالب نمودارها و تحلیل‌ها بیان می‌شود.

### متدولوژی:

برای انجام این پروژه، رویکردی تحلیلی و آماری اتخاذ شد تا اثرات طرح رجیستری بر بازار آیفون به‌طور دقیق و معتبر مورد ارزیابی قرار گیرد. روش‌شناسی به‌طور عمده شامل مراحل زیر است:

1. **جمع‌آوری داده‌:** داده‌های مورد نیاز از منابع مختلف مانند توئیت‌ها، سایت‌های خبری (Zoomit.ir و Digikala)، و اطلاعات اقتصادی مرتبط با قیمت ارز و گوشی‌ها جمع‌آوری شدند. این داده‌ها به‌صورت کیفی و کمی مورد استفاده قرار گرفتند تا تاثیرات اقتصادی، اجتماعی و فنی طرح رجیستری مورد بررسی قرار گیرد
2. **پیش پردازش داده:** پس از جمع‌آوری داده‌ها، فرآیند پاک‌سازی و آماده‌سازی انجام میگیرد که با بررسی کیفیت و کمیت داده های جمع آوری شده، دیتاست هایی برای مرحله بعدی آماده میشود
3. **تحلیل داده‌ها:** پس از جمع‌آوری دیتاست ها و با استفاده از تحلیل‌های آماری، روند تغییرات قیمت‌ها و تقاضا قبل و بعد از اجرای طرح رجیستری مورد ارزیابی قرار گرفت. همچنین، تحلیل احساسات (Sentiment Analysis) بر روی نظرات کاربران در توئیت‌ها و سایت‌ها انجام شد تا واکنش‌های اجتماعی به طرح رجیستری بررسی شوند.
4. **مدلسازی و تفسیر نتایج:** با استفاده از مدل‌های آماری و تحلیل روندهای زمانی، تاثیرات مختلف طرح رجیستری بر بازار و همچنین واکنش‌های اجتماعی به‌طور دقیق مورد بررسی قرار گرفت.
5. **مسائل اقتصادی**

* **پیش پردازش داده های مربوط به نرخ ارز:**

هدف از این مرحله تبدیل داده های ارزی جمع آوری شده از سایت [www.navasan.net](http://www.navasan.net) به data frame های تمیز و قابل استفاده هستش.

مراحل انجام این کار به صورت زیر بوده است:  
1- در فایل (clear dollar data.py) ابتدا با دستورات پایتون، فایل csv دیتاست مربوطه خوانده شده و سپس سطر به سطر دیتاست خوانده شده و فرایند های زیر روی هر سطر داده اتفاق افتاده است:

* با استفاده از کتابخانه persiantools ستون تاریخ میلادی، تبدیل به فرمت تاریخ شمسی شده است.
* با استفاده از الگو های regex و کتابخانه re موارد اضافی داخل ستون description مانند: شکلک ها و نمادها، کلمات اضافی از جمله "تومان" و سایر موارد، شناسایی و حذف شدند تا تنها نام ارز و قیمت مربوط به آن باقی بماند.
* در مرحله آخر نیز نام هر شاخص و قیمت مربوط به آن جدا شده اند و همگی در یک dict قرار گرفته اند و تمام دیکشنری های به دست آمده داخل یک لیست قرار گرفته شده است. (این کار به این دلیل انجام شده است که در این دیتاست تعداد شاخص هایی که در هر سطر آورده شده بود با سطر دیگری، متفاوت بود؛ برای مثال در یک ساعت خاص (یا درواقع یک سطر) قیمت یورو ذکر شده بود، اما در ساعت دیگر این شاخص ذکر نشده بود و با این کار در دیتافریم خروجی تمام شاخص ها به عنوان یک ستون قرار گرفت و اگر در ساعت خاصی مقدار اون شاخص ذکر نشده بود، مقدار null براش لحاظ شد.)

2- در مرحله آخر لیست جمع آوری شده از تمام سطر ها به یک dataframe در کتابخانه pandas تبدیل شده و داخل فایل اکسل ریخته شد.

3- با بررسی ای که در دیتای نهایی صورت گرفت و به دلیل null بودن تعداد زیادی از سطر های مربوط به شاخص یورو، دیتای این شاخص از سایت tgju.org در یک فایل اکسل جمع آوری شد و با دستورات پایتون، دیتای به دست آمده با دیتاست قبلی ترکیب یا به اصطلاح merge شد.(merge euro)

4- به دلیل اینکه قیمت شاخص ها به صورت ساعتی در دیتاست قرار گرفته بود و همچنین مقادیر برای بعضی از ساعات روز وجود نداشت و null بود؛ در نتیجه برای داشتن دیتای تمیز تر و تحلیل بهتر، فرایند زیر انجام شد:

* فایل دیتاست در فایل (keep last dollar price per day.ipynb) خوانده شد.
* با ترکیب دو ستون date و time و ستون datetime ساخته شد و دیتافریم بر اساس ستون datetime سورت شد.
* سپس با دستورات پایتون و کتابخانه pandas آخرین قیمت مربوط به هر روز فقط نگه داری شد.
* در نهایت دیتافریم نهایی که نشان دهنده آخرین قیمت هر شاخص در هر روز بود، داخل فایل اکسل ریخته شد.

5- دیتاست نهایی در برنامه power bi وارد شد و کارهای زیر روی آن انجام شد:

* ستون تاریخ شمسی split شد و ستون های عدد روز، عدد ماه و عدد سال داخل یک ستون جدا قرار گرفت.
* یک جدول واسط به نام MonthsNumber ساخته شد که عدد ماه را به نام ماه شمسی تبدیل میکرد و با استفاده از این جدول واسط در واقع در دیتاست به جای عدد ماه، نام ماه شمسی قرار گرفت.
* **پیش پردازش داده های مربوط به آگهی های موبایل:**

هدف از این مرحله تبدیل داده های آگهی های موبایل جمع آوری شده از سایت دیوار به data frame های تمیز و قابل استفاده هستش.

مراحل انجام این کار به صورت زیر بوده است:  
**1-** در فایل (exported\_cleaning.ipynb) ابتدا با دستورات پایتون، فایل csv دیتاست مربوطه خوانده شده و سپس فرایند های زیر بر روی این دیتاست اتفاق افتاده است:

* **تبدیل تاریخ:** با استفاده از کتابخانه jdatetime ستون تاریخ میلادی، تبدیل به فرمت تاریخ شمسی شده است و همچنین، ستون‌هایی که نیاز نیستند، مانند converted\_created\_at حذف شدند.
* **شناسایی برند ها:** با دستور split متن ستون category را جداکرده و دیتا های مربوط به brand و model گوشی در هر آگهی شناسایی و جدا شده است و هر کدام در یک ستون جدید قرار گرفته است.

**تمیز کردن ستون brand و model:** به دلیل اینکهنامبرند هر گوشی در ستون brand به فرمت های مختلفی وارد شده بود، برای مثال برای گوشی های برند apple نام های مختلفی مانند(" آیفون"، " اپل"، "iphone" و...) نوشته شده بود، اما در این مرحله تمام حالات مربوط به برند اپل، به لیبل apple تغییر پیدا کرد تا دسته بندی دقیق تری بر اساس نام برند اتفاق بیوفتد.( این فراند برای برند های Samsung و Xiaomi نیز انجام شد و برای سایر آگهی ها، لیبل "others" به عنوان برند، قرار گرفت.)

همچنین در ستون model نیز، نام مدل های تکراری و یا اشتباه مربوط به هر برند حذف شده است.

* **نرمال سازی ستون price:** در مورد مدل های اصلی برند آیفون، حداقل و حداکثر قیمت منطقی برای هر مدل با جست و جو، جمع آوری شده است و آگهی هایی که خارج از بازه حداقل و حداکثر قیمت مدل خود بودند، حذف شدند.

و به طور کلی نیز آگهی های تمام برند ها که قیمت آنها خارج از بازه منطقی بود، حذف شدند.

سپس با استفاده از متد **IQR** مقادیر outliers هر برند جداگانه شناسایی و حذف شد.

در قدم بعدی شناسایی شد که برای قیمت بعضی آگهی ها فقط "1" وارد شده و برای رفع این مشکل، قیمت این دسته از آگهی ها را برابر میانگین قیمت برند و مدل آن گوشی قرار داده شد.

**2- تبدیل نام شهر، به استان:** به دلیل اینکه در آگهی ها، نام شهرها ذکر شده بود اما نام استان کاربردی تر به نظر میرسید، فرایند زیر انجام شد:

• در فایل (get data city.ipynb) نام هر شهر کشور ایران به همراه استان آن شهر از یک صفحه سایت و با استفاده از کد های html آن صفحه، با استفاده از فانکشن BeautifulSoup کرال شد و داخل یک فایل اکسل قرار گرفت.

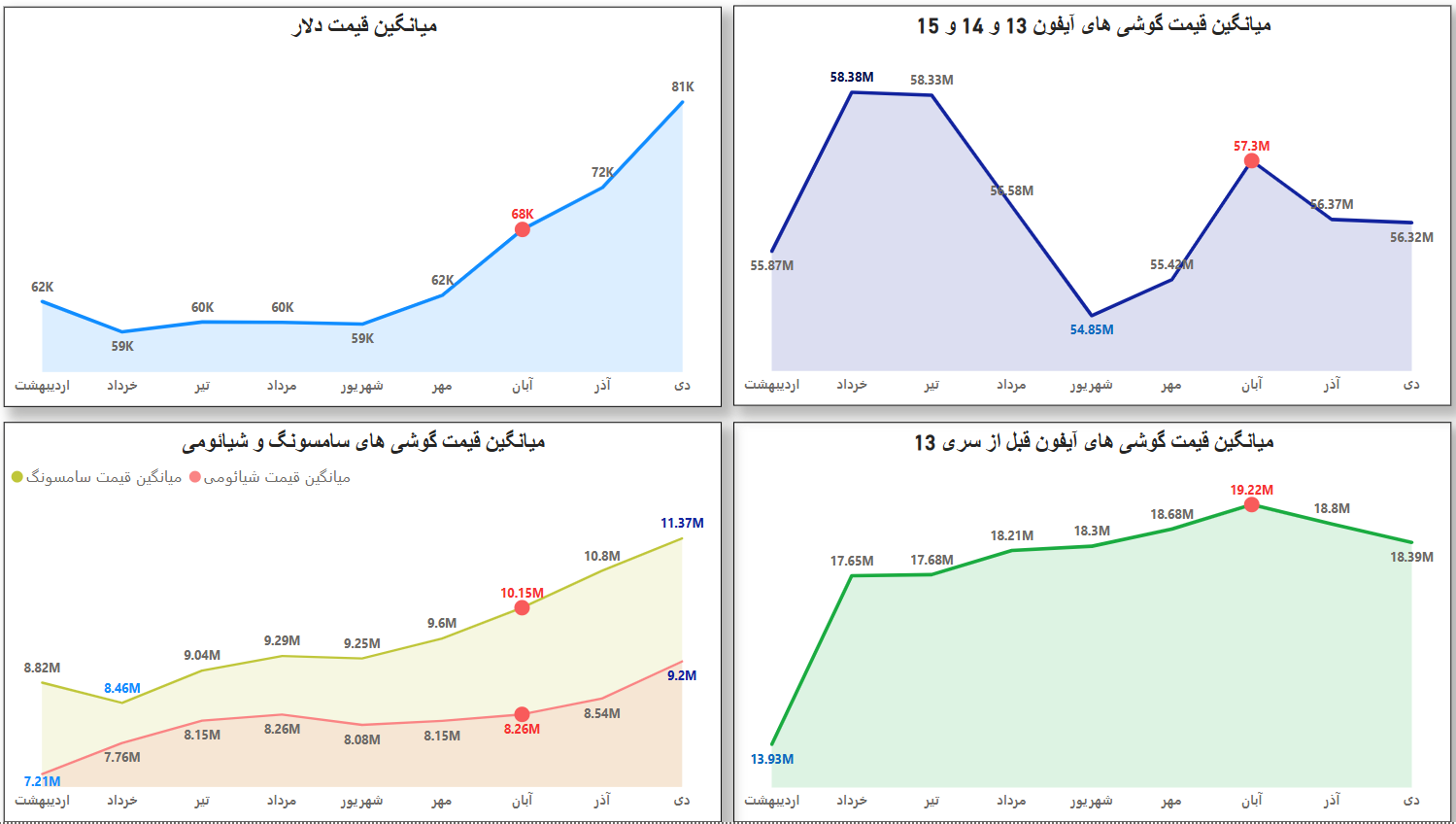
• سپس در فایل (merge city.ipynb) ابتدا دیتاست اصلی آگهی های دیوار خوانده شد و نام بعضی از شهرها که به اشتباه وارد شده بود تصحیح شد.

• سپس با فایل دیتاست آگهی های دیوار با فایل اکسل نام شهر به استان merge شد و به دیتاست، نام استان هر آگهی نیز اضافه شد.

**3-** دیتاست نهایی در برنامه power bi وارد شد و کارهای زیر روی آن انجام شد:

* ستون تاریخ شمسی split شد و ستون های عدد روز، عدد ماه و عدد سال داخل یک ستون جدا قرار گرفت.
* یک جدول واسط به نام MonthsNumber ساخته شد که عدد ماه را به نام ماه شمسی تبدیل میکرد و با استفاده از این جدول واسط در واقع در دیتاست به جای عدد ماه، نام ماه شمسی قرار گرفت.
* چندین measure برای داده های مختلف از جمله میانگین قیمت گوشی های آیفون، میانگین قیمت گوشی های سامسونگ و... و همچنین تعداد آگهی های برند آیفون، تعداد آگهی های برند شیائومی و... ایجاد شد.
* دیتاست به group های کوچک تری تبدیل شد.
* **در نتیجه دو فرایند بالا، پیش پردازش روی داده های قیمت ارز ها و همچنین داده های آگهی های دیوار صورت گرفت و این دو دیتاست، آماده بررسی و تحلیل و همچنین فرایند مصورسازی در نرم افزار power bi شدند.**
* **بررسی روند تغییرات قیمت و تاثیر نرخ ارز بر بازار گوشی ها**

**نمایش روند کلی قیمت ها:**

****

(تصویر 1)

* **توضیح نمودار ها(تصویر 1):**
* **نکات:**

**\*** در تمامی نمودار های تصویر 1 محور افقی(x) نشان دهنده ماه های شمسی از اردیبهشت تا دی هستش.

**\*** هدف از این تصویر نمایش روند کلی بازار ارز و گوشی های موبایل هست.

* **تشریح هر نمودار(تصویر 1):**

1. **میانگین قیمت دلار**:

**\*** یک نمودار خطی(line chart) هست که در محور عمودی این نمودار میانگین قیمت ارزِ دلار در هر ماه رسم شده است.

\*در این نمودار میتوان مشاهده کرد که قیمت دلار در ماه های ابتدایی یعنی خرداد تا شهریور، روند ثابت با تغییرات کمی را داشته است، اما از مهر روند رشد خودش رو با سرعت زیادی شروع کرده به طوری که از مهر تا دی بازه مبلغی 62 هزار تومان تا 81 هزار تومان یا به عبارتی رشد 30.6 درصدی را طی کرده است؛ که به طبع تاثیرات به ویژه ای بر بازار موبایل گذاشته است.

1. **میانگین قیمت گوشی های برند سامسونگ و شیائومی:**

\* یک نمودار خطی(line chart) هست که در محور عمودی این نمودار دو خط رسم شده است که رنگ سبز میانگین قیمت گوشی های سامسونگ و رنگ قرمز میانگین قیمت گوشی های شیائومی را در هر ماه نشان میدهد.

\* در این نمودار به وضوح میتوان مشاهده کرد که قیمت گوشی های دو برند اصلی یعنی سامسونگ و شیائومی دقیقا مانند نمودار دلار حرکت کرده و این دو نیز از مهر ماه روند رشد خودشون رو همزمان با افزایش قیمت دلار، شروع کرده اند.

\* این دو خط و شباهت روند آنها با روند حرکتی قیمت دلار، به ما این نکته رو نشون میدهد که گوشی های برند سامسونگ و شیائومی در تمام ماه ها با ضریب همبستگی بالایی، از قیمت دلار پیروی میکنند و بسیار تحت تاثیر قیمت دلار هستن.

\* همچنین نشان میدهد که طرح رجیستری گوشی های آیفون که در آبان ماه که با نقطه قرمز رنگ در نمودار مشخص شده است، تاثیر چندانی روی قیمت این دو برند اصلی بازار نداشته است و قیمت این دو برند روند پیش بینی شده خودشون رو بر اساس روند قیمت دلار، طی کرده اند.

1. **میانگین قیمت گوشی های آیفون 13 و 14 و 15**:

**\*** یک نمودار خطی(line chart) هست که در محور عمودی این نمودار میانگین قیمت سه سری آخر آیفون یعنی 13 و 14 و 15 که بیشترین تاثیر رو از مصوبه رجیستری میگرفتن در هر ماه رسم شده است.

**\*** در این نمودار میتوان مشاهده کرد که تا قبل از آبان ماه، میانگین قیمت این سه سری آیفون نیز با ضریب همبستگی به نسبت بالایی با نمودار قیمت دلار حرکت میکند.

\* در آبان ماه که مصادف هست با تصویب طرح رجیستری آیفون، بین روند حرکت قیمت آیفون و قیمت دلار، واگرایی اتفاق افتاده است.

\* در واقع از آبان ماه تا دی ماه که قیمت دلار رشد 30.6 درصدی رو تجربه کرد و به دلیل ضریب همبستگی بالا بین این دو نمودار، پیش بینی میشد قیمت سری های آیفون نیز رشد زیادی داشته باشد، اما این مساله برعکس شد و میانگین قیمت گوشی های آیفون 13 و 14 و 15 بعد از آبان ماه و بعد از تصویب طرح رجیستری، نه تنها رشد نکرد، بلکه با کاهش کمی نیز مواجه شد.

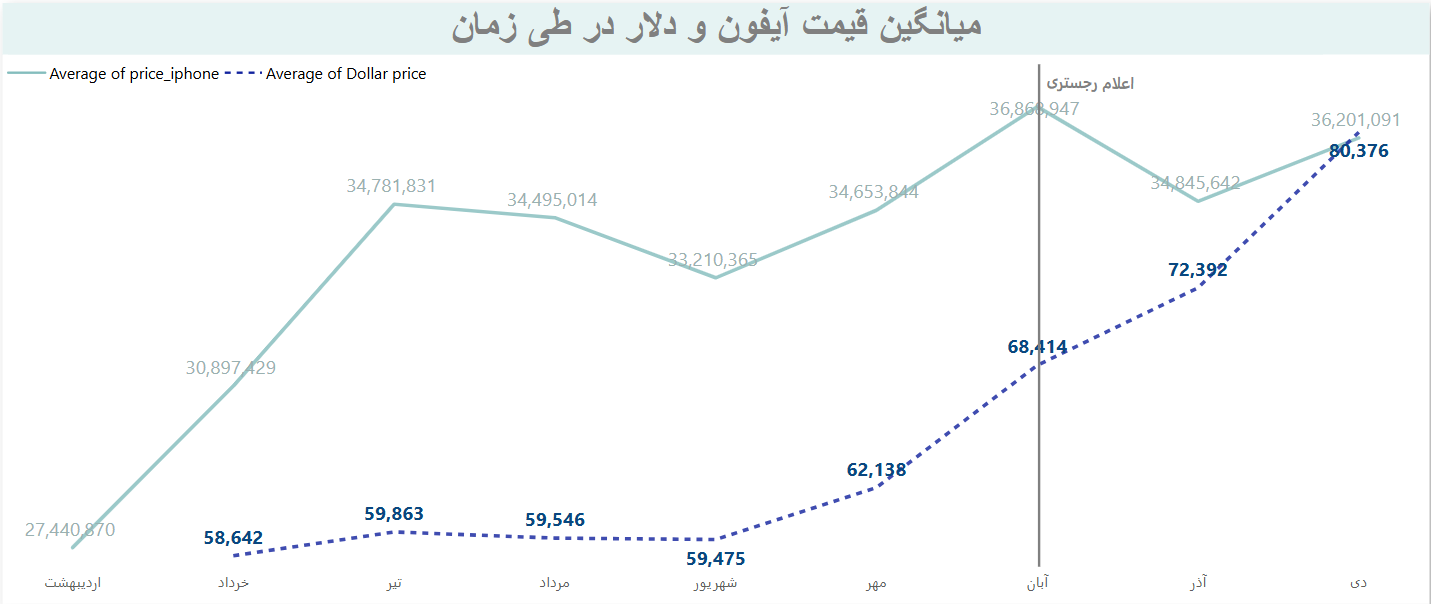
1. **میانگین قیمت گوشی های آیفون قبل از سری 13**:

**\*** یک نمودار خطی(line chart) هست که در محور عمودی این نمودار میانگین قیمت سری های قبل از آیفون 13 که بحث رجیستری شدن براشون مطرح نبوده در هر ماه رسم شده است.

**\*** در این نمودار نیز میتوان مشاهده کرد که تا قبل از آبان ماه، میانگین قیمت این سری آیفون ها نیز با ضریب همبستگی خوبی با نمودار قیمت دلار حرکت میکند.

\* مجدد در آبان ماه که مصادف هست با تصویب طرح رجیستری آیفون، بین روند حرکت قیمت سری های پیشین آیفون و قیمت دلار، واگرایی اتفاق افتاده است.

\* در واقع از آبان ماه تا دی ماه که قیمت دلار رشد 30.6 درصدی رو تجربه کرد و به دلیل ضریب همبستگی خوب بین این دو نمودار، پیش بینی میشد قیمت سری های پیشین آیفون نیز رشد زیادی داشته باشد، اما این مساله برعکس شد و میانگین قیمت گوشی های آیفون قبل از سری 13 بعد از آبان ماه و بعد از تصویب طرح رجیستری، نه تنها رشد نکرد، بلکه با کاهش کمی نیز مواجه شد.



(تصویر 2)

* **توضیح نمودار (تصویر 2):**

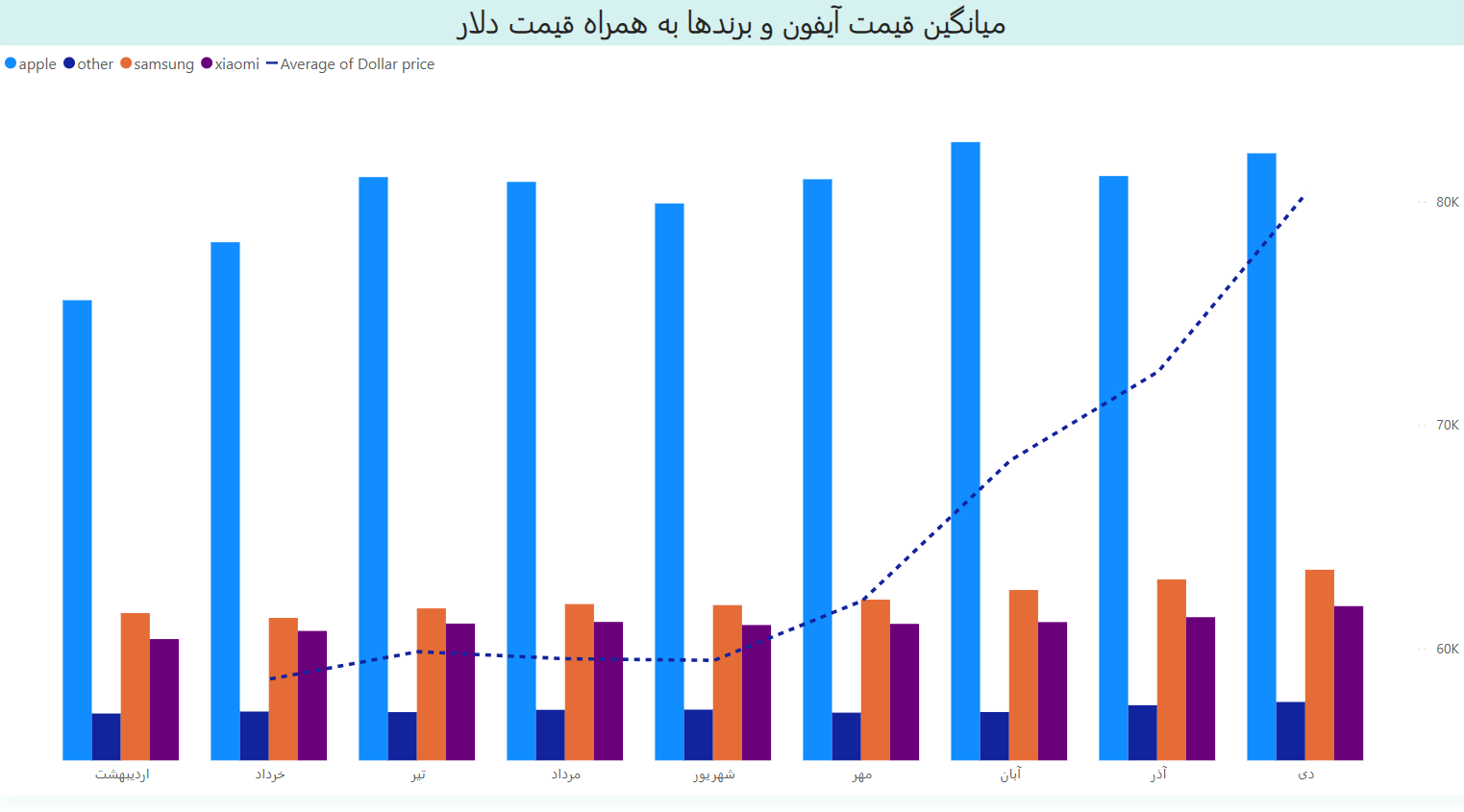
\* یک نمودار خطی(line chart) رسم شده است که در محور افقی(x) آن ماه های شمسی قرار گرفته است و در محور عمودی(y) آن دو خط رسم شده است که خط سبز رنگ نشان دهنده میانگین قیمت کل سری های آیفون و خط آبی رنگ نیز نشان دهنده میانگین قیمت دلار در هر ماه هست.

\* خط عمودی کشیده شده بر روی چارت نشان دهنده ماه اعلام رجیستری می باشد.

* **تشریح نمودار (تصویر 2):**

\* در این نمودار هم به وضوح مشخص هست که تا قبل از آبان ماه، میانگین قیمت کل سری های آیفون با ضریب همبستگی بالایی حرکت میکند و قیمت گوشی ها تاثیر مستقیم و زیادی از قیمت دلار دریافت میکنند.

\* اما بعد از آبان ماه و بعد از تصویب طرح رجیستری آیفون، روند حرکت قیمت کل سری های آیفون تغییر پیدا کرده است و برخلاف دلار که رشد 30.6 درصدی را تجربه کرده، با کاهش محسوسی همراه بوده است.



(تصویر 3)

* **توضیح نمودار (تصویر 3):**

\* یک نمودار line and clustered column chart رسم شده است که در محور افقی(x) آن ماه های شمسی قرار گرفته است و در محور عمودی(y) آن چهار نوع column رسم شده هست به این صورت که:  
 - رنگ آبی کمرنگ نشان دهنده میانگین قیمت تمام گوشی های برند اپل

- و رنگ نارنجی نشان دهنده میانگین قیمت گوشی های برند سامسونگ

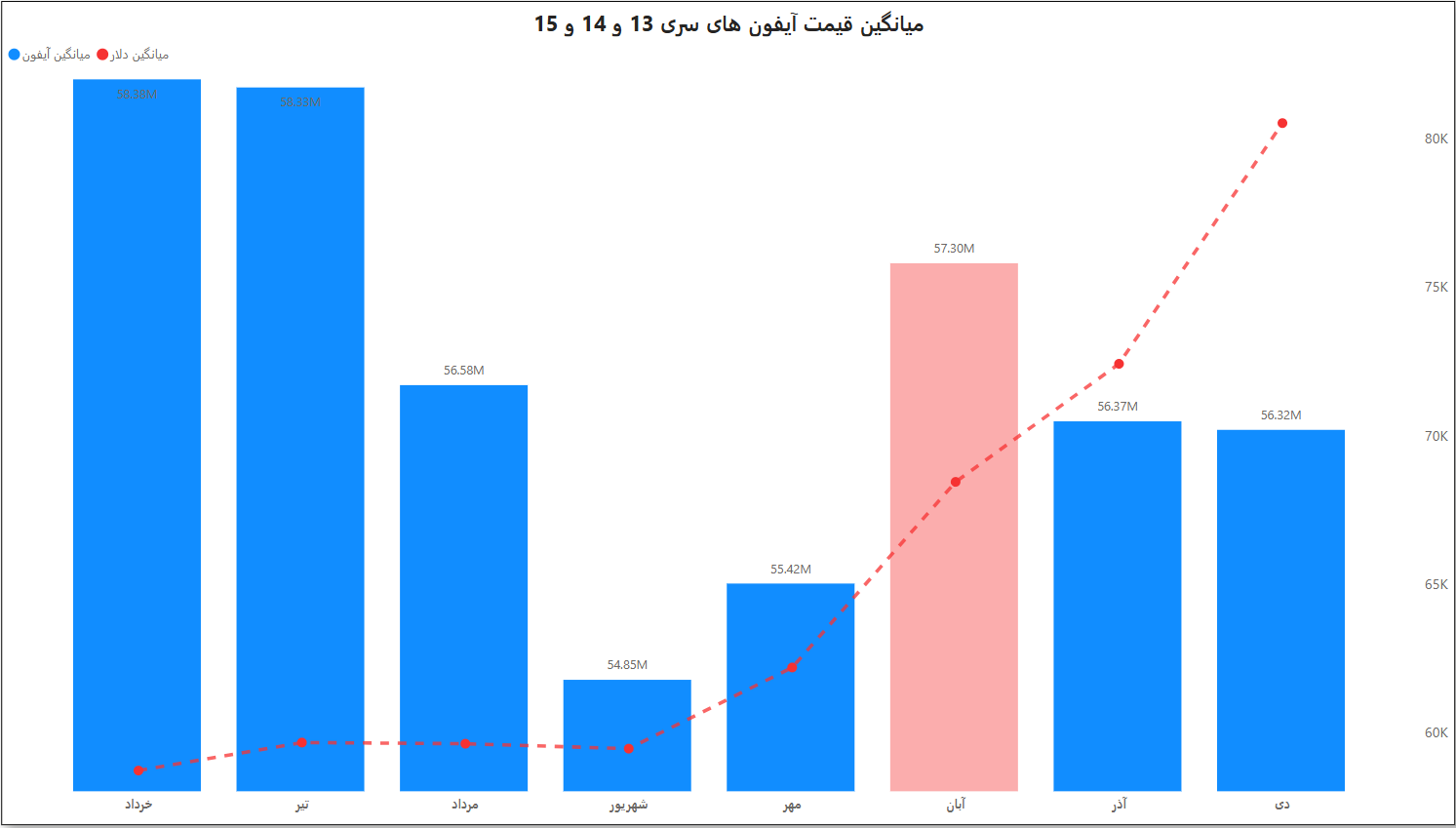
- و رنگ بنفش نشان دهنده میانگین قیمت گوشی های برند شیائومی

- و رنگ آبی پررنگ نیز نشان دهنده میانگین قیمت گوشی های سایر برند ها در هر ماه هستند.

\* خط آبی رسم شده در چارت نیز میانگین قیمت دلار را در هر ماه نشان میدهد.

* **تشریح نمودار (تصویر 3):**

\* با دقت در این نمودار هم میتوان به این نکته که قیمت گوشی های برند سامسونگ، شیائومی و سایر برند ها تحت تاثیر دلار حرکت میکنند و حتی بعد از آبان ماه و اعلام طرح رجیستری کماکان به حرکت خود ادامه داده اند، اما قیمت گوشی های برند اپل که تا قبل از آبان از روند قیمت دلار پیروی میکردند، پس از اعلام طرح رجسیتری آیفون، روند حرکتشون دست خوش تغییر شد و با کمی کاهش مواجه شدند.



(تصویر 4)

* **توضیح نمودار (تصویر 4):**

\* یک نمودار line and clustered column chart رسم شده است که در محور افقی(x) آن ماه های شمسی قرار گرفته است و در محور عمودی(y) آن یک نوع column رسم شده که نشان دهنده میانگین قیمت آیفون های سری 13 و 14 و 15 در هر ماه هست و یک خط چین قرمز رنگ رسم شده که نشان میانگین قیمت دلار در هر ماه هستش.

* **تشریح نمودار (تصویر 4):**

\* در این نمودار تاثیر قیمت دلار و همچنین طرح رجیستری بر قیمت آخرین سری های آیفون با وضوح بهتری نمایان شده است.

\* با نگاه به این نمودار نیز میتوان مشاهده کرد که قیمت این سری های آیفون پس از تصویب طرح رجیستری بر خلاف قیمت دلار، کاهش داشته است.

* **نتیجه گیری:**

\* میانگین قیمت برند های سامسونگ، شیائومی و حتی آیفون، تا قبل از آبان، همگی روندی همانند روند میانگین قیمت دلار را در هر ماه طی میکردند؛ درواقع میتوان گفت تا قبل از تصویب طرح رجیستری آیفون، روند حرکت قیمت گوشی ها تا نسبت زیادی، تحت تاثیر روند قیمت دلار بوده است.

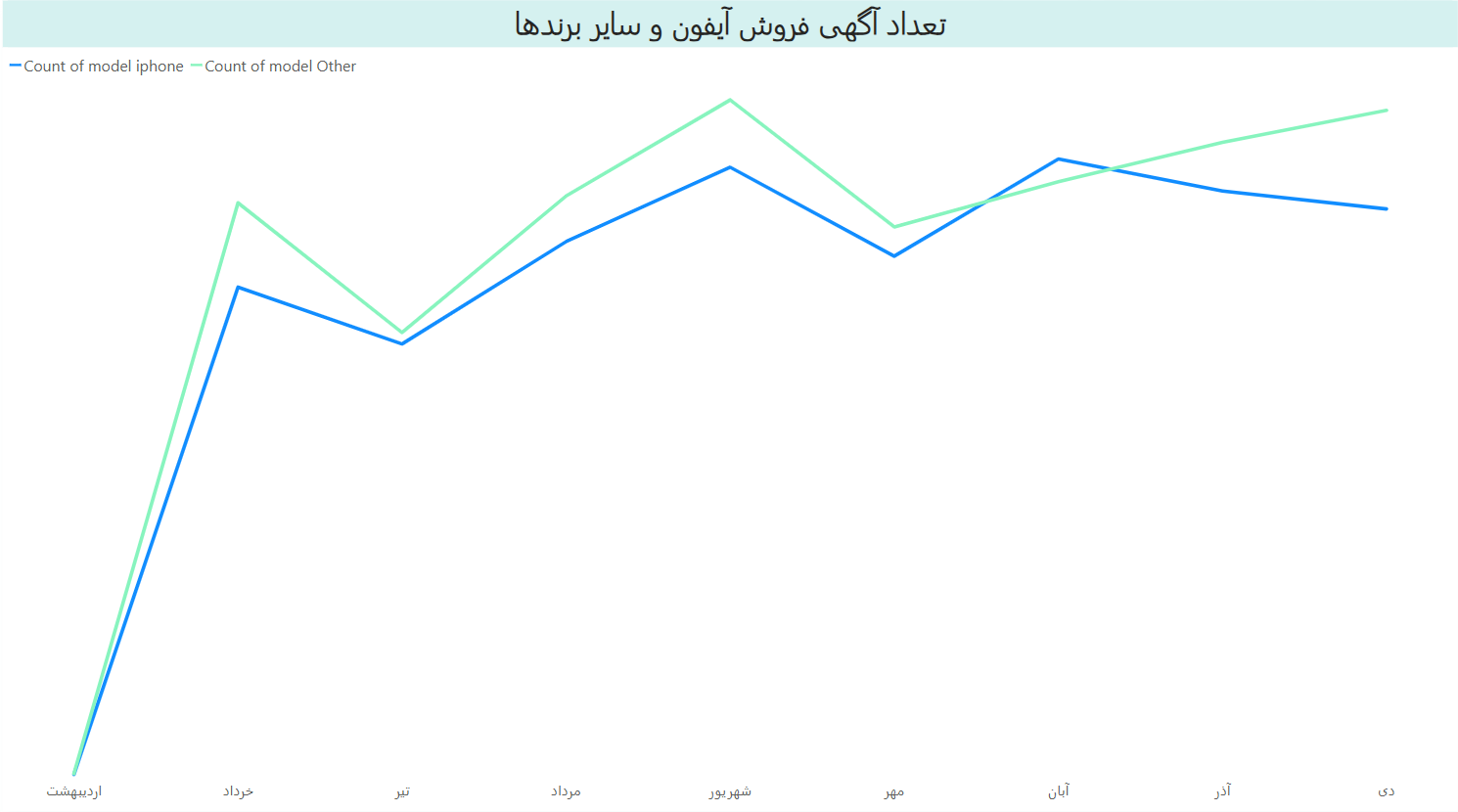
\* گوشی های برند سامسونگ و شیائومی بعد از آبان ماه نیز روند افزایش قیمت خودشون رو به نسبت با دلار پیش گرفتند و طرح رجیستری آیفون، تاثیر چندانی بر قیمت برند های رقیب آیفون ایجاد نکرد.

\* اما تصویب طرح رجیستری آیفون در 9 آبان و سپس اصلاح این طرح در 17 آبان، تغییر چشمگیر و جالبی را در بازار گوشی های آیفون ایجاد کرد.

\* قیمت گوشی های برند آیفون، چه گوشی های سری 13 و 14 و 15 (که بیشترین تاثیر رو از طرح رجیستری میگرفتن) و چه سری های قبل از 13 (که رجیستری برای آنها از قبل وجود داشت) بعد از آبان ماه، حتی با اینکه قیمت دلار رشد زیادی رو سپری کرده بود و انتظار میرفت قیمت این گوشی ها هم افزایش زیادی داشته باشد، اما تصویب این طرح توانست علاوه بر اینکه از رشد زیاد قیمت گوشی آیفون جلوگیری کند، بازار گوشی های آیفون را نیز به ثبات خوبی برساند.

**\*در کل میتوان این نکته را ذکر کرد که طرح رجیستری در کنترل قیمت گوشی های آیفون، کاملا موفق بوده است!**

* **بررسی روند عرضه گوشی ها در بازار**

****

(تصویر 5)

* **توضیح نمودار (تصویر 5) :**

\* یک نمودار خطی(line chart) هست که در محور افقی(x) آن ماه های شمسی قرار گرفته است و محور عمودی(y) آن از دو خط با رنگ های مختلف تشکیل شده است؛ رنگ آبی نشان دهنده تعداد تعداد آگهی گوشی های برند اپل و رنگ سبز نشان دهنده تعداد آگهی گوشی های سایر برند ها در هر ماه هست.

* **تشریح نمودار (تصویر 5):**

\* در این نمودار حجم عرضه و تعداد آگهی های برند اپل با سایر برند ها به طور جامع بررسی شده است و در تصویر 6، به صورت جزئی تر این بررسی اتفاق افتاده است.

\* همانطور که در تصویر مشاهده میکنید حجم آگهی ها و درواقع عرضه گوشی های برند اپل در آبان ماه که مصادف هست با تصویب طرح رجیستری، بیشتر از حجم عرضه گوشی های سایر برند ها بوده، اما در آذر و دی این روند برعکس شده و تعداد آگهی های سایر برند ها به سرعت از تعداد آگهی های گوشی های برند اپل پیشی گرفته اند.

**A graph of a graph

AI-generated content may be incorrect.**

(تصویر 6)

* **توضیح نمودار (تصویر 6) :**

\* یک نمودار خطی(line chart) هست که در محور افقی(x) آن ماه های شمسی قرار گرفته است و محور عمودی(y) آن از سه خط با رنگ های مختلف تشکیل شده است؛ رنگ آبی نشان دهنده تعداد تعداد آگهی گوشی های آیفون سری 13 و 14 و 15 ، رنگ آبی پررنگ نشان دهنده تعداد آگهی گوشی های سامسونگ و رنگ سبز نیز نشان دهنده تعداد آگهی های گوشی های شیائومی در هر ماه هست.

* **تشریح نمودار (تصویر 6):**

\* همانطور که در تصویر مشاهده میکنید حجم آگهی ها و درواقع عرضه گوشی های سامسونگ و شیائومی، پس از آبان ماه و تصویب طرح رجیستری گوشی های آیفون یک روند صعودی کمی را داشته است و کمی بیشتر از ماه قبل از تصویب این طرح یعنی مهرماه شده است.

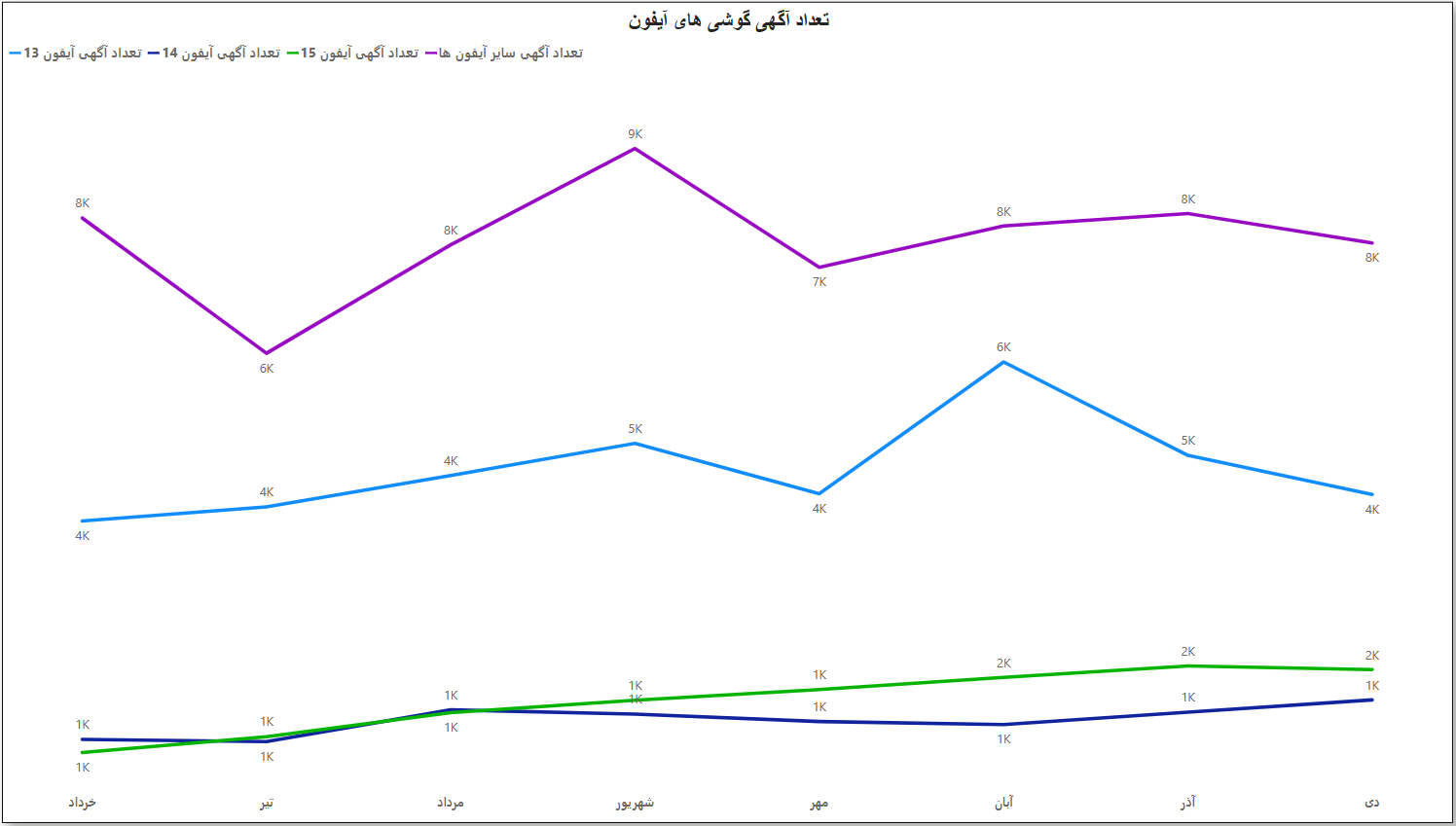
\* و همچنین مشاهده میشود که میانگین آگهی های گوشی های آیفون که تا قبل از آبان تقریبا 6 هزارتا در هر ماه بوده، پس از آبان ماه و اعلام رجیستری آیفون کمی بیشتر شد و به میانگین 7 هزار و 500 آگهی در هر ماه رسید.

* **نتیجه گیری:**

\* تصویب طرح رجیستری علاوه بر اینکه طبق توضیحات بخش قبلی، توانست قیمت گوشی های آیفون رو متعادل نگه دارد و از رشد اون جلوگیری کند، در بخش عرضه و تعداد آگهی های گوشی های آیفون هم تا حدودی توانست ثبات ایجاد کند.

\* میانگین تعداد آگهی گوشی های آیفون سری 13 به بعد که تا قبل از آبان با مشکل عدم رجیستری مواجه بودند، بعد از آبان ماه و با تصویب طرح رجیستری، کمی بیشتر شد و میتوان گفت گوشی های آیفون 14 و 15 و حتی 16 هم بعد از رجیستری، بیشتر وارد بازار عرضه شدند و خریدار ها از خرید گوشی های آیفون پس از این طرح منصرف نشدند.

* **بررسی روند گوشی های آیفون**



(تصویر 7)

* **توضیح نمودار (تصویر 7) :**

\* یک نمودار خطی(line chart) هست که در محور افقی(x) آن ماه های شمسی قرار گرفته است و محور عمودی(y) آن از چهار خط با رنگ های مختلف تشکیل شده است و هرکدام تعداد آگهی های سری خاصی از گوشی های برند آیفون را در هر ماه به نشان میدهد به صورتی که:

- رنگ آبی کم رنگ برای نمایش سری 13، رنگ آبی پررنگ برای نمایش سری 14، رنگ سبز برای نمایش سری 15 و رنگ بنفش نیز برای نمایش سایر سری های آیفون( سری های قبل از 13 ) استفاده شده است.

A screenshot of a graph

AI-generated content may be incorrect.

(تصویر 8)

* **توضیح نمودار (تصویر 8) :**

\* یک نمودار clustere column chart هست که محور افقی آن به دو دسته قبل از رجیستری و بعد از رجیستری تقسیم شده و در محور عمودی(y) آن پنج نوع column رسم شده که هر کدام میانگین قیمت گوشی های مدل خاصی از برند آیفون رو در دو بازه زمانی قبل از رجیستری و بعد از رجیستری نشان میدهد به این صورت که:  
 - رنگ آبی کمرنگ برای گوشی های سری 13 برند آیفون،

- و رنگ آبی پررنگ برای گوشی های سری 14 برند آیفون،

- و رنگ نارنجی برای گوشی های سری 15 برند آیفون،

- و رنگ بنفش برای گوشی های سری 16 برند آیفون،

- و رنگ صورتی نیز برای گوشی های سایر مدل های برند آیفون استفاده شده است.

* **نتیجه گیری:**

\* اصلی ترین نتیجه ای که میشه از حجم عرضه و همچنین میانگین قیمت سری های مختلف برند آیفون قبل و بعد از طرح رجیستری داشت به این صورت هست:

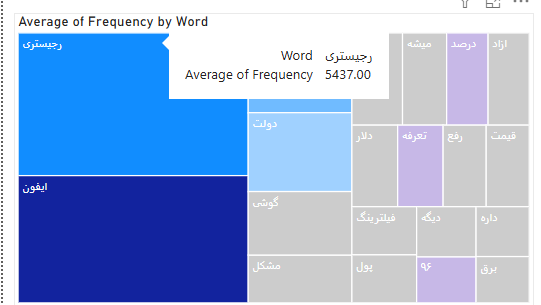
- در دوره قبل از اجرای طرح رجیستری( قبل از آبان ماه) که گوشی های سری 14 و 15 و 16 در ایران رجیستر نمیشدند، سری آیفون 13 حتی با اینکه سری های جدیدتر هم به بازار آمده بود، عرضه بسیار زیاد داشت و همچنین به دلیل تقاضای بالا برای این سری، قیمت به نسبت بالایی هم داشت؛

- اما در دوره بعد از تصویب طرح رجیستری( بعد از آبان ماه) که دیگه گوشی های سری 14 و 15 و 16 هم رجیستر میشدن و مردم میتونستن ازشون استفاده کنن، به تدریج عرضه سری آیفون 13 کاهش پیدا کرد و همچنین به دلیل کاهش تقاضا برای این گوشی، قیمت آن نیز روند کاهش ای رو طی کرد؛

- و دقیقا بر خلاف آیفون 13، سری 14 و 15 جون تازه گرفت و عرضه و تقاضای آن در بازار موبایل بیشتر از قبل شد.

1. **مسائل اجتماعی**

* پیش پردازش داده های مربوط به توییتر:
* هدف از این مرحله تبدیل داده های جمع آوری شده از توییتر به data frame های تمیز و قابل استفاده هستش.
* مراحل انجام این کار به صورت زیر بوده است:
* ابتدا با دستورات پایتون دیتاست خوانده شده است، ستون های اضافی حذف شده اند و دیتا های null با مقادیر مناسبی پر شدند و غلط های املایی شناسایی شده و تصحیح شدند( به دلیل اضافه شدن دیتاست جدید، یک مرحله ترکیب دو دیتاست توییتر نیز انجام شده است).
* با استفاده از کتابخانه persiantools تاریخ میلادی به تاریخ شمسی تبدیل شده است.
* از دیتای توییت ها مواردی مانند لینک ها، هشتگ ها و منشن ها شناسایی شده و در دسته بندی های جداگانه ای قرار گرفته اند.
* مراحل Preprocess دیتا با استفاده از کتابخانه hazm و متد هایی مانند Normalizer و WordTokenizer و Lemmatizer روی تمام توییت ها انجام شده است.
* با استفاده از فانکشن Counter، 20 کلمه پر تکرار شناسایی شده است.
* با استفاده از کتابخانه transformers مدل های مختلف تشخیص احساسات روی دیتا ها تست شده است تا نتایج مدل های مختلف بررسی شود.
* به دلیل نتایج ضعیف مدل ها برای متون فارسی، در لیبل گذاری و تشخیص احساسات، با پرامپت های متعدد و حرفه ای، از ChatGPT استفاده شده است تا توییت ها لیبل گذاری شوند.
* مراحل متعددی برای دسته بندی کاربران و شناسایی احساسات آنها انجام شده است،
* مانند: دسته بندی توییت ها بر اساس متن توییت به زیر مجموعه های کوچک تر مانند (کوچیک شمردن رجیستری، مشکل فنی-تعلل رجیستری، گرانی-رجیستری و...) و شناسایی نوع احساس از بین سه احساس مثبت، منفی و خنثی برای هر توییت.
* به دست آوردن احساس غالب هر فرد نسبت به رجیستری.
* همچنین داده ها برای رسم گراف و استفاده در برنامه گفی، به حالت های مورد نیاز تبدیل شده اند.
* تمام مراحل Preprocess به دلیل اینکه در جاهایی مختلفی استفاده میشد، تبدیل به کتابخانه شده است.



این نمودار که به نظر می‌رسد از نوع "Treemap" است، برای نمایش فراوانی کلمات مختلف در داده‌های متنی (مانند توئیت‌ها یا نظرات کاربران) استفاده می‌شود. در این نوع نمودار، ابعاد مستطیل‌ها به تناسب فراوانی کلمات تغییر می‌کند، به‌طوری که کلمات پرتکرار یا بیشتر مورد استفاده در داده‌ها با ابعاد بزرگتر نمایش داده می‌شوند. در ادامه تحلیل این نمودار آورده شده است:

### تحلیل نمودار:

**کلمات پرکاربرد**: در این نمودار، کلمات "رجیستری" و "آیفون" به‌عنوان کلمات پرکاربرد در داده‌ها مشاهده می‌شوند. این کلمات به‌طور بارزی در نمودار با ابعاد بزرگتر نمایش داده شده‌اند که نشان‌دهنده این است که در توئیت‌ها و نظرات کاربران، این دو کلمه بیشترین فراوانی را دارند. این می‌تواند نشان‌دهنده این باشد که کاربران بیشتر درباره طرح رجیستری و تاثیرات آن صحبت کرده‌اند.

**الگوی کلمات دیگر**: در کنار این دو کلمه پرکاربرد، کلمات دیگری مانند "دولت"، "قیمت"، "مشکل" و "گزارش" در ابعاد کوچکتری قرار دارند. این نشان می‌دهد که سایر موضوعات به‌طور کمتری در داده‌ها ذکر شده‌اند، اما همچنان حائز اهمیت هستند.

**نقاط توجه**: این نمودار به‌طور خاص می‌تواند به شناسایی کلمات کلیدی کمک کند که بیشترین توجه را در یک دوره زمانی خاص به خود جلب کرده‌اند. برای مثال، اگر کلمه "رجیستری" بیش از سایر کلمات در نمودار ظاهر شده باشد، نشان‌دهنده اهمیت بالای این موضوع برای کاربران در آن زمان است.

**استفاده از نمودار Treemap**: انتخاب نمودار "Treemap" به‌عنوان ابزار نمایش این داده‌ها منطقی است، زیرا این نمودار به‌طور مؤثری کلمات را بر اساس فراوانی آن‌ها مقایسه می‌کند و می‌تواند به‌طور سریع و واضح، مهم‌ترین کلمات در داده‌ها را نشان دهد. این نمودار در تحلیل‌هایی که نیاز به مقایسه فراوانی کلمات دارند، بسیار مفید است و به تحلیلگر امکان می‌دهد به‌راحتی و با یک نگاه کلی، الگوهای متنی موجود در داده‌ها را شناسایی کند.

### نتیجه‌گیری:

نمودار Treemap به‌طور موثری به تحلیل فراوانی کلمات در داده‌ها پرداخته و مهم‌ترین کلمات را در مقیاس‌های مختلف به نمایش می‌گذارد. این نوع نمودار به تحلیلگر کمک می‌کند تا کلمات یا موضوعات کلیدی که بیشترین توجه کاربران را به خود جلب کرده‌اند، شناسایی کند و روندهای اصلی را در تحلیل‌های بعدی درک کند.

A graph with numbers and a triangle

AI-generated content may be incorrect.

**محور افقی (ماه‌ها)**: محور افقی نشان‌دهنده ماه‌های سال است، که از خرداد تا بهمن ادامه دارد. این بازه زمانی به‌طور خاص نشان‌دهنده این است که احساسات کاربران نسبت به موضوعات مختلف (مثلاً طرح رجیستری) در این مدت زمانی چطور تغییر کرده است.

**محور عمودی سمت چپ (تعداد احساسات)**: محور عمودی سمت چپ به تعداد دفعاتی که هر احساس (مثبت، منفی، یا خنثی) توسط کاربران بیان شده است اختصاص داده شده است. این مقیاس به‌صورت دسته‌بندی‌شده برای هر احساس (مثبت، منفی، یا خنثی) استفاده می‌شود و نشان می‌دهد که هر کدام از این احساسات در هر ماه چه تعداد بروز داشته است.

**محور عمودی سمت راست (تعداد کاربران)**: محور عمودی سمت راست نمایانگر تعداد کاربران یا توئیت‌هایی است که در هر ماه ثبت شده‌اند. نمودار خطی این محور را نشان می‌دهد و می‌توان فهمید که در هر ماه تعداد کاربران یا توئیت‌ها چگونه تغییر کرده است.

**توزیع احساسات**:

در ماه آبان، به‌ویژه یک افزایش بزرگ در تعداد احساسات مشاهده می‌شود، به‌طوری که بیشتر این احساسات منفی (خط آبی رنگ) هستند. این می‌تواند به‌دلیل افزایش نگرانی‌ها یا نارضایتی‌ها از طرح رجیستری یا تغییرات مرتبط با آن باشد.

احساسات مثبت (خط سبز رنگ) نسبت به احساسات منفی در ماه‌های دیگر به‌طور نسبتاً برابر یا کمی کمتر است، به‌خصوص در ماه‌هایی مانند تیر و مرداد که تعداد احساسات مثبت و خنثی (خط خاکی رنگ) بیشتر مشاهده می‌شود.

از ماه‌های آبان و آذر می‌توان فهمید که مردم در این ماه‌ها بیشتر نسبت به مسائل اجتماعی (احتمالاً با توجه به طرح رجیستری و تاثیرات آن) واکنش نشان داده‌اند.

**انتخاب نمودارهای ترکیبی**: استفاده از نمودار ترکیبی (میله‌ای و خطی) برای این داده‌ها بسیار مناسب است زیرا این نمودار به‌طور مؤثری دو جنبه مختلف از داده‌ها را نمایش می‌دهد: توزیع احساسات در هر ماه و همچنین تعداد کلی کاربران یا توئیت‌ها در همان ماه. این به تحلیلگر این امکان را می‌دهد که درک بهتری از این داشته باشد که چگونه تغییرات در تعداد کاربران یا توئیت‌ها با تغییرات در احساسات (مثبت یا منفی) مرتبط است.

**نتیجه‌گیری**:

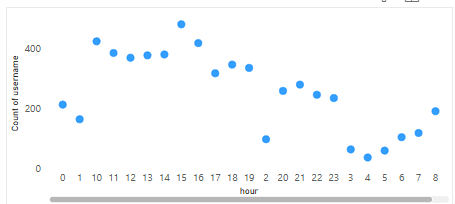
از این نمودار می‌توان به این نتیجه رسید که احساسات کاربران در رابطه با طرح رجیستری در برخی ماه‌ها به‌ویژه در آبان به اوج خود می‌رسد و بیشتر این احساسات منفی است.

این نمودار نشان می‌دهد که تعاملات کاربران و احساسات نسبت به طرح رجیستری در ماه‌های مختلف تغییر کرده است، به‌ویژه بعد از اعمال طرح رجیستری.

استفاده از نمودارهای ترکیبی به تحلیلگران این امکان را می‌دهد که تغییرات در احساسات و تعداد کاربران را در کنار هم مشاهده کنند و ارتباط بین این دو را بررسی کنند.

**نتیجه‌گیری کلی:**

این نمودار با ترکیب نمودار میله‌ای و خطی اطلاعات ارزشمندی را برای تحلیل احساسات کاربران در رابطه با طرح رجیستری فراهم می‌آورد. روندهای زمانی در تغییرات احساسات و تعداد کاربران می‌تواند به‌طور دقیق‌تر بررسی شود و نقاط اوج و کاهش احساسات منفی یا مثبت شناسایی گردد.



**تحلیل نمودار:**

**محور افقی (ساعت‌ها)**: محور افقی نشان‌دهنده ساعت‌های شبانه‌روز است که از ساعت ۰ تا ۲۳ به صورت پیوسته نمایش داده شده است. این ترتیب زمانی نشان‌دهنده این است که در هر ساعت از شبانه‌روز، چه تعداد کاربران در داده‌ها فعال بوده‌اند.

**محور عمودی (تعداد کاربران)**: محور عمودی نشان‌دهنده تعداد کاربران یا توئیت‌هایی است که در هر ساعت خاص ثبت شده‌اند. این مقیاس به تعداد کاربران یا توئیت‌های ثبت شده در هر ساعت اشاره دارد.

**نقاط پراکنده**: نمودار نقاط پراکنده را در طول ساعات مختلف شبانه‌روز نمایش می‌دهد که نشان می‌دهند در هر ساعت خاص چند کاربر یا توئیت ثبت شده است. این نقاط پراکنده به‌وضوح روندهای زمانی فعالیت کاربران را نشان می‌دهند.

**روند کلی**: در این نمودار می‌توان دید که تعداد کاربران در ساعات اولیه صبح (ساعت‌های ۰ تا ۸) به شدت کاهش یافته است و در ساعات عصر (ساعت‌های ۱۵ تا ۱۹) به بالاترین حد خود می‌رسد. این روند می‌تواند نشان‌دهنده این باشد که بیشترین میزان فعالیت کاربران در ساعات بعدازظهر و عصر رخ می‌دهد، در حالی که ساعات شب و صبح فعالیت کمتری دارند.

**استفاده از نمودار Scatter Plot**: انتخاب نمودار پراکندگی (Scatter Plot) به‌طور مؤثری به نمایش تغییرات زمانی و توزیع داده‌ها در هر ساعت از روز کمک می‌کند. این نوع نمودار به‌خصوص برای نمایش نوسانات و الگوهای زمانی در داده‌ها کاربرد دارد. به‌ویژه در اینجا که قصد داریم ساعات پیک فعالیت کاربران را شناسایی کنیم، این نمودار گزینه مناسبی است.

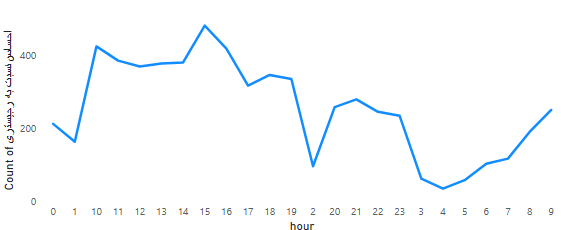
**نتیجه‌گیری**:

این نمودار نشان می‌دهد که کاربران در ساعات بعدازظهر و عصر بیشترین فعالیت را دارند.

نمودار پراکندگی به تحلیلگر این امکان را می‌دهد که به‌راحتی الگوهای فعالیت کاربران را بر اساس زمان مشاهده کند و روندهای زمانی مشخصی را شناسایی کند.

**نتیجه‌گیری کلی:**

این نمودار پراکندگی (Scatter Plot) با نمایش توزیع تعداد کاربران در هر ساعت از روز، به تحلیل‌گر کمک می‌کند تا ساعات پیک فعالیت و ساعات کم‌فعالیت کاربران را شناسایی کند. استفاده از این نمودار برای شناسایی الگوهای زمانی در داده‌های کاربران بسیار مفید است، به‌ویژه زمانی که می‌خواهیم زمان‌هایی که بیشترین تعاملات در آن‌ها صورت می‌گیرد را تحلیل کنیم.



**تحلیل نمودار:**

**محور افقی (ساعت‌ها)**: محور افقی نشان‌دهنده ساعت‌های شبانه‌روز است که از ساعت ۰ تا ۲۳ به صورت پیوسته نمایش داده شده است. این ساعت‌ها بیانگر زمان‌های مختلفی از روز هستند که کاربران احساسات مختلف خود را بیان کرده‌اند.

**محور عمودی (تعداد احساسات)**: محور عمودی نشان‌دهنده تعداد احساسات ثبت‌شده است که در هر ساعت خاص توسط کاربران به اشتراک گذاشته شده است. این احساسات می‌تواند مثبت، منفی یا خنثی باشد و به‌طور خاص تعداد دفعاتی را نشان می‌دهد که احساسات مختلف در آن ساعت بیان شده‌اند.

**روند کلی (خط آبی)**: خط آبی در نمودار نشان‌دهنده تغییرات تعداد احساسات در هر ساعت است. این خط بیانگر نوسانات در تعداد احساسات در طول روز است و به‌وضوح می‌توان مشاهده کرد که در ساعات مختلف شبانه‌روز، تعداد احساسات تغییر کرده است.

**نقاط اوج و افت**:

در ساعت‌های خاص، مانند ساعت ۱۷ تا ۱۸، مشاهده می‌شود که تعداد احساسات به اوج خود می‌رسد. این می‌تواند نشان‌دهنده یک پیک زمانی برای بیان احساسات کاربران باشد.

در ساعت‌های ۲۱ و ۳، افت قابل‌توجهی در تعداد احساسات مشاهده می‌شود. این کاهش می‌تواند به دلیل کم‌فعالیت‌تر بودن کاربران در این ساعات شبانه‌روز باشد.

**استفاده از نمودار خطی**: استفاده از نمودار خطی برای نمایش تغییرات زمانی احساسات بسیار مناسب است، زیرا این نوع نمودار به‌خوبی نوسانات داده‌ها را در طول زمان نشان می‌دهد و برای شناسایی روندها و الگوهای زمانی بسیار مؤثر است. در اینجا، هدف این است که به تحلیلگر کمک کند تا شناسایی کند که در چه ساعاتی از روز بیشترین تغییرات در احساسات کاربران رخ می‌دهد و این تغییرات چگونه با زمان‌بندی روزانه کاربران ارتباط دارند.

**نتیجه‌گیری**:

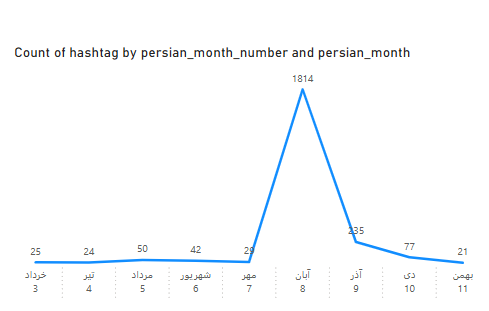
نمودار نشان می‌دهد که احساسات کاربران در ساعات مختلف از روز نوسان زیادی دارند.

پیک‌های بالاتر در ساعات خاص می‌تواند نشان‌دهنده ارتباط آن ساعات با تعاملات اجتماعی بیشتر یا بیشتر شدن نظرات کاربران نسبت به یک رویداد خاص باشد.

ساعات اوج فعالیت (مثلاً ساعت‌های ۱۷ تا ۱۸) می‌تواند نشان‌دهنده این باشد که افراد تمایل دارند نظرات و احساسات خود را در زمان‌های خاصی از روز (احتمالاً بعد از ساعات کاری یا در ساعات استراحت) بیشتر بیان کنند.

**نتیجه‌گیری کلی:**

این نمودار خطی با نمایش تغییرات تعداد احساسات در هر ساعت از روز، به تحلیلگر کمک می‌کند تا رفتار کاربران را در زمان‌های مختلف شناسایی کند. استفاده از این نمودار برای شناسایی ساعت‌هایی که بیشترین تعاملات احساسی و احساسات منفی یا مثبت وجود دارد بسیار مفید است. این تحلیل می‌تواند در تنظیم استراتژی‌های بازاریابی و تحلیل نظرات کاربران در بسترهای مختلف مفید واقع شود.



**تحلیل نمودار:**

**محور افقی (ماه‌ها)**: محور افقی نشان‌دهنده ماه‌های سال است که به ترتیب از خرداد تا بهمن نمایش داده شده‌اند. برای هر ماه، تعداد هشتگ‌های استفاده‌شده در آن ماه به‌طور جداگانه بررسی شده است.

**محور عمودی (تعداد هشتگ‌ها)**: محور عمودی نشان‌دهنده تعداد هشتگ‌ها است که در هر ماه به‌طور خاص استفاده شده‌اند. این مقیاس به تعداد هشتگ‌های استفاده‌شده در هر ماه اشاره دارد.

**روند کلی (خط آبی)**: خط آبی در نمودار نشان‌دهنده تغییرات تعداد هشتگ‌ها در هر ماه است. این خط نوسانات تعداد هشتگ‌ها را نشان می‌دهد و می‌توان مشاهده کرد که در برخی ماه‌ها تعداد هشتگ‌ها به‌طور چشمگیری افزایش پیدا کرده است.

**نقاط اوج و افت**:

در ماه مهر، تعداد هشتگ‌ها به‌طور قابل‌توجهی افزایش یافته است (۱۸۱۴ هشتگ)، که این افزایش چشمگیر ممکن است به‌دلیل یک رویداد یا کمپین خاص در آن ماه باشد.

ماه آذر نیز شاهد افزایش تعداد هشتگ‌ها بوده است، اما این افزایش به اندازه ماه مهر نبوده است.

سایر ماه‌ها (خرداد تا دی) تعداد هشتگ‌های کمتری دارند و میزان استفاده از هشتگ‌ها در این ماه‌ها نسبتا ثابت و پایین است.

**استفاده از نمودار خطی**: انتخاب نمودار خطی (Line Chart) به‌طور مؤثری به نمایش تغییرات ماهیانه استفاده از هشتگ‌ها کمک می‌کند. نمودار خطی به‌ویژه برای نمایش روندهای زمانی و شناسایی تغییرات و الگوهای فصلی یا مقطعی بسیار مناسب است. در اینجا، استفاده از این نمودار به‌خوبی می‌تواند الگوهای زمانی هشتگ‌ها را در طول ماه‌های مختلف سال نشان دهد و به تحلیل‌گر کمک کند تا ماه‌هایی که بیشترین فعالیت یا توجه به یک موضوع خاص وجود دارد را شناسایی کند.

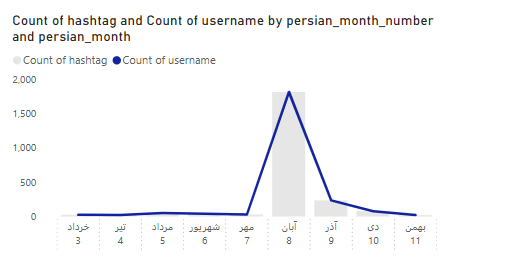
**نتیجه‌گیری**:

ماه مهر و آذر دارای بالاترین تعداد هشتگ‌ها هستند که می‌تواند نشان‌دهنده اوج فعالیت یا توجه به یک موضوع خاص در این ماه‌ها باشد.

سایر ماه‌ها، به‌ویژه خرداد، تیر، شهریور، و بهمن، تعداد هشتگ‌های کمتری دارند و فعالیت در این ماه‌ها نسبت به مهر و آذر کاهش یافته است.

**نتیجه‌گیری کلی:**

این نمودار خطی به تحلیل‌گر کمک می‌کند تا روند استفاده از هشتگ‌ها را در طول ماه‌های مختلف سال شناسایی کند. استفاده از این نوع نمودار برای تحلیل توزیع زمانی هشتگ‌ها و شناسایی زمان‌های اوج فعالیت در موضوعات خاص مفید است. در اینجا، بررسی ماه‌های خاص که در آن‌ها استفاده از هشتگ‌ها به طور قابل‌توجهی افزایش می‌یابد، می‌تواند نشانه‌ای از زمان‌های پیک توجه یا رویدادهای خاص باشد که موجب افزایش تعاملات و توجه کاربران به یک موضوع خاص می‌شود.



**ویژگی‌های نمودار:**

**محور افقی (ماه‌ها)**:

محور افقی ماه‌های سال را از خرداد تا بهمن نشان می‌دهد. این ماه‌ها بر اساس تقویم فارسی شماره‌گذاری شده‌اند و هر ماه با نام خود نمایش داده می‌شود.

**محور عمودی (تعداد هشتگ‌ها و تعداد کاربران)**:

دو خط مختلف در این نمودار به‌طور همزمان نمایش داده شده‌اند: یکی برای تعداد هشتگ‌ها (نشان‌دهنده تعداد هشتگ‌های استفاده‌شده در هر ماه) و دیگری برای تعداد کاربران (نشان‌دهنده تعداد کاربران که در هر ماه به‌طور فعال از هشتگ‌ها استفاده کرده‌اند).

از این‌رو، این نمودار به‌طور همزمان اطلاعاتی را در مورد فعالیت کاربران و استفاده از هشتگ‌ها در طول ماه‌های مختلف فراهم می‌کند.

**روند کلی (خط آبی و خط خاکی)**:

خط آبی نشان‌دهنده تعداد هشتگ‌ها است، در حالی که خط خاکی نشان‌دهنده تعداد کاربران است.

این دو خط نشان‌دهنده یک روند زمانی هستند که در آن هر دو متغیر در طول ماه‌ها تغییر می‌کنند. مشخص است که در ماه مهر، تعداد هشتگ‌ها و کاربران به‌شدت افزایش یافته است.

**نقاط اوج و افت**:

همانطور که در نمودار مشاهده می‌شود، در ماه مهر (آبان) تعداد هشتگ‌ها و کاربران به‌طور چشمگیری افزایش می‌یابد. این احتمالاً نشان‌دهنده یک رویداد یا کمپین خاص است که در این ماه توجه زیادی را جلب کرده است.

در سایر ماه‌ها، تعداد هشتگ‌ها و کاربران به‌طور نسبی ثابت یا پایین است.

**استفاده از نمودار خطی**:

انتخاب نمودار خطی برای نمایش این داده‌ها بسیار مؤثر است، زیرا این نوع نمودار به‌طور مؤثر تغییرات و روندهای زمانی را نشان می‌دهد. در اینجا، روند تعاملات کاربران و استفاده از هشتگ‌ها در طول زمان قابل مشاهده است.

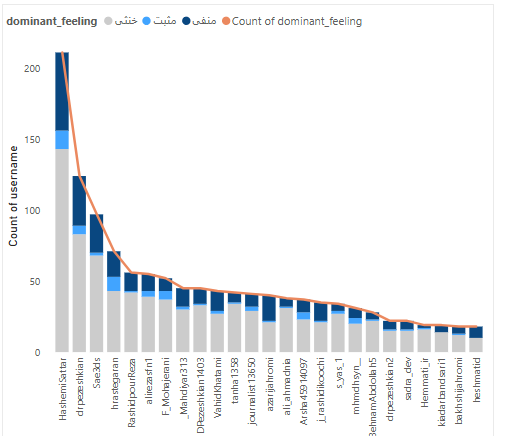
همچنین، خط آبی و خاکی به‌طور همزمان تعاملات مختلف (هشتگ‌ها و کاربران) را در کنار هم قرار می‌دهند و به تحلیل‌گر امکان مقایسه آن‌ها در هر ماه را می‌دهند.

**نتیجه‌گیری:**

**ماه مهر و آبان** بیشترین استفاده از هشتگ‌ها و بیشترین تعامل کاربران را نشان می‌دهند. این نشان‌دهنده یک پیک یا رویداد خاص است که موجب افزایش توجه به هشتگ‌ها و تعاملات کاربران در این ماه‌ها شده است.

**سایر ماه‌ها** به‌ویژه خرداد و تیر، کمتر از هشتگ‌ها استفاده کرده‌اند و تعاملات کاربران در این ماه‌ها کم‌تر از سایر ماه‌ها بوده است.

این نوع تحلیل و استفاده از نمودار خطی می‌تواند به تحلیل‌گران کمک کند تا زمان‌های اوج فعالیت کاربران در موضوعات خاص را شناسایی کنند و به‌طور بهتری درک کنند که چه ماه‌هایی از سال کاربران بیشتر به موضوعاتی خاص توجه کرده‌اند.



**ویژگی‌های نمودار:**

**محور افقی (نام‌های کاربری)**:

محور افقی نشان‌دهنده نام‌های کاربری مختلف است که بیشترین تعداد احساسات غالب (dominant feeling) را ثبت کرده‌اند.

برخی از کاربران مانند "HashemStar" بیشترین تعاملات و احساسات را نشان داده‌اند که موجب ظهور آن‌ها در بالای نمودار شده است.

**محور عمودی (تعداد کاربران و احساسات غالب)**:

محور عمودی در دو بخش تقسیم شده است:

**نوار خاکی‌رنگ**: نشان‌دهنده تعداد کاربران برای هر نام کاربری است.

**نوار آبی‌رنگ**: نشان‌دهنده احساسات غالب کاربران (منفی، مثبت، یا خنثی) در مورد موضوع یا رویداد خاص است.

**روند تغییرات در احساسات**:

در این نمودار، بسیاری از کاربران احساسات غالب منفی را ابراز کرده‌اند، زیرا نوار آبی‌رنگ نسبت به نوار خاکی‌رنگ در بسیاری از موارد بیشتر است.

همچنین برخی کاربران نیز احساسات مثبت یا خنثی خود را ابراز کرده‌اند که در این نمودار مشخص است.

**استفاده از نمودار ترکیبی**:

نمودار ترکیبی به‌ویژه برای نمایش اطلاعات پیچیده‌ای که شامل چندین متغیر است، بسیار مفید است. در اینجا، این نمودار به‌طور مؤثر تعداد کاربران و احساسات غالب آن‌ها را در یک زمان نمایش می‌دهد.

انتخاب این نوع نمودار مناسب بوده است زیرا دو مجموعه داده مختلف (تعداد کاربران و احساسات غالب) نیاز به نمایش همزمان دارند.

**نمودار هیستوگرام (Bar Chart)**:

برای هر نام کاربری، از یک هیستوگرام استفاده شده است که تعداد کاربران را نشان می‌دهد.

این نمودار به‌طور مؤثر توزیع کاربران را در طول نام‌های کاربری مختلف نشان می‌دهد.

**روند اوج احساسات**:

در بسیاری از نام‌های کاربری، شاهد یک اوج بزرگ در تعداد احساسات در ماه‌های خاص هستیم. این اوج‌ها ممکن است نشان‌دهنده دوره‌های خاص یا موضوعاتی باشند که کاربران نسبت به آن‌ها واکنش شدیدتری نشان داده‌اند.

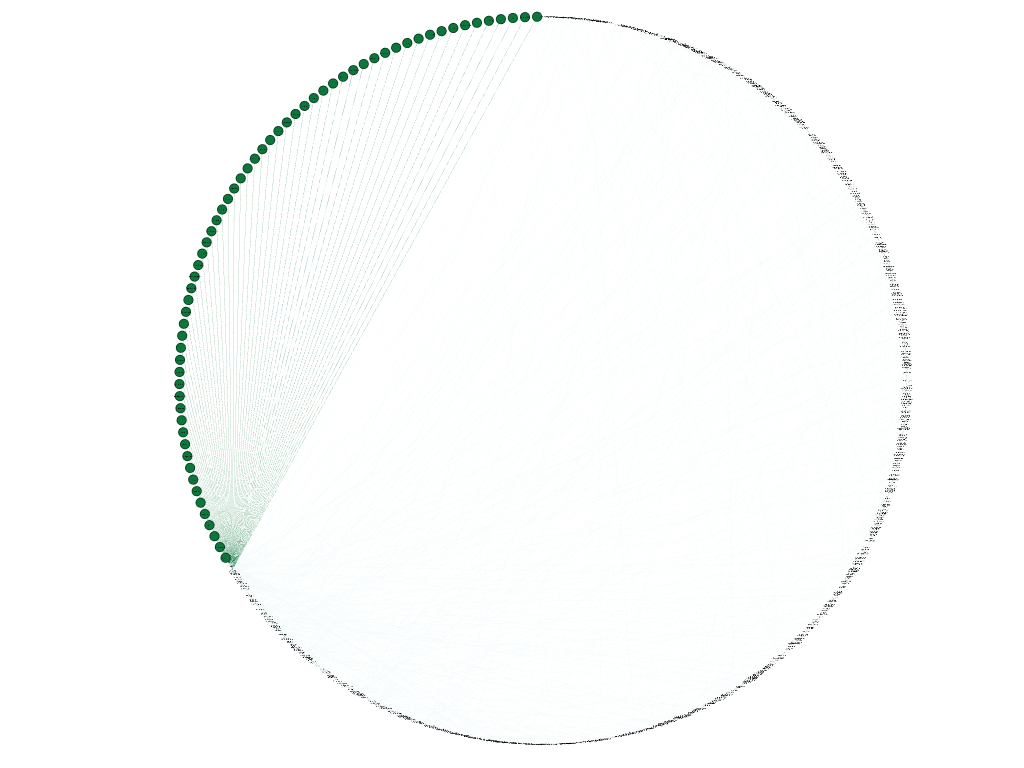
**نتیجه‌گیری:**

**نام کاربری "HashemStar"** به‌عنوان کاربر با بیشترین تعداد احساسات در نمودار دیده می‌شود. این کاربر ممکن است به‌طور خاص در واکنش به یک رویداد خاص یا موضوع بسیار مورد توجه قرار گرفته باشد.

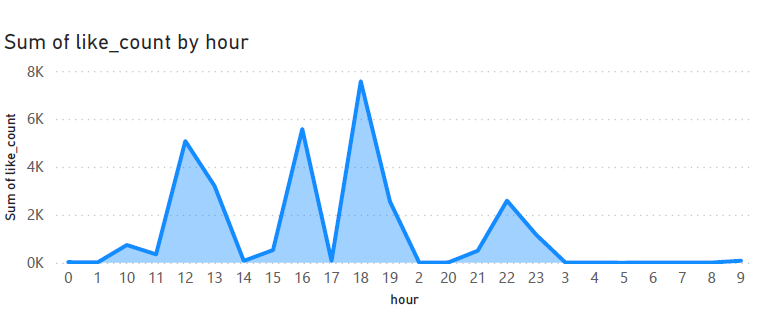
**احساسات غالب** نشان می‌دهد که بسیاری از کاربران واکنش‌های منفی را در مورد این موضوع خاص داشته‌اند.

نمودار ترکیبی به‌طور مؤثر نشان‌دهنده توزیع احساسات و تعداد کاربران در رابطه با هر نام کاربری است و به تحلیل‌گران این امکان را می‌دهد که روندهای خاص را در احساسات کاربران و تعداد آنان شناسایی کنند.

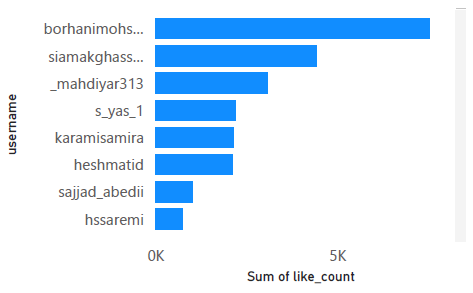
این نوع نمودار در تحقیقات اجتماعی و تحلیل‌های بازار می‌تواند بسیار مفید باشد، زیرا به‌طور همزمان توزیع احساسات و میزان تعاملات کاربران را بررسی می‌کند.



در این نمودار که در برنامه گفی رسم شده است، به ترتیب تمام هاب ها رسم شده اند که نود ها نشان دهنده هر یوزر و یال ها بر اساس پاسخ کاربران به هم رسم شده است.



این نمودار نشان میدهد به طور میانگین چه ساعتی پست های مربوط به این حوزه لایک بیشتری دریافت کرده اند.



در این نمودار کاربران با بیشترین تعداد لایک نشان داده شده اند.

**مسائل قانونی و حقوقی:**

این کد یک فرآیند کامل تحلیل متن را از ابتدا تا انتها برای دسته‌بندی نظرات کاربران در مورد آیفون با استفاده از الگوریتم KMeans و تحلیل خوشه‌بندی انجام می‌دهد. هدف اصلی این کد گروه‌بندی نظرات کاربران بر اساس شباهت‌های متنی است. در ادامه، تحلیل مراحل انجام‌شده آورده شده است:

**بارگذاری داده‌ها و پیش‌پردازش داده‌ها:**

ابتدا داده‌ها از فایل CSV به‌نام cleaned\_iphone.csv بارگذاری می‌شود.

تمام ردیف‌های دارای مقادیر خالی (NaN) حذف و فهرست بازسازی می‌شود.

سپس، فقط ردیف‌هایی که در ستون description کلمه "آیفون" را دارند، فیلتر می‌شوند.

**پیش‌پردازش متن:**

برای هر متن موجود در ستون description، از تابع preprocess\_text استفاده می‌شود که:

تمامی علائم نگارشی حذف می‌شوند.

کلمات متوقف (stopwords) که در فهرست از پیش تعریف‌شده آمده‌اند، از متن حذف می‌شوند.

سپس کلمات باقی‌مانده به هم پیوسته و به‌عنوان متن تصفیه‌شده ذخیره می‌شوند.

**. نمایش ویژگی‌های متنی:**

از TfidfVectorizer برای استخراج ویژگی‌های متنی (TF-IDF) استفاده می‌شود.

این ویژگی‌ها در فضای n-gram با محدوده n از 1 تا 3 استخراج می‌شوند و فقط کلماتی که در بیش از 2 درصد اسناد ظاهر می‌شوند (با حداکثر فراوانی 80%) به‌عنوان ویژگی‌های معتبر در نظر گرفته می‌شوند.

**خوشه‌بندی با KMeans:**

ابتدا مدل KMeans با تعداد خوشه‌های 2 آزمایش می‌شود.

معیار ارزیابی مدل، نمره silhouette است که برای ارزیابی کیفیت خوشه‌بندی استفاده می‌شود.

پس از آزمایش، بهترین تعداد خوشه‌ها با بالاترین نمره silhouette انتخاب می‌شود. در اینجا، مدل با 2 خوشه مورد ارزیابی قرار می‌گیرد.

**استخراج کلمات کلیدی برای هر خوشه:**

پس از اعمال خوشه‌بندی، کلمات کلیدی هر خوشه استخراج می‌شوند. این کلمات از میان ویژگی‌های TF-IDF استخراج‌شده، برای هر خوشه با توجه به نزدیکی مرکز خوشه‌ها به ویژگی‌ها انتخاب می‌شوند.

این کلمات به‌عنوان ویژگی‌های اصلی هر خوشه معرفی می‌شوند.

**تخصیص نام خوشه‌ها:**

خوشه‌ها با توجه به کلمات کلیدی شایع در هر خوشه نام‌گذاری می‌شوند:

خوشه اول با نام "Commercial Import & Legal Procedures" (واردات تجاری و فرآیندهای قانونی) و خوشه دوم با نام "iPhone Registration & Regulations" (ثبت آیفون و مقررات) شناخته می‌شوند.

**کاهش ابعاد با PCA:**

برای تجسم داده‌ها و نمایش خوشه‌ها، از روش کاهش ابعاد PCA استفاده می‌شود.

فضای ویژگی‌های استخراج‌شده به دو بعد کاهش یافته و داده‌ها در این دو بعد به‌صورت گرافیکی نمایش داده می‌شوند.

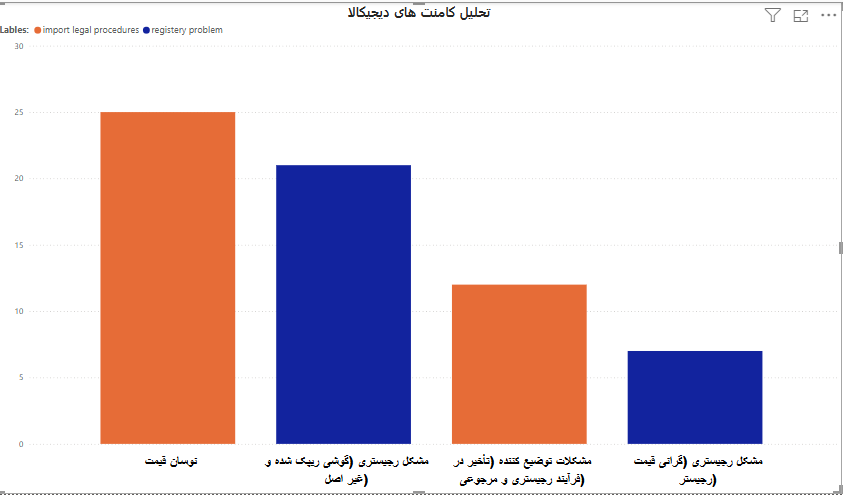
**نمودار PCA:**

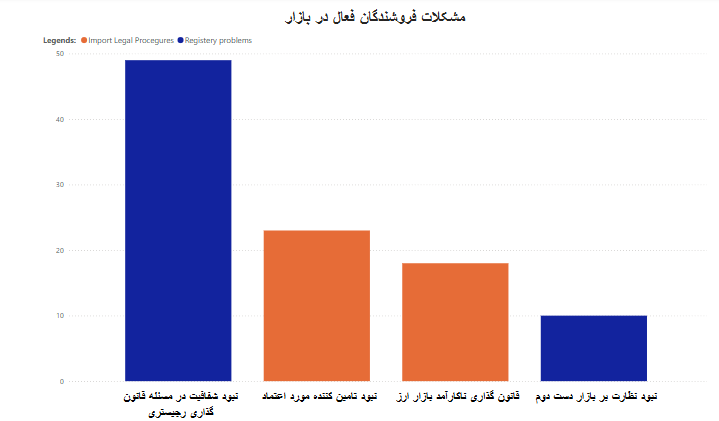
در پایان، نموداری با استفاده از داده‌های PCA برای نمایش خوشه‌بندی نهایی تولید می‌شود. در این نمودار، هر نقطه نمایانگر یک نظر است که به یکی از خوشه‌ها تعلق دارد.

از رنگ‌های مختلف برای تمایز هر خوشه استفاده می‌شود.

**نتیجه‌گیری:**

این کد به تحلیل متنی نظرات کاربران در مورد آیفون پرداخته و آن‌ها را به دو خوشه اصلی تقسیم کرده است. خوشه اول بیشتر مربوط به واردات تجاری و فرآیندهای قانونی است، در حالی که خوشه دوم به مسائل مربوط به ثبت و مقررات آیفون می‌پردازد. این نوع تحلیل می‌تواند برای تحقیقات اجتماعی یا نظرسنجی‌ها در صنعت فناوری و موبایل بسیار مفید باشد، زیرا می‌تواند موضوعات داغ و نقاط درد مشترک را در میان کاربران شناسایی کند



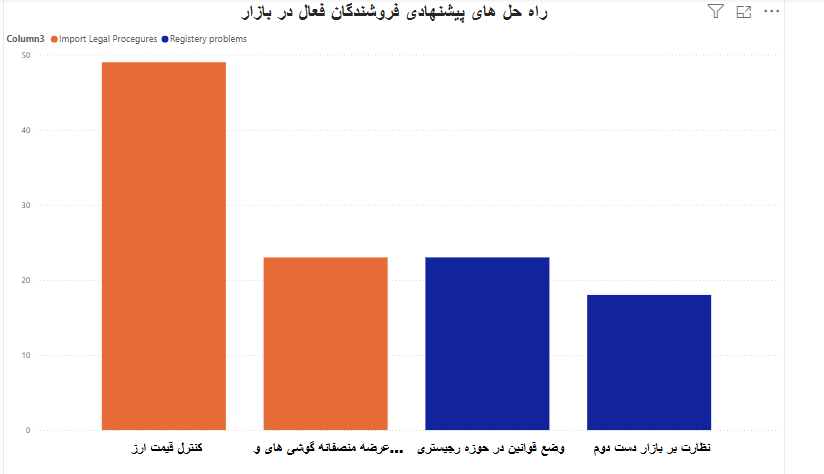


این نمودارها و تجزیه و تحلیل‌هایی که به اشتراک گذاشته‌اید نشان‌دهنده روندهای مختلف مرتبط با مشکلات فروشندگان در بازار، نظرها، و روند خوشه‌بندی متن‌ها برای تحلیل‌های مربوط به آیفون است. در ادامه، توضیحاتی درباره این نمودارها و آنچه که از آنها می‌توان استخراج کرد، ارائه می‌شود:

**نمودارهای اول و دوم: مشکلات فروشندگان در بازار**

این نمودارها نشان‌دهنده مشکلات مختلف فروشندگان در بازار و چگونگی تقسیم‌بندی آنها به دو دسته اصلی هستند:

* **مشکلات مرتبط با واردات قانونی**
* **مشکلات ثبت‌نام و مشکلات مربوط به رجیستری** .

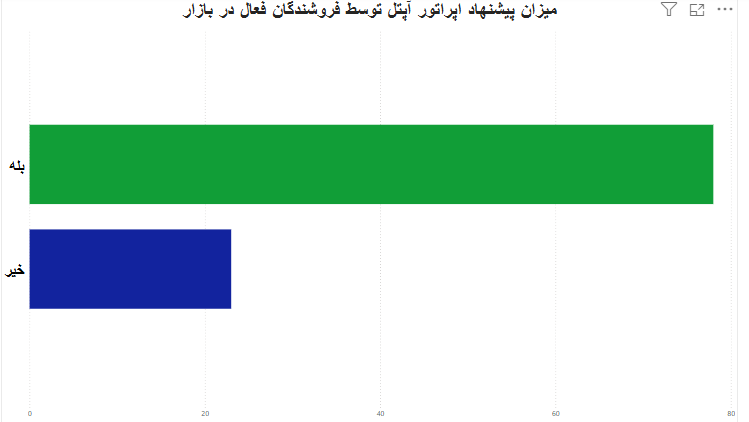
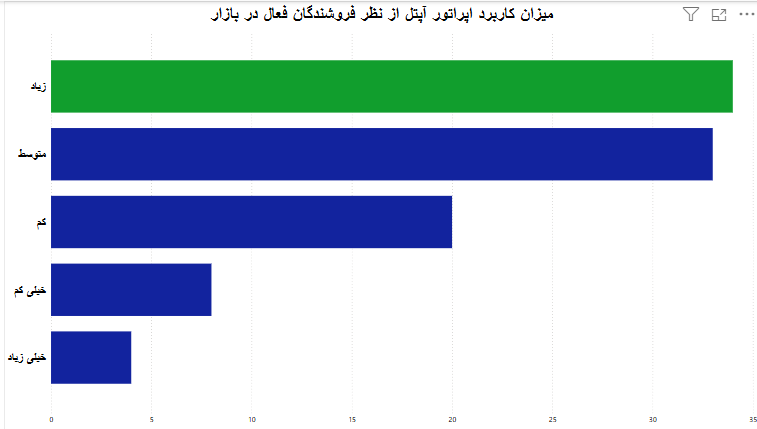


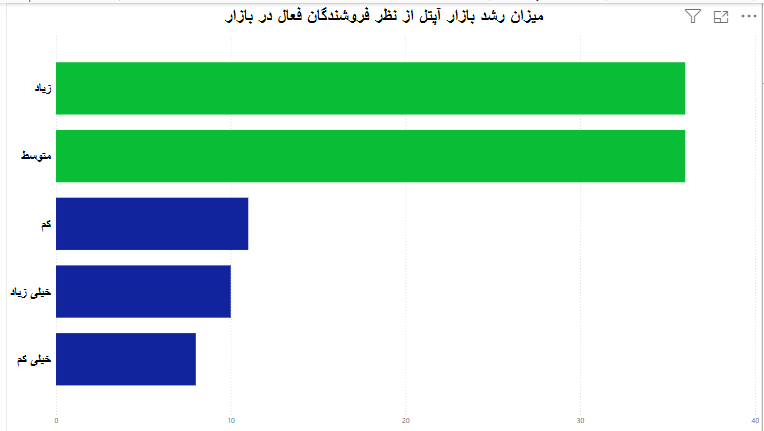
**نمودار سوم: پیشنهادات برای رفع مشکلات**

این نمودارها به بررسی پیشنهادات فروشندگان برای حل مشکلات موجود در بازار می‌پردازد:

* **پیشنهادات برای حل مشکلات واردات قانونی** در رنگ آبی
* **پیشنهادات برای حل مشکلات ثبت‌نام** در رنگ نارنجی

از این نمودار می‌توان نتیجه گرفت که در بیشتر ماه‌ها، پیشنهادات مربوط به حل مشکلات واردات قانونی بیشتر از مشکلات ثبت‌نام است.

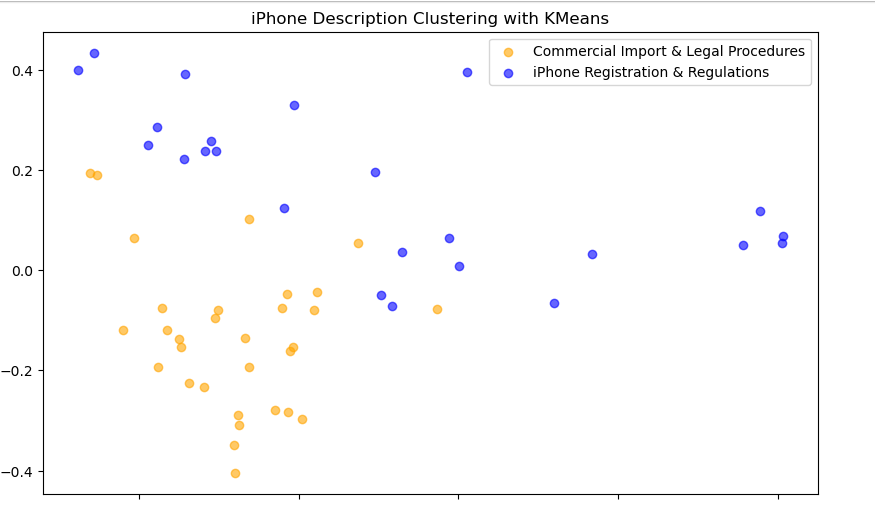
 



**نمودارهای چهارم و پنجم و ششم: بررسی وضعیت اپراتور آپتل**

این نمودارها به بررسی پیشنهادات فروشندگان و همچنین میزان کاربرد و رشد بازار اپراتور آپتل میپردازد:

* **با توجه به نمودار چهارم مشخص است که از نظر فروشندگان فعال در بازار این اپراتور به کاربران پیشنهاد میشود**
* **با توجه به نمودارهای پنجم و ششم مشخص است که میزان کاربرد و رشد بازار این اپراتور با رشد قابل ملاحظه ای روبرو بوده است**



**4. نمودارهای خوشه‌بندی (Clustering)**

در این نمودارها، خوشه‌بندی‌های مختلفی که با استفاده از الگوریتم KMeans روی متن‌های توصیفی مربوط به آیفون انجام شده، نمایش داده می‌شود. در این نمودارها:

* **خوشه اول** (رنگ نارنجی) به مشکلات واردات قانونی و رویه‌های تجاری اشاره دارد.
* **خوشه دوم** (رنگ آبی) به مشکلات ثبت‌نام و مشکلات مربوط به رجیستری اشاره دارد.

با استفاده از تحلیل خوشه‌بندی، می‌توانیم به وضوح ببینیم که چه مشکلاتی بیشتر توسط فروشندگان مطرح شده‌اند و کدامیک از دسته‌ها بیشترین تمرکز را در نظرات فروشندگان دارند. از این رو، خوشه‌بندی به ما کمک می‌کند تا الگوهای تکراری و مشکلات متداول را شناسایی کنیم.

**5. تجزیه و تحلیل کلمات پرکاربرد**

در این بخش، از نمودارهای «کلمات پرکاربرد» استفاده شده که کلمات کلیدی مرتبط با مشکلات واردات و رجیستری را در مقیاس‌های مختلف دسته‌بندی می‌کنند. این نمودارها به کمک TF-IDF و KMeans به تحلیل دقیق‌تری از ویژگی‌های متنی و موضوعات داغ بازار آیفون می‌پردازند.

### جمع‌بندی:

در کل، این نمودارها به کمک تجزیه و تحلیل داده‌ها و خوشه‌بندی متن‌ها، چشم‌انداز خوبی از وضعیت فروشندگان، مشکلات و پیشنهادات آنها برای بازار آیفون ارائه می‌دهند. نمودارهای مختلف نشان‌دهنده تعامل بین تغییرات در بازار، وضعیت ثبت‌نام و مسائل قانونی مربوط به واردات است. همچنین، استفاده از خوشه‌بندی و مدل‌های یادگیری ماشینی کمک می‌کند تا مشکلات و پیشنهادات به طور دقیق‌تری دسته‌بندی شده و تجزیه و تحلیل شوند.