# Task 1

## Definition of GA4 Events.

In Google Analytics 4 (GA4) events is a concept that allows tracking of all the interactions or activities done by users, in a particular website or mobile application (Chandan, 2023). There are 4 types of GA4 events namely automatically collected events, enhanced events, recommended events and custom events. Some examples of GA4 events are given below.

1. Clicking on a button.
2. Submitting a form.
3. Scrolling through a page.
4. Viewing a page

## Function And Role of Events.

Events can be considered as the foundation of GA4 (Duncan, 2023). There are several roles and functions of events and they are described in the table below.

|  |  |
| --- | --- |
| **Function/Role** | **Description** |
| Track user interactions, measuring engagement of users | GA4 Events track user interactions and engagement, providing businesses with insights into the most interacted elements of their websites. By measuring the frequency of occurrences, such as clicking a button, duration a user stays in website etc. organizations can understand user behavior in mobile applications or websites (Duncan, 2023).  . |
| Help create personalized experiences. | As mentioned above, GA4 provides insights into user interactions, which can help understand trends and patterns followed by users. This information can be used to understand customer segments' preferences, enabling marketing campaigns and personalized content, such as targeted ads (McGuirk, 2023). |
| Providing support to increase conversions. | GA4 tracks conversion-related information, helping businesses understand marketing campaign success, user flows, and areas for improvement. For example, if users abandon a form half-way, it could indicate a blockage in submitting the form (Salon, 2023). |

Table 1:Function/Roles of GA4 Events

## Event Properties and How They Are Tied to User Interactions.

In GA4 event properties provides more additional information that helps get a better understanding on an event. Some of the event properties and how they are tied with user interactions are given below.

|  |  |
| --- | --- |
| Property name | How it is tied with user interactions |
| Event Name | This is tied with user interactions since it provides an understanding on exact type of event/user interaction that happened. For example, the event name can be click\_button. |
| Event Parameters | This property provides additional information about user events that happened and provides deeper insights on to user interactions (Sharma, 2023) for example parameters such as page\_title, page\_location etc helps understand the name of page, location user interacted with. |

Table 2:Event properties of GA4 and how they are tied with user interaction

## Events in relation to Microsite as A Case-Study.

As shown in figure 1 Google analytics has tracked user interactions done in the site such as page view, scrolling. In addition, as shown in figure 2 additional information on events such as the page views of each page are counted. Other than that event parameters such as page location (refer to figure 3) are tracked as well. Through tracking of these, GA4 provides insights on the microsite such as pages that are most visited, user paths etc. giving a better understanding on user behavior, engagement etc. To get more insight custom events could also be added. For example, there is a button “Subscribe” in my home page and if this was tracked it would show the users who were interested to subscribe.

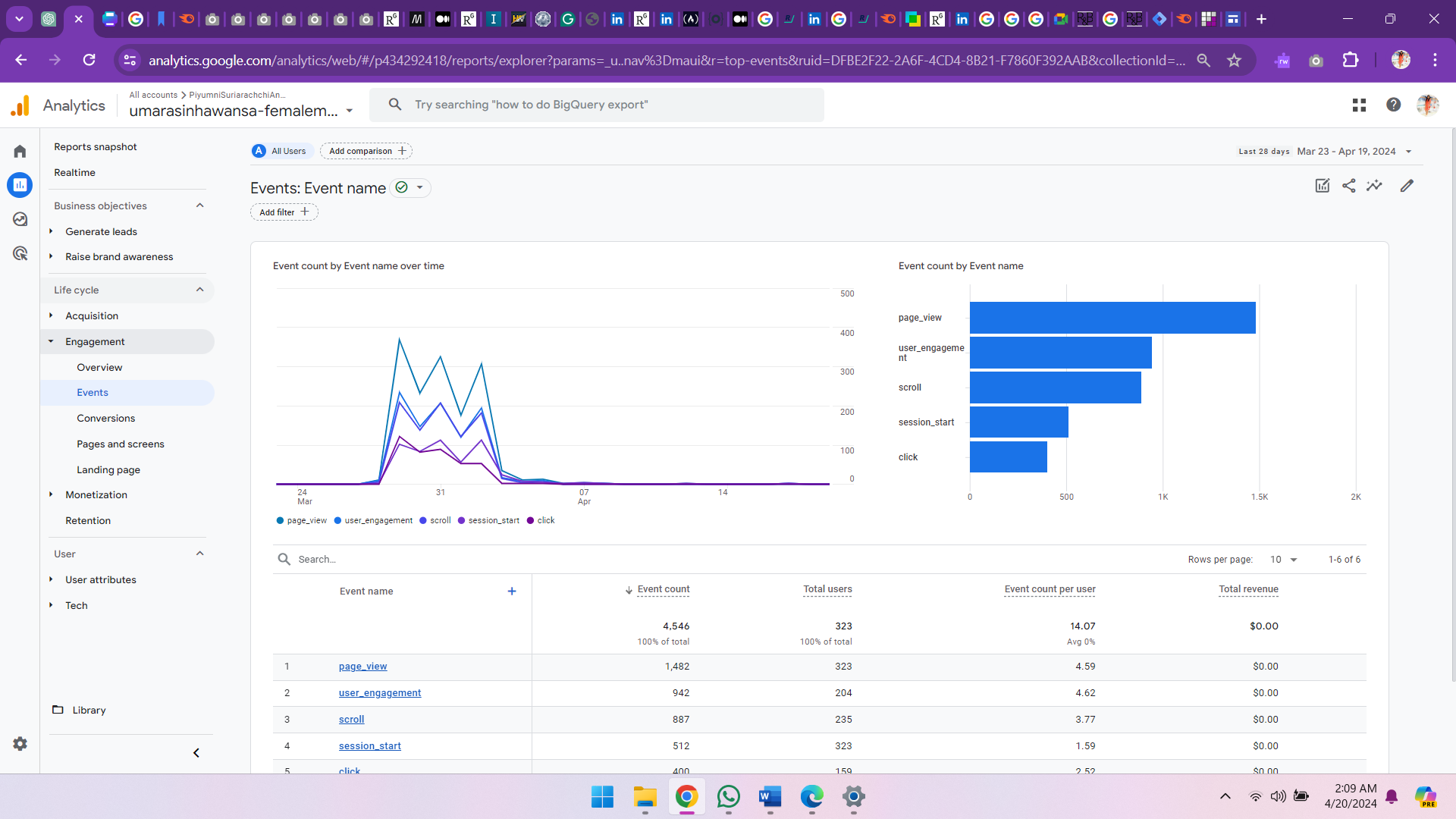


Figure 1:Events types

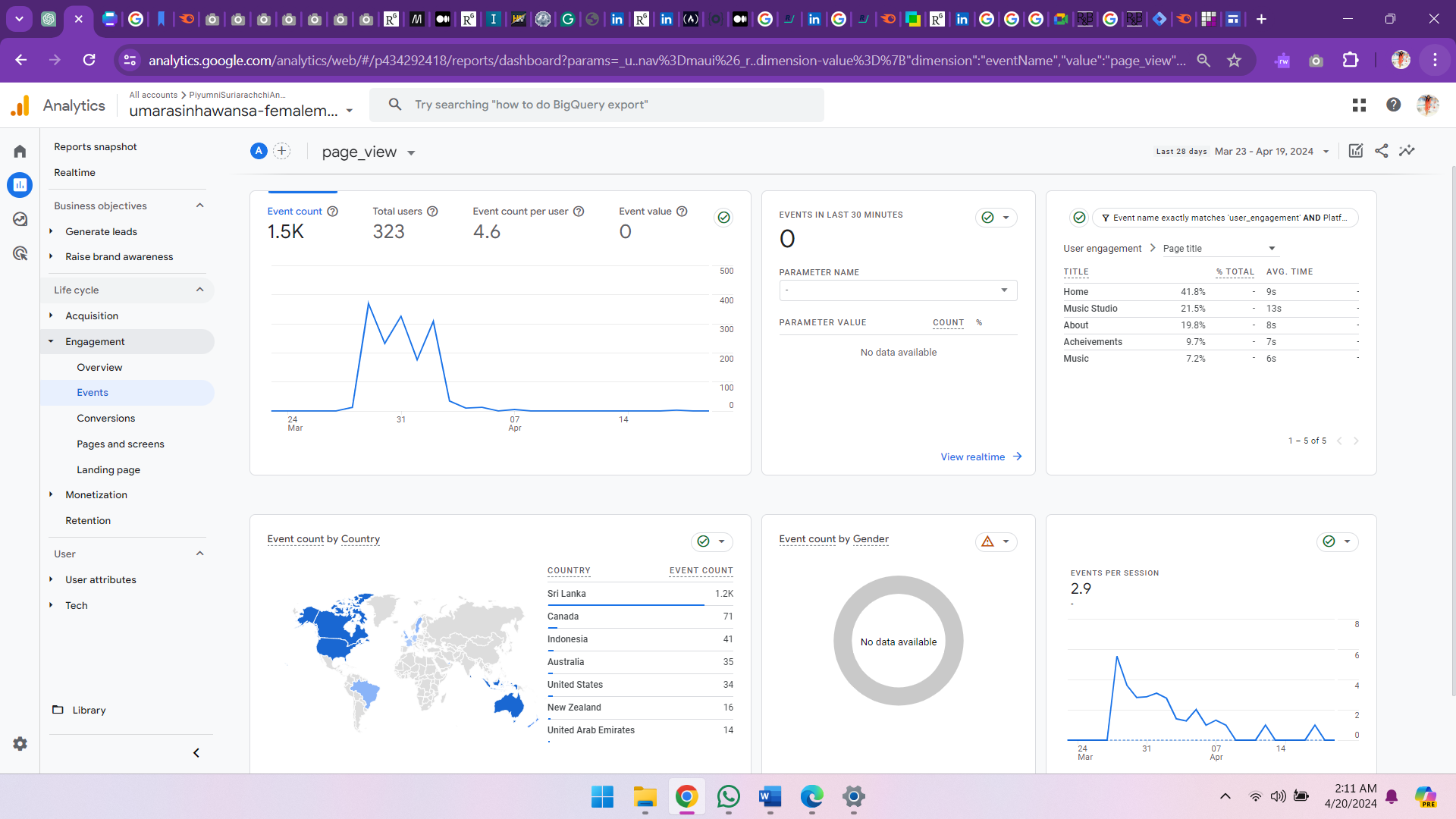


Figure 2: Page View event detailed report

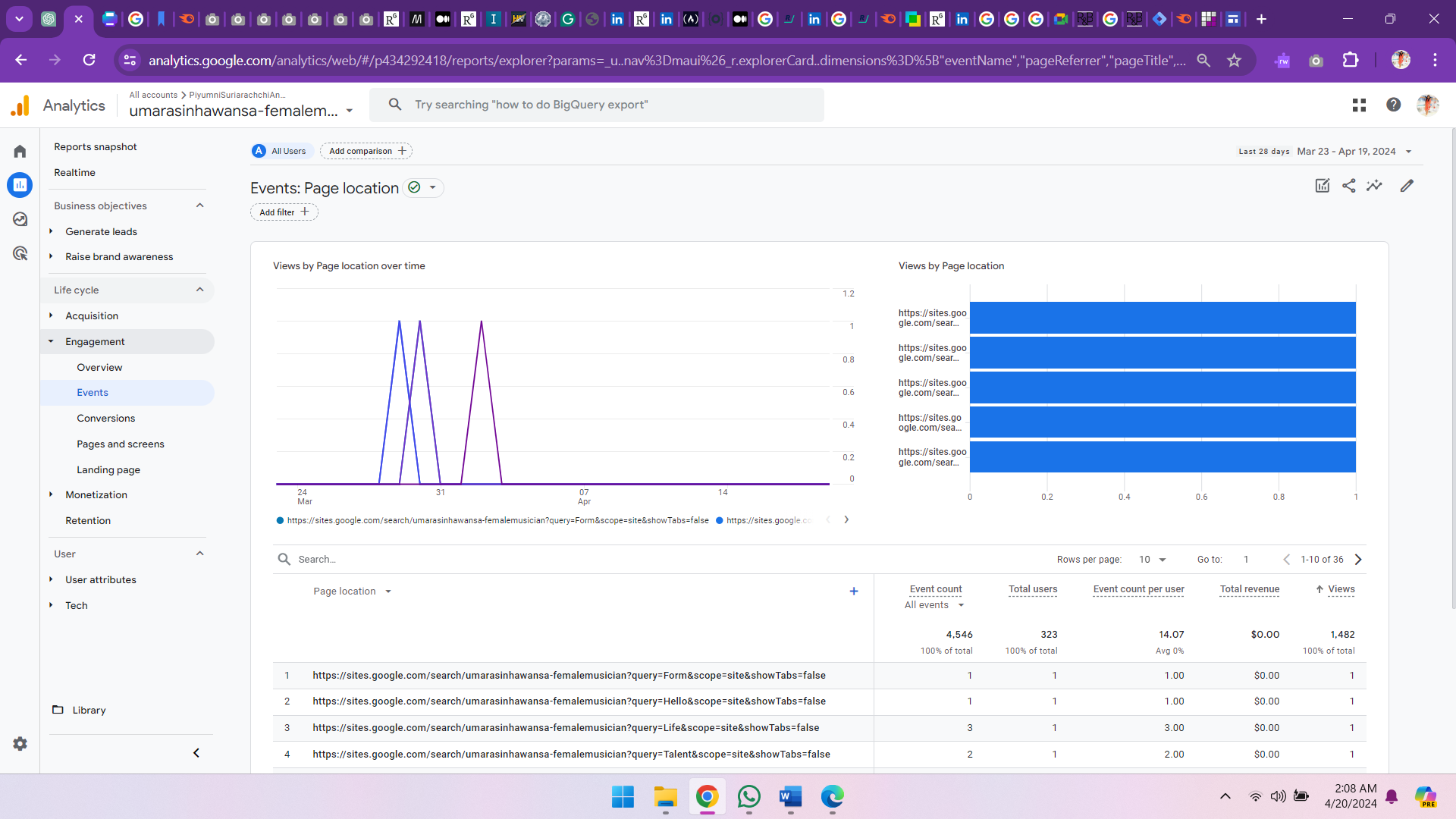
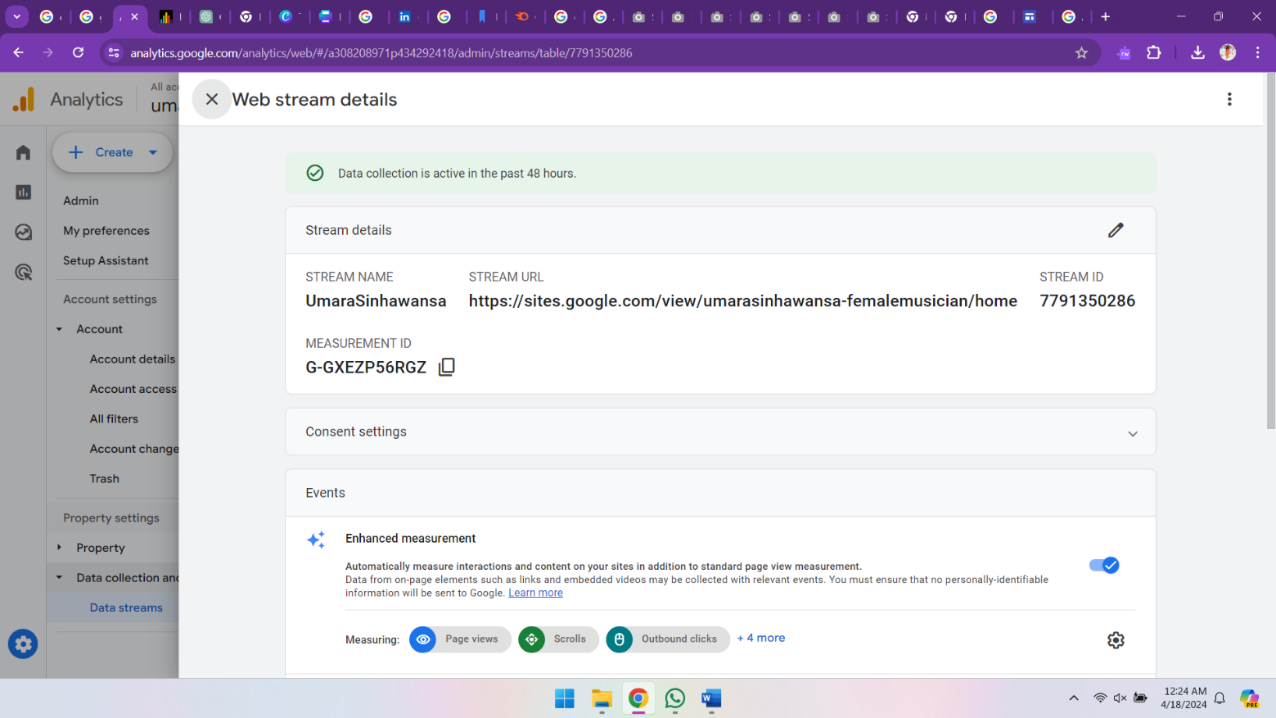


Figure 3:Page\_location parameter

# Task 2

## Steps Followed to Create Google Analytics Account and Connect It to the Microsite.

1. Selected “Get Started button” in google marketing platform to start account creation process.
2. Then selected on “start measuring” button and completed the 5 sections mentioned below.
   1. account creation – Section to add account name and confirm account data sharing settings.
   2. property creation – Section to add property name, select time zone and currency.
   3. Business details – Section to select industry category, size of organization
   4. Business objectives – Section to select objectives of the business.
   5. Data collection - Section to select type of platform to collect data. In this scenario “Web” platform was selected.
3. After completing above, in set up web stream page, I added the website URL, stream name, kept the enhanced analytics option as on and clicked on “Create stream” button to create the web stream. Once web stream was created, web stream details were displayed as shown below in figure 1.



1

Figure 4:Web stream details including measurement Id

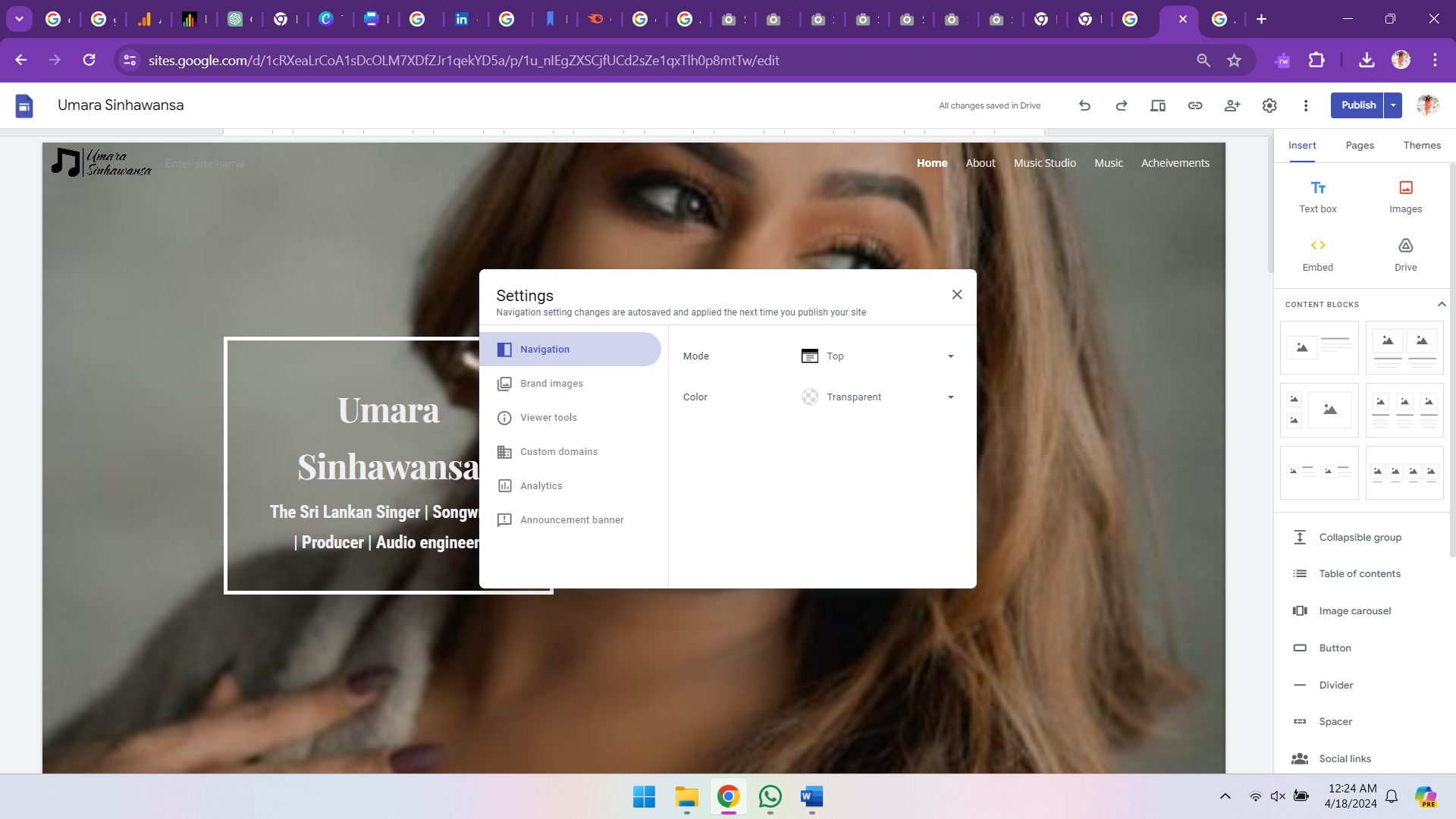
1. I copied the measurement Id (Refer to section 1 in figure 4) from web stream detail page.
2. Opened my microsite in google sites and clicked on settings icon (outlined in red in figure 5).

Figure 5: Settings tab in Google sites

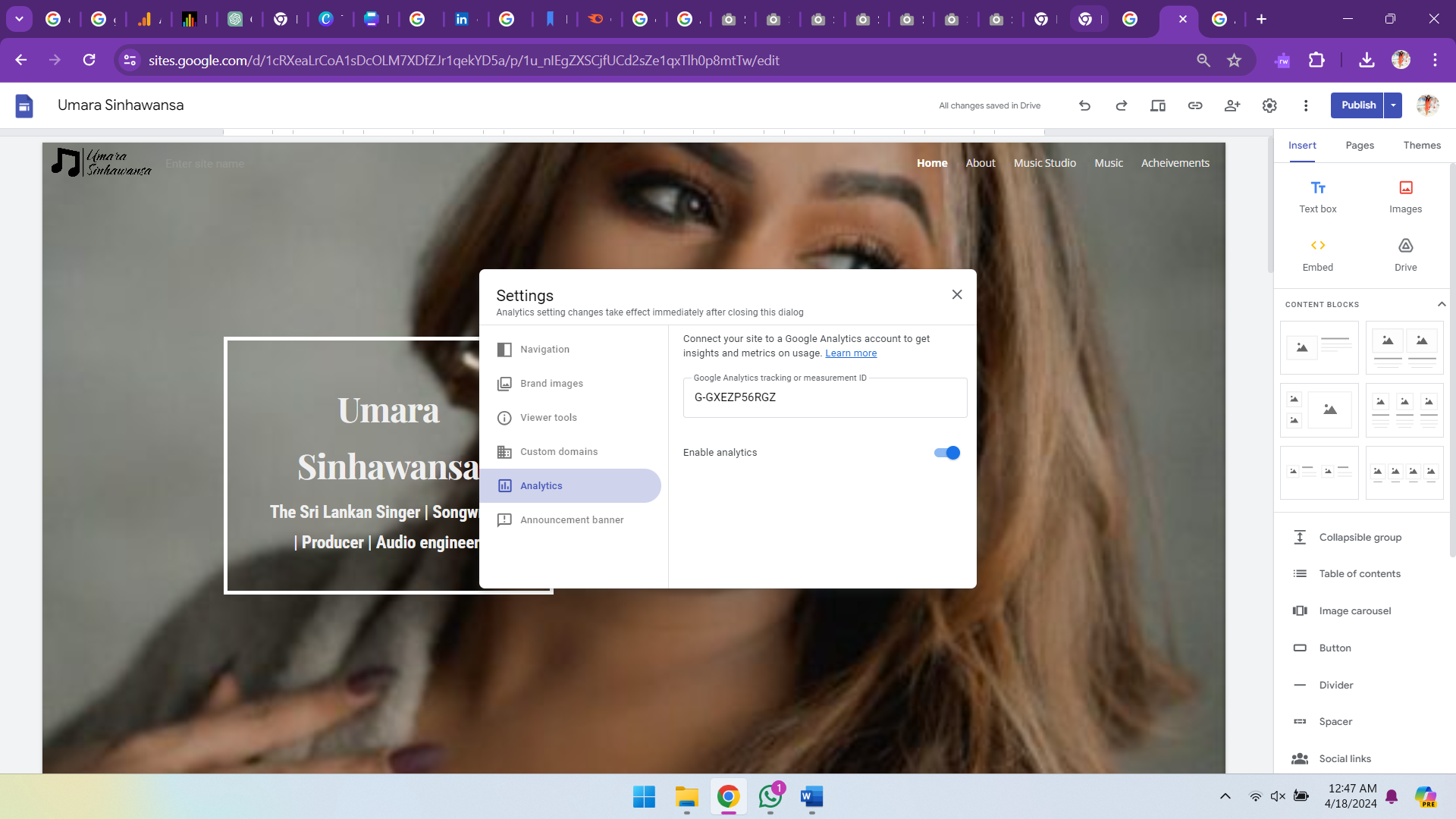
1.  Then from settings, I selected analytics option and added the measurement Id, enabled analytics from toggle button to connect the microsite to google analytics as shown in figure 6.

Figure 6: Analytics settings in Google sites:

# Task 3

## Definition of HTTP Cookies.

HTTP cookie is certain information sent by a server to a user's web browser when they access a website, and is available mostly in text form (Pantelic, Jovic and Krstovic, 2022). They enable web browsers to manage user session, recognize users. For example, cookies will help web browser in recognizing users by identify whether a request came from the same browser or not. In addition, cookies also help track user behavior and personalization (Cahn et al., 2016).

## Problems Overcome by Cookies

|  |  |
| --- | --- |
| Problem | How Cookies overcome the problem |
| Authentication and session management issues. | The stateless nature of HTTP protocol prevents saving data like password, items added to cart but cookies overcome this by storing a unique id referencing to a session in browser files allowing servers to recognize users, automatically authenticate and save information like shopping cart details (Khu-smith and MitchellI, 2001). |
| Providing personalized experiences. | * Through placement of text files in the browser files in client end, data is collected when user queries webpages. This allows the web browser collect information on user behavior/preferences and create a more personalized experience for the users based on information collected (Nkomalago, 2023). |

Table 3:Problems Overcome by Cookies

## Limitations of Cookies.

|  |  |
| --- | --- |
| **Limitation** | **Description** |
| Life span | Cookies has a set life span and expires after that. Hence cookies are not a viable option to store data for longer periods (Verma, 2024). |
| Security | Cookies are vulnerable to security risks, including cross-site scripting (XSS) attack, cross-site request forgery (CSRF) attack, and these attacks can allow attackers to control the affected party's cookies, execute malicious code, and manipulate the user's browsing history etc. leading to a compromise of security in cookies (Dodt, 2020). |
| Size and storage | Cookies has limited limit around 4KB. This makes it harder to save larger amount of data in a cookie (Verma, 2024). |

Table 4:Limitations of cookies

## Stateless Web and HTTP Cookies

When a user visits a website, a HTTP request is sent to the server. Server considers each of these requests as unique. Hence by default the server does not identify unique users who visits a website (client), since requests that are sent to the server does not have any reference to the past requests. This is called the stateless web or stateless nature of HTTP protocol. Cookies are used to overcome this problem through building state between the server and client (Patil, 2022). HTTP cookies overcome stateless nature by collecting data in name value pairs into small text files that are saved in clients browser files (Demir et al., 2024). And when sending requests to server, cookies are also included enabling the server retrieve data from cookies and recognize the users across multiple requests (Singh, 2020).

## Diagram Showing the Key Stages in The HTTP Process and Data Flows Between Client and Server.

# Task 4

To complete task 4 traffic was generated for the microsite over 5 consecutive days from 29th March 2024 to April 2nd 2024.

## Users by Browser Over Time.

Time period selection

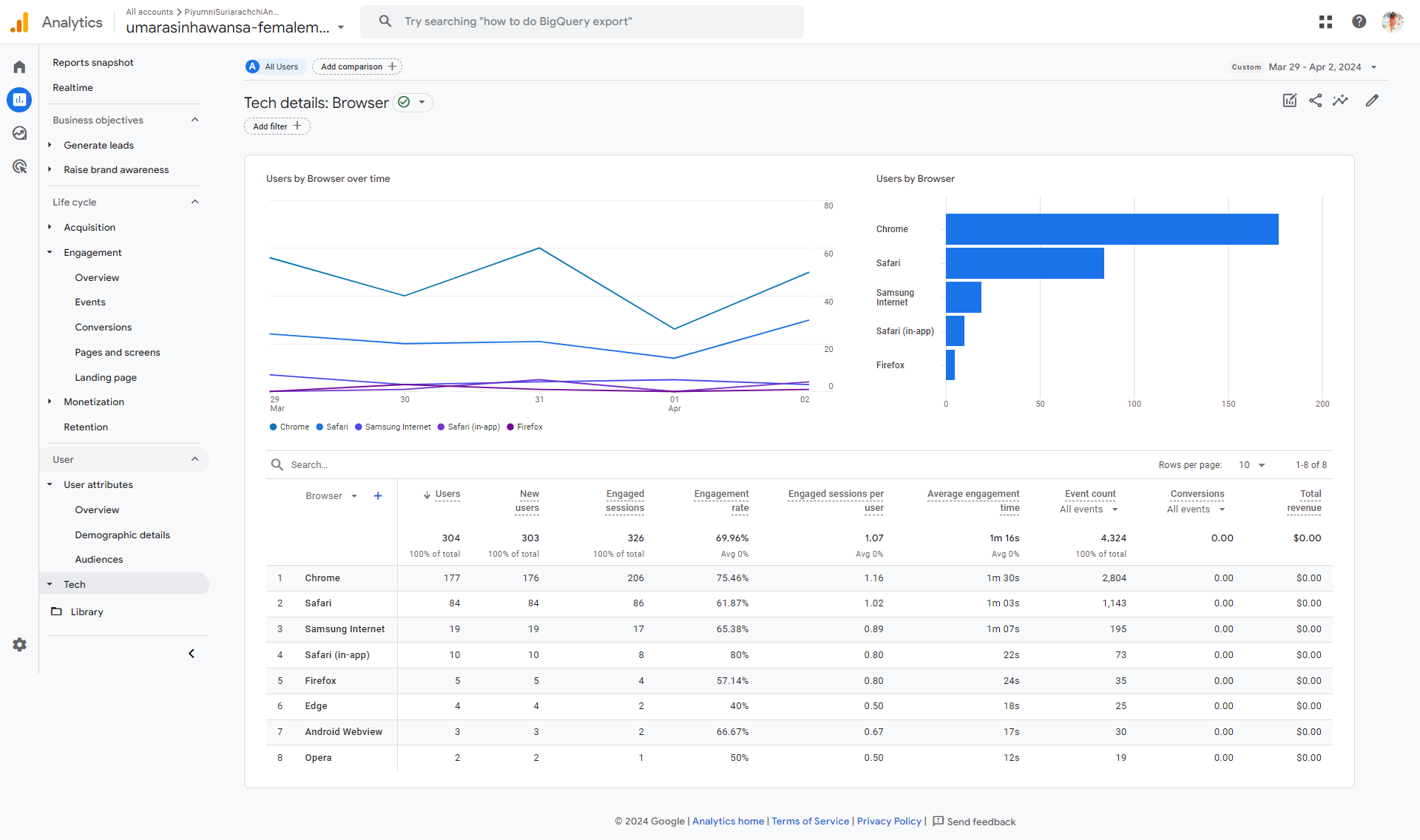


Figure 7: Google Analytics Report for Users By Browser Over Time

The above report shows that Google Chrome is the most used browser with 177 users, accounting for over half of total users. This is followed by Safari and Samsung Internet browsers, with 84 and 19 users respectively. In addition, Chrome also has the highest average engagement time of 1m 30s followed by Safari with 1m 03s suggesting that users have engaged with the site well using Chrome hence optimizing website for web browsers Chrome and safari could be prioritized.

## Views by Page Title and Screen Class Over Time

Time period selection

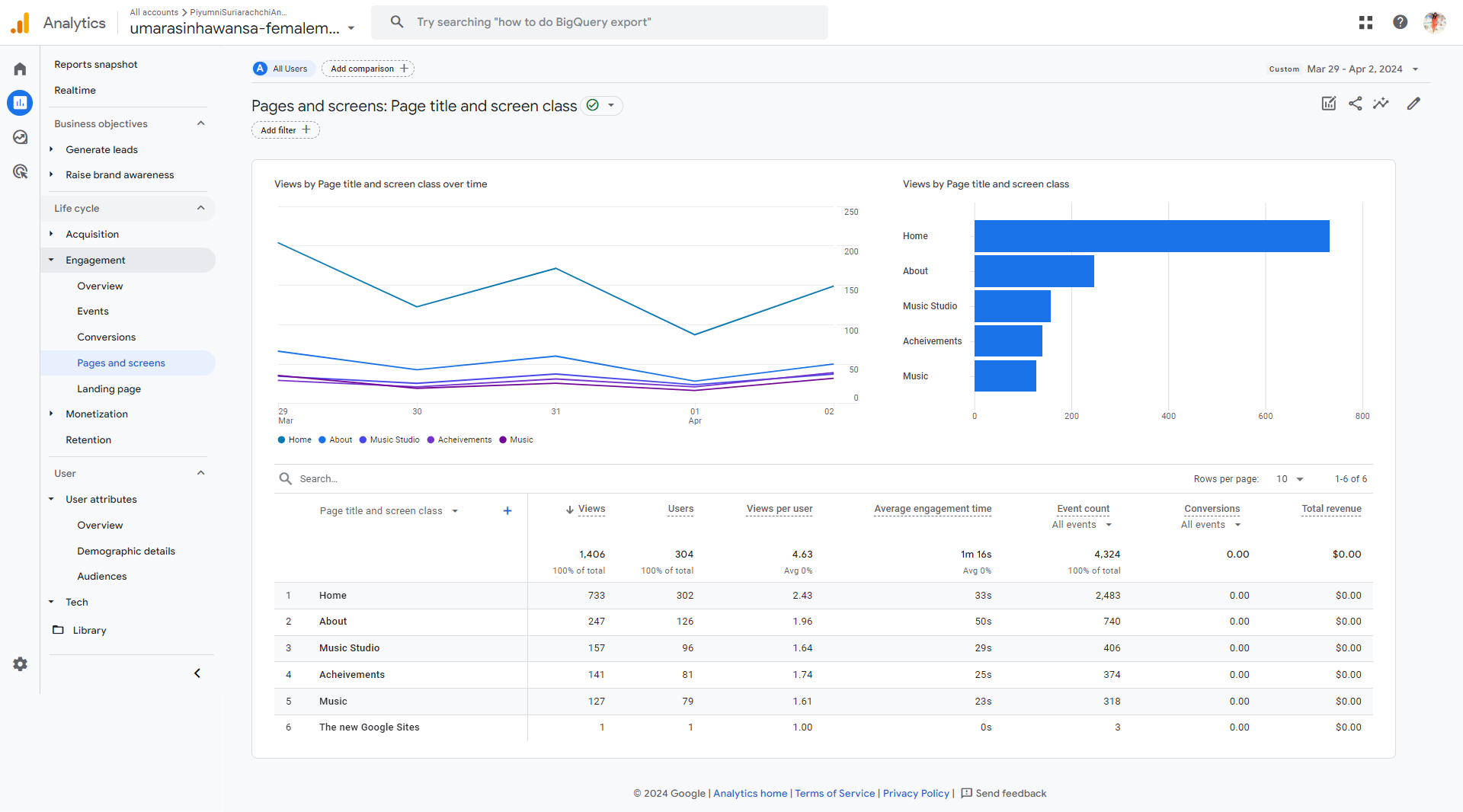


Figure 8:Google analytics Report for Views By Page Title and Screen Class Over Time

The above report shows that the most viewed page is home page with 733 views, followed by About page with 247 views. In addition to this “Home” page takes precedence in the event count as well. One of reason for having increased views and event count for home page could be because when sharing the website, it was home page link that was shared. However, the decrease in page views, events count after the home page is quite significant, hence in future certain improvement can be done in terms of user interfaces content etc., optimize website to increase traffic.

## Sessions by ‘Session Default Channel Grouping’

Time period selection

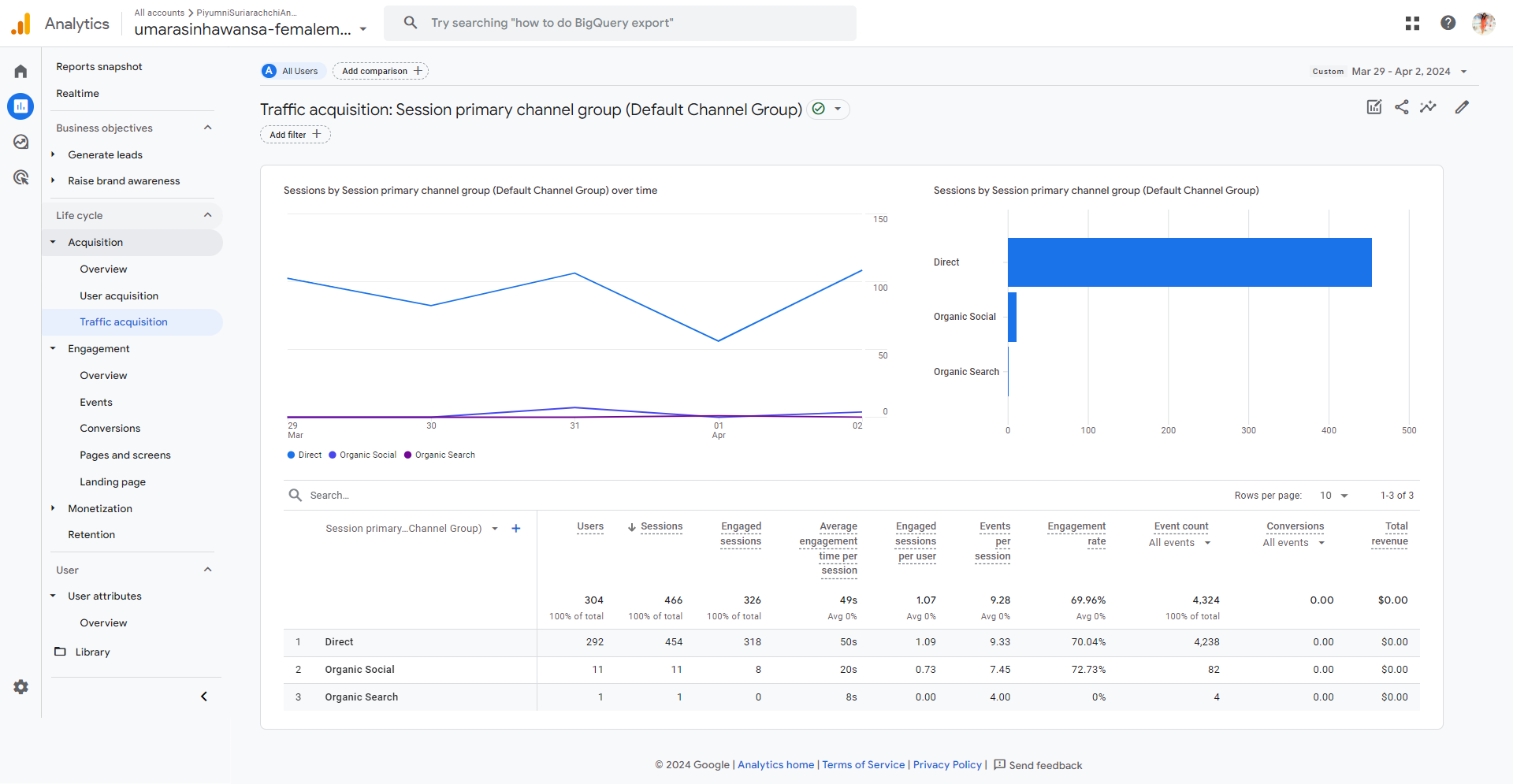
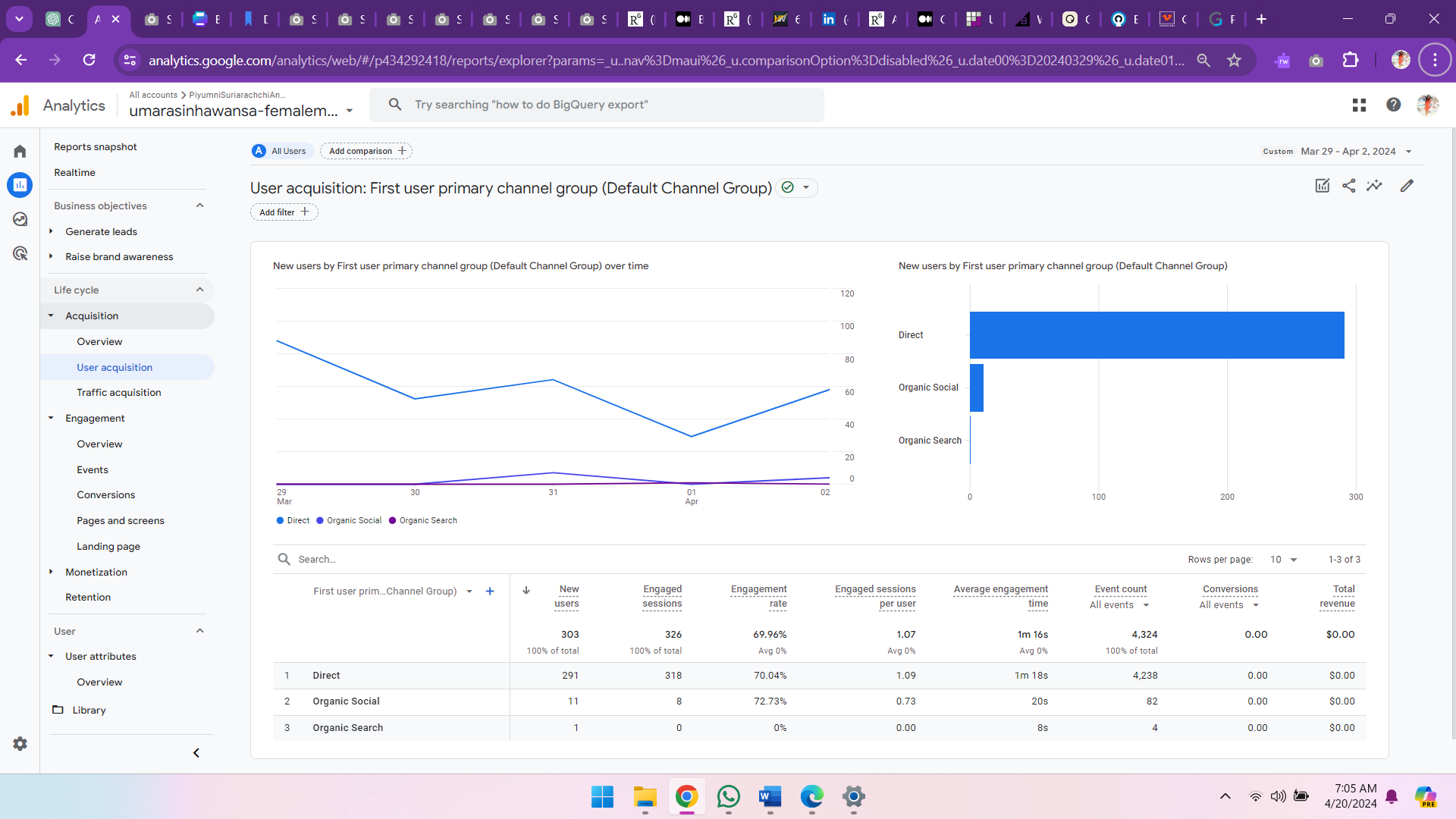


Figure 9:Google Analytics Report for Sessions By ‘Session Default Channel Grouping’

The above shows that direct channel was used by most user sessions that is 454 sessions accounting to over three-quarters of total sessions. This channel also has the highest average engagement time of 50s and highest events per session that is 9.33. Organic social and organic search channels didn’t have a significant number of sessions, suggesting that direct channel was preferred by most users and in future usage of good marketing strategies in organic social channels is also important in website promotion.

## New Users by ‘First User Default Channel Grouping’



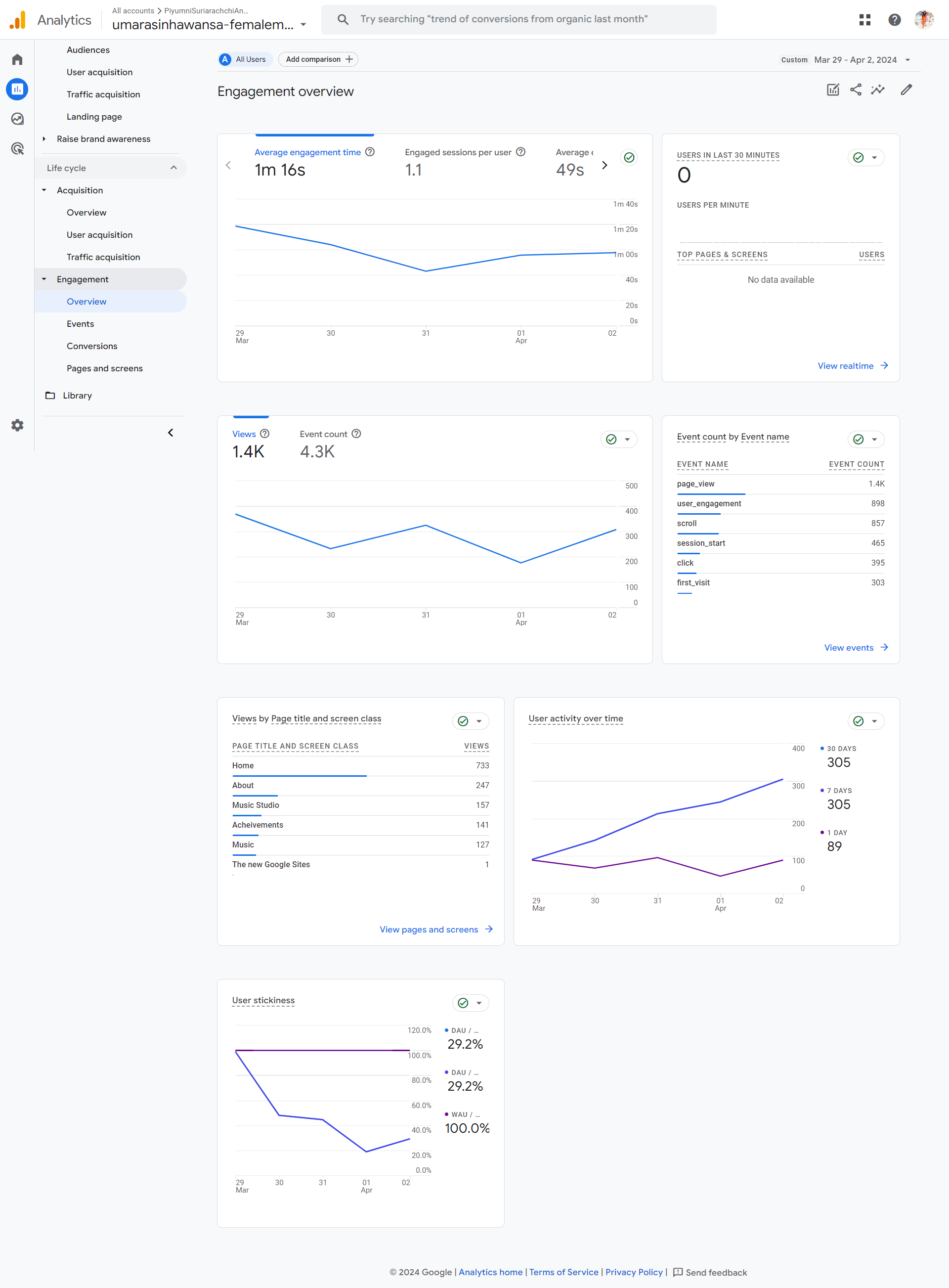
Time period selection

Figure 10: Google Analytics Report for New Users By ‘First User Default Channel Grouping’

The above report shows that direct channel is the most successful channel for user acquisition, with 291 new users gained, accounting for over 3 quarters of the total new user count suggesting its high success rate. Other channels like organic social and organic search weren’t used by a significant number of users. Hence in future direct channel can be considered important when considering user acquisition campaigns, however it is also important to ensure usage of proper strategies in marketing in other channels to gain users through them as well.

## User Engagement (Overview)

Time period selection



6

4

5

3

2

1

Figure 11: Google Analytics Report for User Engagement Overview

The above report provides an overview of user engagement over the period. There are several charts in this report. Here chart 1 shows an average user engagement time of 1m 16s and 1.1 sessions per user, with a decline in engagement in the third day but a gradual increase in the last two days. Charts 2 show 1.4K views and 4.3K event counts, with a decline in views and events over time while chart 3 shows that Page views is the most frequent event. Chart 4 shows page views by page title and screen class, and to get a better understanding on this refer to section 4.2. Chart 5 shows an upward trend in user activity with day 4 having the lowest activity. The upward trend suggests that strategies taken toward having good user activity was successful while Chart 6 compares active user engagement over a short period and a broader time period, showing a 29.2% DAU/WAU, suggesting moderate level of user retention or sticking. Considering all these findings it can be said that user engagement is good but continuous monitoring and adjustments are needed to sustain and enhance engagement.

## User Retention (Overview)

Time period selection

5

6

4

3

2

1

Figure 12:Google Analytics Report for User Retention Overview

5

The above report provides an overview of user retention over the period. There are several charts in this report. Here chart 1 shows that a upward trend in retention over the period with day 4 having the lowest returning users. In addition, the total of 81 returning users is 26.64% of total user count. Charts 2 shows user retention by cohort . Chart 4 shows page views by page title and screen class, and to get a better understanding on this refer to section 4.2. Chart 5 shows an upward trend in user activity with day 4 having the lowest activity. The upward trend suggests that strategies taken toward having good user activity was successful while Chart 6 compares active user engagement over a short period and a broader time period, showing a 29.2% DAU/WAU, suggesting moderate level of user retention or sticking. Considering all these findings it can be said that user engagement is good but continuous monitoring and adjustments are needed to sustain and enhance engagement.

## Users by City

Time period selection

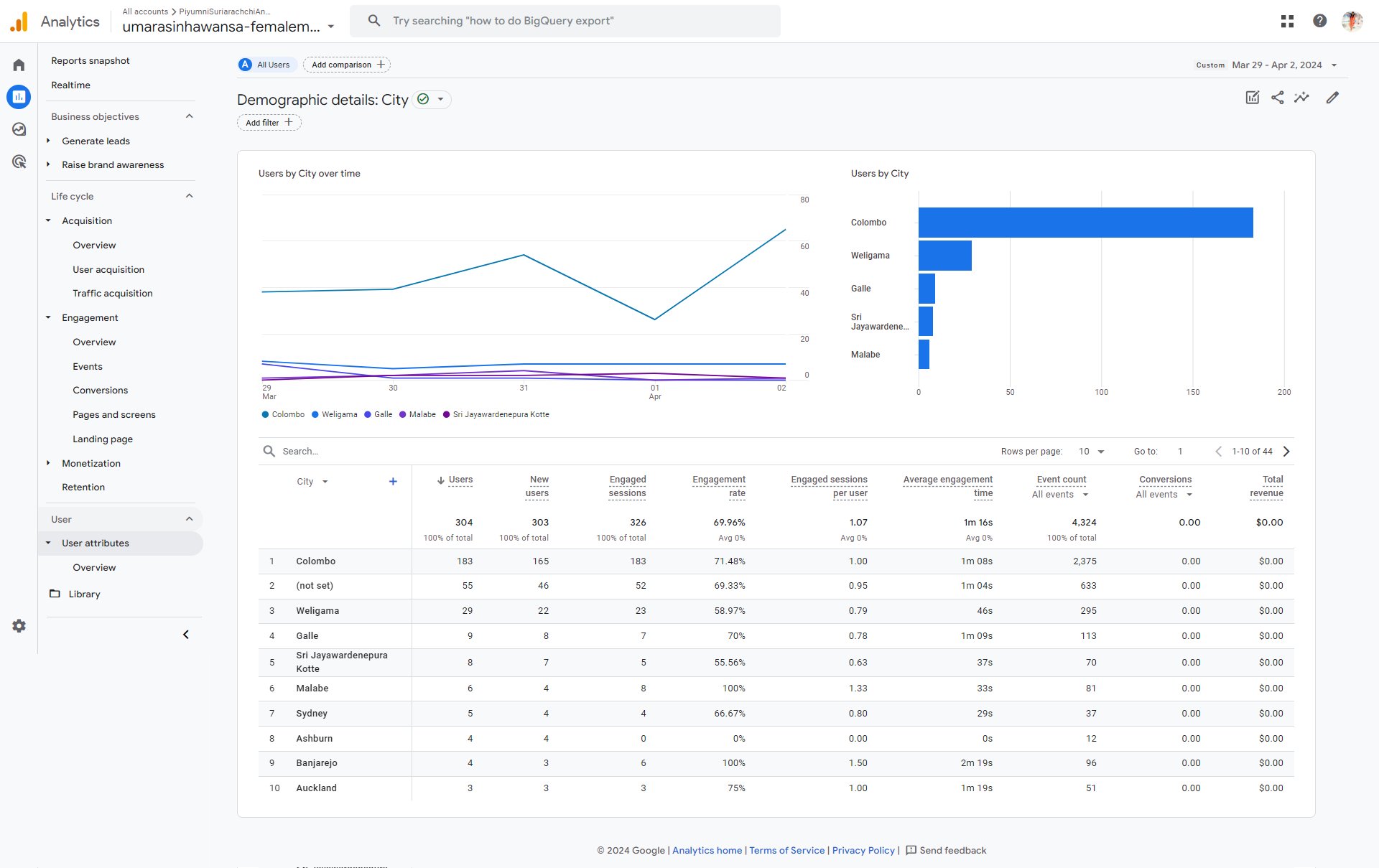


Figure 13:Google Analytics Report for Users By City

The above report provides insights into the cities the users are from. As per the report over the period highest number of users are recorded from Colombo city with 183 users followed by users from Weligama When considering internationally not including local cicties, Sydney in Australia takes precedence however users from Sydney is not a significant amount compared to total users. Hence in overall, local users have contributed the most towards generation. Here “not set” wording refers to was unable to capture the IP address to identify the user’s location and because of that the cities of 55 users are not identified.

## Returning Users by Device Category

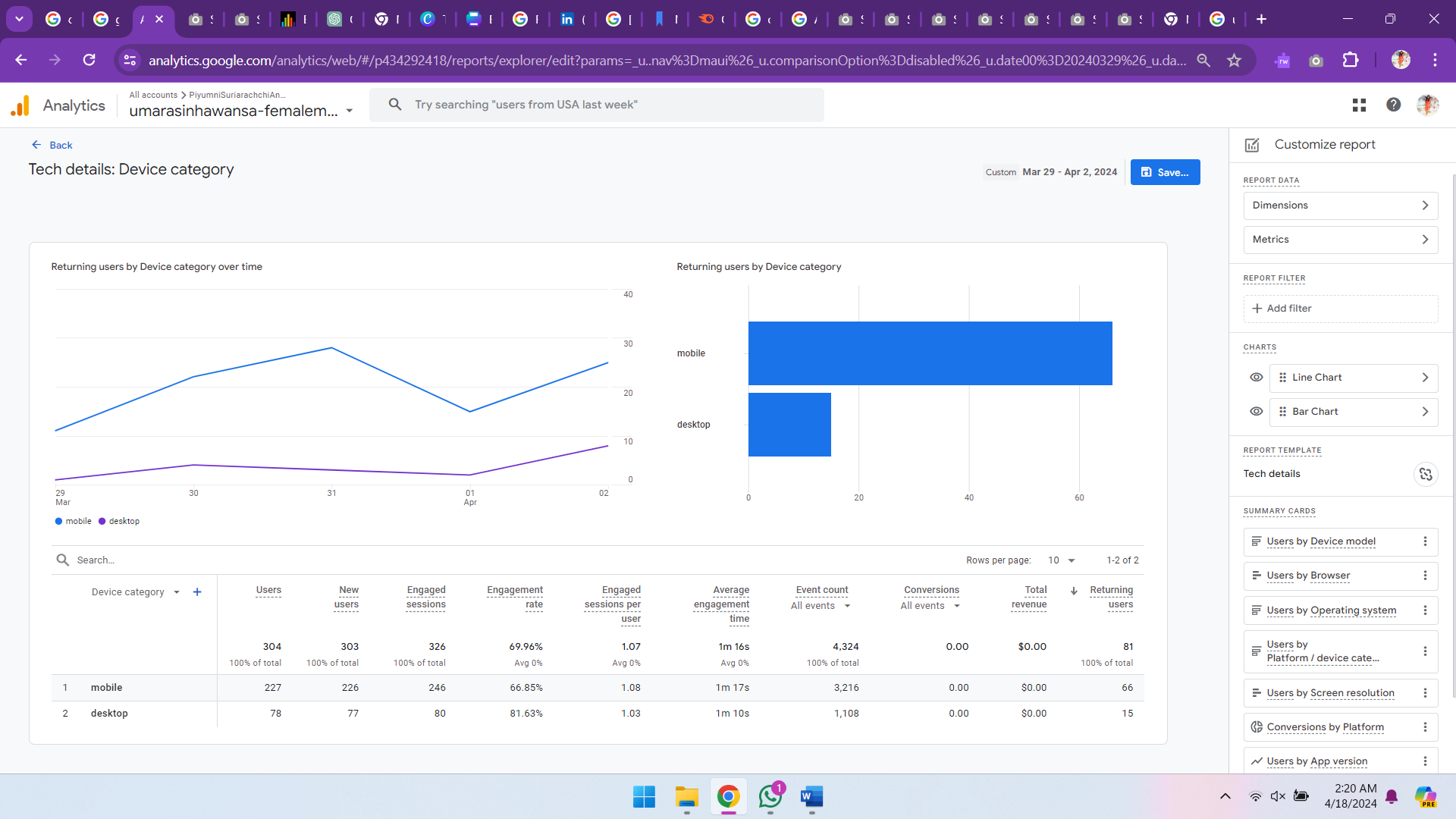


Figure 14: Google Analytics Report for Returning Users by Device Category

The above report gives insights into the devices used by returning users when visiting the sights. As per the report the most used device category is mobile with 66 returning users. This shows that the user retention through mobile is high. As a suggestion based on this report it can be said that when doing optimizations focus on the mobile view as well.

## New Users by Screen Resolution

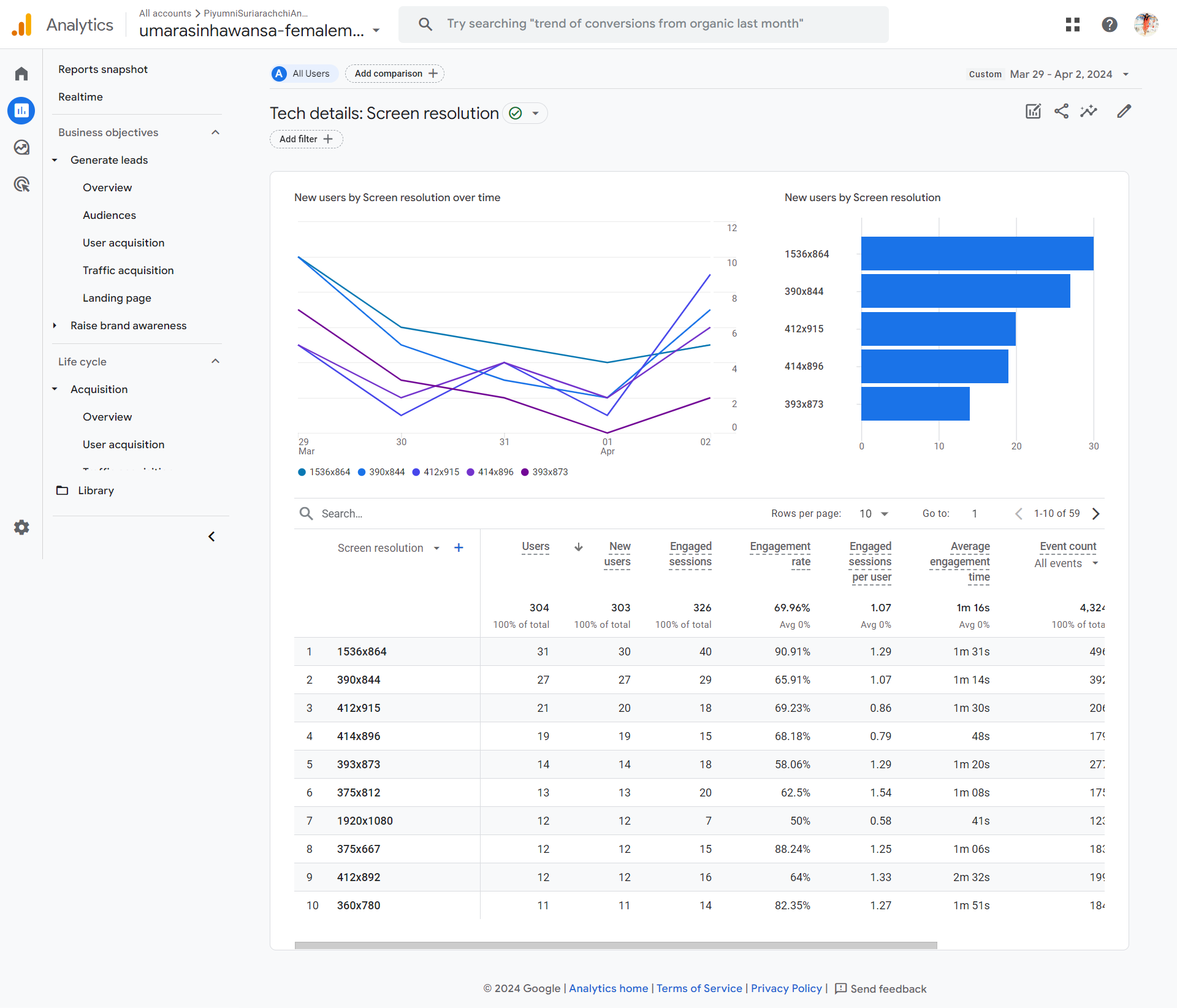


Figure 15:Google Analytics Report for New Users by Screen Resolution

The above report gives insights into the screen resolution of the devices used by new users when visiting the site. As per the report most acquisition of new users is done through devices having 153 x 864 screen resolution.

# Task 5

## State KPI

Improve average time a user spent on site by 15% by the end of 6 months.

The KPI should be monitored end of every two weeks changes that will be taken will be should include optimizing web pages, improve user design and experience of site and do enhanced marketing.

## Justification based on Task 4

One of the goals in this microsite is to ensure that the users are engaging with the content, hence a stakeholder goal would be to see improved engagement, In the task 4 analysis it was identified that user engagement has declined over the period. This was mainly identified through the decline of average engagement time and averaged engaged time per session over the time in the User Engagement overview report. This clearly highlights that certain actions should be taken to increase the time spent by users engaging with the site, hence justifying the reason for forming the KPI.

## Justification of the sections of KPI

* Description –Improving average time spent by a user in the site is justified due to the there is a decline in engagement time that is justified in section 5.2 as well.
* Exact change – 10% increase in average time spent in site. Currently average session timing is around 49, where the lowest is around 36s and highest around 1 minute hence a 10 percent increase will be ideal to stop the declining of engaging time.
* Time – 6 months, this is justified because currently there the average session timing is 49s and with a 10 percent increase it will not be difficult if the changes are properly performed.
* Changes to be done - The changes planned are optimizing the web pages, improving the Ui design and conduct marketing. All these such as having good quality content with easy navigation are vital towards increasing the time user spends in the site and marketing aspects can further promote the site and make the visitors more engaged with the site.
* Monitoring period: Monitoring period is kept twice a week. Over the period of 6 months if monitoring is done every 2 weeks the, if there are any concerns they can be addressed without any delay of time.

.

# Task 6

# Task 7

## Challenges related to the calculation of the number of unique visitors to a website.

* **Deleting or blocking of cookies** – There is an option to block cookies in most sites, and users has a tendency to use this option to block them (Dainow, 2017) This causes an issue in calculating unique visitors because as mentioned above HTTP protocol is stateless and without cookies analytics tools will not be able to calculate unique visitors properly, resulting in an incorrect unique visitor count (Meredith, 2023).
* **Using of multiple devices** – users can have more than one device with them, and because of this a single user can access a particular site from more than one device, leading to showing unique users more than what is should be, creating an incorrect unique visitor count (Jansen, 2022).
* **Using shared devices** – In above there is an incorrect increased amount in visitors, while in this scenario it is an incorrect lower number of users because, if multiple users use the same device analytics tools will not be able to calculate the unique visitor count properly (Jansen, 2022)
* **Dynamic IP address** – there is a possibility of IP addresses changing hence, this creates a possibility of a user being calculated more than once. This too will cause an incorrect increase in unique visitors.

## Differences between “visitors”, “new visitors” and “unique visitors”

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Visitors** | **New Visitors** | **Unique Visitor** |
| **Definition** | Total number of users who visited the site including repeated users over a selected time period. | users who visit and interacts with website for the first time within a particular time frame. | Distinct Users who visit and interacts with the site more than once during a particular time frame and a unique visitor is calculated only once during the reporting period. |
| **How the visitor count is taken** | Can be counted more than once in a period of time. For example, if a user visits a site in 2 days it will be counted as 2 visitors | Counted only once during a period of time and the returning users are not included. | Counted only once even if they visit a number of times during a particular time frame. |

## **Challenges** associated with measuring session duration and quantifying the amount of time spent on an individual web page

|  |  |
| --- | --- |
| Challenge | Description |
| **Misleading session timing due to calculation process of bounce rate.** | When calculating the engagement (bounce rate) of users’ analytics platform,if a user visits a page on a website but does not go and interact with another page, the session duration is recorded as Zero considering the visit to be a single visit. This causes an underestimation of engagement and session timing calculations (Jansen, Jung and Salminen, 2022). |
| **Users working on different tabs and being inactive in pages.** | Currently users have the option to open a number of tabs, and there can be options where the users have opened the tabs but not properly engaging or else they can be going from one tab to another and keep the pages running. This can cause an overestimation or underestimation of time users spent in sessions or in a page (Lehmann et al., 2013) |
| **Issues in identifying session duration** | It is not possible to say if a user session is ended from transactional log files, this is mainly due to users usually do not log out from sites and just keep it as it is. This makes it difficult to find when the session exactly ended (Huntington, Nicholas and Jamali, 2008). |
| **Session Time Out delimiter** | Session have set timeout settings to end sessions. If a session time is long, it will combine 2 sessions that happened separately. Hence these concerns related with session time out can cause an overestimation or underestimation of session durations. (Huntington, Nicholas and Jamali, 2008). |
| **Using proxy server and shared IP addresses** | Due to the use of shared IPs, in the session longs multiple users can be shown hence this can create difficulties in accurately track sessions (Huntington, Nicholas and Jamali, 2008). |

## Explain what is meant by the “hotel problem”

The hotel problem is a hypothetical situation that is used to describe the misinterpretations that can happen in web analytics and the difference of unique visitors and total visitors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Day 1** | **Day 2** | **Day 3** | **Total Unique visitors** |
| **Room A** | Anne | Anne | Sarah | 2 |
| **Room B** | Sarah | Katy | Katy | 2 |
| **Total visitors** | 2 | 2 | 2 | **3** |

In the below table it can be seen that in each day hotel had two visitors, and over the period, room A had two unique visitors and room B had two unique visitors.

Here if unique visitor per day is summed up the total would be four. And if the total visitors per day is summed up will be six. However, in reality during this 3-day period the hotel had only three guests. This shows that

* Unique visitors + unique visitors are not equal to total number of unique visitors
* New visitors + repeat visitors are not equal to total number of visitors.

The reasons for these incorrect calculations that led to duplication of visitors is because when calculating total visitors and unique visitors certain factors are not considered. They are

* Visitor switching rooms, for example Sarah has switched to Room B in day 3.
* Visitor staying in more than one day in a room. For example, Anne has stayed in Room A for 2 days

Similarly, this could happen in web analytics as well. For example, if a person who visits a site switch device like the hotel visitors switched room or if multiple users visit website from one device like a room being used by more than one person over a time the web analytics will not be properly calculated.

Hence in overall hotel problem shows how data can be misinterpreted due to certain situations and highlights importance of knowing how analytics tools track data, the difference between total visitors and unique visitors, related challenges that cause misinterpretation which will help to properly understand the user engagement, user count and related web analytics.

## Other Problems related to data misinterpretation in web analytics.

* Sampling errors – When analyzing data most analytics tools sample the data after reaching a certain limit of actions. In addition, these analyzing of sample sets can cause accuracies and hence this can lead to misinterpretation since the business or the people who need the analytics to arrive at decisions based on analytics of sampled data (Jansen, Jung and Salminen, 2022).
* Misleading engagement Rate: The user behavior of visitors is different from each other, For example some users take a long period to read pages. In this sort of scenario, just because the engagement time is high it cannot be assumed that the user engagement is good. This shows that the engagement rate can be misleading (Agarwal, 2016).
* Limited data accuracy due to ad blockers – Due to certain technical limitations such as browser settings, ad blockers or all the data will not be collected properly. This is because ad blockers can stop the pages from rendering properly making it difficult for the analytics tool to do tracking through asset blocking and element hiding. Hence the reports generated will not show accurate data. This will result in misinterpretation of data (Donica, 2022)