# Introduction

E- commerce has resulted in organizations investing significant resources in online strategies to extend business processes on to the World Wide Web. Traditional methods of measuring Web usage fall short of the richness of data required for the effective evaluation of such strategies. Web analytics are an approach that may meet organizational demand for effective evaluation of online strategies.

A website user is a person who is accessing, browsing or interacting with a website, and user behavior refers to how people use a website. Behaviors include everything from the journey they take through the site to interactions such as clicks. When it comes to optimizing a website, simply monitoring behavior can only get you so far. The real value comes from analyzing users' actions to get to the bottom of what makes them behave as they do.

Behaviour is complex and varies across different websites depending on the target audience. This means you need to learn specifically about *your* users. Who are they? What are their needs? Which browsers and devices do they prefer? How often do they purchase? Answering these questions is crucial if you want to have a competitive edge, meet consumer needs and retain your customers. By researching online behaviours, you can get an idea of what users are trying to achieve, the factors driving certain behaviours, where they experience friction and areas where user experience can be better. Ultimately, learning how visitors behave on your website allows you to provide an enhanced experience that's in line with user needs, which in turn will ensure your business continues to grow.

One of the main tools we use to carry out research is Google Analytics. This is free and fairly easy to set up. Once installed, you'll have access to valuable data about how users behave on your site, including where they land and go next, where they drop off and what they interact with. You can also use it to discover overall trends and patterns and source opportunities for growth.

# Web Application

E-commerce websites are online portals that facilitate online transactions of goods and services through means of the transfer of information and funds over the Internet. In the early days, e-Commerce was done partially through emails and phone calls. Now, with a single website, anything and everything that a transaction needs, can be executed online.

I have used following steps to create my E-commerce website:

- Firstly, I have taken domain name (infinity.cloudaccess.host) and hosting from cloudaccess.net, which also provided me access to the wordpress.

- Then, I use a pre-existent theme Astra from wordpress and edit the same to suit my requirements.

- My website contains following webpages:

    a) **Home**- It is the homepage of my website.
    b) **Store**- It contains various products categorized into different categories.
    c) **Account-** It contains account details of the user.
    d) **Contact Us-** It contains the contact details of the company.
    e) **About-** It contains team details.

- Website contains various features:

    a) View a list of products
    b) View product details
    c) Search products
    d) Use filters to change the product list (eg. Category, price range, etc.)
    e) Add a product to the cart.
    f) Payment Gateway.

- **UserGuiding**- I have applied user guiding tool named WalkMe that will guide a new user through orderly steps on my website to use it in the most simple way to buy a product.

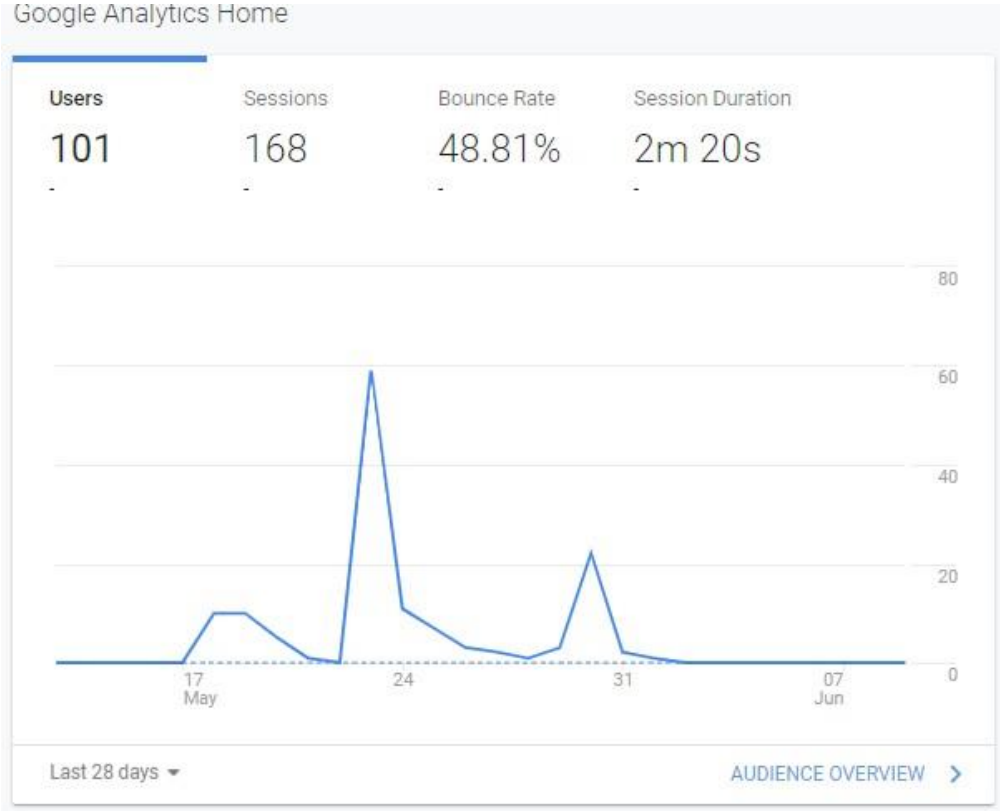
# Analytics

User Behavior Analytics or UBA focuses on what the user is doing: apps launched, network activity, and, most critically files accessed (when the file or email was touched, who touched it, what was done with it and how frequently).

UBA technology searches for patterns of usage that indicate unusual or anomalous behavior — regardless of whether the activities are coming from a hacker, insider, or even malware or other processes. While UBA won't prevent hackers or insiders from getting into your system, it can quickly spot their work and minimize damage.

I have integrated my website to analytics platform named Google Analytics. This tool is able to track the users action and the time spent by the users on the various actions of the website. I have accumulated the analytics data from various users and stored it.
I have collected data of around 103 users with various data points such as:

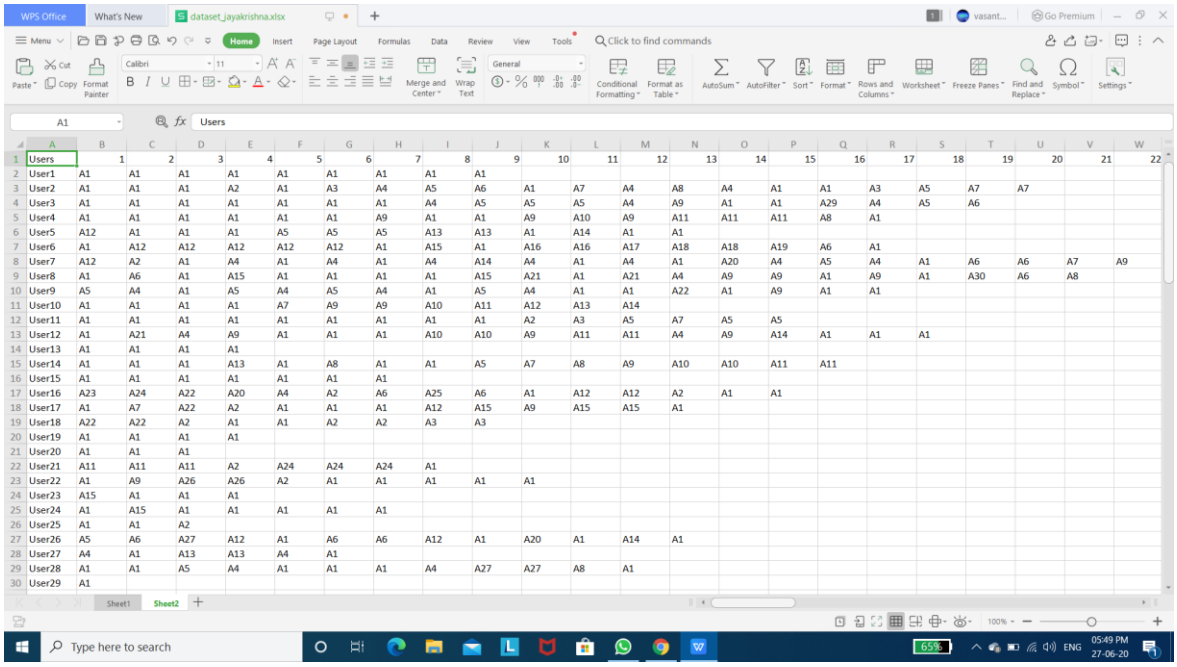
- Time spent by users on the different pages on website,
- Filters being used by the users ,
- Time spent by users on particular product page, etc.




**Audience Overview**

| Page | Pageviews | Unique Pa | Avg. Time | Entrances | Bounce Ra | % Exit |
|---|---|---|---|---|---|---|
| / | 210 | 129 | 50.91 | 128 | 40.63% | 38.57% |
| /?post_type=product | 96 | 62 | 28.39 | 1 | 0.00% | 7.29% |
| /?page_id=39 | 65 | 51 | 26.63 | 10 | 70.00% | 38.46% |
| /?product=a-hundred-little-flame | 32 | 13 | 5.25 | 1 | 100.00% | 12.50% |
| /?product=bheem | 21 | 9 | 7.50 | 0 | 0.00% | 14.29% |
| /?page_id=40 | 17 | 17 | 26.17 | 4 | 100.00% | 64.71% |
| /?page_id=43 | 17 | 11 | 48.43 | 0 | 0.00% | 17.65% |
| /?product=black-holes-the-reith-l | 17 | 6 | 15.06 | 0 | 0.00% | 0.00% |
| /?page_id=414 | 15 | 12 | 60.00 | 3 | 33.33% | 26.67% |
| /?product=everyone-has-a-story | 15 | 5 | 28.79 | 0 | 0.00% | 6.67% |
| /?product=half-girlfriend | 12 | 7 | 17.67 | 0 | 0.00% | 25.00% |
| /?post_type=product&paged=2 | 9 | 6 | 15.25 | 0 | 0.00% | 11.11% |
| /?product=harry-potter-and-the-g | 9 | 3 | 6.67 | 0 | 0.00% | 0.00% |
| /?product=flirting-with-stocks-stc | 7 | 2 | 2.71 | 0 | 0.00% | 0.00% |
| /?product=half-torn-hearts | 5 | 3 | 5.60 | 0 | 0.00% | 0.00% |
| /?product=the-intelligent-investc | 5 | 2 | 7.00 | 0 | 0.00% | 0.00% |
| /?page_id=41 | 4 | 4 | 7.50 | 0 | 0.00% | 0.00% |
| /?product=harappa-curse-of-the- | 4 | 1 | 4.50 | 0 | 0.00% | 0.00% |
| /?product=harry-potter-and-the-c | 4 | 2 | 8.00 | 0 | 0.00% | 25.00% |
| /?product=life-is-what-you-make- | 4 | 1 | 9.50 | 0 | 0.00% | 0.00% |
| /?product=the-alchemist | 4 | 1 | 4.75 | 0 | 0.00% | 0.00% |
| /?product=will-you-still-love-me | 4 | 1 | 4.50 | 0 | 0.00% | 0.00% |
| /?product=harry-potter-and-the-p | 3 | 1 | 11.33 | 0 | 0.00% | 0.00% |
| /?product=one-indian-girl | 3 | 2 | 9.00 | 0 | 0.00% | 0.00% |

Example – Analytics All Website Data Pages




Example- Action Perform By Users

# ML Model and Data Analysis

The elusive clickstream data. Many platforms, like Facebook rely on these generated data from what a user clicks and what doesn't. To start analyzing clickstream data, we need first to be able to capture step by step a user's activity across a web page or application. And that is of great value in the hands of any Internet marketer. Getting a 360-degree view of a customer by knowing what he is clicking and what he is not can get you a huge improvement in both your products and your customers' experience.

## Data Collection

Either you have your data in your data warehouse, or you need to enrich it with more data sources you need to have a way to collect and store data consistently into a database.

## Data Preparation

Raw data is like a rough diamond; It requires some refinement before being truly valuable.

In the data world, refinement includes data processing, cleaning, and transformation of the initial data into something convenient for the analysis you are going to carry out.

In this case, we would like to have our data grouped into users. It would be good too, we could arrange the events of each user in time order before moving to actual analysis.

In contrast to other data sequences, clickstream data can have varying length for every different users.

In order to transform the initially collected event log into clickstream data we need to:

- Identify events/actions performed by the same user and group them together
- Split them further into subgroups of events based on which of those were performed during the same session according to the session's definition given above.

At this point the dataset we are going to use for the rest of the analysis should look like this:

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| User5 | A12 | A1 | A1 | A1 | A5 | A5 | A5 | A13 | A13 | A1 | A14 | A1 | A1 | | | | | | | | | | |
| User6 | A1 | A12 | A12 | A12 | A12 | A12 | A1 | A15 | A1 | A16 | A16 | A17 | A18 | A18 | A19 | A6 | A1 | | | | | | |
| User7 | A12 | A2 | A1 | A4 | A1 | A4 | A1 | A4 | A14 | A4 | A1 | A4 | A1 | A20 | A4 | A5 | A4 | A1 | A6 | A6 | A7 | A9 | |
| User8 | A1 | A6 | A1 | A15 | A1 | A1 | A1 | A1 | A15 | A21 | A1 | A21 | A4 | A9 | A9 | A1 | A9 | A1 | A30 | A6 | A8 | | |
| User9 | A5 | A4 | A1 | A5 | A4 | A5 | A4 | A1 | A5 | A4 | A1 | A1 | A22 | A1 | A9 | A1 | A1 | | | | | | |
| User10 | A1 | A1 | A1 | A1 | A7 | A9 | A9 | A10 | A11 | A12 | A13 | A14 | | | | | | | | | | | |
| User11 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A2 | A3 | A5 | A7 | A5 | A5 | | | | | | | | |
| User12 | A1 | A21 | A4 | A9 | A1 | A1 | A1 | A10 | A10 | A9 | A11 | A11 | A4 | A9 | A14 | A1 | A1 | A1 | | | | | |
| User13 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User14 | A1 | A1 | A1 | A13 | A1 | A8 | A1 | A1 | A5 | A7 | A8 | A9 | A10 | A10 | A11 | A11 | | | | | | | |
| User15 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | |
| User16 | A23 | A24 | A22 | A20 | A4 | A2 | A6 | A25 | A6 | A1 | A12 | A12 | A2 | A1 | A1 | | | | | | | | |
| User17 | A1 | A7 | A22 | A2 | A1 | A1 | A1 | A12 | A15 | A9 | A15 | A15 | A1 | | | | | | | | | | |
| User18 | A22 | A22 | A2 | A1 | A1 | A2 | A2 | A3 | A3 | | | | | | | | | | | | | | |
| User19 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User20 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | | |
| User21 | A11 | A11 | A11 | A2 | A24 | A24 | A24 | A1 | | | | | | | | | | | | | | | |
| User22 | A1 | A9 | A26 | A26 | A2 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | |
| User23 | A15 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User24 | A1 | A15 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | |
| User25 | A1 | A1 | A2 | | | | | | | | | | | | | | | | | | | | |
| User26 | A5 | A6 | A27 | A12 | A1 | A6 | A6 | A12 | A1 | A20 | A1 | A14 | A1 | | | | | | | | | | |

In this representation, each line corresponds to a user. The first field is the user's name while the next fields the actions performed by the user during this session.

## Model Construction

As in most cases, the methods we can deploy for solving this problem are many.

## Markov Chains

The type of data Markov Chains work with are sequential data.

Markov process is a stochastic process that satisfies the Markov Property of memorylessness. A Markov chain is, in fact, a Markov process too in either discrete or continuous time with a countable state space.

In clickstream analysis, we usually utilize these Markov Chains. The process takes the state from a finite set at each time .
The order of a Markov Chain is derived from the number of recent states on which the current state, we assume, depends. Based on this, zero-order chains imply that the probability of being in a state in the next step is independent of all previous states.

Higher order Markov Chain introduced by the Raftery (1985) will lead to more realistic models. At the same time, the parameters needed for the representation increase exponentially and so it is important to find a right balance between these two.

## Fitting a Markov Chain

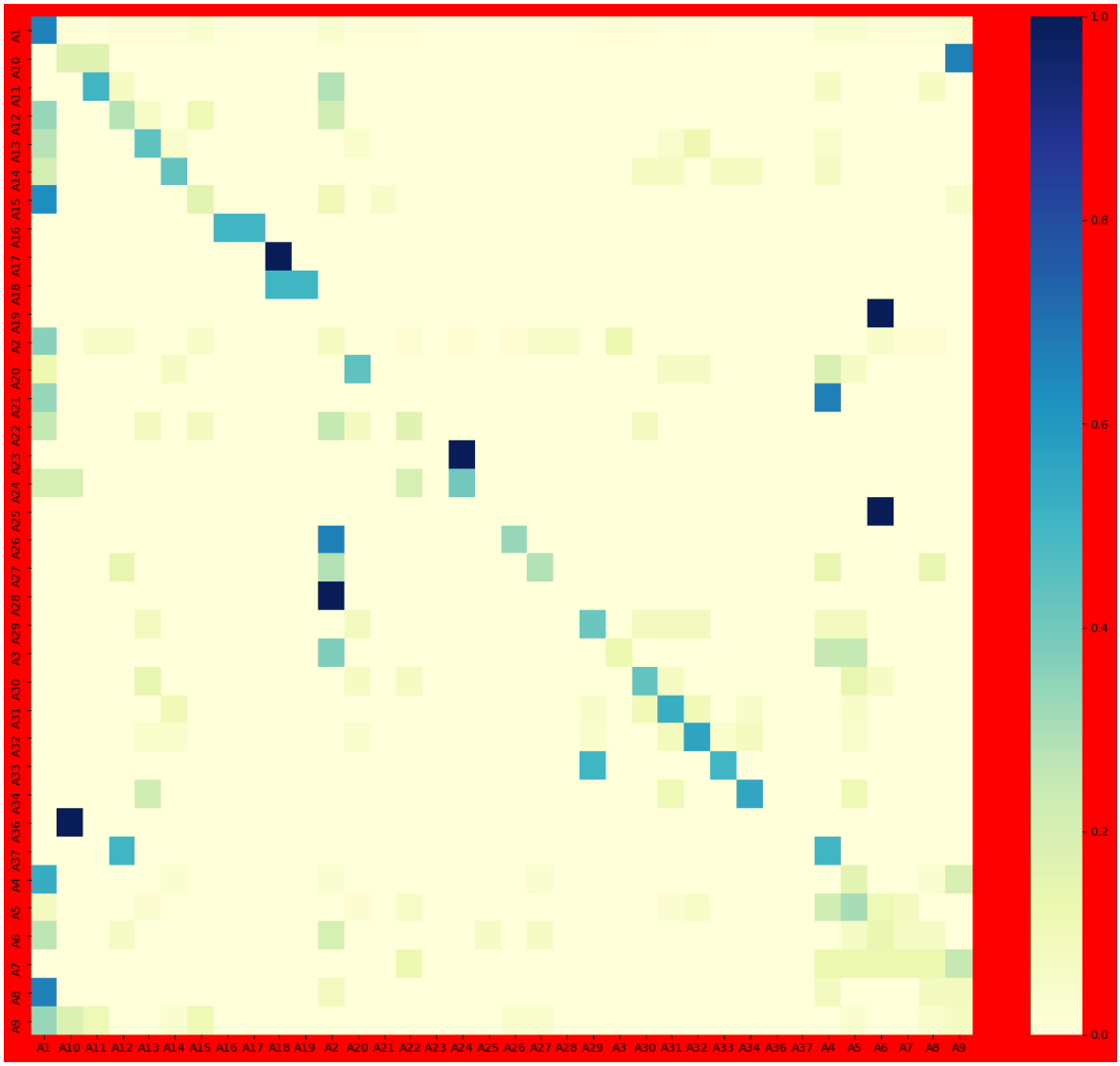
As mentioned before at this point our dataset looks like:

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| User5 | A12 | A1 | A1 | A1 | A5 | A5 | A5 | A13 | A13 | A1 | A14 | A1 | A1 | | | | | | | | | | |
| User6 | A1 | A12 | A12 | A12 | A12 | A12 | A1 | A15 | A1 | A16 | A16 | A17 | A18 | A18 | A19 | A6 | A1 | | | | | | |
| User7 | A12 | A2 | A1 | A4 | A1 | A4 | A1 | A4 | A14 | A4 | A1 | A4 | A1 | A20 | A4 | A5 | A4 | A1 | A6 | A6 | A7 | A9 | |
| User8 | A1 | A6 | A1 | A15 | A1 | A1 | A1 | A1 | A15 | A21 | A1 | A21 | A4 | A9 | A9 | A1 | A9 | A1 | A30 | A6 | A8 | | |
| User9 | A5 | A4 | A1 | A5 | A4 | A5 | A4 | A1 | A5 | A4 | A1 | A1 | A22 | A1 | A9 | A1 | A1 | | | | | | |
| User10 | A1 | A1 | A1 | A1 | A7 | A9 | A9 | A10 | A11 | A12 | A13 | A14 | | | | | | | | | | | |
| User11 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | A2 | A3 | A5 | A7 | A5 | A5 | | | | | | | | |
| User12 | A1 | A21 | A4 | A9 | A1 | A1 | A1 | A10 | A10 | A9 | A11 | A11 | A4 | A9 | A14 | A1 | A1 | A1 | | | | | |
| User13 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User14 | A1 | A1 | A1 | A13 | A1 | A8 | A1 | A1 | A5 | A7 | A8 | A9 | A10 | A10 | A11 | A11 | | | | | | | |
| User15 | A1 | A1 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | |
| User16 | A23 | A24 | A22 | A20 | A4 | A2 | A6 | A25 | A6 | A1 | A12 | A12 | A2 | A1 | A1 | | | | | | | | |
| User17 | A1 | A7 | A22 | A2 | A1 | A1 | A1 | A12 | A15 | A9 | A15 | A15 | A1 | | | | | | | | | | |
| User18 | A22 | A22 | A2 | A1 | A1 | A2 | A2 | A3 | A3 | | | | | | | | | | | | | | |
| User19 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User20 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | | |
| User21 | A11 | A11 | A11 | A2 | A24 | A24 | A24 | A1 | | | | | | | | | | | | | | | |
| User22 | A1 | A9 | A26 | A26 | A2 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | |
| User23 | A15 | A1 | A1 | A1 | | | | | | | | | | | | | | | | | | | |
| User24 | A1 | A15 | A1 | A1 | A1 | A1 | A1 | | | | | | | | | | | | | | | | |
| User25 | A1 | A1 | A2 | | | | | | | | | | | | | | | | | | | | |
| User26 | A5 | A6 | A27 | A12 | A1 | A6 | A6 | A12 | A1 | A20 | A1 | A14 | A1 | | | | | | | | | | |

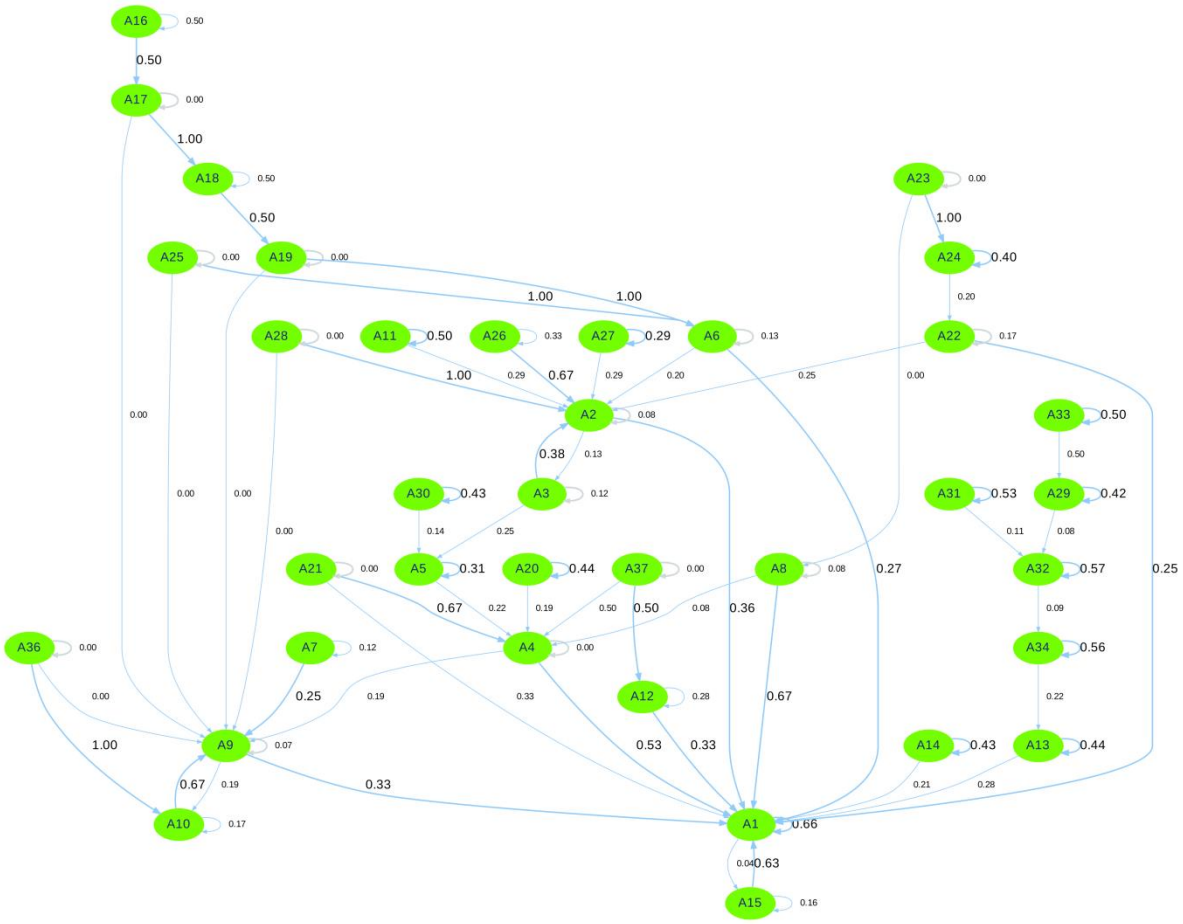
Fitting the Markov Chain model gives us the transition probabilities matrices and the lambda parameters of the chain for each one of the three lags along with the Start and End Probabilities.

Start and End probabilities correspond to the probability that a clickstream will start or end with this specific event.

The transition probability matrix can be represented as a heat map with the y-axis representing the current state and x-axis the next one. The more blu-ish the color, the more probable the indicated transition will occur.




**Heat Map**




# Transaction diagram

# Actions Details-

A1 -- /
A2 -- /?product_cat=men
A3 -- /?product=dnk-red-shoes
A4 -- /?post_type=product
A5 -- /?product=anchor-bracelet
A6 -- /?product_cat=accessories
A7 -- /?page_id=41
A8 -- /?product=dark-brown-jeans
A9 -- /?page_id=39
A10 -- /?page_id=40
A11 -- /?product=black-hoodie
A12 -- /?product=dnk-blue-shoes
A13 -- /?product=blue-denim-jeans
A14 -- /?product=blue-denim-shorts
A15 -- /?product=dnk-yellow-shoes
A16 -- /?s=no-results:Women jackets&cat=no-results
A17 -- /?s=no-results:Women headwear&cat=no-results
A18 -- /?s=no-results:Women cap&cat=no-results
A19 -/?orderby=price&paged=1&product_cat=accessories
A20 -- /?product=basic-gray-jeans
A21 -- /?post_type=product&paged=2
A22 -- /?product_cat=women
A23 -- /?page_id=43
A24 -- /?page_id=414
A25/?orderby=popularity&paged=1&product_cat=accessories
A26 -- /?product=faint-washed-denim-blue-jeans
A27 -- /?product=dark-blue-denim-jeans
A28 -- /?product=dnk-green-shoes
A29 -- /?product=gray-pattern-tshirt
A30 -- /?product=oranges-tshirt
A31 -- /?product=flamingo-tshirt
A32 -- /?product=lemons-tshirt
A33 -- /?product=slim-fit-blue-jeans
A34 -- /?product=dnk-black-shoes
A35 -- /?product=white-underground-tshirt
A36 --/?page_id=41&lost-password=
A37 -- /?product=dark-gray-jeans

# Suggestion to UX Designer

The following are the insights from the Heat Map generated by the Markov chain model:

- The transaction from Action 10 to Action 26, Action 16 to Action 15 , and Action 24 to Action 26 is more co-related to each other.
- Action 1 to Action 7, and Action 2 to Action 19 is not co-related to each other.

Suggestion taken from transaction diagram along with the corresponding probability are:

- All the users must start from A1 (homepage).
- The user from action A25, A26, A13, A14, A22, A23, A24 (all are products) are occasionally not going to A5 (cart). They should go to A5 (cart) so the UX designer should guide the user to go to the A5 from these actions.
- The UX designer should guide the user to go to the A6 (Checkout) from A5 (cart).
- Users are either going back to A5 (cart) or close the application at A6 (Checkout) without proceeding to payment, so the UX designer should guide the user to place order and make payment.
- The actions A2, A18, A17 and A20, A3 (all are products) instead of going forward comes back to A1 (homepage). So the UX designer should guide the user to go to next actions to place order.