

Aim -

Study web analytics using open-source tools like Matomo, Open Web Analytics, AWStats, Countly, and Plausible.

Theory -

Web analytics is the process of analyzing the behavior of visitors to a website. This involves tracking, reviewing and reporting data to measure web activity, including the use of a website and its components, such as web pages, images and videos.

Data collected through web analytics may include traffic sources, referring sites, page views, paths taken and conversion rates. The compiled data often forms a part of customer relationship management analytics (CRM analytics) to facilitate and streamline better business decisions.

Web analytics enables a business to retain customers, attract more visitors and increase the dollar volume each customer spends.

Analytics can help in the following ways:

Determine the likelihood that a given customer will repurchase a product after purchasing it in the past.

Personalize the site to customers who visit it repeatedly.

Monitor the amount of money individual customers or specific groups of customers spend.

Observe the geographic regions from which the most and the least customers visit the site and purchase specific products.

Predict which products customers are most and least likely to buy in the future.

The objective of web analytics is to serve as a business metric for promoting specific products to the customers who are most likely to buy them and to determine which products a specific customer is most likely to purchase. This can help improve the ratio of revenue to marketing costs.

In addition to these features, web analytics may track the clickthrough and drill down behavior of customers within a website, determine the sites from which customers most often arrive, and communicate with browsers to track and analyze online behavior. The results of web analytics are provided in the form of tables, charts and graphs.

Plausible -

Plausible Analytics is an open-source project dedicated to making web analytics more privacy-friendly. Our mission is to reduce corporate surveillance by providing an alternative web analytics tool which doesn't come from the AdTech world.

Apache Jena-

Apache Jena is an open source Semantic Web framework for Java. It provides an API to extract data from and write to RDF graphs. The graphs are represented as an abstract "model". A

model can be sourced with data from files, databases, URLs or a combination of these. A model can also be queried through SPARQL 1.1.

Jena is similar to RDF4J (formerly OpenRDF Sesame); though, unlike RDF4J, Jena provides support for OWL (Web Ontology Language). The framework has various internal reasoners and the Pellet reasoner (an open source Java OWL-DL reasoner) can be set up to work in Jena.

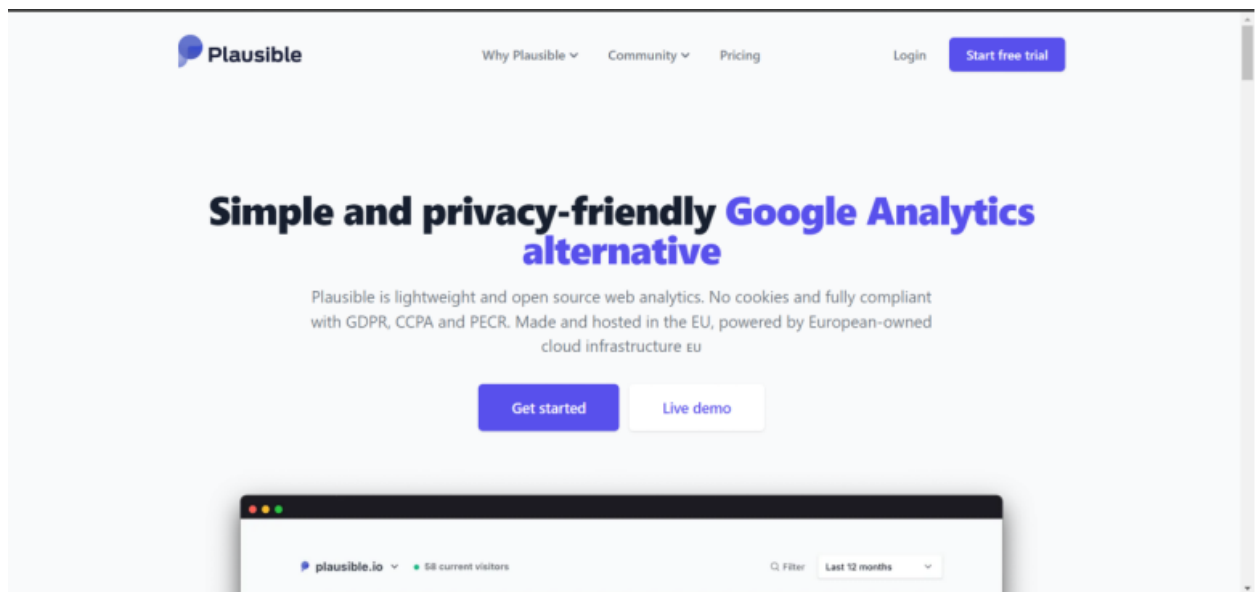
Jena supports serialization of RDF graphs to:

a relational database

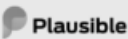
- RDF/XML
- Turtle
- TriG
- Notation 3
- JSON-LD

Output -

1. Create an account on Plausible.



2. Login to your account and provide the name of the site whose analytics you want.



Your website details

Domain
Just the naked domain or subdomain without 'www'

https://

example.com

Reporting Timezone
To make sure we agree on what 'today' means

(GMT+05:30) Asia/Calcutta

Add snippet →

✓ Register


✓ Activate account

● Add site info

■ Install snippet

©2023 Plausible Analytics. All rights reserved.

3. Add the code snippet to the head of your website.



Add JavaScript snippet

Paste this snippet in the <head> of your website.

```
<script defer data-domain="nikilshani.netlify.app"
src="https://plausible.io/js/script.js"></script>
```

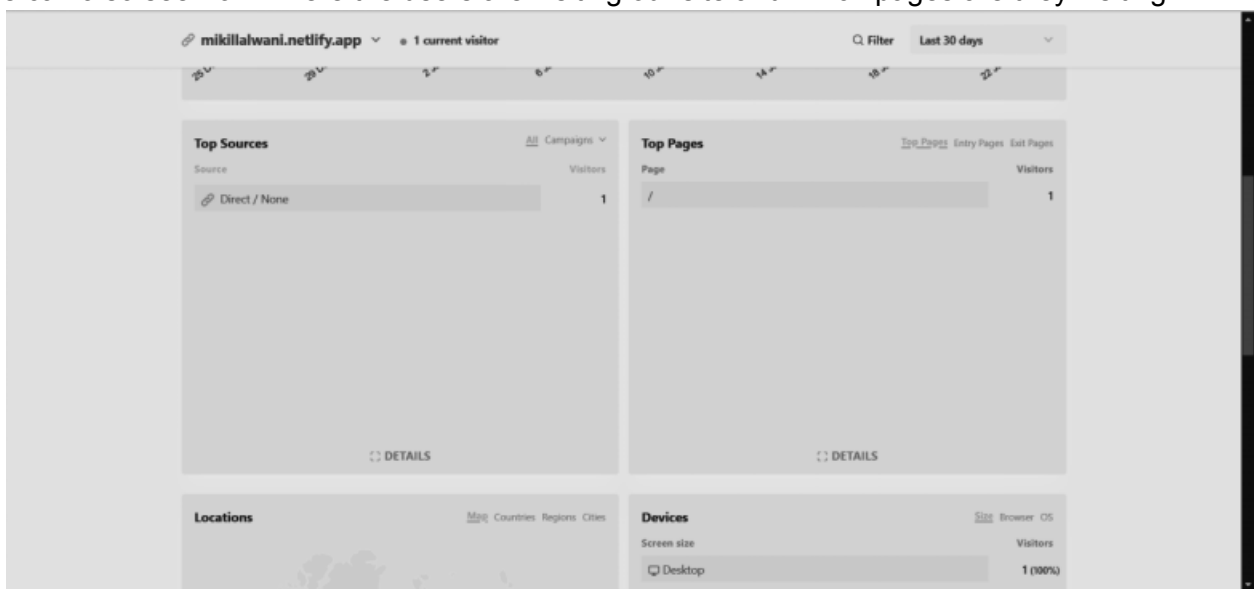
Start collecting data →

©2023 Plausible Analytics. All rights reserved.

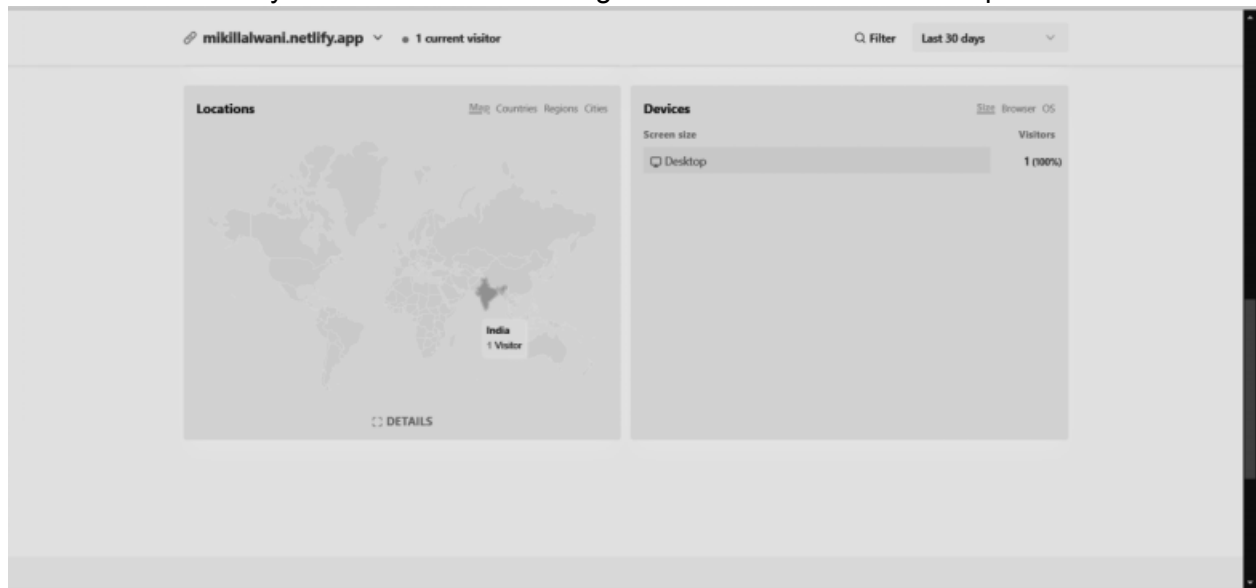
4. We can see the number of viewers as time passes.



5. We can also see from where the users are visiting our site and which pages are they visiting.



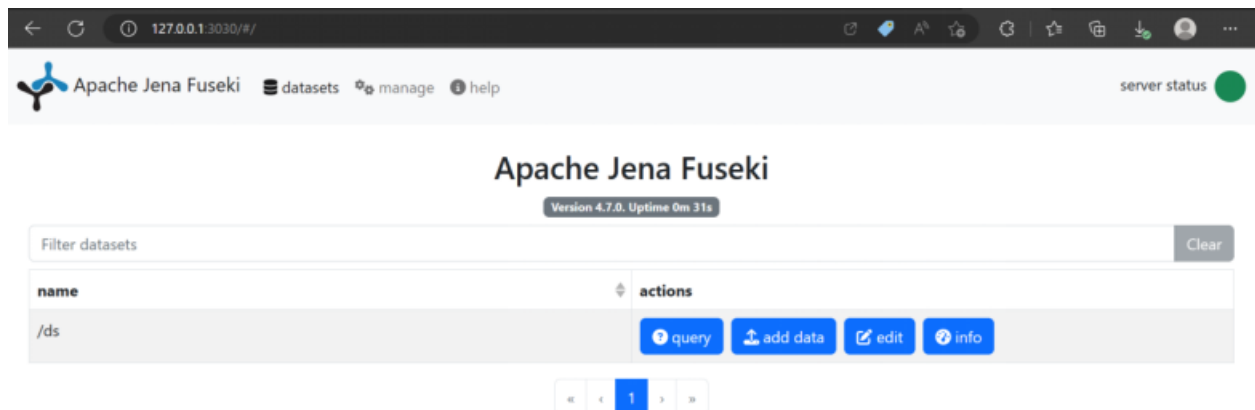
6. Also the device used by the user while accessing our site and their location is present.



Apache Jena

Run server using

Fuseki-server -update -mem /ds



/ds

[query](#) [add data](#) [edit](#) [info](#)

SPARQL Query

To try out some SPARQL queries against the selected dataset, enter your query here.

Example Queries

[Selection of triples](#) [Selection of classes](#)

Prefixes

[rdf](#) [rdflib](#) [owl](#) [xsd](#)

SPARQL Endpoint

/ds/

Content Type (SELECT)

JSON

Content Type (GRAPH)

Turtle



/ds

[query](#) [add data](#) [edit](#) [info](#)

Available Services

Graph Store Protocol (Read)	/ds/get
Graph Store Protocol	/ds/data
Graph Store Protocol	/ds/
SPARQL Query	/ds/sparql
SPARQL Query	/ds/query
SPARQL Update	/ds/
SPARQL Update	/ds/update

Dataset size

count triples in all graphs

graph name	triples
------------	---------

Statistics

Endpoint	Requests	Good	Bad
Graph Store Protocol (Read) (get)	0	0	0
Graph Store Protocol (data)	0	0	0
Graph Store Protocol	0	0	0
SPARQL Query	0	0	0
SPARQL Query (sparql)	0	0	0
SPARQL Query (query)	0	0	0
SPARQL Update	0	0	0
SPARQL Update (update)	0	0	0
Overall	0	0	0

Conclusion -

Thus we analyzed our site using plausible and we can use this data to improve the site. We also used apache jena.