**EXPERIMENT 1**

**AIM:** Study web analytics using open source tools like Matomo, Open Web Analytics,

AWStats, Countly, Plausible.

**THEORY:**

**Web Analytics**

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.

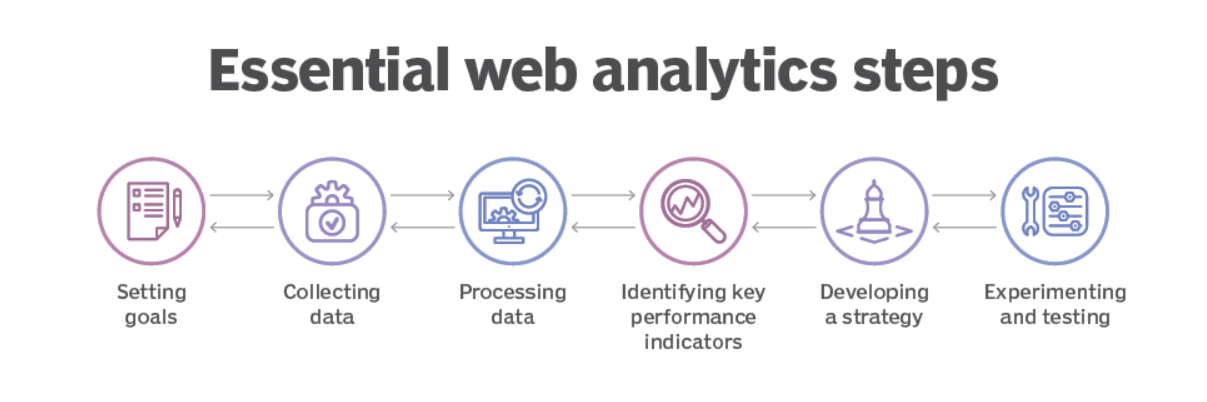
Data collected through web analytics may include traffic sources, referring sites, page views, paths taken and conversion rates. The compiled data often forms a part of customer relationship management analytics (CRM analytics) to facilitate and streamline better business decisions.

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

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

**Web Analytics Process**

The web analytics process involves the following steps:



1. Setting goals:

The first step in the web analytics process is for businesses to determine goals and the end results they are trying to achieve. These goals can include increased sales, customer satisfaction and brand awareness. Business goals can be both quantitative and qualitative.

1. Collecting data:

The second step in web analytics is the collection and storage of data. Businesses can collect data directly from a website or web analytics tool, such as Google Analytics. The data mainly comes from Hypertext Transfer Protocol requests -- including data at the network and application levels -- and can be combined with external data to interpret web usage. For example, a user's Internet Protocol address is typically associated with many factors, including geographic location and clickthrough rates.

1. Processing data:

The next stage of the web analytics funnel involves businesses processing the collected data into actionable information.

1. Identifying key performance indicators (KPIs):

In web analytics, a KPI is a quantifiable measure to monitor and analyze user behavior on a website. Examples include bounce rates, unique users, user sessions and on-site search queries.

1. Developing a strategy:

This stage involves implementing insights to formulate strategies that align with an organization's goals. For example, search queries conducted on-site can help an organization develop a content strategy based on what users are searching for on its website.

1. Experimenting and testing:

Businesses need to experiment with different strategies in order to find the one that yields the best results. For example, A/B testing is a simple strategy to help learn how an audience responds to different content. The process involves creating two or more versions of content and then displaying it to different audience segments to reveal which version of the content performs better.

**Importance of web analytics**

* Refine your marketing campaigns
* Understand your website visitors
* Analyze website conversions
* Improve the website user experience
* Boost your search engine ranking
* Understand and optimize referral sources
* Boost online sales

**Web Analytics Tools**

Web analytics tools are software platforms that help website owners and marketers track, analyze, and interpret data about website visitors and their behavior. Web analytics tools provide insights into website traffic, user behavior, and other important metrics that can help website owners optimize their website and marketing efforts.

Examples of web analytics tools include:

* Google Analytics: This is one of the most popular web analytics tools and is available for free. It provides insights into website traffic, user behavior, and conversions.
* Adobe Analytics: This is a paid web analytics tool that provides a comprehensive view of website performance, user behavior, and customer journeys.
* Mixpanel: This is a web analytics tool that focuses on user behavior and provides insights into user engagement, retention, and conversion.
* Kissmetrics: This is another web analytics tool that focuses on user behavior and provides insights into user engagement, retention, and conversion.
* Piwik PRO: This is a web analytics tool that provides insights into website performance, user behavior, and conversions, with a focus on data privacy and security.
* Matomo: This is a free and open-source web analytics tool that provides insights into website traffic, user behavior, and conversions, with a focus on data privacy.

**Semantic web**

Semantic Web is a vision for the future of the World Wide Web where information is organized in a way that is machine-readable and meaningful to computers. The idea is to enable computers to understand the content of Web pages and provide more accurate and relevant results to users. The Semantic Web builds on existing Web technologies, such as HTML, HTTP, and URI, but adds a layer of metadata that describes the meaning of the content.

The Semantic Web is based on the Resource Description Framework (RDF), a flexible and extensible data model for representing information in the Web. RDF allows for the creation of structured data that can be easily shared and reused across different applications and domains. The Semantic Web also includes ontology languages, such as OWL and RDF Schema, which provide a way to define the relationships between concepts and entities in a domain.

The benefits of the Semantic Web include improved search results, more accurate and relevant recommendations, and better integration of data from different sources. However, the Semantic Web is still a work in progress, and its full potential has yet to be realized.

**Standards of Semantic Web**

The standards of Semantic Web provide a framework for representing and exchanging information on the World Wide Web in a standardized, machine-readable format. These standards include:

1. Resource Description Framework (RDF): RDF is a standard model for data interchange on the Web. It provides a way to describe resources and their relationships, and enables data to be shared across different applications and domains.
2. RDF Schema (RDFS): RDFS is an ontology language that provides a way to define vocabularies and classes of resources, and their relationships.
3. Web Ontology Language (OWL): OWL is an ontology language that provides a way to define more complex relationships and rules, and supports automated reasoning and inference.
4. SPARQL Protocol and RDF Query Language (SPARQL): SPARQL is a query language that provides a way to query RDF data and retrieve information from it.
5. Uniform Resource Identifier (URI): URI is a standard way of identifying resources on the Web, and is used to reference and link to resources.

The learning curve for using Semantic Web technologies is perceived to be steep because few educational resources currently exist for users new to the concepts, and still fewer resources can be found that discuss when and how to apply the technologies to real world scenarios.

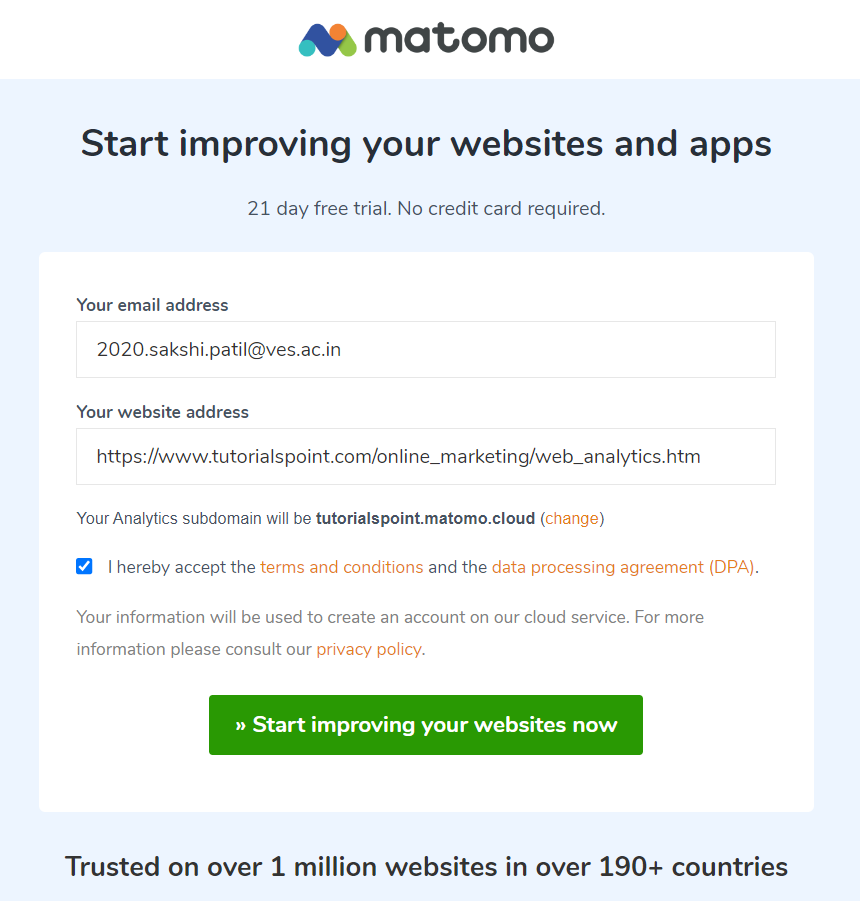
**STUDY:**

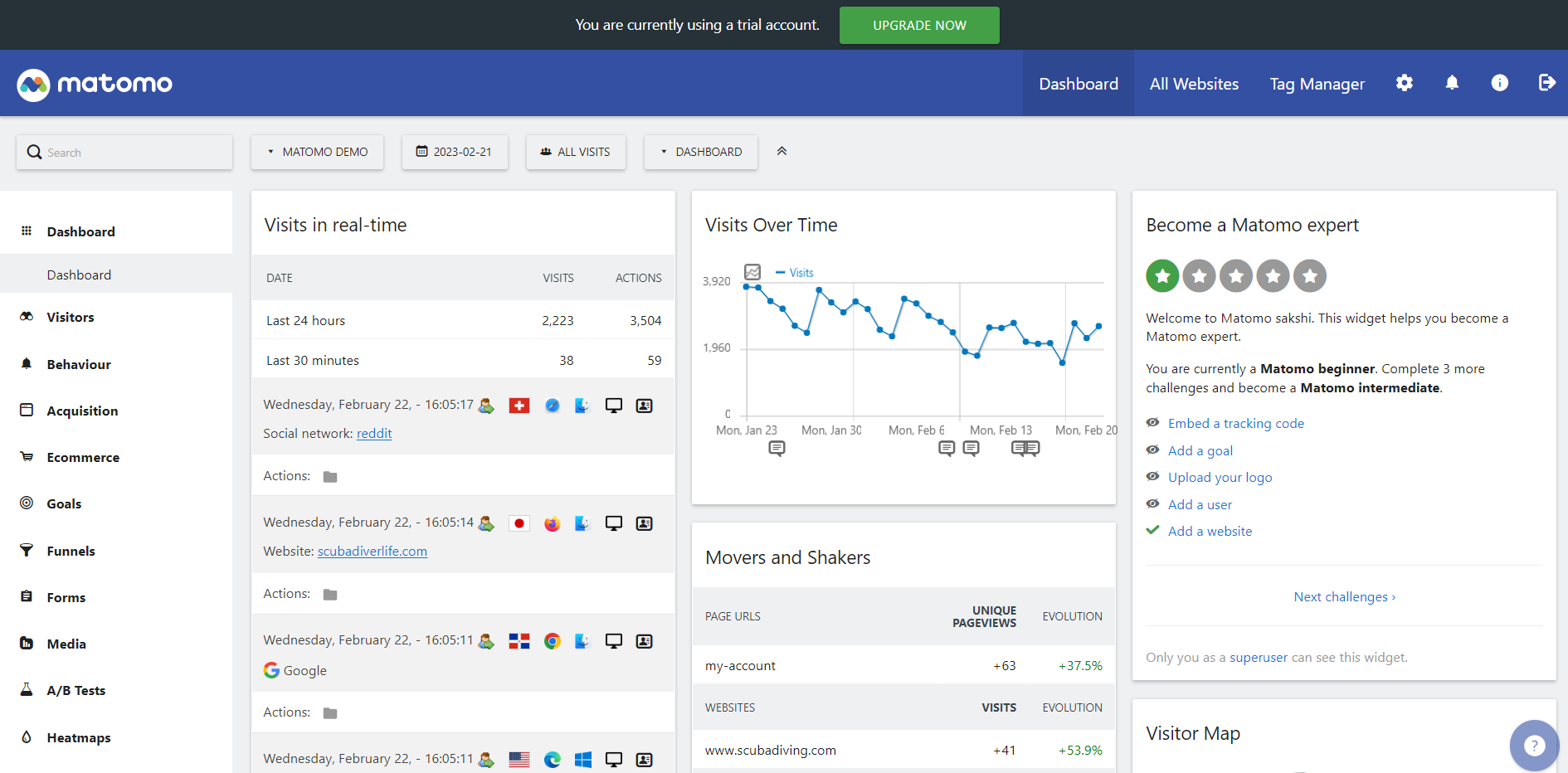
**Matomo**

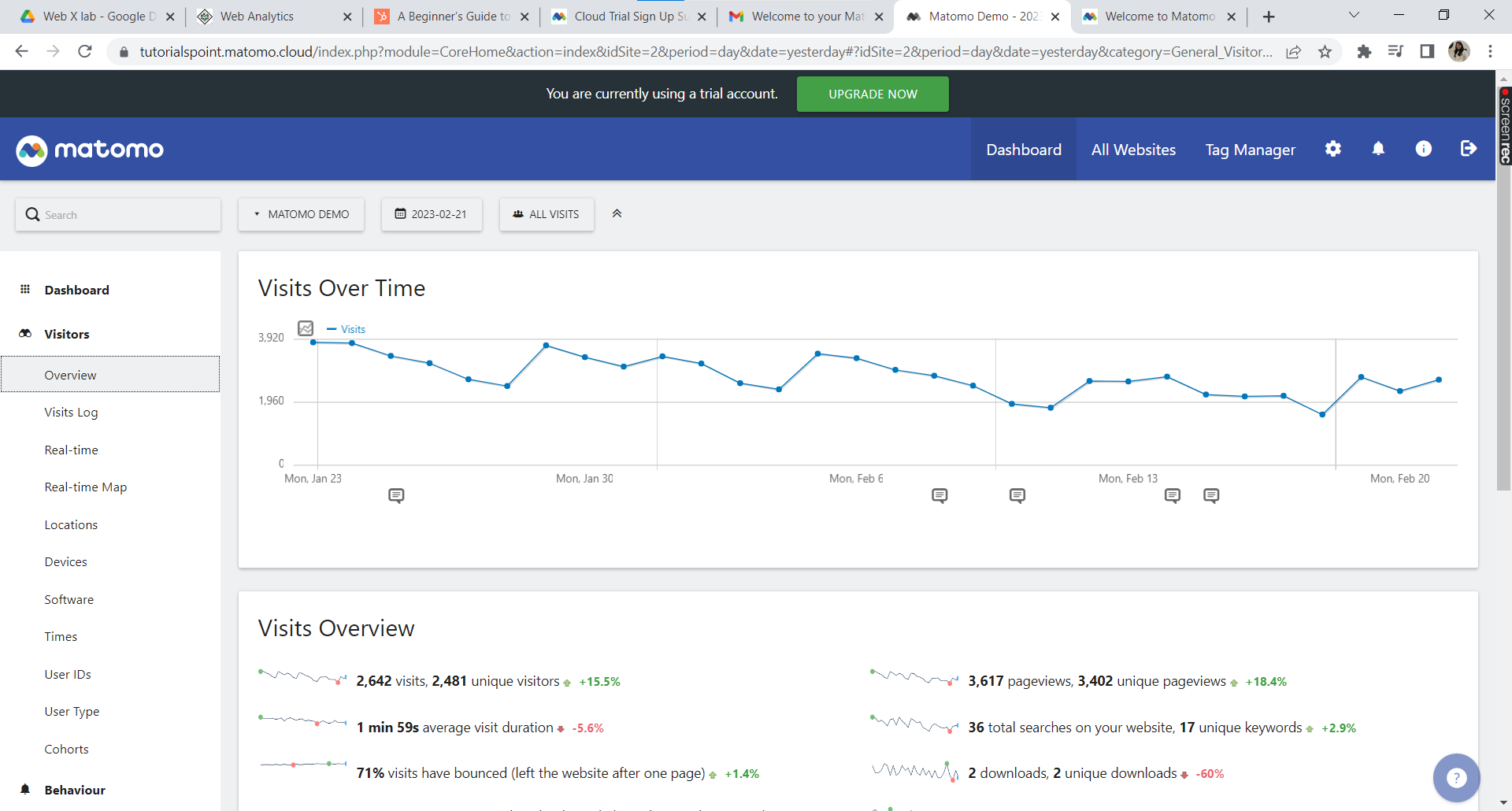
Matomo is an open-source web analytics tool that provides insights into website traffic, user behavior, and conversions. It is based on the principles of data privacy, security, and transparency. Matomo allows website owners to track and analyze their website traffic and user behavior, including metrics such as pageviews, bounce rate, time on site, and conversion rates. It also provides features such as real-time tracking, custom reporting, and goal tracking, which allow website owners to measure the effectiveness of their marketing campaigns and optimize their website accordingly.

One of the key features of Matomo is its emphasis on data privacy and security. Matomo allows website owners to keep their data on their own servers, rather than relying on third-party data collection and analysis. This means that website owners have complete control over their data and can ensure that it is protected from unauthorized access.

Matomo also provides a range of tools for data visualization, reporting, and segmentation, allowing website owners to drill down into specific aspects of website performance and user behavior. By understanding how visitors interact with their website, website owners can optimize their website to improve user experience, increase engagement, and ultimately drive more conversions.







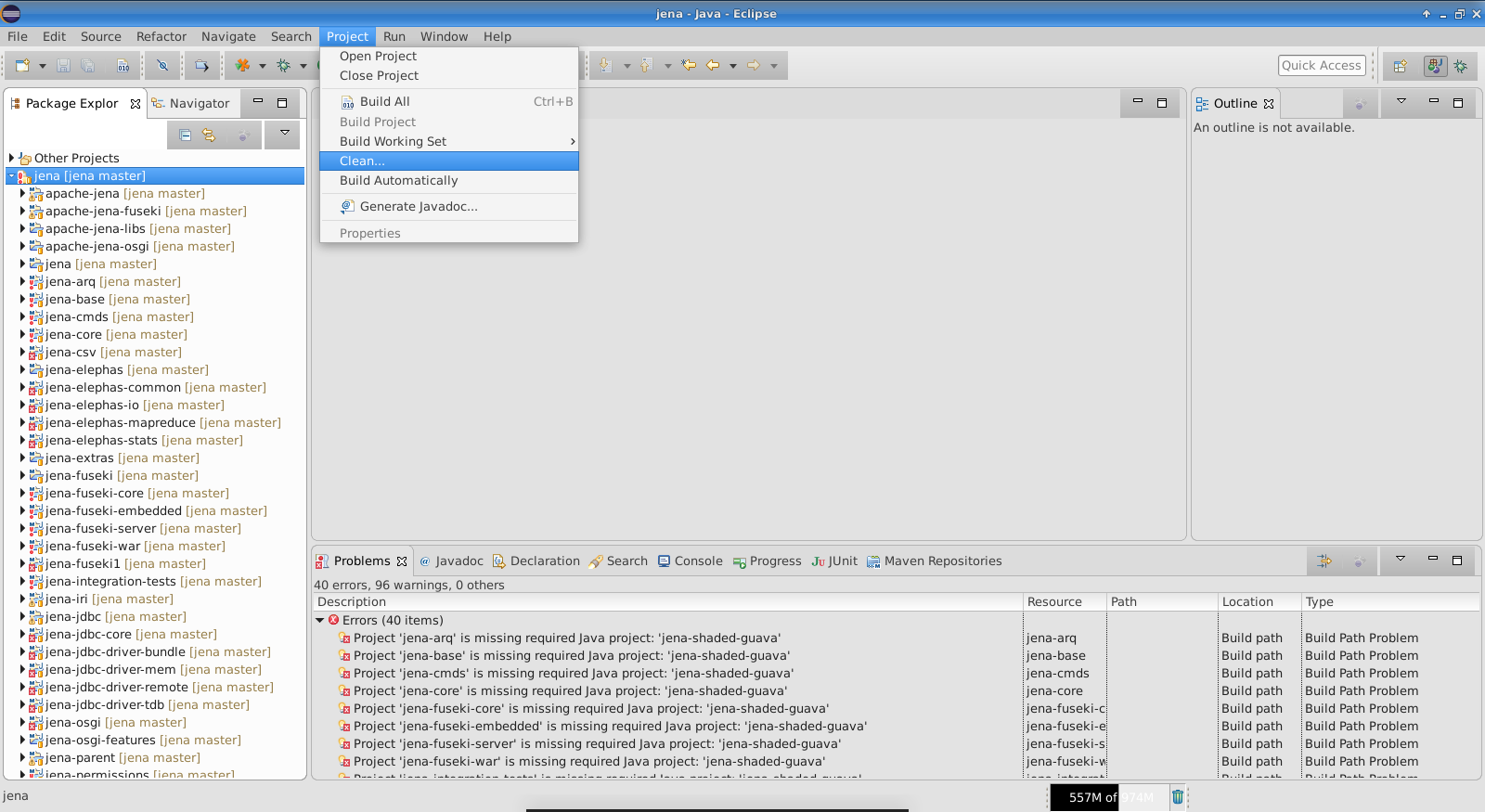
**Apache Jena**

Apache Jena is not a traditional web analytics tool, it can be used for analytics on semantic web and linked data. Apache Jena provides a range of tools for parsing, querying, and reasoning over RDF (Resource Description Framework) data, which can be used to analyze data in a semantic web context.

For example, Apache Jena can be used to analyze data in RDF format to gain insights into relationships between entities and properties, and to identify patterns and trends in the data. This could be useful in various domains, such as healthcare, finance, or e-commerce, where there is a need to analyze large amounts of structured and unstructured data.

Apache Jena can also be used for data integration and aggregation, allowing data from multiple sources to be combined and analyzed in a semantic web context. This could be useful in applications such as data warehousing, where there is a need to integrate data from multiple sources to provide a comprehensive view of the data.

Overall, while Apache Jena is not a traditional web analytics tool, it can be used for analytics on semantic web and linked data, providing a range of tools for parsing, querying, and reasoning over RDF data, and for data integration and aggregation.



**CONCLUSION:**

We have successfully studied the Matomo tool & Apache Jena tool for web analytics &

Semantic Web respectively