**Website Traffic Analysis**

**OBJECTIVE:**

Website traffic analysis is a crucial aspect of understanding and improving the performance of a website. It involves collecting and analyzing data related to user interactions, page views, and other relevant metrics to gain insights into how visitors engage with the site. This information can be used to optimize content, user experience, and marketing strategies.The problem at hand is to assess and improve the website traffic analysis process for a specific website or a set of websites. The primary objective is to gather meaningful data, derive actionable insights, and make data-driven decisions to enhance the overall online presence and user experience.Define what data needs to be collected. This includes web traffic data such as page views, unique visitors, session duration, referral sources, and more.Determine the tools and technologies used for data collection, such as Google Analytics, Adobe Analytics, or custom tracking solutions.Specify how the collected data will be processed, cleaned, and prepared for analysis.Identify any data sources or integrations necessary for a comprehensive analysis.Analysis Goals:Define clear objectives for the analysis. What insights are you trying to gain from the website traffic data?Set measurable KPIs (Key Performance Indicators) to track progress and success.Audience Segmentation:Determine how to segment website visitors based on relevant criteria, such as demographics, behavior, or traffic sources.Identify key user personas and their distinct characteristics.Content Assessment:Evaluate the performance of website content, including pages, articles, images, videos, etc.Identify high-performing and low-performing content.User Experience (UX) Analysis:Assess the user journey on the website, including navigation, load times, and user interactions.Identify pain points and areas for improvement in the UX.Traffic Source Analysis:Analyze the sources of website traffic, including organic search, social media, referral sites, and direct traffic.Determine which sources are driving the most valuable traffic.Conversion Analysis:Evaluate conversion goals, such as sign-ups, purchases, or downloads.Identify bottlenecks in the conversion funnel and strategies to optimize them.Competitor Benchmarking:Compare website traffic and performance metrics with key competitors.Identify opportunities to gain a competitive edge.Reporting and Visualization:Determine how the analysis results will be presented and shared with stakeholders.Choose appropriate visualization tools and reporting formats.Actionable Insights:Translate data into actionable recommendations and strategies for improving website performance.Prioritize recommendations based on their potential impact.Continuous Improvement:Establish a process for ongoing website traffic analysis and optimization.Ensure that the analysis is integrated into the website management workflow.

**DESIGN THINKING:**

Data Collection:

The Dataset contain the many columns

Data Preprocessing:

Data Cleaning: Clean the data to remove duplicates, missing values, and outliers.

Feature Engineering:

The feature engineering means, how to select the features for the build the model like chi-square, f-test ,etc..

Model Selection:

The product demand prediction project build by the ML models like supervised and unsupervised algorithms. The algorithms are LOGISTIC REGRESSION, LINEAR REGRESSION, XGBoost ,SVM, RANDOM FOREST , and, etc…

Model Training:

The dataset is split into training and testing set for the model evaluation . the training set size is 80% and testing size 20%.

Model Evaluation:

The all models measure the accuracy score, f1- score, pression and recall and etc.. the high accurate model select the testing data.

**PRE-PROCESSING:**

**STEPS:**

1. **DATA CLEANING**
2. **HANDLE MISSING VALUES**
3. **CATEGORICAL TO NUMERICAL REPRESENTATIONS.**

**DATA CLEANING:**

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled.

**HANDLE MISSING VALUES:**

1. Deleting Rows with missing values
2. Impute missing values for continuous variable
3. Impute missing values for categorical variable
4. Other Imputation Methods
5. Using Algorithms that support missing values
6. Prediction of missing values
7. Imputation using Deep Learning Library — Datawig

**CATEGORICAL TO NUMERICAL REPRESENTATIONS:**

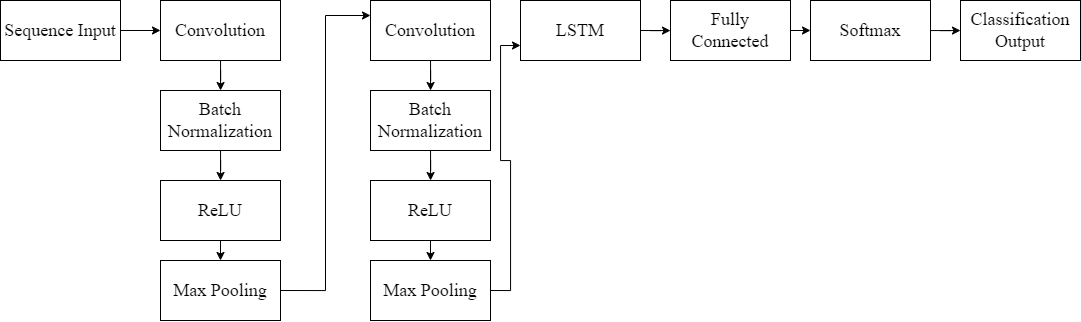
1. cat.codes Attribute
2. replace
3. Label Encoder

**ALGORITHM:**

1. DEEP LEARNING TECHNIQUE : CNN-LSTM
2. ATTENTION MECHANISMS

**CNN-LSTM:**

 CNN-LSTM network use convolutional and LSTM layers to learn from the training data. To train a CNN-LSTM network with audio data, you extract auditory-based spectrograms from the raw audio data and then train the network using the spectrograms.

****

**ATTENTION MECHANISM:**

An attention mechanism is an Encoder-Decoder kind of neural network architecture that allows the model to focus on specific sections of the input while executing a task. It dynamically assigns weights to different elements in the input, indicating their relative importance or relevance.

**PROJECT WORKFLOW:**

**Hits Management**

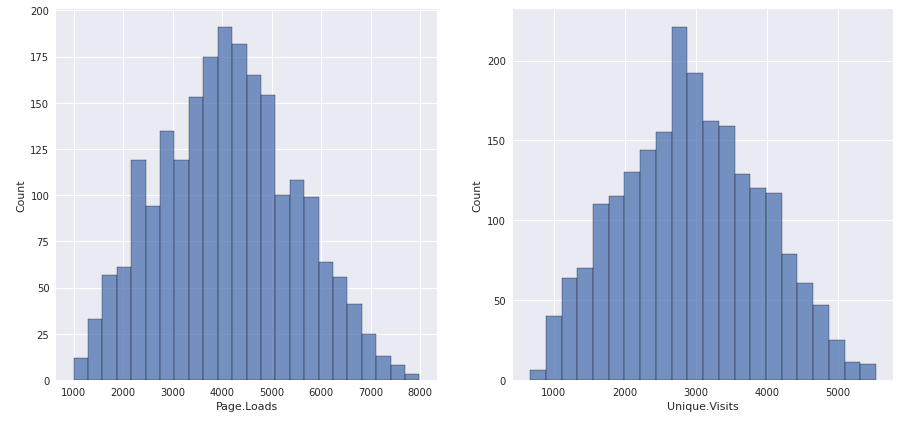
**Login Management**

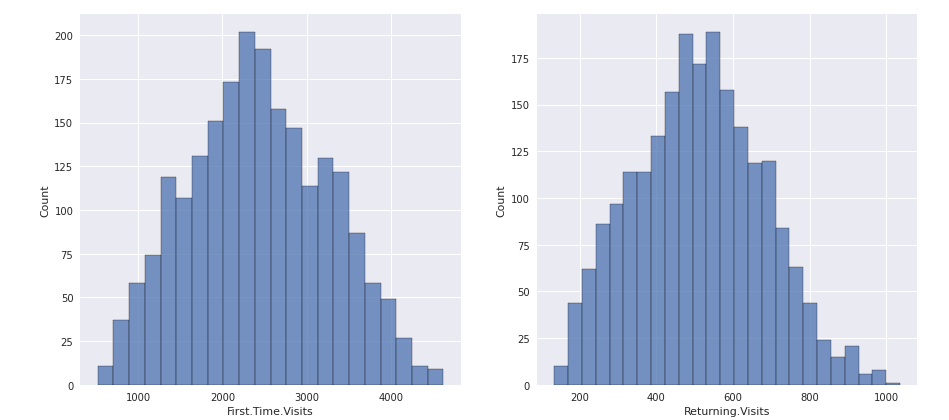
**Weekly Stat Page Management**

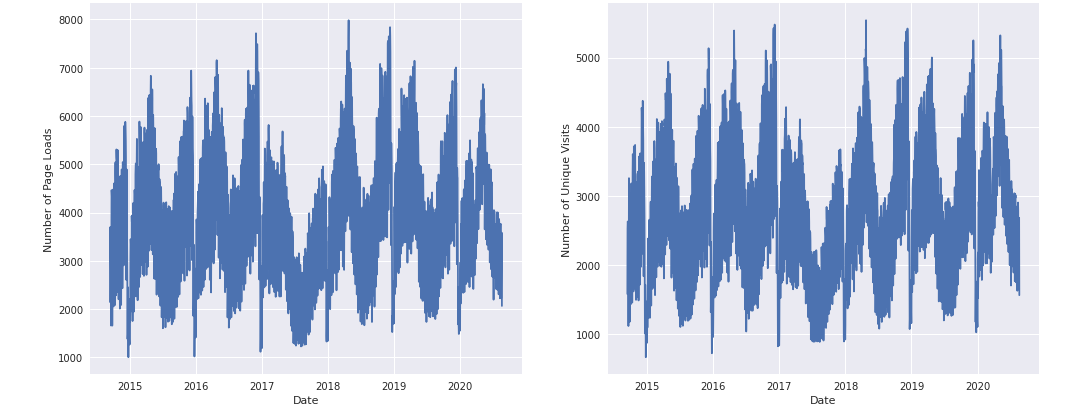
**System** **user management**

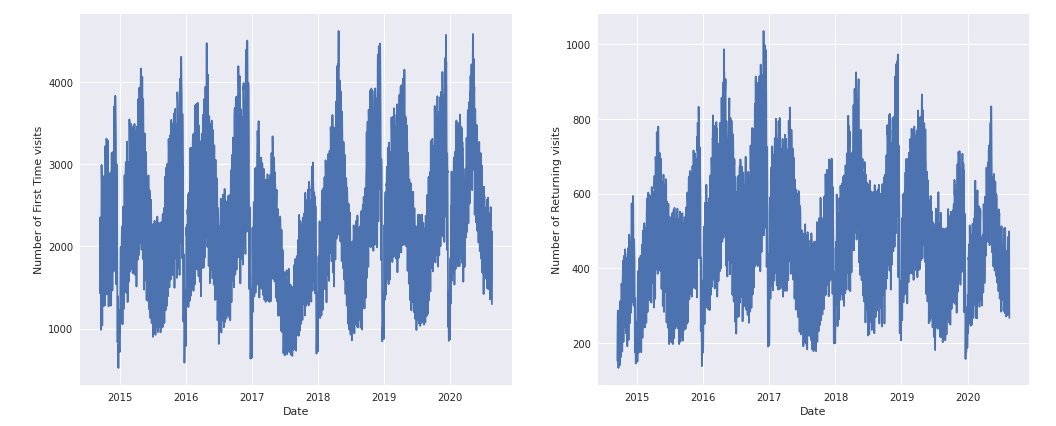
**Hosts Stat Page Management**

**Traffic Management**

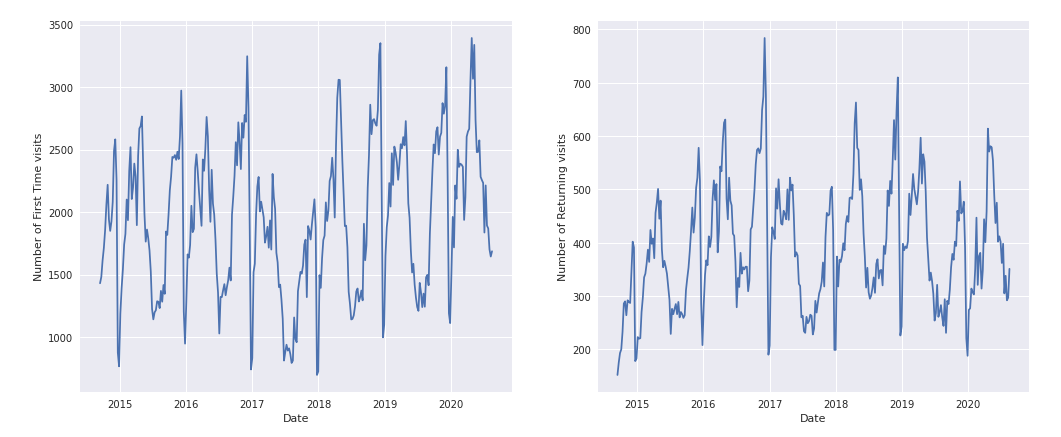


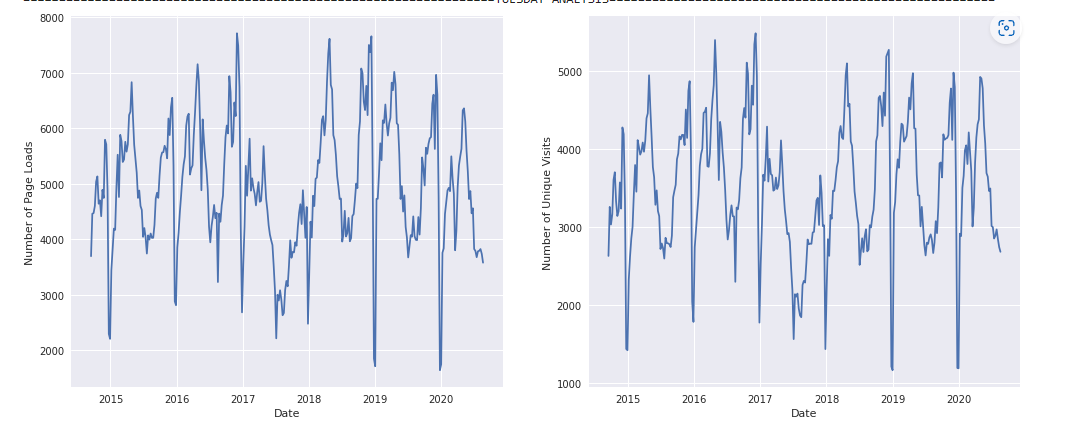


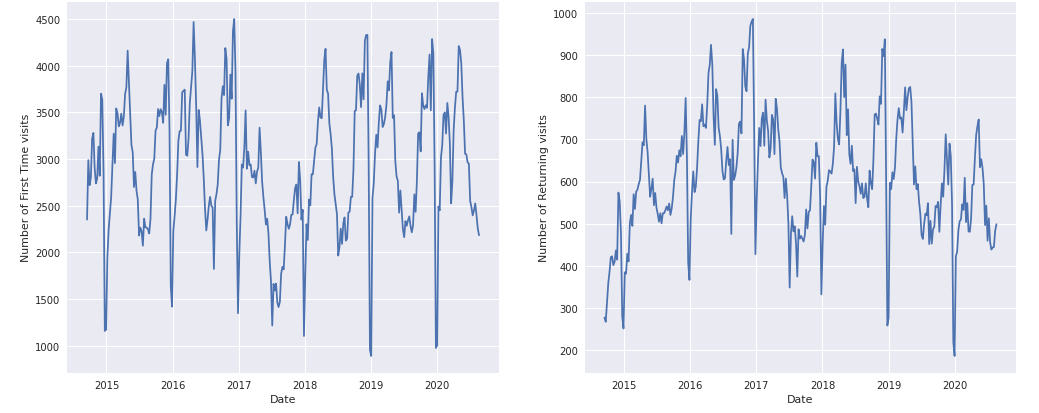


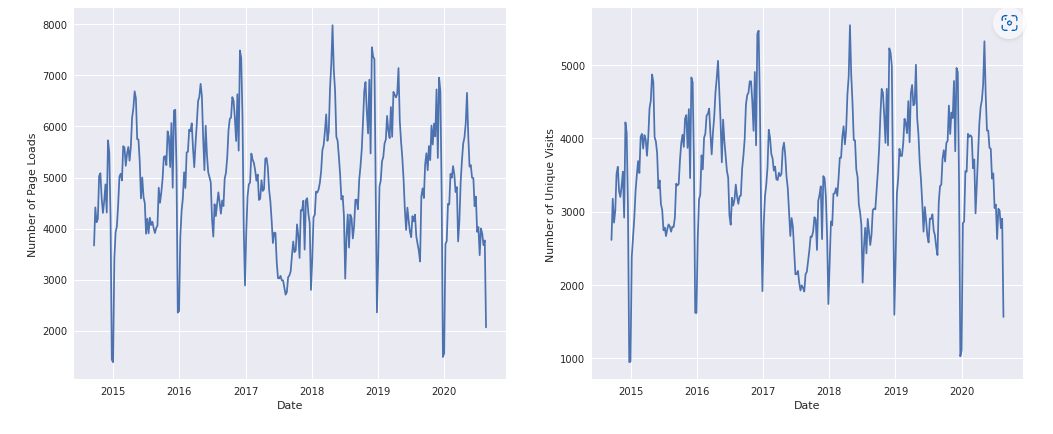


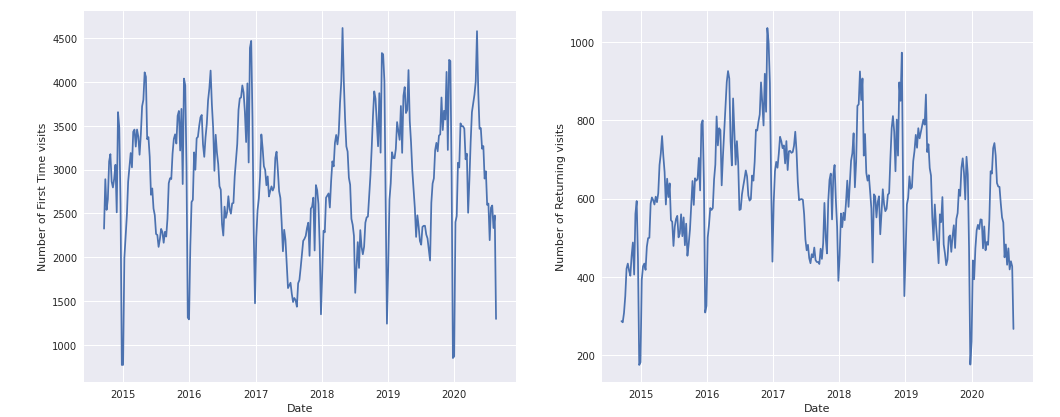


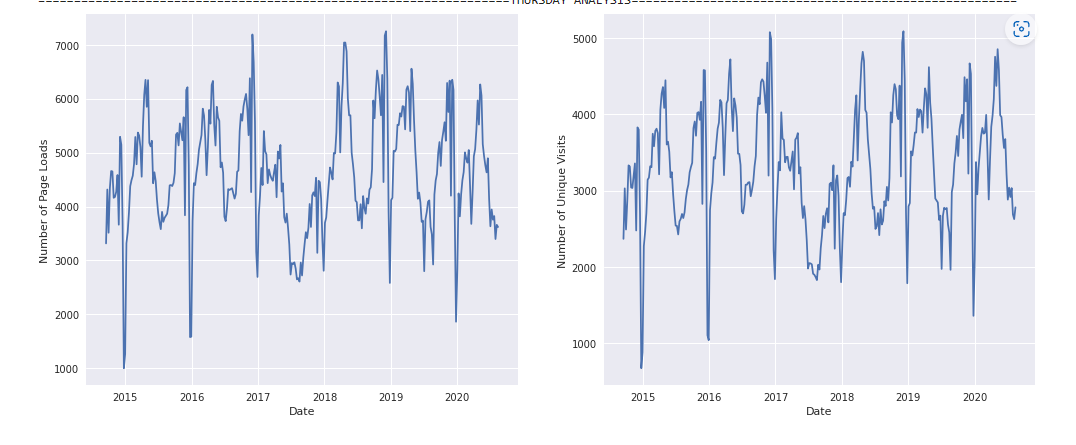


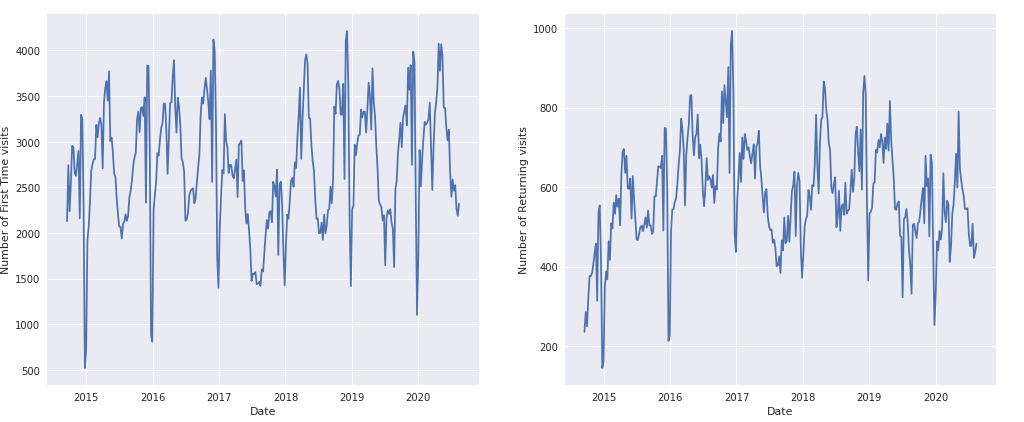


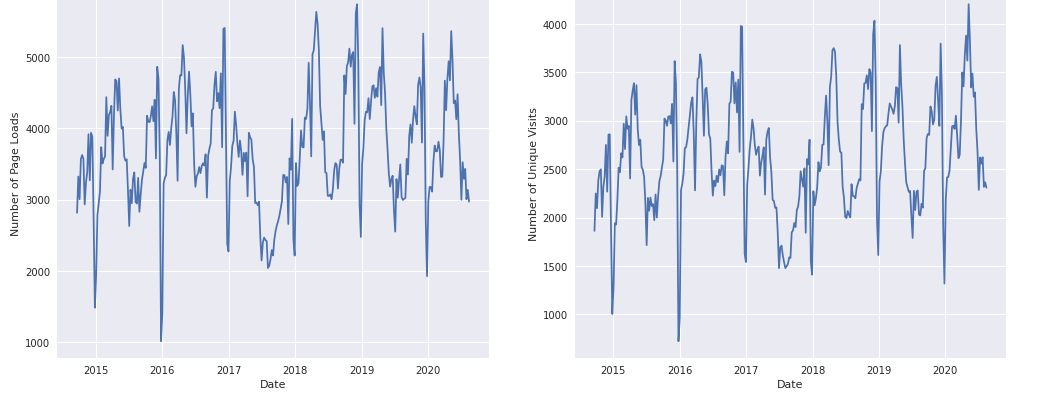


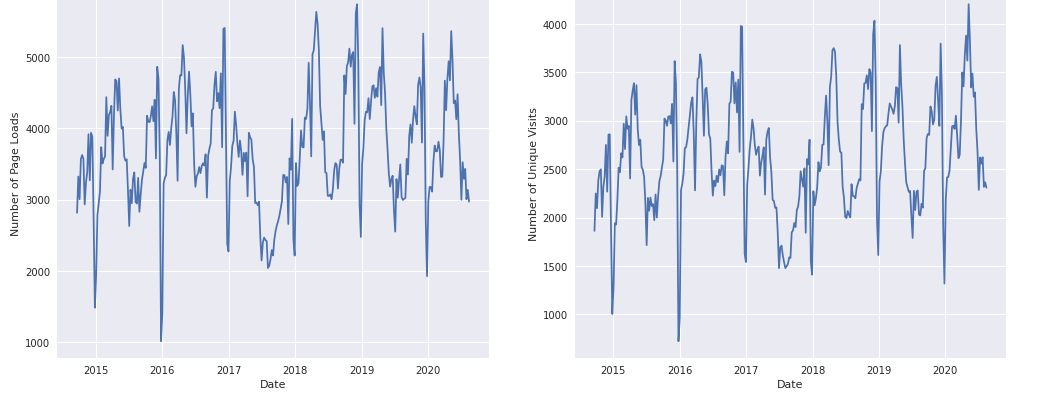


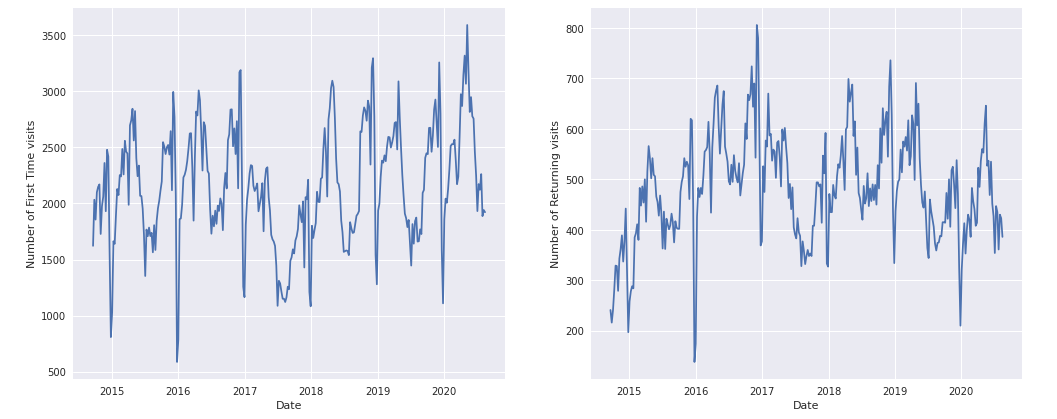


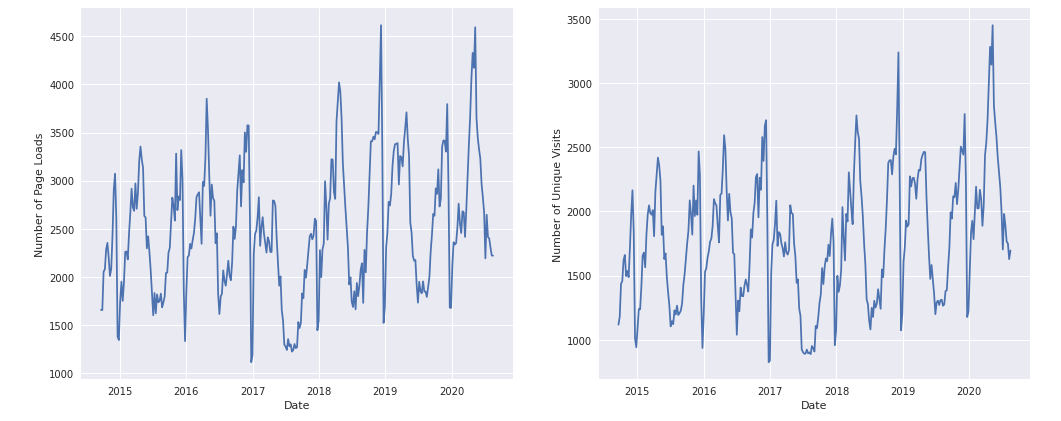


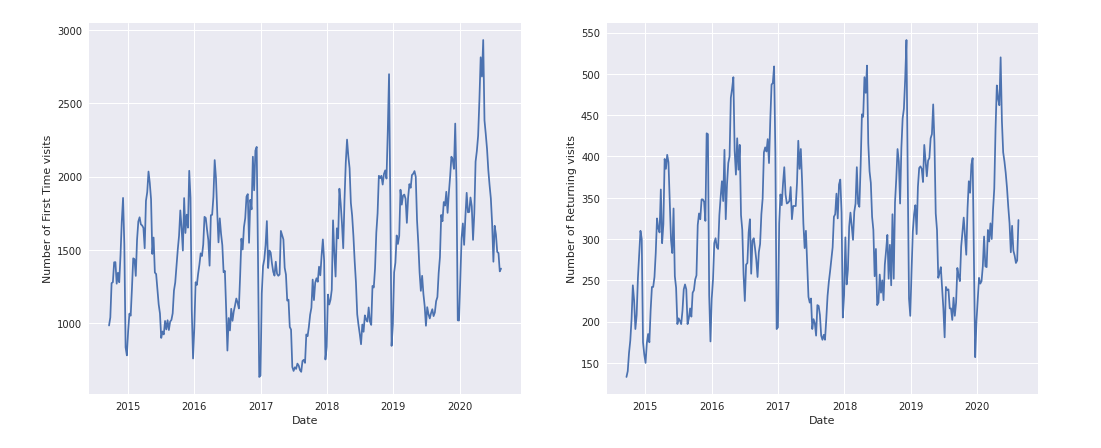












REFERENCE:

* 1. *Jeffay, Kevin.*[*"Tracking the Evolution of Web Traffic: 1995-2003\*"*](http://www.cs.unc.edu/~jeffay/papers/MASCOTS-03a.pdf)*(PDF). UNC DiRT Group's Publications. University of North Carolina at Chapel Hill.*[*Archived*](https://web.archive.org/web/20120513161722/http:/www.cs.unc.edu/~jeffay/papers/MASCOTS-03a.pdf)*(PDF) from the original on 2012-05-13. Retrieved 2012-02-20.*
  2. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-2) *Miller, Rich (2004-10-26).*[*"Bush Campaign Web Site Rejects Non-US Visitors"*](http://news.netcraft.com/archives/2004/10/26/bush_campaign_web_site_rejects_nonus_visitors.html)*.*[*Archived*](https://web.archive.org/web/20110219131545/http:/news.netcraft.com/archives/2004/10/26/bush_campaign_web_site_rejects_nonus_visitors.html)*from the original on 2011-02-19. Retrieved 2004-10-28.*
  3. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-3) [*"Denial of Service"*](http://www.cert.org/tech_tips/denial_of_service.html)*. Cert.org.*[*Archived*](https://web.archive.org/web/20120607061124/http:/www.cert.org/tech_tips/denial_of_service.html)*from the original on 7 June 2012. Retrieved 28 May 2012.*
  4. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-4) *Vranica, Suzanne (23 March 2014).*[*"A 'Crisis' in Online Ads: One-Third of Traffic Is Bogus"*](https://www.wsj.com/articles/the-secret-about-online-ad-traffic-one-third-is-bogus-1395684863)*. Wall Street Journal.*[*Archived*](https://web.archive.org/web/20170916182318/https:/www.wsj.com/articles/the-secret-about-online-ad-traffic-one-third-is-bogus-1395684863)*from the original on 16 September 2017. Retrieved 3 May 2017.*
  5. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-5) [*"36% Of All Web Traffic Is Fake"*](http://www.businessinsider.com/bots-and-fake-traffic-2014-3)*. Business Insider.*[*Archived*](https://web.archive.org/web/20170407161639/http:/www.businessinsider.com/bots-and-fake-traffic-2014-3)*from the original on 7 April 2017. Retrieved 3 May 2017.*
  6. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-6) [*"We're Halfway to Encrypting the Entire Web"*](https://www.eff.org/deeplinks/2017/02/were-halfway-encrypting-entire-web)*. Electronic Frontier Foundation. 21 February 2017.*[*Archived*](https://web.archive.org/web/20210331093749/https:/www.eff.org/deeplinks/2017/02/were-halfway-encrypting-entire-web)*from the original on 31 March 2021. Retrieved 3 May 2017.*
  7. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-7) *Finley, Klint (31 January 2017).*[*"Half the Web Is Now Encrypted. That Makes Everyone Safer"*](https://www.wired.com/2017/01/half-web-now-encrypted-makes-everyone-safer/)*. WIRED.*[*Archived*](https://web.archive.org/web/20210303145936/https:/www.wired.com/2017/01/half-web-now-encrypted-makes-everyone-safer/)*from the original on 3 March 2021. Retrieved 1 May 2017.*
  8. [**^**](https://en.wikipedia.org/wiki/Web_traffic#cite_ref-8) [*"What is Hypertext Transfer Protocol Secure (HTTPS)?"*](https://www.techtarget.com/searchsoftwarequality/definition/HTTPS)*. SearchSoftwareQuality.*[*Archived*](https://web.archive.org/web/20220808115342/https:/www.techtarget.com/searchsoftwarequality/definition/HTTPS)*from the original on 2022-08-08. Retrieved 2022-08-08.*