WEBSITE TRAFFIC ANALYSIS

Abstract:

The "Website Traffic Analysis Project" is a comprehensive endeavor aimed at understanding, evaluating, and optimizing web traffic to enhance the online presence and performance of a website. In today's digital age, web traffic is a critical metric for businesses and organizations seeking to engage with their target audience, increase conversions, and make data-driven decisions.

This project employs a multifaceted approach to accomplish its objectives. It begins with data collection using advanced web analytics tools, tracking user interactions, page views, referral sources, and various other relevant metrics. Through this data, the project seeks to uncover valuable insights into visitor behavior, preferences, and demographic information, enabling a deeper understanding of the website's audience.

Objective

The primary objectives of the "Website Traffic Analysis Project" are multifaceted. First and foremost, the project seeks to comprehensively understand the website's audience by delving into their demographics, preferences, and online behavior. This understanding will serve as the foundation for tailoring content and strategies to meet the specific needs of different user segments. The analysis also focuses on assessing the effectiveness of various traffic sources, such as search engines, social media, and referrals, in order to allocate resources more efficiently and improve each source's performance. Furthermore, the project aims to evaluate the popularity and engagement levels of different website content and pages, thereby guiding content creation and optimization efforts.

Another critical objective is to analyze the website's conversion funnels, identifying areas where visitors drop off or successfully convert. This analysis is vital for enhancing the user journey and increasing conversion rates. The project will also segment website visitors into distinct groups or personas based on their behavior and characteristics, enabling the development of personalized marketing and content strategies. In terms of technical aspects, the project will address search engine optimization (SEO) opportunities to improve the website's visibility in search results and attract more organic traffic. Additionally, the project will evaluate the user experience by examining factors such as page load times, mobile-friendliness, and navigation to ensure a seamless and engaging experience for visitors.

Importantly, this project promotes data-driven decision-making, harnessing insights to make informed choices regarding website improvements, content creation, and marketing strategies. It establishes a framework for continuous monitoring and analysis, allowing for adaptation to evolving digital trends and user behaviors. Finally, the definition of key performance metrics and goals will enable the ongoing measurement of success and performance tracking, ensuring that the website consistently advances toward its objectives in terms of traffic, engagement, and conversions.

CODE

import pandas as pd

import matplotlib.pyplot as plt

import plotly.express as px

import plotly.graph\_objects as go

from statsmodels.tsa.seasonal import seasonal\_decompose

from statsmodels.graphics.tsaplots import plot\_pacf

from statsmodels.tsa.arima\_model import ARIMA

import statsmodels.api as sm

data = pd.read\_csv("Thecleverprogrammer.csv")

print(data.head())

**Date Views**

**0 01/06/2021 7831**

**1 02/06/2021 7798**

**2 03/06/2021 7401**

**3 04/06/2021 7054**

**4 05/06/2021 7973**

data["Date"] = pd.to\_datetime(data["Date"],

format="%d/%m/%Y")

print(data.info())

**<class 'pandas.core.frame.DataFrame'>**

**RangeIndex: 391 entries, 0 to 390**

**Data columns (total 2 columns):**

**# Column Non-Null Count Dtype**

**--- ------ -------------- -----**

**0 Date 391 non-null datetime64[ns]**

**1 Views 391 non-null int64**

**dtypes: datetime64[ns](1), int64(1)**

**memory usage: 6.2 KB**

**None**

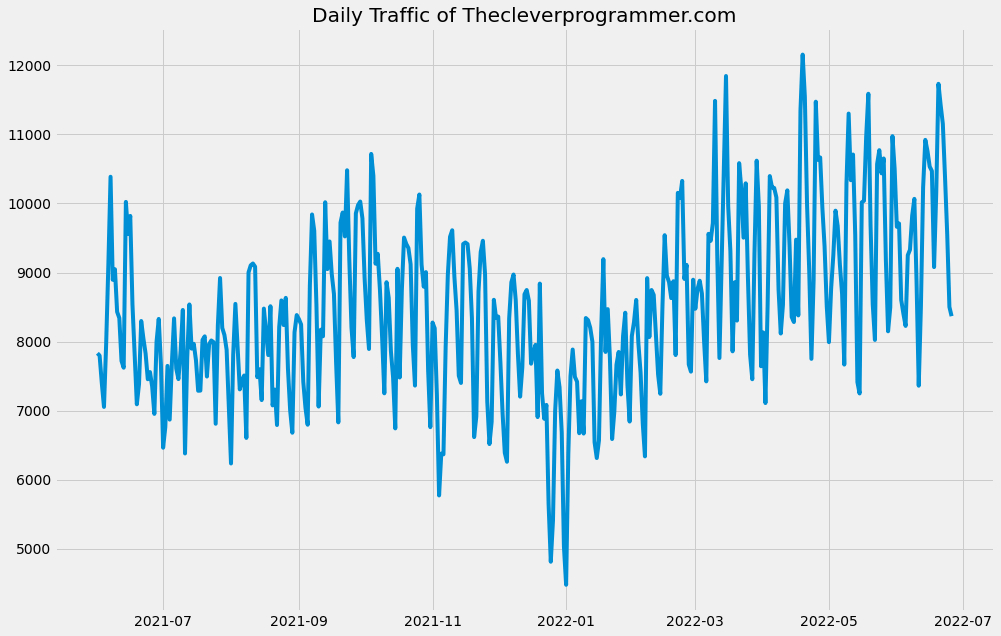
plt.style.use('fivethirtyeight')

plt.figure(figsize=(15, 10))

plt.plot(data["Date"], data["Views"])

plt.title("Daily Traffic of Thecleverprogrammer.com")

plt.show()



result = seasonal\_decompose(data["Views"],

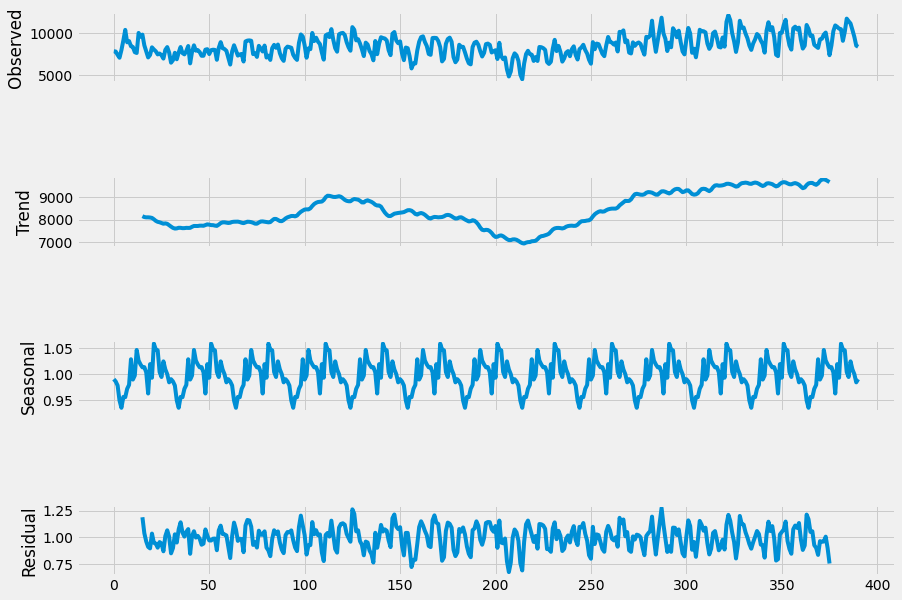
model='multiplicative',

freq = 30)

fig = plt.figure()

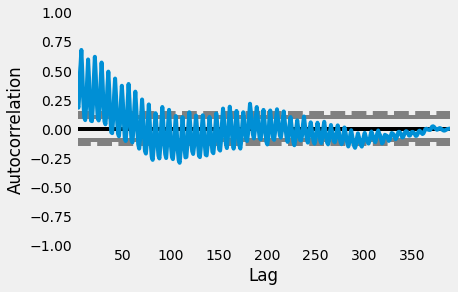
fig = result.plot()

fig.set\_size\_inches(15, 10)



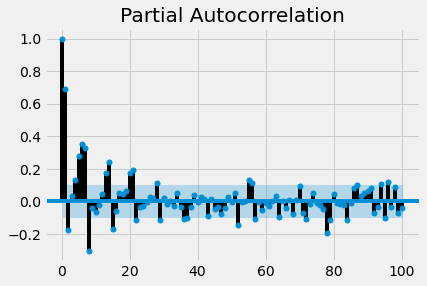
1

pd.plotting.autocorrelation\_plot(data["Views"])

**p = 5**

1

plot\_pacf(data["Views"], lags = 100)

**q = 2**

1

p, d, q = 5, 1, 2

2

model=sm.tsa.statespace.SARIMAX(data['Views'],

3

order=(p, d, q),

4

seasonal\_order=(p, d, q, 12))

5

model=model.fit()

6

print(model.summary())

**Statespace Model Results**

**==========================================================================================**

**Dep. Variable: Views No. Observations: 391**

**Model: SARIMAX(5, 1, 2)x(5, 1, 2, 12) Log Likelihood -3099.402**

**Date: Tue, 28 Jun 2022 AIC 6228.803**

**Time: 07:01:10 BIC 6287.827**

**Sample: 0 HQIC 6252.229**

**- 391**

**Covariance Type: opg**

**==============================================================================**

**coef std err z P>|z| [0.025 0.975]**

**------------------------------------------------------------------------------**

**ar.L1 0.7808 0.134 5.836 0.000 0.519 1.043**

**ar.L2 -0.7973 0.135 -5.920 0.000 -1.061 -0.533**

**ar.L3 -0.1442 0.170 -0.850 0.395 -0.477 0.188**

**ar.L4 -0.1833 0.151 -1.210 0.226 -0.480 0.114**

**ar.L5 -0.1548 0.139 -1.117 0.264 -0.426 0.117**

**ma.L1 -1.1826 0.094 -12.515 0.000 -1.368 -0.997**

**ma.L2 0.8856 0.078 11.304 0.000 0.732 1.039**

**ar.S.L12 -0.2606 4.608 -0.057 0.955 -9.293 8.772**

**ar.S.L24 0.0428 0.781 0.055 0.956 -1.488 1.573**

**ar.S.L36 -0.1880 0.246 -0.764 0.445 -0.670 0.294**

**ar.S.L48 -0.2151 0.959 -0.224 0.823 -2.095 1.664**

**ar.S.L60 0.0127 0.986 0.013 0.990 -1.920 1.946**

**ma.S.L12 -0.6902 4.611 -0.150 0.881 -9.728 8.348**

**ma.S.L24 -0.0994 3.637 -0.027 0.978 -7.228 7.029**

**sigma2 1.257e+06 1.59e+05 7.914 0.000 9.46e+05 1.57e+06**

**===================================================================================**

**Ljung-Box (Q): 102.98 Jarque-Bera (JB): 1.32**

**Prob(Q): 0.00 Prob(JB): 0.52**

**Heteroskedasticity (H): 1.03 Skew: 0.14**

**Prob(H) (two-sided): 0.85 Kurtosis: 3.01**

**===================================================================================**

1

predictions = model.predict(len(data), len(data)+50)

2

print(predictions)

**391 9874.390136**

**392 10786.957398**

**393 10757.445305**

**394 9863.890552**

**395 8765.031698**

**396 8212.310651**

**397 8929.181869**

**398 9685.809771**

**399 10270.622236**

**400 10625.904093**

**401 9854.870630**

**402 9362.193417**

**403 9040.021193**

**404 9081.558484**

**405 10538.993124**

**406 11003.816870**

**407 10897.859601**

**408 10083.291284**

**409 9445.806523**

**410 8629.901288**

**411 9184.420361**

**412 10392.770399**

**413 10593.941868**

**414 10788.128238**

**415 10263.101427**

**416 9449.467789**

**417 9040.226113**

**418 9168.972091**

**419 9887.094079**

**420 10218.658067**

**421 10715.657122**

**422 9899.224399**

**423 9541.622897**

**424 9065.810941**

**425 8825.335634**

**426 10137.936392**

**427 10839.866240**

**428 10905.862922**

**429 10411.640309**

**430 9451.211368**

**431 8698.339931**

**432 8725.534103**

**433 10060.678587**

**434 10506.263524**

**435 10842.515622**

**436 10485.387495**

**437 9335.244813**

**438 9175.122336**

**439 9357.034382**

**440 10295.910655**

**441 11162.934817**

**dtype: float64**

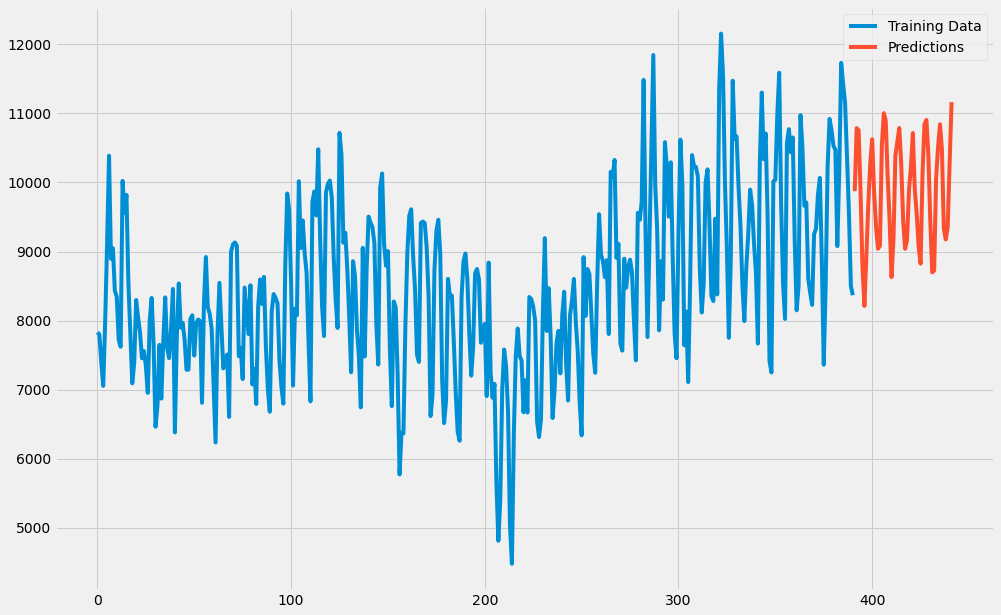
1

data["Views"].plot(legend=True, label="Training Data",

2

figsize=(15, 10))

3

predictions.plot(legend=True, label="Predictions")

Conclusion

In conclusion, the "Website Traffic Analysis Project" serves as a pivotal endeavor for website owners, businesses, and organizations operating in the digital landscape. Through the accomplishment of its multifaceted objectives, this project empowers stakeholders to make informed decisions, enhance their online presence, and ultimately achieve their digital goals.