الفصل الرابع

تحليل بيانات السلاسل الزمنية

Decomposing Time Series Data

محتويات الفصل الرابع

- مقدمة Introduction
- Time Series Decomposition تحليل السلاسل الزمنية
- أنواع نمذجة السلاسل الزمنية Time Series Modeling Types
 - و مقارنة بين خصائص النمذجة الجمعية و النمذجة الضربية
- أنواع الاتجاه و الموسمية Types of Trends and Seasonality
 - إنشاء الأنماط الأربعة لنمذجة السلاسل الزمنية
 - ملخص Summary

لتحميل ملحقات هذا الفصل من الأكواد البرمجية والشروحات التوضيحية زيارة المستودع التالي على صفحتي في موقع GitHub :

السلاسل الزمنية باللغة العربية

مقدمة Introduction

حتى الآن ، قد أخذنا نظرة عامة على تحليل السلاسل الزمنية ، وتعرفنا على أساسيات معالجة بياناتها باستخدام مكتبة Pandas كما تطرقنا إلى موضوعات جوهرية مثل التشويش الأبيض White Noise والسير العشوائي Random Walk.

في هذا الفصل سنطبق عمليًا تحليل السلاسل الزمنية إلى مكوناتها الأساسية الاتجاه Trend والموسمية Seasonality والبواقي Residuals وسنفهم لماذا ومتى يجب علينا القيام بذلك.

تحليل السلاسل الزمنية Time Series Decomposition

ما هو تحليل السلاسل الزمنية؟

ببساطة ، تحليل السلاسل الزمنية هي عملية تفكيك سلسلة زمنية إلى مكوناتها الأساسية التالية :

- الاتجاه Trend: هي الحركة العامة للسلسلة الزمنية بمرور الوقت
- الموسمية Seasonality: سلوكيات (أنماط) يتم التقاطها في فترات موسمية فردية.
- البواقي Residuals: كل شيء لا يتم التقاطه بواسطة مكونات الاتجاه والموسمية (الجزء المتبقي بعد فصل كل من الاتجاه والموسمية).

غالبًا ما تستخدم هذه التقنية لتحليل بيانات السلاسل الزمنية التاريخية. كما أنها تستخدم أحيانًا للتنبؤ. قد يكون عمل نماذج تنبؤية على الاتجاه والموسمية في وقت واحد مهمة صعبة للغاية ، لذلك قد يكون التعامل مع المكونات بشكل فردي أسلوبًا أفضل. في هذا الفصل سنركز فقط على التحليل و سنترك التنبؤ لفصل آخر !

أنواع نمذجة السلاسل الزمنية Time Series Modeling Types

سيتطلب تحليل السلاسل الزمنية تحديد نوع النمذجة والذي يخبر بايثون بكيفية فصل أو دمج المكونات الأساسية (الاتجاه ، الموسمية، و البواقي) لإنتاج السلسلة الزمنية الأصلية. ستتعرف على طريقتين :

1- نمذجة جمعية (جمع المكونات) Additive - ونمذجة ضربية (ضرب المكونات) Multiplicative.

1- السلسلة الزمنية الجمعية Additive Time Series : المكونات الفردية (الاتجاه ، الموسمية ، و البواقي) تجمع معًا لإنتاج السلسلة الزمنية

$$y_t = T_t + S_t + R_t$$

2- السلسلة الزمنية الضربية Multiplicative Time Series : المكونات الفردية (الاتجاه ، الموسمية ،والبواقي) يتم ضربها معًا لإنتاج السلسلة الزمنية

$$\mathbf{y_t} = \mathbf{T_t} * \mathbf{S_t} * \mathbf{R_t}$$

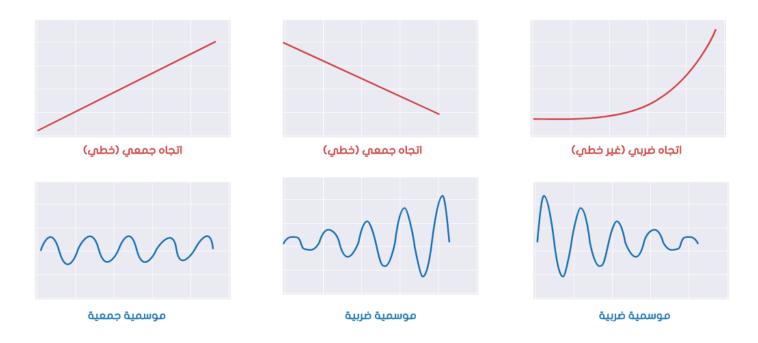
مقارنة بين خصائص النمذجة الجمعية و النمذجة الضربية

1- خصائص السلسلة الزمنية الجمعية Additive Time Series :

- يعتبر الاتجاه الجمعى Additive Trend دالة خطية ا
- تعتبر الموسمية الجمعية Additive Seasonality دالة تحمل نفس التردد (عرض الموجة) والسعة (الارتفاع) للدورات الموسمية.

2- خصائص السلسلة الزمنية الضربية Multiplicative Time Series

- · يعتبر الاتجاه الضربى دالة غير خطية non-linear (خط الاتجاه منحنى).
- · تشير الموسمية الضربية إلى تزايد / تناقص التردد (عرض الموجة) و / أو السعة (الارتفاع) للدورات الموسمية.



أنواع الاتجاه و الموسمية Types of Trends and Seasonality



فيما يلي سنقوم بتشكيل الأنماط الأربعة لنمذجة السلاسل الزمنية!

- اتجاه جمعی وموسمیة جمعیة Additive Trend & Additive Seasonality
- اتجاه جمعی وموسمیة ضربیة Additive Trend & Multiplicative Seasonality
- اتجاه ضربی وموسمیة جمعیة Multiplicative Trend & Additive Seasonality
- اتجاه ضربی وموسمیة ضربیة Multiplicative Seasonality فربی وموسمیة ضربیة •

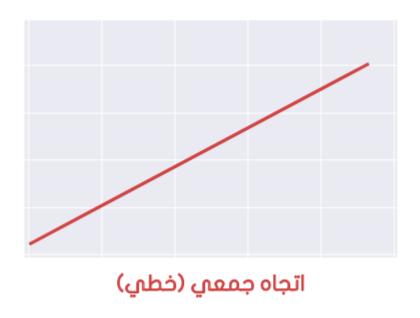
```
# استيراد المكتبات #
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

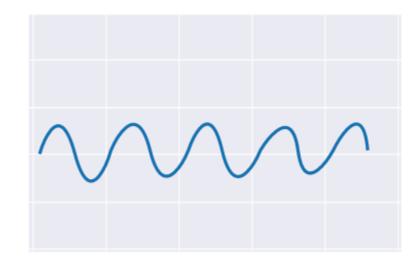
# نمط التصوير البياني #
plt.style.use('ggplot')
```

المكتبات التى سنستخدمها فى عملية النمذجة

1- سلسلة زمنية مكونة من اتجاه جمعي وموسمية جمعية Additive Trend & Additive Seasonality Time Series

الاتجاه الجمعي يشير إلى اتجاه خطي (خط مستقيم) ، وتعني الموسمية الجمعية عدم وجود أي تغييرات في عروض أو ارتفاعات الفترات الموسمية بمرور الوقت :





موسمية جمعية

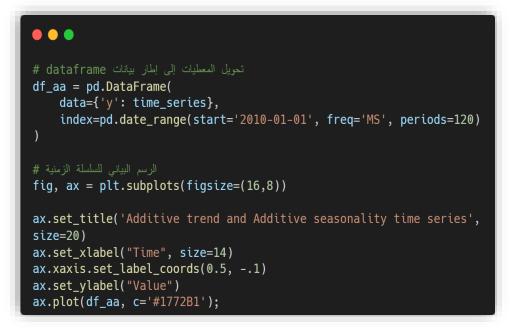
تحتوي 12 قيمة لـ 12 شهرًا (سنة) ثم يضاف إليها الحدود الـ 12 للموسمية كما هو موضح بالـ for loop

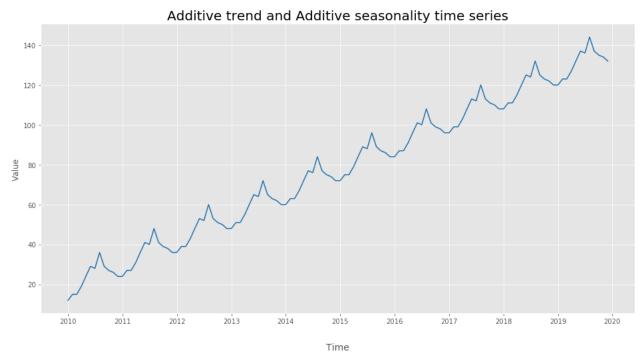
1- سلسلة زمنية مكونة من اتجاه جمعي وموسمية جمعية Additive Trend & Additive Seasonality Time Series

سننشئ سلسلة زمنية لمدة ٦٥ سنوات ذات اتجاه جمعي وموسمية جمعية !

```
array([ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13,
14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
                                                                                               27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39,
                                                                                               40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52,
إنشاء 120 نقطة (تمثل 10 * 12 شهر) أو 10 سنوات #
                                                                                               53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65,
                                                                                               66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78,
x_range = np.arange(1, 121)
                                                                                               79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91,
                                                                                               92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104,
عدد الحدود = 12) حدود تستخدم لإضافة تأثير الموسمية) #
                                                                                              105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117,
                                                                                              118, 119, 120])
seasonal_adders = [11, 13, 12, 15, 19, 23, 21, 28, 20, 17, 15, 12]
                                                                                             ا 120 نقطة تمثل قيم السلسلة لمدة 120 شمراً أو 10 سنوات!
قيم السلسلة الزمنية #
time_series = []
                                                                                          [12,15,15,19,24,29,28,36,29,27,26,24,
                                                                                            24,27,27,31,36,41,40,48,41,39,38,36,
تقسيم الـ 120 نقطة إلى 10 مجموعات كل مجموعة تشكل سنة #
                                                                                            36,39,39,43,48,53,52,60,53,51,50,48,
for chunk in np.split(x_range, 10):
    for i, val in enumerate(chunk):
                                                                                            108,111,111,115,120,125,124,132,125,123,122,120,
        time_series.append(val + seasonal_adders[i])
                                                                                            120,123,123,127,132,137,136,144,137,135,134,132]
                                                                                             ىتم تقسيم الـ 120 نقطة إلى 10 مجموعات كل مجموعة
```

التصوير البياني لسلسلة زمنية مكونة من اتجاه جمعي وموسمية جمعية





هل بإمكانك استنتاج كل من الاتجاه و الموسمية من خلال التصوير البياني ؟ إذا أجبت بنعم فلابد من أنك لاحظت أن الاتجاه العام للسلسلة هو تصاعدي خطي و الموسمية تتكرر كل سنة (٦2 شهر) بنمط معين ثابت مما يشير إلى اتجاه جمعي و موسمية جمعية ، وفي حال قد أجبت بـ لا فسنقوم بالشريحة التالية بتحليلها عن طريق مكتبة statsmodels

تحليل السلسلة الزمنية بواسطة مكتبة statsmodels

```
from statsmodels.tsa.seasonal import seasonal_decompose

# تحليل السلسلة الزمنية إلى كل من الاتجاه والموسمية والبواقي 
ss_decomposition = seasonal_decompose(df_aa, model='additive', period=12)
estimated_trend = ss_decomposition.trend
estimated_seasonal = ss_decomposition.seasonal
estimated_residual = ss_decomposition.resid
```

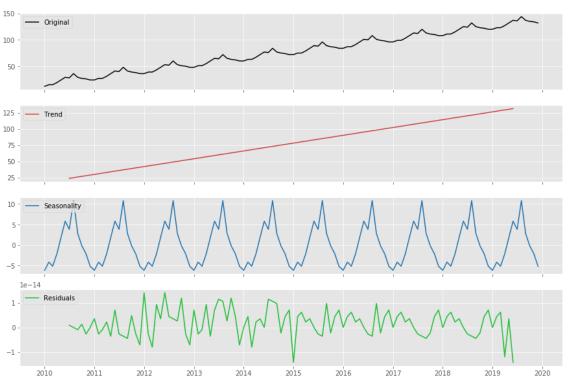
تم استدعاء الدالة seasonal_decompose من مكتبة statsmodels والمسؤولة عن عمل **تقدير** لتحليل السلسلة الزمنية إلى أجزائها الأساسية ونلاحظ أنه تم تمرير ثلاثة arguments وهي كالتالي :

- قيم السلسلة الزمنية المراد تحليلها time_series
- نوع النمذجة (نمذجة الموسمية) : جمعية additive
- الفترة : 12 شهر وهي تمثل دورة الموسمية period

نلاحظ أنه يوجد ثلاثة trend, seasonal, resid) : methods) وهي مسؤولة عن استخراج الأجزاء الثلاثة و قد تم تعيين متغير لكل منها لتصويرها لاحقًا

تصوير السلسلة الزمنية ومكوناتها الأساسية!

```
fig, axes = plt.subplots(4, 1, sharex=True, sharey=False)
fig.set_figheight(10)
fig.set_figwidth(15)
axes[0].plot(df_aa, label='Original', c='#0000000')
axes[0].legend(loc='upper left');
axes[1].plot(estimated_trend, label='Trend', c='#D14444')
axes[1].legend(loc='upper left');
axes[2].plot(estimated_seasonal, label='Seasonality', c='#1772B1')
axes[2].legend(loc='upper left');
axes[3].plot(estimated_residual, label='Residuals', c='#01BF21')
axes[3].legend(loc='upper left');
```

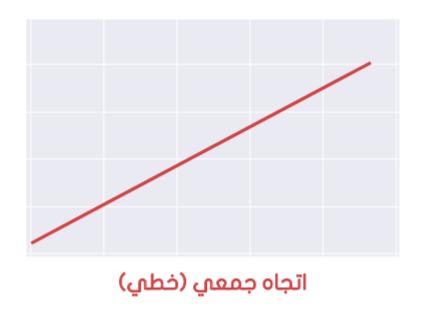


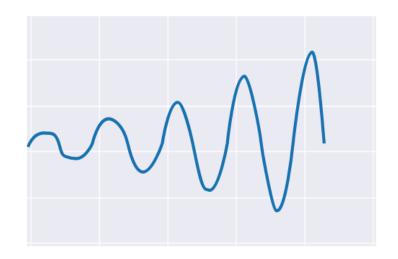
نلاحظ أن الأمور تجلت الآن بشكل أفضل وقد توضح كل من الاتجاه Trend على أنه اتجاه خطي متصاعد (جمعي) و الموسمية وseasonality على أنها جمعية كما هو موضح من النمط الذي يتكرر كل دورة موسمية (12 شهر) أما بالنسبة للبواقي Residuals فنلاحظ أنها تتذبذب بين -1 و 1 وتتمركز حول الصفر

ملاحظة : في الفصول اللاحقة سيكون تركيزنا أكبر على مفهوم البواقي لأنها تستخدم في عملية التنبؤ

2- سلسلة زمنية مكونة من اتجاه جمعى وموسمية ضربية Multiplicative Seasonality عصمى وموسمية ضربية -2

الاتجاه الجمعي يشير إلى اتجاه خطي (خط مستقيم) ، وتشير الموسمية الضربية إلى وجود تزايد أو تناقص في التردد (عرض الموجة) و / أو السعة (الارتفاع) للدورات الموسمية بمرور الوقت :





موسمية ضربية

2- سلسلة زمنية مكونة من اتجاه جمعي وموسمية ضربية Additive Trend & Multiplicative Seasonality

سننشئ سلسلة زمنية لمدة 10 سنوات ذات اتجاه جمعى وموسمية جمعية!

قمنا بتغيير كل من الحدود و العملية الحسابية إلى ضرب لإضافة تأثير الموسمية الضربية

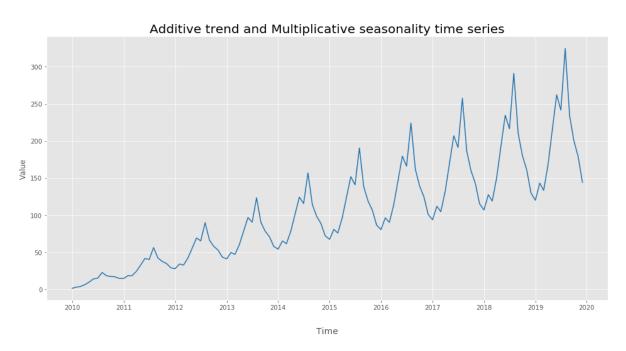
```
array([ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120])
```

120 نقطة زمنية تمثل قيم السلسلة لمدة 120 شهراً أو 10 سنوات!

يتم تقسيم الـ 120 نقطة إلى 10 مجموعات كل مجموعة تحتوي على 12 قيمة لـ 12 شهرًا (سنة) ثم يضرب كل عنصر بما يقابله في مصفوفة الحدود الـ 12 للموسمية كما هو موضح بالـ for loop

التصوير البياني لسلسلة زمنية مكونة من اتجاه جمعي وموسمية ضربية





هل بإمكانك استنتاج كل من الاتجاه و الموسمية من خلال التصوير البياني ؟ الاتجاه العام للسلسلة هو تصاعدي خطي أما الموسمية فهي تتضاعف مع كل دورة موسمية (كل سنة) لكنها تحمل نفس النمط، في الشريحة التالية سنقوم بتحليل المكونات الأساسية للسلسلة عن طريق مكتبة statsmodels

تحليل السلسلة الزمنية بواسطة مكتبة statsmodels

```
from statsmodels.tsa.seasonal import seasonal_decompose

# تحليل السلسلة الزمنية إلى كل من الاتجاه والموسمية والبواقي ss_decomposition = seasonal_decompose(df_am, model='multiplicative', period=12)
estimated_trend = ss_decomposition.trend
estimated_seasonal = ss_decomposition.seasonal
estimated_residual = ss_decomposition.resid
```

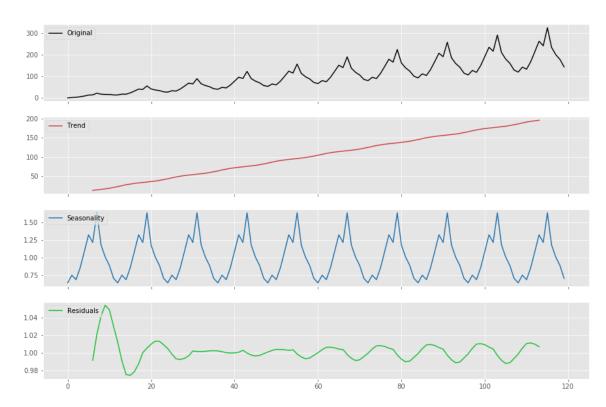
تم استدعاء الدالة seasonal_decompose من مكتبة statsmodels والمسؤولة عن عمل تقدير لتحليل السلسلة الزمنية إلى أجزائها الأساسية ونلاحظ أنه تم تمرير ثلاثة arguments وهي كالتالي :

- قيم السلسلة الزمنية المراد تحليلها time_series
- نوع النمذجة (نمذجة الموسمية) : ضربية multiplicative
 - الفترة : 12 شهر وهي تمثل دورة الموسمية period

نلاحظ أنه يوجد ثلاثة trend, seasonal, resid) : methods) وهي مسؤولة عن استخراج الأجزاء الثلاثة و قد تم تعيين متغير لكل منها لتصويرها لاحقًا

تصوير السلسلة الزمنية ومكوناتها الأساسية!

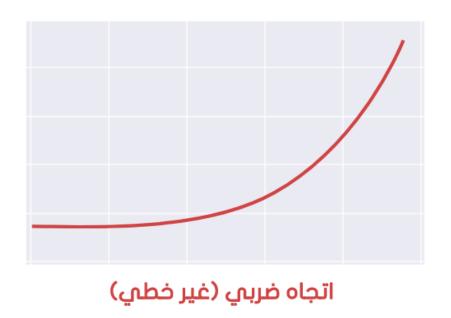
```
fig, axes = plt.subplots(4, 1, sharex=True, sharey=False)
fig.set_figheight(10)
fig.set_figwidth(15)
axes[0].plot(df_am, label='Original', c='#0000000')
axes[0].legend(loc='upper left');
axes[1].plot(estimated_trend, label='Trend', c='#D14444')
axes[1].legend(loc='upper left');
axes[2].plot(estimated_seasonal, label='Seasonality', c='#1772B1')
axes[2].legend(loc='upper left');
axes[3].plot(estimated_residual, label='Residuals', c='#01BF21')
axes[3].legend(loc='upper left');
```

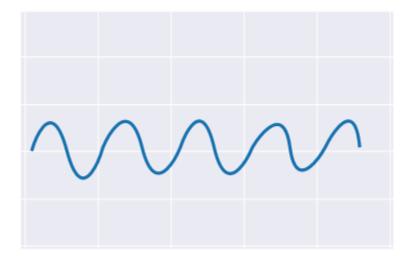


نلاحظ من التصوير البياني أن الاتجاه Trend خطي متصاعد (جمعي) أما بالنسبة للموسمية Seasonality فالنموذج قادر على استخراج النمط السائد و غير قادر على توضيح فيما إذا كانت جمعية أم ضربية إنما يمكن استنتاج ذلك من السلسلة الأصلية بسهولة !

3- سلسلة زمنية مكونة من اتجاه ضربي وموسمية جمعية Multiplicative Trend & Additive Seasonality

الاتجاه الضربي يشير إلى اتجاه غير خطي nonlinear ، وتعني الموسمية الجمعية عدم وجود أي تغييرات في عروض أو ارتفاعات الفترات الموسمية بمرور الوقت :





موسمية جمعية

3- سلسلة زمنية مكونة من اتجاه ضربي وموسمية جمعية Multiplicative Trend & Additive Seasonality

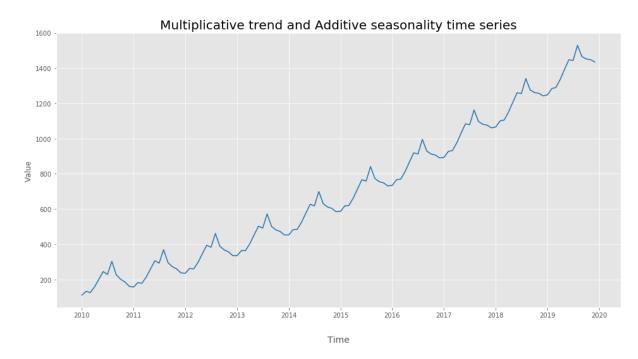
سننشئ سلسلة زمنية لمدة ٦٥ سنوات ذات اتجاه ضربي وموسمية جمعية !

```
• • •
إنشاء 120 نقطة (تمثل 10 * 12 شهر) أو 10 سنوات #
x_range = np.arange(1, 121)
(سنقوم برفع كل قيمة في إلى القوى 1.5 لجعل الاتجاه ضربي (نمو أسى غير خطى #
x_{range} = np.array([x**1.5 for x in xs])
حدود تستخدم لإضافة تأثير الموسميةالجمعية #
seasonal adders = [110, 130, 120, 150, 190, 230, 210, 280, 200, 170, 150, 120]
time_series = []
تقسيم الـ 120 نقطة إلى 10 مجموعات كل مجموعة تشكل سنة #
for chunk in np.split(x_range, 10):
    for i, val in enumerate(chunk):
         time_series.append(val + seasonal_adders[i])
```

نلاحظ أنه تم رفع كل قيمة في الـ 120 نقطة للقوة 1.5 وذلك لفرض إضافة تأثير اللاخطية non-linearity أي تحويل الاتجاه إلى اتجاه ضربي ثم قمنا بإضافة حدود الموسمية إلى السلسلة الزمنية كما فعلنا سابقًا!

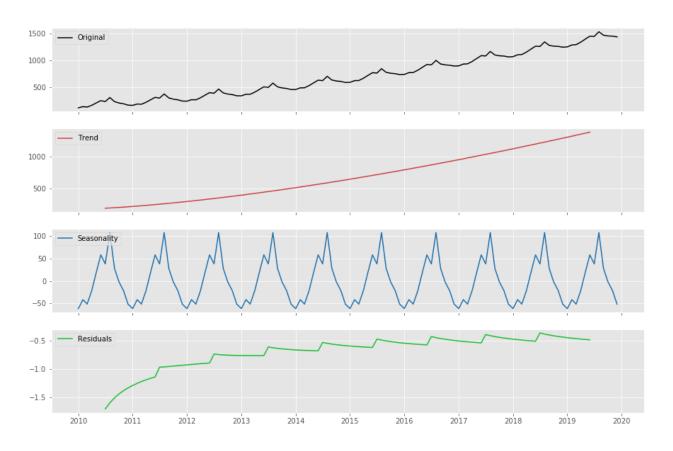
التصوير البياني لسلسلة زمنية مكونة من اتجاه جمعي وموسمية جمعية





في هذه المرة فإن الاتجاه غير واضح تمامًا على أنه ذو طبيعة غير خطية (ضربي) بينما الموسمية فهي واضحة على أنها جمعية لذا سنقوم بالشريحة التالية بتحليل هذه السلسلة عن طريق مكتبة statsmodels للتأكد من ذلك

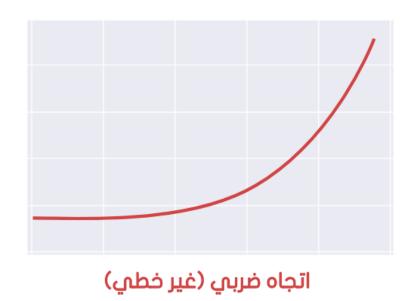
تصوير السلسلة الزمنية ومكوناتها الأساسية!

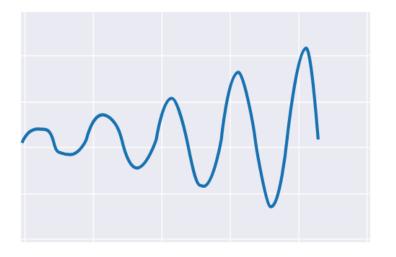


نلاحظ أن الاتجاه يتمثل بدالة غير خطية (اتجاه ضربي) بينما الموسمية يمكن استخلاصها من السلسلة الأصلية على أنها جمعية

4- سلسلة زمنية مكونة من اتجاه ضربي وموسمية ضربية Multiplicative Trend & Multiplicative Seasonality

الاتجاه الضربي يشير إلى اتجاه غير خطي ، وتشير الموسمية الضربية إلى وجود تزايد أو تناقص في التردد (عرض الموجة) و / أو السعة (الارتفاع) للدورات الموسمية بمرور الوقت :





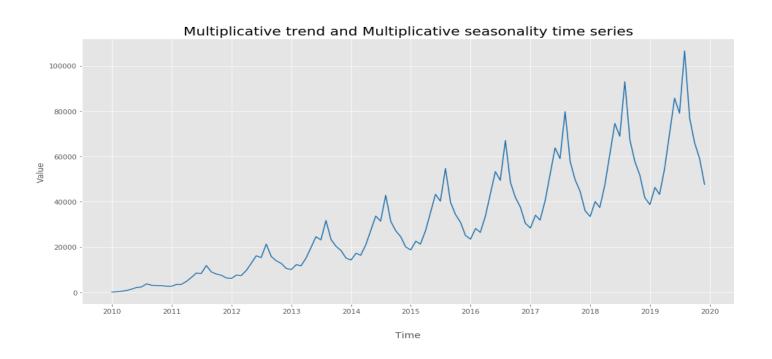
موسمية ضربية

4- سلسلة زمنية مكونة من اتجاه ضربي وموسمية ضربية Multiplicative Trend & Multiplicative Seasonality

سننشئ سلسلة زمنية لمدة 10 سنوات ذات اتجاه ضربي وموسمية جمعية !

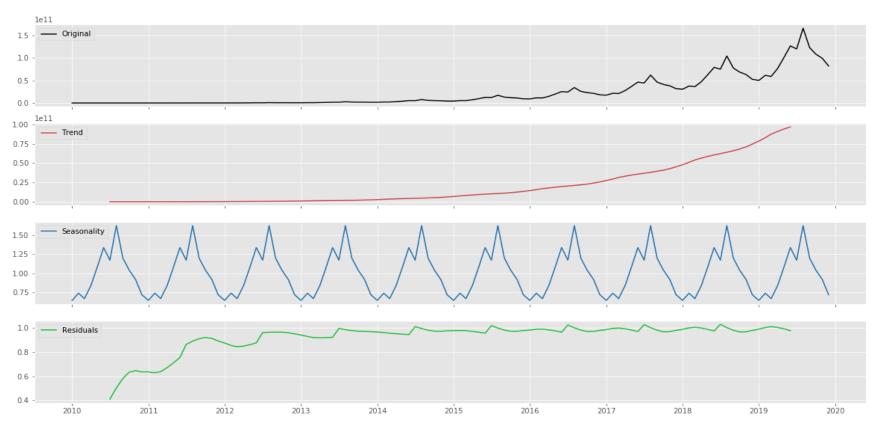
نلاحظ أنه تم رفع كل قيمة في الـ 120 نقطة للقوة 1.25 وذلك لفرض إضافة تأثير اللاخطية non-linearity أي تحويل الاتجاه إلى اتجاه ضربي ثم قمنا بضرب حدود الموسمية في قيم الاتجاه لإنتاج السلسلة الزمنية.

التصوير البياني لسلسلة زمنية مكونة من اتجاه جمعي وموسمية جمعية



الاتجاه غير واضح تمامًا على أنه ذو طبيعة غير خطية (ضربي) بينما الموسمية فهي واضحة على أنها ذات طبيعة ضربية لذا بالشريحة التالية بتحليل هذه السلسلة عن طريق مكتبة statsmodels للتأكد من ذلك

تصوير السلسلة الزمنية ومكوناتها الأساسية!



نلاحظ أن الاتجاه يتمثل بدالة غير خطية (اتجاه ضربي) و أيضًا الموسمية تعتبر ضربية كما هو موضح في السلسلة الأصلية أما بالنسبة للبواقي فهي تتمركز بشكل كبير عند القيمة 1 مما يدل على انحراف معياري صغير أي مدى أقل من القيم

ملخص Summary

يُعد تحليل السلاسل الزمنية أداة رائعة لمعرفة البيانات التاريخية Historical Data ونمذجتها وتفكيكها إلى مكوناتها الأساسية الفردية وذلك لاكتشاف الأنماط المتكررة والسلوك العام في السلسلة الزمنية لتسهيل عملية النمذجة والتنبؤ فيما بعد حيث أن استكشاف نوعية الموسمية على سبيل المثال سيساعد في استخدام اختيار النموذج كاختيار النماذج الموسمية والتي سنتعرف عليها في فصل التنبؤ في السلاسل الزمنية!