محتويات الفصل 2. التنبؤ البسيط

فصل 2. التنبؤ البسيط 37
1. طرق بسيطة في التنبؤ 37
1-1. التنبؤ البسيط أو الساذج 37
1-2. التنبؤ باستخدام المتوسط 39
1-3. طريقة الانزلاق 39
1-4. طريقة الخط المستقيم 40
1-5. التنبؤ البسيط في R 41
2. تقييم التقدير ومؤشرات الخطأ 47
2-1. عينة البناء وعينة التقييم 47
3-1. مؤشرات الخطأ: MAE، MSE وMAPE 47
3-3. حساب مؤشرات التقدير في EXCEL 50
3-4. التمثيل البياني للتقدير وحساب مؤشراته في R 51
3. خلاصة 52
4. ملحق. مؤشرات أخرى للخطأ والتشتت 53
*4-1. متوسط خطأ التنبؤ **MFE** 53*
4-2. مؤشر تتبع الإشارة (TS Tracking Signal). 54
4-3. التباين والانحراف المعياري 54
5. سلسلة تمارين: التنبؤ البسيط 55
5-1. التمارين 55
5-2. الحلول 57
6. مراجع الفصل 57

فصل 2. التنبؤ البسيط

طرق التنبؤ البسيط – مؤشرات الخطأ – خلاصة – ملحق – تمارين

توطئة. المتوسطات المتحركة هي أداة بسيطة وشائعة الاستخدام تفيد في استخراج التوجه العام من خلال تمهيد (طمس) تذبذبات المدى القصير. يفيد استخراج التوجه العام في فهم حركة المتغيرة في المدى البعيد وبالتالي التنبؤ بمسارها. تستخدم هذه التقنية مثلا في البورصة لطمس التذبذبات اليومية لسعر السهم لاستخراج التوجه العام من أجل القرار بالاستثمار أم لا.

نتعرف في هذا الفصل على كيفية استخدام هذه الطريقة وأنواعها، وكيفية الاستفادة منها لاتخاذ القرار الاستثماري، ونتطرق لاستخدام Excel لإجراء الحسابات وللتمثيل البياني. في الأخير نتطرق أيضا لعدد من المؤشرات المستخدمة في تقييم دقة التقدير وتقييم شدة التمهيد.

الغرض هو أن يصبح الطالب قادرا على حساب المتوسطات المتحركة للتحليل وللتنبؤ يدويا وباستخدام برنامج Excel والاستفادة منها في اتخاذ القرار، مع تقييم دقة التقدير وشدة التمهيد المتحصل عليهما.

1. طرق بسيطة في التنبؤ

التنبؤ البسيط

التنبؤ باستخدام المتوسط

طريقة الانزلاق

طريقة الخط المستقيم

1-1. التنبؤ البسيط أو الساذج

تتمثل الطريقة البسيطة أو الساذجة للتنبؤ (simple/naïve) في اعتماد آخر قيمة مشاهدة كتوقع للقيمة المقبلة:

$$\widehat{y}_T(1) = y_T$$

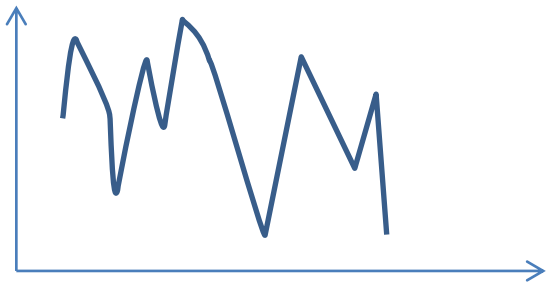
حيث y_T هي القيمة الحقيقية للسلسلة في الزمن T.

و$\hat{y}_T(1)$ القيمة المتوقعة للسلسلة بعد فترة واحدة إنطلاقا من الزمن T.

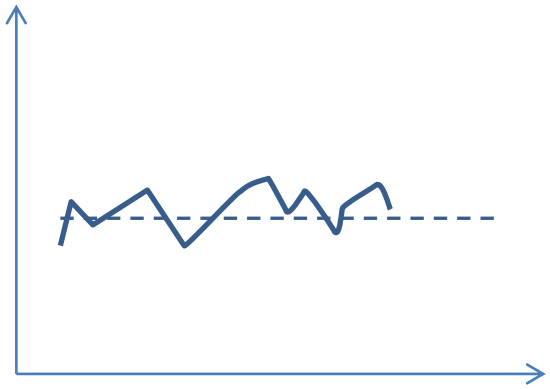
تسمى T "أصل التنبؤ"، غالبا (لكن ليس دوما) هي زمن آخر قيمة معروفة للسلسلة $\{y_1, y_2, \ldots, y_n\}$.

مثال. مبيعات الأحد 120 وحدة، إذن نتوقع أن تكون مبيعات الاثنين 120 أيضا؛ ثم إذا جاءت مبيعات الاثنين 130، نتوقع أن تكون مبيعات الثلاثاء 130؛ وهكذا.

تكون هذه الطريقة في التنبؤ معقولة عندما يكون الأفق قريبا، مثلا اليوم أو الأسبوع الموالي. تستخدم هذه الطريقة مثلا في الأسواق المالية، حيث يكون التوقع لسعر السهم غدا هو سعره اليوم.



رسم توضيحي 1. سلسلة شديدة التباين، حالة لا يصلح لها التنبؤ البسيط.



رسم توضيحي2 سلسلة بتباين ضئيل وبدون توجه أو موسمية

الطريقة البسيطة الموسمية: يمكن بتعديل بسيط أخذ الموسمية في الحسبان في التنبؤ البسيط (seasonal naive method)، وذلك بأخذ قيمة المتغيرة في الموسم نفسه من الفترة الماضية. مثلا، للتنبؤ لشهر فيفري نأخذ القيمة في شهر فيفري للسنة السابقة، وللتنبؤ لشهر مارس المقبل نأخذ القيمة المسجلة في شهر مارس السابق له، وللتنبؤ للخميس المقبل نأخذ القيمة في الخميس الحالي، وهكذا...

$$\hat{y}_T(\mathbf{h}) = y_{T+h-p}$$

حيث $\hat{y}_T(h)$: القيمة المتوقعة بعد h فترة انطلاقا من الزمن T،

و p نافذة الموسمية أي عدد المشاهدات في النمط الواحد.

تمسى h أفق التنبؤ، أي المدة بين أصل التنبؤ T واللحظة المستهدفة (T + h).

للتنبؤ لعدة مواسم مقبلة نستخدم قيم آخر نافذة معلومة.

التنبؤ البسيط والتوجه: يمكن أيضا تعديل الطريقة البسيطة لاستيعاب الزيادة أو الانخفاض الحاصل بين فترتين متتاليتين من خلال إضافة الفرق بين القيمة الحالية والسابقة (Ritzman Larry et al., 2004).

مثال: إذا كانت مبيعات شهر جانفي 100 ومبيعات فيفري 110، يمكن أن ندخل قيمة هذه الزيادة (10) في توقعنا للشهر المقبل، فنتوقع مبيعات قدرها 120 في مارس. ثم إذا جاءت مبيعات مارس 125، أي بزيادة قدرها 15 عن فيفري بدلا من 10، نتوقع أن تكون مبيعات أفريل بنفس الزيادة (15)، فنتوقع 140 (125+15).

عموما تصلح طرق التنبؤ البسيط عندما تكون السلسلة خالية من التوجه والموسمية أو يكون لها موسمية مستقرة أو توجه خطي، وبتغيرات عشوائية ضئيلة، وبدون انتقالات مفاجئة، وإلا فإن خطأ التقدير يكون كبيرا (Ritzman Larry et al., 2004).

2-1. التنبؤ باستخدام المتوسط

في هذه الطريقة نستخدم متوسط عدد من القيم الماضية للتنبؤ بالقيمة المستقبلية:

$$\hat{y}_T(1) = \frac{1}{k}(y_{T-k+1} + \cdots + y_{T-1} + y_T)$$

مثال: لديك مبيعات تسع أشهر من السنة. قم بالتنبؤ للشهر 10 إنطلاقا من 6 أشهر ماضية.

T	1	2	3	4	5	6	7	8	9	10
y_t	100	120	110	130	128	130	97	115	97	122

$\hat{y}_9(1)$ = (130+128+130+97+115+97)/6 = 122

كما التنبؤ البسيط تفترض الطريقة عدم وجود توجه. إستخدام المتوسط يسمح بإطالة أفق التنبؤ قليلا في حالة الحاجة إلى ذلك. يتحكم المحلل من خلال k في مقدار الرجوع إلى الوراء، فيستخدم درجة k صغيرة إن قدر أن الوضع الحالي الجديد هو الذي سيسود في المستقبل؛ والعكس، يستخدم k كبيرة إذا اعتبر أن التغيرات الأخيرة آنية وأن الأمور ستعود إلى مجراها السابق.

3-1. طريقة الانزلاق

في حالة وجود توجه عام في السلسلة، يمكن استخدام طريقة الانزلاق (drift method)، وتسمى أيضا (random walk). تسمح هذه الطريقة للتنبؤ لأفق أبعد من 1: نحسب متوسط الانتقالات أي متوسط الفروق بين الفترات المتتالية السابقة ل T، ثم نضرب هذا المتوسط في الأفق h، ثم نضيفه إلى قيمة السلسلة في T. يمكن اختصار هذه الصيغة لأن مجموع الفروق هو ذاته الفرق بين آخر وأول قيمة.

$$\hat{y}_T(\mathbf{h}) = y_T + h\frac{\sum_2^T(y_t - y_{t-1})}{T-1} = y_T + h\frac{(y_T - y_1)}{T-1}$$

تسمح هذه الطريقة باحتساب التوجه ليس فقط من آخر تغيير وإنما منذ بداية السلسلة.

مثال: لدينا المبيعات من جانفي إلى سبتمبر، ونريد التنبؤ لديسمبر.

t	1	2	3	4	5	6	7	8	9	10	11	12
y_t	100	120	110	130	128	130	132	115	132			144

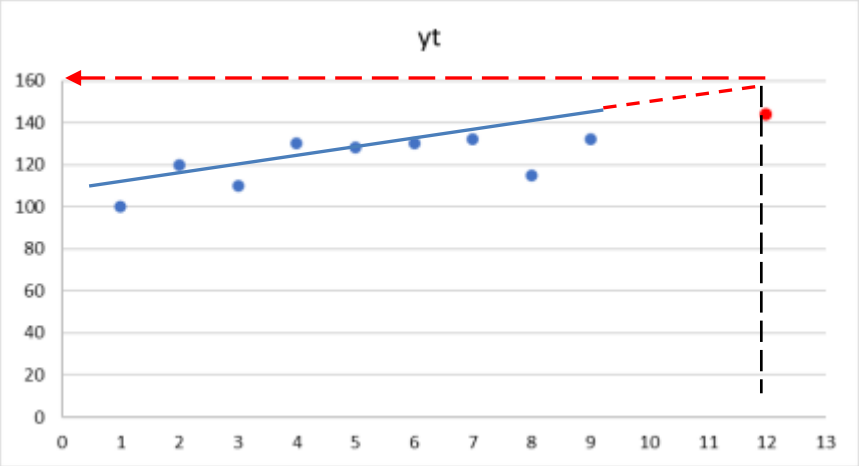
$$\hat{y}_T(\mathbf{h}) = y_T + h\frac{(y_T - y_1)}{T-1}$$

T = 9, h = 3, y_T = 132

$\hat{}y_9(3) = 132 + 3((132 - 100)/(9 - 1)) = 144$

4-1. طريقة الخط المستقيم

تسمح هذه الطريقة باحتساب التوجه من خلال رسم خط على التمثيل البياني للسلسلة بين أول وآخر قيمة في السلسلة، ثم تمديده إلى غاية اللحظة المستهدفة. الإسقاط على المحور العمودي يعطي التنبؤ.



رسم بياني 1. طريقة الخط المستقيم

لاحظ.

الطرق المذكورة تتميز بسهولتها وليس بدقتها، قد تكون إحدى هذه الطرق كافية للتنبؤ في حالات ما لكنها تستخدم أيضا كمرجع للمقارنة، بحيث تستخدم طرق أخرى للتنبؤ وتقرن مع الطرق البسيطة فإن لم تعط تقديرا أفضل لم تكن تستحق الجهد. من بين الطرق الأخرى للتنبؤ:

التنبؤ بطريقة دالة التوجه: نستخرج دالة تربط y بمتغيرة الزمن أو بمتغيرات أخرى. تسمح هذه الدالة بالتنبؤ بالمتغيرة باستخدام قيم محتملة لمتغيرة الزمن أو المتغيرات الأخرى. يمكن إدراج الموسمية في التنبؤ بعد حساب معاملات الموسمية المشاهدة في السلسلة (الفصلين المقبلين).

التنبؤ بطريقة التمهيد الأسي: نستخدم دالة في القيم السابقة مع إعطاء أكبر أهمية في التنبؤ للقيم الأحدث. يمكن أن تستوعب هذه الطرق وجود موسمية أو توجه أو كلاهما (فصول التمهيد الأسي البسيط والثنائي والثلاثي).

التنبؤ باستخدام نماذج ARIMA: هي عائلة من الطرق تعتمد على نمذجة حركة السلسلة والتنبؤ لها بمجال ثقة معين، من خلال دالة تربط المتغيرة بماضيها وفي حالات بمتغيرات أخرى.

يتضمن الرمز ARIMA المكونات التالية: AR: Auto regression التنبؤ بالسلسلة من خلال قيم سابقة مسجلة في t-1، t-2، مكون ...MA: Moving Average المتوسطات المتحركة، يتم التنبؤ بقيم مستقبلية من خلال الأخطاء السابقة؛ و I: التفريق الذاتي للسلسلة.

فيما يلي ندرس المتوسطات المتحركة بأنواعها وكيفية استخدامها في التنبؤ.

5-1. التنبؤ البسيط في *R*

يوفر R من خلال الحزمة forecast عددا من الدوال لغرض التنبؤ:

- للتنبؤ بالطريقة الساذجة نستخدم الدالة naive() أو الدالة rwf()
- للتنبؤ بالطريقة الساذجة الموسمية نستخدم الدالة snaive()،
- للتنبؤ باستخدام المتوسط نستخدم الدالة meanf()،
- وللتمثيل البياني نستخدم الدالة autoplot()

فيما يلي نستخدم سلسلة عدد المسافرين عبر المطارات الأمريكية (AirPassengers) رغم عدم ملائمتها للتنبؤ البسيط لكونها تتضمن موسمية غير مستقرة وتوجه، لكن ذلك يساعد على فهم أهمية الاشتراطات.

1- التنبؤ البسيط naive()

نستدعي الحزمة Forecast وإن لم تكن مثبتة نثبتها ثم نستدعها لتنفيذ الدوال المذكورة أعلاه. نستخدم الدالة naive() للتنبؤ بأفق قدره 24، ثم نمثل التنبؤ والسلسلة بالدالة autoplot():

```
library(forecast)
naive(AirPassengers,4)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 1961	432	388.7984	475.2016	365.9288	498.0712
Feb 1961	432	370.9037	493.0963	338.5612	525.4388
Mar 1961	432	357.1726	506.8274	317.5613	546.4387
Apr 1961	432	345.5967	518.4033	299.8576	564.1424

لاحظ أن قيمة التوقع ثابتة بينما يعطي R مجالي ثقة 80 بالمائة و95 بالمائة يتوسعان مع زيادة الأفق.

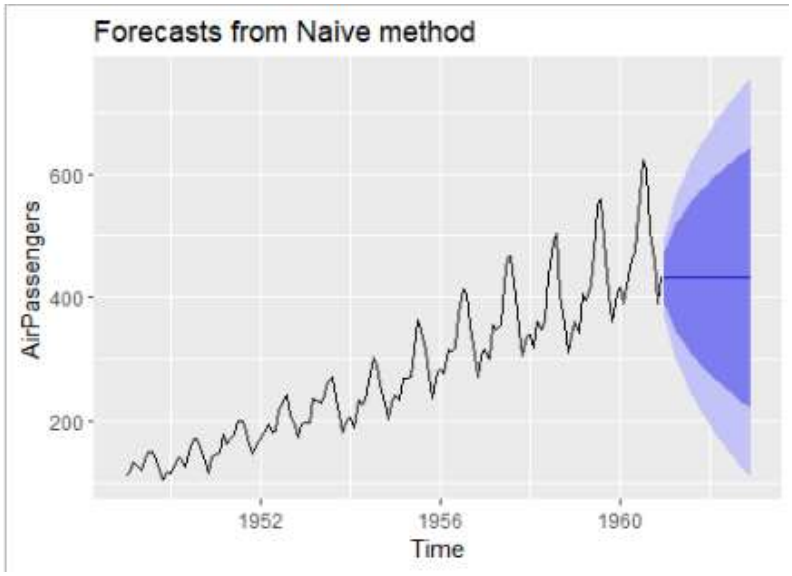
(Lo) تعني الحد الأدنى للمجال، و Hi الحد الأقصى.

للقيام بالتمثيل البياني للسلسلة مع التنبؤ نستخدم هنا autoplot() ونطيل الأفق إلى 24 شهرا.

```
autoplot(naive(AirPassengers,24))
```

الخط في الوسط يعبر عن التنبؤ البسيط وهو ببساطة آخر قيمة متوفرة. التوقع البسيط هنا غير ملائم لأنه – كما يبين الرسم – يهمل الموسمية والتوجه.

الهالة حول خط التنبؤ هي لمجالي الثقة 80 بالمائة و95 بالماءة على التوالي، أي أن هناك احتمال 0.8 أن يأتي عدد المسافرين في المجال باللون الداكن، و0.95 أن يأتي ضمن المجال الكلي الملون. مجال الثقة مهم جدا لأنه يظهر مدى الثقة في التنبؤ، كلما كان المجال ضيقا كلما دل على دقة في التقدير.



2– التنبؤ البسيط مع احتساب الموسمية snaive()

باستخدام snaive() نحصل على تنبؤ بسيط مع احتساب الموسمية. بما أن الحزمة مفعلة نكتب مباشرة:

```
snaive(AirPassengers,4)
```

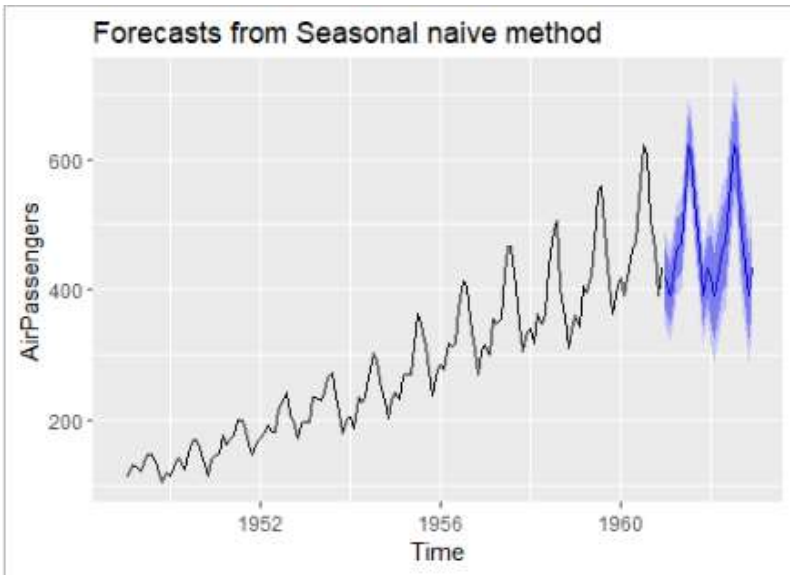
النتيجة هي كالتالي:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 1961	417	370.4595	463.5405	345.8224	488.1776
Feb 1961	391	344.4595	437.5405	319.8224	462.1776
Mar 1961	419	372.4595	465.5405	347.8224	490.1776
Apr 1961	461	414.4595	507.5405	389.8224	532.1776

البيانات تعطي التنبؤ (Point forecast) مع حدي مجالي الثقة 80 بالمائة و95 بالمائة. لاحظ أن القيم ه نا تتغير، وذلك تتبعا للموسمية.

لتمثيل البيانات مع التنبؤ نستخدم الدالة autoplot():

```
autoplot(snaive(AirPassengers,24))
```



الرسم يبين القيم المتوقعة بهذه الطريقة مع مجالي الثقة حوله، ولاحظ أن هذه الطريقة تحتسب الموسمية لكنها تهمل التوجه، لذلك فهي بالنسبة لهذه البيانات غير مناسبة.

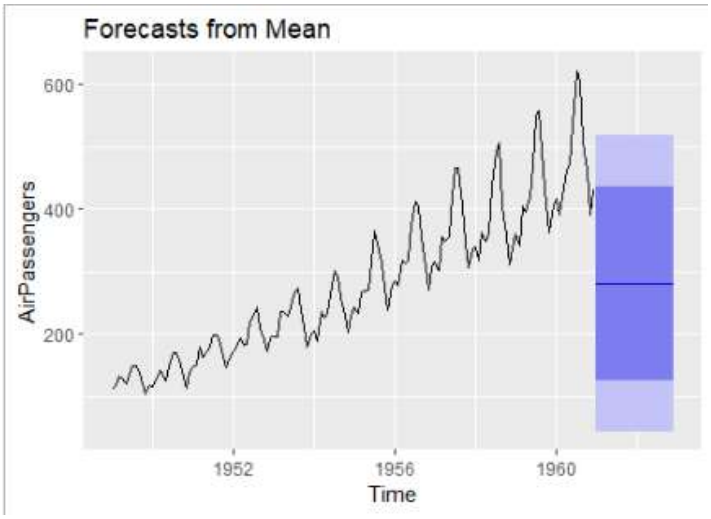
3- التنبؤ باستخدام المتوسط meanf()

الدالة meanf() أيضا هي ضمن حزمة forecast. لكوننا استدعيناها من قبل لا نعيد طلبها، فنكتب مباشرة: الدالة تعطي التنبؤ باستخدام المتوسط الإجمالي للسلسلة.

```
meanf(AirPassengers,4)
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Jan 1961       280.2986 125.3066 435.2906 42.34016 518.2571
Feb 1961       280.2986 125.3066 435.2906 42.34016 518.2571
Mar 1961       280.2986 125.3066 435.2906 42.34016 518.2571
Apr 1961       280.2986 125.3066 435.2906 42.34016 518.2571
```

لطلب تمثيل بياني للسلسلة مع التنبؤ باستخدام المتوسط:

```
autoplot(meanf(AirPassengers,24))
```



هنا أيضا تظهر طريقة المتوسط غير مناسبة لعدم احتسابها التوجه ولا الموسمية.

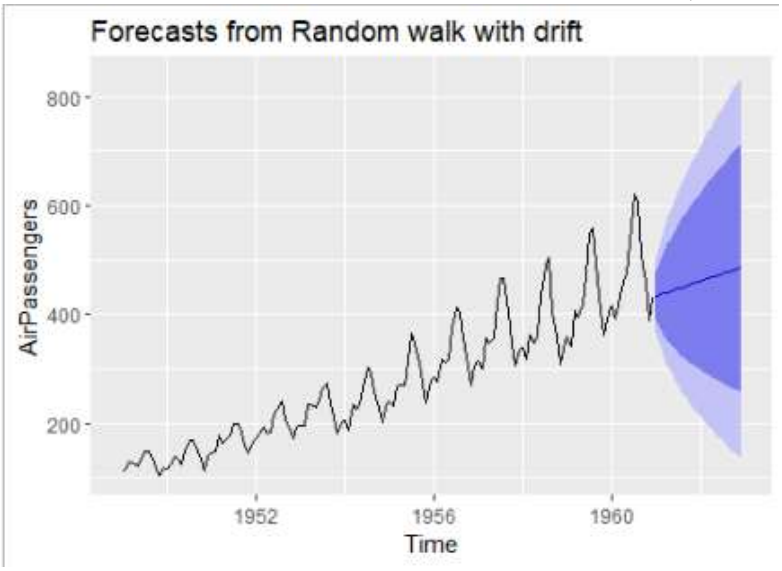
4- طريقة الانزلاق

تسمح طريقة الانزلاق باحتساب التوجه إن وجد من خلال إضافة أو طرح متوسط التغير في y. الدالة المستخدمة في r هي rwf(). فيما يلي تطبيق بسيط لها على ذات السلسلة السابقة.

```
rwf(AirPassengers,4,drift = T)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 1961	434.2378	390.9799	477.4956	368.0806	500.3949
Feb 1961	436.4755	375.0862	497.8649	342.5886	530.3625
Mar 1961	438.7133	363.2664	514.1602	323.3272	554.0994
Apr 1961	440.9510	353.5325	528.3696	307.2560	574.6461

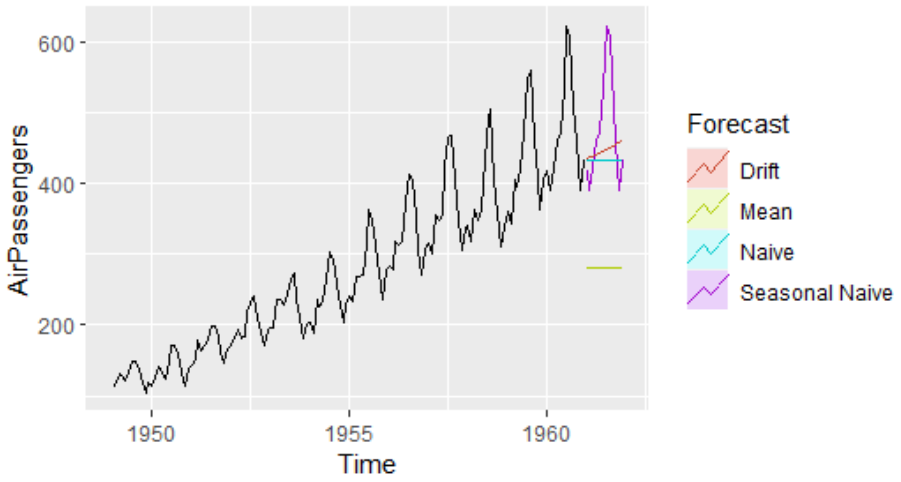
نتفحص الآن بالرسم البياني التنبؤ الذي تعطيه الدالة rwf().



لاحظ أن القيم المتوقع هي عبارة عن منحنى صاعد خطيا بدون احتساب الموسمية، وبالتالي فطريقة الانزلاق هي الأخرى لا تصلح مع هذه السلسلة.

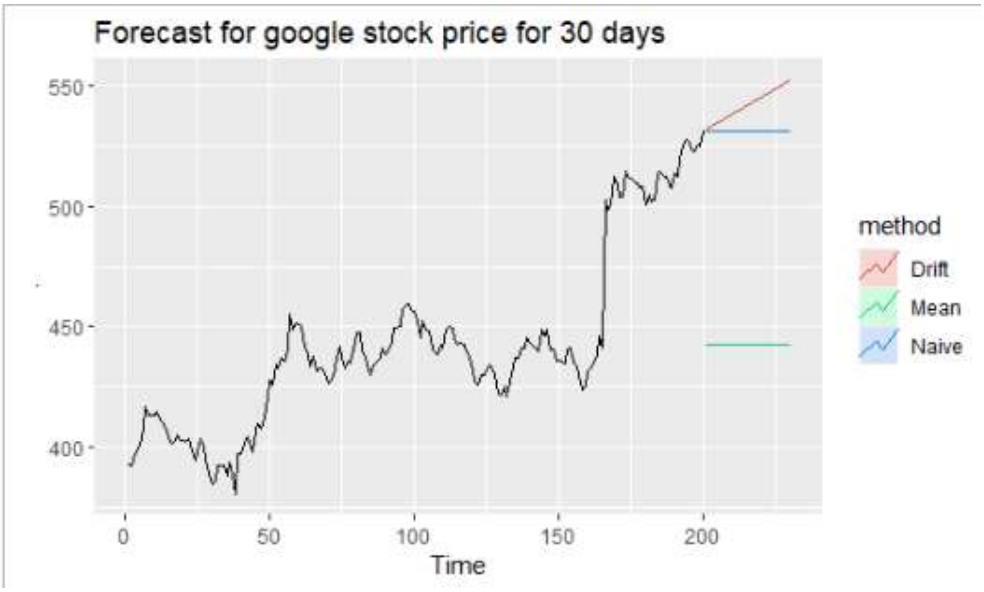
يمكن في R إظهار تنبؤات مختلفة على ذات الرسم باستخدام autolayer():

```
library(ggplot2) # guides() needs it.
autoplot(AirPassengers)+
  autolayer(snaive(AirPassengers,12),series = "Seasonal
Naive",PI=F)+
  autolayer(naive(AirPassengers,h=12),series = "Naive",PI=F)+
  autolayer(meanf(AirPassengers,h=12),series = "Mean",PI=F)+
  autolayer(rwf(AirPassengers,12,drift=TRUE),series = "Drift",PI=F)+
  guides(colour=guide_legend(title="Forecast"))
```



كل من الطرق الأربع غير ملائمة لبيانات AirPassengers، لكن طبعا يمكن أن نجد حالات كثيرة يصلح فيها التنبؤ البسيط، وذكرنا أن من بينها التنبؤ بسعر السهم. فيما يلي نطبق طريقة التنبؤ البسيط على بيانات سعر سهم قوقل من خلال بيانات goog200 (متضمنة في R).

```
goog200%>%autoplot()+
  autolayer(naive(goog200,30), series="Naive",PI=F)+
  autolayer(rwf(goog200,h=30,drift=T), series="Drift",PI=F)+
  autolayer(meanf(goog200,30), series="Mean",PI=F)+
  guides(colour=guide_legend(title="method"))+
  ggtitle("Forecast for google stock price for 30 days")
```



طريقة أخرى للتنبؤ في R هي باستخدام الدالة forecast() . تختار هذه الدالة اتوماتيكيا طريقة التنبؤ، لكن غالبا يبني المحلل المتمرس نموذجه بنفسه ويستخدم الدالة للتنبؤ اعتمادا عليه.

```
forecast(AirPassengers,h=4)
```

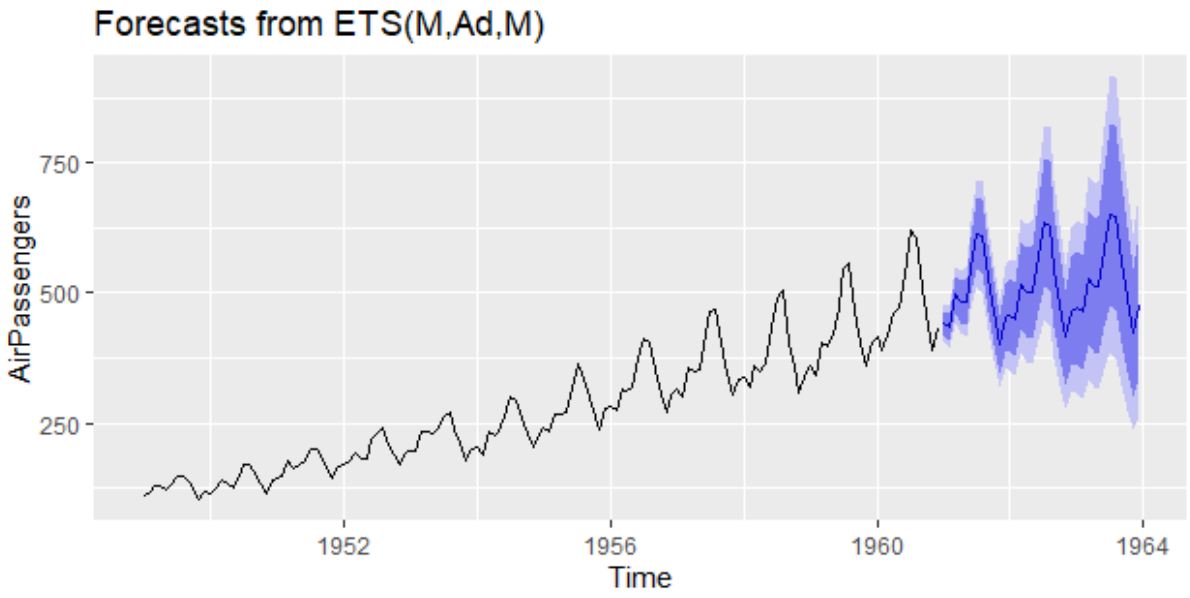
```
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Jan 1961       441.8018 419.6256 463.9780 407.8863 475.7174
Feb 1961       434.1186 407.1668 461.0704 392.8994 475.3379
Mar 1961       496.6300 460.6291 532.6310 441.5714 551.6887
```

```
Apr 1961       483.2375 443.6210 522.8539 422.6493 543.8256
```

لإظهار نتيجة التنبؤ في الرسم البياني:

```
library(forecast)
autoplot(forecast(AirPassengers,h=36))
```

Output :



2. تقييم التقدير ومؤشرات الخطأ

عينة البناء وعينة التقييم

مؤشرات الخطأ

2-1. عينة البناء وعينة التقييم

لتقييم دقة التقدير الذي تعطيه أي طريقة أو للمفاضلة بين طريقتين أو أكثر في التنبؤ يمكن أن ننظر إلى "دقة" التقدير الذي تعطيه الطريقة لتواريخ سابقة، حيث القيم الحقيقية معلومة. هذا يعني تقسيم السلسلة إلى جزئين، جزء لبناء النموذج يشكل عادة 80 بالمائة من البيانات، ويمكن أن نسميه عينة البناء (training data)، والجزء الباقي وهو الأخير لتقييم الطريقة أو النموذج، ونسميه عينة التقييم(test data)، بحيث نقارن فيه بين القيم المتوقعة بالنموذج وبين القيم الحقيقية. حجم كل من الجزئين يختلف بحسب حجم البيانات المتوفرة وبحسب طول الأفق المراد التنبؤ له، وفي كل الأحوال لا ينبغي أن تقل عينة التقييم عن طول الأفق (Hyndman, 2014). إذا جاءت التقديرات للقيم الماضية دقيقة وكنا نتوقع أن المستقبل لن يختلف كثيرا، عندها يمكن أن نأمل أن الطريقة ستعطي تنبؤا دقيقا في المستقبل أيضا.

2-1. مؤشرات الخطأ: *MAE*، *MSE* و*MAPE*

تستخدم مؤشرات الخطأ لاختيار القيم الأمثل لمعاملات هذه الطريقة أو تلك، مثلا لاختيار معامل المتوسطات المتحركة أو الأوزان المعطاة للمتوسطات المتحركة المرجحة. على سبيل المثال نتساءل هل نعطي أهمية أكبر للقيم القريبة لكي نأخذ في الحسبان الظروف الحالية أم نولي الأهمية للقيم البعيدة في الماضي، لألا نهمل التوجه العام للظاهرة أو لا تتأثر تقديراتنا بأحداث عابرة.

متوسط القيم المطلقة للخطأ MAE

يقيس MAE (Mean Absolute Error) خطأ التقدير. المؤشر يحسب كما يلي:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|E_t|, \quad E_t = y_t - \hat{y}_t$$

حيث t هي مؤشر لعينة التقييم أي للمشاهدات المستخدمة في تقييم النموذج، و n هو حجم عينة التقييم. القيمة المطلقة تمنع إلغاء القيم لبعض بسبب الإشارة، والمتوسط يمنع تأثر المؤشر بعدد المشاهدات.

مثال: يقدر مالك مطعم يوميا استهلاك الخبز بطريقة التنبؤ الساذج، ولديك استهلاك أخر 7 أيام.

Day	81	82	83	84	85	86	87
Y	126	125	119	133	129	127	131

أكتب المبيعات المقدرة، وقيم التقدير من خلال MAE.

Day	t	Y	^y	|E | = |y – ^y|
81	1	126	/	/
82	2	125	126	|125 – 126| = 1
83	3	119	125	|119 – 125| = 6
84	4	133	119	|133 – 119| = 14
85	5	129	133	|129 – 133| = 4
86	6	127	129	|127 – 129| = 2
87	7	131	127	|131 – 127| = 4
				∑ = 31
				MAE = 31/6 = 5.17

لاحظ أن القسمة على 6 لا على 7، لأننا نجمع 6 قيم. يتراوح الخطأ إذن حول 5 خبزات زيادة أو نقصانا، مقارنة مع حجم المبيعات اليومي يمكن أن نعتبر هذا مقبولا، ونقبل إذن الطريقة البسيطة.

ميزة MAE سهولة تفسيره، فقيمته تدل على القيمة التي يتراوح حولها الخطأ بالقيمة المطلقة، كلما كانت قيمته أقل دل على أن طريقة التقدير ملائمة ودقيقة. يحتفظ MAE بالوحدة لذلك يفيد عندما نكون بصدد المقارنة بين طرق تنبؤ لذات السلسلة أو لسلاسل لها نفس الوحدة (Hyndman, 2014). عيب MAE أنه يعنى بالمتوسط العام للخطأ بينما قد تكون هناك قيم متطرفة قليلة كفيلة بتوليد نتائج كارثية على المؤسسة حتى لو كان المتوسط العام ضئيلا أو مقبولا، لذلك يجب عدم الاكتفاء بالنظر إلى قيمة المؤشر وإنما تفحص عموم عمود الخطأ. المؤشر الذي يتحسس القيم المتطرفة إن وجدت هو MSE.

متوسط مربعات الخطأ MSE

يقيس MSE ((Mean Squared Error مربعات الخطأ، ويحسب كما يلي:

$$\boldsymbol{MSE} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{t=1}}^{\boldsymbol{n}} (\boldsymbol{E_t})^2$$

مثال: قيم التقدير في المثال السابق من خلال MSE.

Day	Y	^y	$(y - \hat{y})^2$
81	126	/	/
82	125	126	$(125 - 126)^2 = 1$
83	119	125	$(119 - 125)^2 = 36$
84	133	119	$(133 - 119)^2 = 196$
85	129	133	16
86	127	129	4
87	131	127	16
			∑ = 269
		MSE =	269/6=44.83

قيمة الخطأ هنا ليس لها وحدة، ولا تقارن مع قيم السلسلة، لذلك قد يصعب تقييم الخطأ بواسطتها إلا إذا كان الأمر يتعلق بالمقارنة بين قيم مختلفة ل MSE، مثلا مقارنة بين MSE التي تعطيها طريقتين للتقدير، أو مقارنة بين قيمة المؤشر في سلسلتين. يمكن أيضا وضع القيمة الناتجة (MSE) تحت الجذر للحصول على مؤشر ثاني له وحدة، وهو RMSE. مثلا هنا سيكون 6.7 خبزة.

يستخدم التربيع لتجنب إلغاء القيم الموجبة والسالبة بعضها لبعض. المتوسط يسمح بعدم تأثر المؤشر بعدد المشاهدات. ميزة هذا المؤشر أنه يعطي أهمية أكبر للقيم المتطرفة، كما أنه يعطي معنى مفهوم التباين المعروف، لكن هذه المرة تباين السلسلة المحولة ليس عن متوسطها وإنما عن السلسلة الأصلية.

متوسط نسبة الخطأ MAPE

يقيس MAPE (Mean Absolute Percentage Error) نسبة الخطأ إلى القيمة الحقيقية، ويحسب كما يلي:

$$\boldsymbol{MAPE = \frac{1}{n}\sum_{t=1}^{n}(|E_t| \times 100/y_t)}$$

مثال: قيم التقدير في المثال السابق من خلال MAPE.

Day	y	^y	| y – ^y |*100/y
81	126	/	/
82	125	126	|125 – 126| x 100/125 = 0,80
83	119	125	|119 – 125| x 100/119 = 5,04
84	133	119	|133 – 119| x 100/133 =10,53
85	129	133	3,10
86	127	129	1,57
87	131	127	3,04
			∑ = 24,09
		MAPE	24.09/6=4,015%

خطأ التقدير هنا في المتوسط 4 بالمائة. هذه النسبة يمكن أن تعد مقبولة، لكن يجب أيضا تفحص مختلف قيم الخطأ، لاحظ مثلا أن نسبة الخطأ وصلت إلى 10 بالمائة في اليوم الرابع. ميزة هذا المؤشر هو سهولة الحكم على مقدار الخطأ وإمكانية المقارنة بين دقة تقدير سلاسل مختلفة وإن اختلفت الوحدة.

مشكلة هذا المؤشر هي مع الصفر ومع القيم السالبة للخطأ. المؤشر لا يستطيع حساب نسبة الخطأ عندما تكون القيمة الحقيقية y_t هي الصفر؛ كذلك، يفترض هذا المؤشر الصفر الحقيقي (meaningful zero)، أي أنه يفترض أن الصفر يعني العدم، لذلك فإن نسبة الخطأ ليس لها معنى في حالات السلم الاصطناعي، مثلا درجة الحرارة على سلم سيلسيوس أو فهرنهايت (Hyndman, 2014). من جهة أخرى فأن هذا المؤشر يضخم الخطأ عندما يكون سالبا، وهناك طرق لتلافي ذلك لكنها جدلية ولا تحل المشكلة نهائيا – أنظر (Hyndman, 2014).

3-3. حساب مؤشرات التقدير في *EXCEL*

في **Excel** الدوال التالية تستخدم في Excel لحساب مؤشرات التقدير ومؤشرات التشتت.

القيمة المطلقة	ABS()
التربيع	…^2
التباين	VAR.P()
الانحراف المعياري للمجتمع	STDEV.P()
الجذر	SQRT()
مجموع المربعات	SUMSQ()
مجموع مربعات الفروق	SUMXMY2(array_x;array_y)

مثال: إستخدم Excel لحساب MSE وRMSE في مثال المطعم، ثم أحسب MAE، ثم MAPE.

- لحساب MAE نحتاج لإدراج عمود لحساب القيم الطلقة للفروق،
- ولحساب MAPE ندرج عمودا لحساب النسب المئوية للفروق بالقيم المطلقة.
- لحساب MSE لا نحتاج لإدراج عمود جديد، يكفي استخدام الدالة =SUMXMY2

الجدول التالي يبين الحسابات المتبعة.

	A	B	C	D	E
1	Day	Y	^y	| y – ^y|	| y – ^y |*100/y
2	81	126	/	/	/
3	82	125	126	=ABS(B3 – C3)	=D3*100/B3
4	83	119	125	=ABS(B4–C4)	=D4*100/B4
5	84	133	119	=ABS(B5 – C5)	=D5*100/B5
6	85	129	133	=ABS(B6 – C6)	=D6*100/B6
7	86	127	129	=ABS(B7 – C7)	=D7*100/B7
8	87	131	127	=ABS(B8 – C8)	=D8*100/B8
9	MSE	=SUMXMY2(B3:B8;C3:C8)/6		=AVERAGE(D3 :D8)	=AVERAGE(E3 :E8)
10	RMSE	=SQRT(SUMXMY2(B3:B8;C3:C8)/6)		MAE	MAPE

النتيجة تأتي كما يلي:

Day	Y	^y	| y – ^y|	| y – ^y |*100/y
81	126	/	/	/
82	125	126	1	0,8
83	119	125	6	5,04202
84	133	119	14	10,5263
85	129	133	4	3,10078
86	127	129	2	1,5748
87	131	127	4	3,05344
MSE	44,833		5,167	4,016
RMSE	6,70		MAE	MAPE

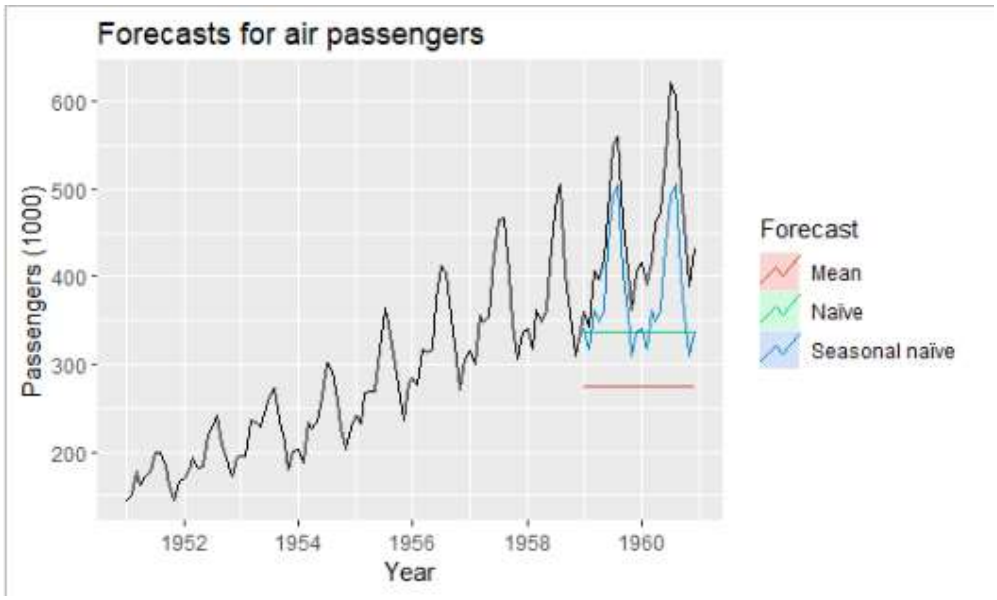
3-4. التمثيل البياني للتقدير وحساب مؤشراته في *R*

لتحديد عينة جزئية من السلسلة نستخدم في R الدالة window() أو الدالة subset(). يسمح R بتمثيل التنبؤ مع السلسلة الأصلية لتفحص مدى مطابقة التقدير للقيم الحقيقية المعلومة.

في المثال التالي نقوم بالتقدير في سلسلة AirPassengers لسنتين بالطرق الثلاث البسيطة: المتوسط، والساذجة والساذجة الموسمية، باستخدام الدوال meanf()، وrwf()، snaive()، ثم نقوم بتمثيل التقديرات الثلاث مع القيم الحقيقية في رسم بياني. قبل ذلك نقوم بتحديد عينة البناء من سنة 1951 إلى سنة 1958.

```
library(fpp2)
AP <- window(AirPassengers,start=1951,end=c(1958,12))
APfit1 <- meanf(AP,h=24)
APfit2 <- rwf(AP,h=24)
APfit3 <- snaive(AP,h=24)
autoplot(window(AirPassengers, start=1951))+
  autolayer(APfit1, series="Mean", PI=FALSE)+
  autolayer(APfit2, series="Naïve", PI=FALSE)+
  autolayer(APfit3, series="Seasonal naïve", PI=FALSE)+
  xlab("Year") + ylab("Passengers (1000)")+
  ggtitle("Forecasts for air passengers")+
  guides(colour=guide_legend(title="Forecast"))
```

Output:

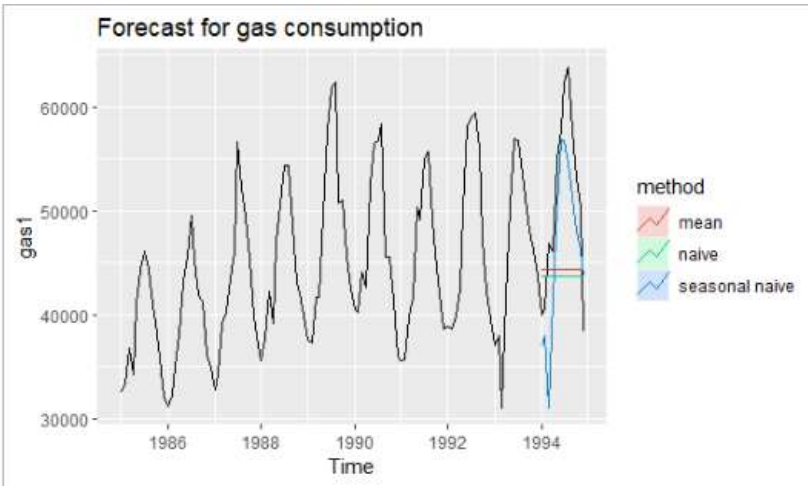


لأن الطرق الثلاثة لا تحتسب التوجه فإنها جميعا غير ملائمة لبيانات AirPassengers كما يتضح ذلك جيدا من الرسم. الطرق البسيطة لا تحتسب الموسمية مع التوجه.

دعونا الآن نأخذ مثالا على بيانات تتضمن الموسمية لكن لا تتضمن توجها ظاهرا، فنأخذ جزءا من بيانات gas من 1985 إلى نهاية 1994، ونسميها gas1. ثم نأخذ من هذه البيانات السنة الأخيرة لنجعلها

عينة تقييم والباقي عينة بناء. يبقى بعد ذلك التمثيل البياني للعينة الكلية ثم تمثيل فوق الرسم ذاته التوقعات للسنة الأخيرة حسب الطرق الثلاث. في الأخير نضبط العنوان العام وعنوان الدليل (guides).

```
gas1<- window(gas,start=1985,end=c(1994,12))+
gas2<- window(gas,start=1985,end=c(1993,12))
autoplot(gas1)+
  autolayer(naive(gas2,12),series="naive",PI=F)+
  autolayer(snaive(gas2,12),series="seasonal naive",PI=F)+
  autolayer(meanf(gas2,12,),series="mean",PI=F)+
  ggtitle("Forecast for gas consumption")+
  guides(colour=guide_legend(title="method"))
```



لحساب مؤشرات الفروق في R نستخدم من حزمة Forecast الدالة accuracy() ونقارن بين نتائج الطرق الثلاث.

```
accuracy(naive(gas2,12))
accuracy(snaive(gas2,12))
accuracy(meanf(gas2,12))
```

مثال. استخرج بيانات goog200 وقم بإظهار التنبؤ الساذج، والمتوسط والانزلاق على آخر 30 يوم.

3. خلاصة

يمكن أن يختار المحلل بين سلة من أدوات التنبؤ البسيطة: التنبؤ البسيط، التنبؤ البسيط مع احتساب الموسمية، التنبؤ البسيط مع احتساب التوجه، التنبؤ باستخدام المتوسط، التنبؤ بطريقة الإنزلاق، وطريقة الخط المستقيم. على بساطتها، تتمتع هذه الأدوات بشهرة كبيرة نظرا لسهولتها وملائمتها لكثير من الحالات.

المعدلات المستخدمة هي كالتالي:

التنبؤ البسيط	$\hat{y}_T(1) = y_T$
التنبؤ البسيط الموسمي	$\hat{y}_T(\mathrm{h}) = y_{T+h-p}$
التنبؤ باستخدام المتوسط	$\hat{y}_T(1) = \frac{1}{k}(y_{T-k+1} + \cdots + y_{T-1} + y_T)$
التنبؤ بطريقة الانزلاق	$\hat{y}_T(\mathrm{h}) = y_T + h\frac{(y_T - y_1)}{T-1}$

تقدم البرمجيات تسهيلات للتنبؤ باستخدام هذه الطرق وتمثيل التنبؤ بيانيا، ورأينا كيفية القيام بذلك في R.

من المفيد لتقييم طريقة التنبؤ استخدام أحد أو بعض مؤشرات الخطأ. تقيس هذه المؤشرات (MAE, MAPE, MSE, RMSE) مقدار الخطأ لدى الطريقة عند تطبيقها على البيانات السابقة. في جميع الأحوال، من المهم تتبع الخطأ في كل السلسلة وليس الاكتفاء بحساب المؤشرات.

الصيغ الرياضية هي كالتالي:

متوسط القيم المطلقة للخطأ:	$MAE = \frac{1}{n}\sum_{t=1}^{n}\lvert E_t\rvert, \quad E_t = y_t - \hat{y}_t$
متوسط مربعات الفروق	$MSE = \frac{1}{n}\sum_{t=1}^{n}(E_t)^2$
متوسط نسبة الخطأ	$MAPE = \frac{1}{n}\sum_{t=1}^{n}(\lvert E_t\rvert \times 100/y_t)$

4. ملحق. مؤشرات أخرى للخطأ والتشتت

هناك مؤشرات أخرى للخطأ أقل شهرة واستخداما، نذكر منها ما يلي:

4-1. متوسط خطأ التنبؤ *MFE*

يقيس هذا المؤشر متوسط الخطأ بدون حذف الإشارة، وله ميزة أنه يحدد ما إذا كانت طريقة التنبؤ تضخم السلسلة – من خلال إشارة موجبة للمؤشر – أم تقللها. عيب المؤشر أن القيم الموجبة تلغي القيم السالبة والعكس، فهو لا يسمح بتقدير حجم الخطأ وإنما فقط اتجاهه.

$$\boldsymbol{MFE = \frac{1}{n}\sum_{t=1}^{n} E_t}$$

2-4. مؤشر تتبع الإشارة *(TS Tracking Signal)*.

يحسب مؤشر تتبع الإشارة عند كل سطر كما يلي:

$$TS_t = \frac{1}{MAE_t}\sum_{t=1}^{t} E_t$$

كقاعدة متبعة: تعد قيمة المؤشر مقبولة بين 4- و 4+.

ميزة هذا المؤشر أنه يبرز الفترات التي يكون فيها خطأ التنبؤ كبيرا، أي خارج المجال المذكور.

مثال: قيم التقدير في المثال السابق من خلال TS.

Day	y	^y	E = y – ^y	∑ (E)	|E|	MAE	TS = ∑ (E)/MAE
81	126	/	/	/	/	/	/
82	125	126	-1	-1	1	1/1 = 1,0	-1/1 = -1,00
83	119	125	-6	-1 – 6 = -7	6	(1 + 6)/2 = 3,5	-7/3.5 = -2,00
84	133	119	14	-7 + 14 = 7	14	(1 + 6 + 14)/3 = 7,0	7/7 = 1,00
85	129	133	-4	3	4	(1 + 6 + 14 + 4)/4 = 6,3	3/6.3 = 0,48
86	127	129	-2	1	2	5,4	0,19
87	131	127	4	5	4	5,2	0,97
			MFE = 1/6				

نلاحظ أن التقدير في نهاية السلسلة أفضل منه في بدايتها، وعموما لم يخرج عن المجال بين 4 و-4.

3-4. التباين والانحراف المعياري

من بين المؤشرات المستخدمة أيضا، مؤشرات التشتت للسلسلة المحولة: التباين، والانحراف المعياري. تستخدم هذه المؤشرات في قياس مستوى التمهيد، فكلما كان التمهيد أقوى يكون تباينه والانحراف المعياري أقل. يحسب التباين V(^y) كما يلي:

$$V(\hat{y}) = \frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - m_{\hat{y}})^2$$

الانحراف المعياري هو جذر التباين:

$$S(\hat{y}) = \sqrt{V(\hat{y})}$$

مثال: يقدر مسير مطعم يوميا كمية الخبز التي يحتاجها ليومه من خلال متوسط استهلاك اليومين السابقين (طريقة المتوسط). لديك الاستهلاك خلال آخر 7 أيام.

أكتب المبيعات المقدرة، وقيم مستوى التمهيد من خلال التباين والانحراف المعياري.

قارن مع مستوى التمهيد عند استخدام الطريقة البسيطة (التنبؤ بالقيمة السابقة).

Day	y	^y	$(\hat{y} - m_{\hat{y}})^2$
81	126	/	/
82	125	(126+125)/2=125,5	(125,5 – 126,92)²= 2,01
83	119	(125+119)/2 = 122	(122 – 126,92)² = 24,17
84	133	126	(126 – 126,92)² = 0,84
85	129	131	(131 – 126,92)² = 16,67
86	127	128	(128 – 126,92)² = 1,17
87	131	129	(129 – 126,92)² = 4,34
	sum	761,5	Sum 49,2
	m^y	761/6 = 126,92	V(^y) 49,2/6 = 8,2

$$S(\hat{y}) = \sqrt{8.2} = 2.86$$

Day	Y	$\hat{y}_t = y_{t-1}$	$(\hat{y} - m_{\hat{y}})^2$
81	126	/	/
82	125	126	0,25
83	119	125	2,25
84	133	119	56,25
85	129	133	42,25
86	127	129	6,25
87	131	127	0,25
		$m_{\hat{y}}$ 759/6 = 126,5	V(^y) = 322.5/6 = 107,5
			S(^y) = 107.5 ½ = 10,368

من الواضح أن التقدير بطريقة المتوسط أكثر تمهيدا وهذا مفهوم لأنه يستخدم متوسطات وليس القيم الحقيقية.

5. سلسلة تمارين: التنبؤ البسيط

5-1. التمارين

تمرين 1. مراجعة

- أكتب صيغ التنبؤ البسيط، والتنبؤ البسيط مع احتساب الموسمية.
- هل يمكن للطريقة البسيطة في التنبؤ أن تحتسب التوجه؟ كيف؟
- ما هي ميزة طريقة **التنبؤ باستخدام** المتوسط عن الطريقة الساذجة؟
- أذكر حالة تكون فيها الطريقة البسيطة أنسب من طريقة المتوسط؟
- أكتب صيغ كل من مؤشرات الفروق والتشتت.
- ما هو الفرق بين مؤشرات الفروق ومؤشرات التشتت؟

تمرين 2. التنبؤ البسيط

لديك البيانات التالية لعدد الغيابات أسبوعيا لشركة وطنية.

الأسبوع	1	2	3	4	5	6	7	8	9	10	11	12
عدد الغيابات	120	127	129	130	110	130	132	120	124	110	132	130

قم بالتمثيل البياني للسلسلة.

- هل تصلح طريقة التنبؤ البسيط لهذه السلسلة؟ برر.
- قم بالتنبؤ للأسبوع المقبل (رقم 13) بطريقة التنبؤ البسيط.

تمرين 3. التنبؤ باستخدام المتوسط

إستخدم بيانات التمرين السابق للتنبؤ ل3 أسابيع بطريقة المتوسط.

قم بالتمثيل البياني للتنبؤ مع السلسلة وناقش إن كانت الطريقة ملائمة للبيانات.

تمرين 4. التنبؤ البسيط مع إدراج الموسمية

لديك البيانات التالية للمبيعات اليومية.

اليوم	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس
المبيعات	129	127	120	114	90	130	132	120	114	95	132	130	122	115	92

- قم بالتنبؤ لأيام الأسبوع المقبل باستخدام التنبؤ البسيط مع إدراج الموسمية.
- قم بالتمثيل البياني للسلسلة مع التنبؤ وناقش ملائمة الطريقة.

تمرين 5. طريقة الانزلاق

لديك السلسلة التالية لعدد المواليد شهريا في عيادة أمومة.

إستخدم طريقة الانزلاق للتنبؤ بعدد المواليد في الأشهر 13 و 14 و 15.

قم بتمثيل السلسلة وناقش صلاحية طريقة الانزلاق للتنبؤ لهذه البيانات.

الأشهر	1	2	3	4	5	6	7	8	9	10	11	12
عدد الولادات	100	107	108	119	119	116	123	134	127	141	131	135

تمرين 6. مؤشرات الخطأ

استخدم بيانات التمرين الثاني (المبيعات اليومية) للإجابة على ما يلي:

قم بتمهيد السلسلة بطريقة التنبؤ البسيط ثم:

- أحسب مؤشرات الخطأ MAE, MSE, RMSE, MAPE

- أحسب مؤشرات التشتت للسلاسل المحولة مع التعليق.
- أظهر السلسلة مع التمهيد في رسم بياني وعلق عليه.
- أظهر الخطأ في رسم بياني وعلق عليه.

2-5. الحلول

6. مراجع الفصل

Anderson, S. W. (2007). *Statistiques pour l'économie et la gestion* (éd. 2). (A. David R., W. Dennis J., & A. Thomas A., Trads.) Bruxelles: De Booeck.

Droesbeck, J. J. (1997). *Eléments de statistique.* Belgique: Ellips.

Hyndman, R. J. (2014). *Forecasting:Principles & Practice.* University of Western Australia. Retrieved from robjhyndman.com/uwa

Malhotra, N., Décaudin, J.-M., & Bouguerra, A. (2007). *Etude Marketing avec SPSS* (éd. 5). Paris: Pearson.

Ritzman Larry et al. (2004). *Management des opérations, principes et application.* Paris: Pearson Education.

zonebourse. (2019). *SAVOIR INVESTIR.* Retrieved 10 18, 2019, from zonebourse: https://www.zonebourse.com