stc tv

التنبؤ بعدد المشاهدات وتحديد أوقات الذروة

بناء نموذج سهل وبسيط للتنبؤ بإجمالي ساعات المشاهدة خلال الشهرين القادمين
وتحديد أوقات الذروة المحتملة (Jawwy Dataset)

تقرير تحليل بيانات وتنبؤ — مؤسسة مسك / مهارات Skills

أولًا: نظرة عامة على البيانات

1. شكل البيانات

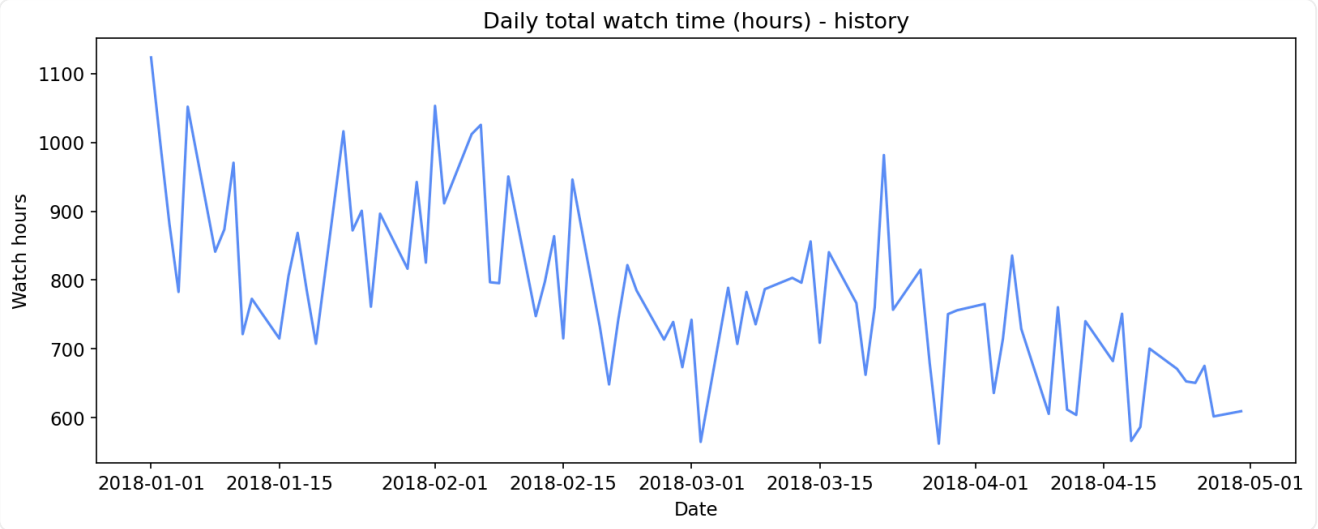
يحتوي ملف البيانات على **86** صفًا في **2** عمودًا، يمثلان إجمالي ساعات المشاهدة اليومية (`Total_watch_time_in_houres`) للفترة من **01-01-2018** إلى **30-04-2018**. تجدر الإشارة إلى أن البيانات تشمل أيام الأسبوع من الاثنين إلى الجمعة فقط (بدون عطلة نهاية الأسبوع)، وهو ما يجب أخذه بالاعتبار عند بناء أي تنبؤ مستقبلي ليبقى متوافقًا مع طريقة جمع البيانات الأصلية.

2. فحص البيانات الفارغة

Column	Null count
date_	0
Total_watch_time_in_houres	0

لا توجد أي قيم فارغة في البيانات.

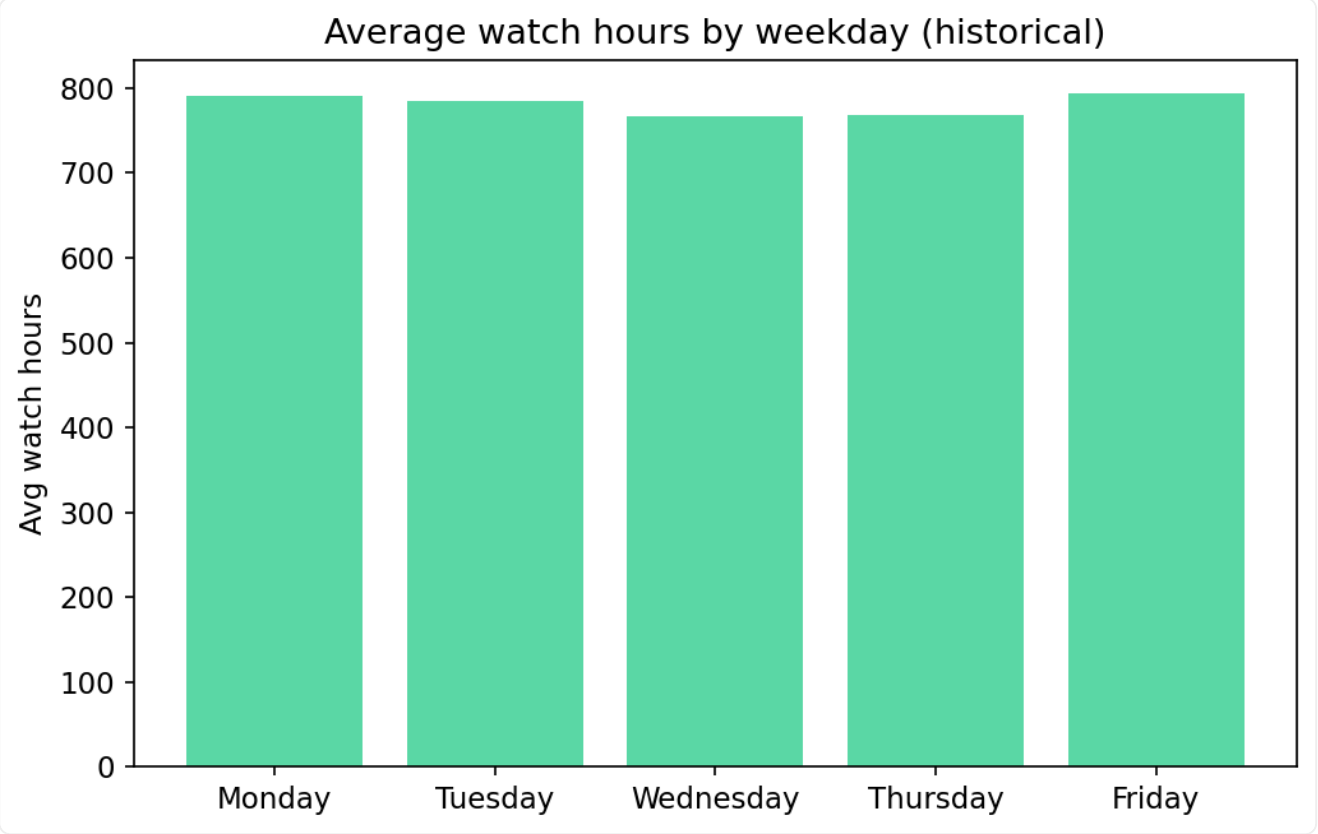
3. السلسلة الزمنية التاريخية



يظهر الرسم اتجاهًا تنازليًا واضحًا في إجمالي ساعات المشاهدة طوال الفترة المرصودة، مع تذبذب أسبوعي منتظم (نمط متكرر صعودًا وهبوطًا كل أسبوع).

4. متوسط المشاهدة حسب يوم الأسبوع

Weekday	Avg watch hours
Monday	790.2
Tuesday	785.2
Wednesday	766.3
Thursday	767.9
Friday	793.9



يوم **Friday** هو الأعلى تاريخيًا بمتوسط **793.9** ساعة، بينما يوم **Wednesday** هو الأدنى بمتوسط **766.3** ساعة. الفروقات بين الأيام طفيفة نسبيًا لكنها منتظمة ومتكررة (نمط أسبوعي ثابت).

ثانيًا: بناء نموذج التنبؤ

تم اختيار نموذج **انحدار خطي (Linear Regression)** لبساطته وسهولة تفسيره: يفترض النموذج أن إجمالي ساعات المشاهدة اليومية = اتجاه عام مع الزمن (Trend) + تأثير ثابت لكل يوم من أيام الأسبوع (Weekly Seasonality). هذا يكفي لرصد الاتجاه العام وأوقات الذروة الأسبوعية المتكررة، وهو نموذج يمكن لأي مسؤول غير تقني فهمه (لكل متغير معامل واحد واضح المعنى).

1. تقييم جودة النموذج

تم اختبار النموذج بشكل صادق على آخر 10 أيام عمل لم يرها أثناء التدريب (Holdout):

R^2 (بيانات التدريب)	RMSE	MAE (خطأ متوسط مطلق)
0.327	**56.09**	**43.57**

خطأ متوسط مطلق (MAE) بقيمة 43.57 ساعة يُعد جيدًا جدًا مقارنة بمتوسط المشاهدة اليومي (≈780 ساعة)، أي بنسبة خطأ نسبية أقل من 6%.

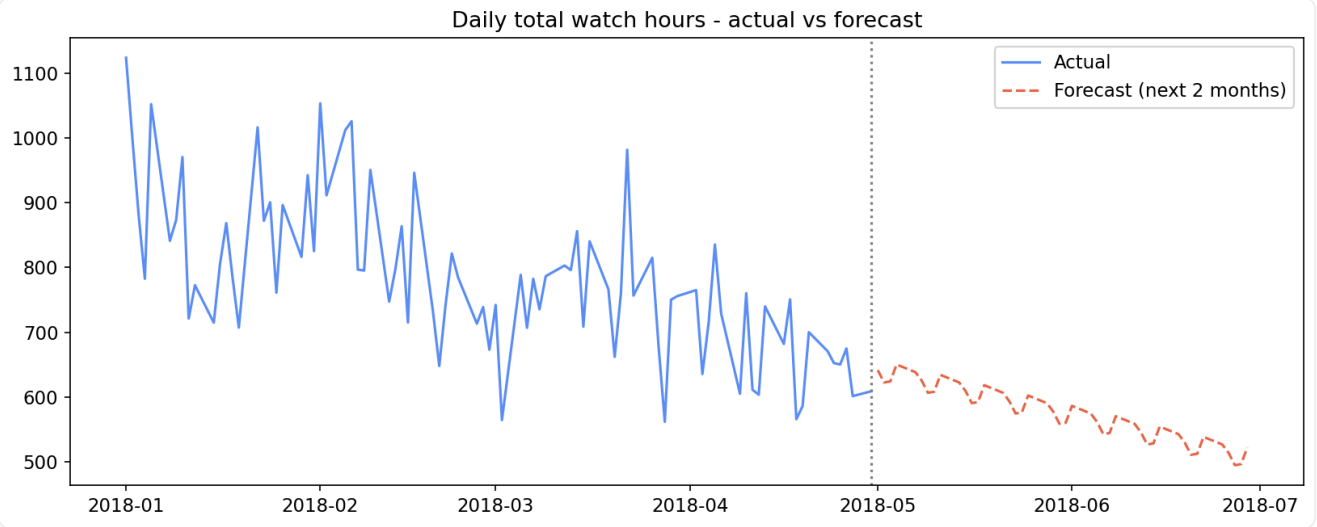
2. معاملات النموذج (Coefficients)

Variable	Effect on watch hours
t	-3.19
wd_1	-9.72
wd_2	-25.49
wd_3	-20.65
wd_4	8.54

القيمة الثابتة (Intercept) = 925.7 تمثل القيمة المتوقعة عند بداية البيانات في يوم الإثنين. معامل `t` السالب يعكس الاتجاه التنازلي العام بمرور الوقت، بينما تمثل بقية المعاملات الفرق عن يوم الإثنين لكل يوم آخر من أيام الأسبوع.

ثالثًا: نتائج التنبؤ للشهرين القادمين

فترة التنبؤ: من **01-05-2018** إلى **29-06-2018** (44 يوم عمل، بما يتوافق مع نمط البيانات التاريخية أيام الاثنين-الجمعة فقط).



متوسط آخر 30 يوم فعلي	متوسط الساعات المتوقع / يوم	إجمالي الساعات المتوقعة (شهرين)
673.7	**573.2**	**25,222**

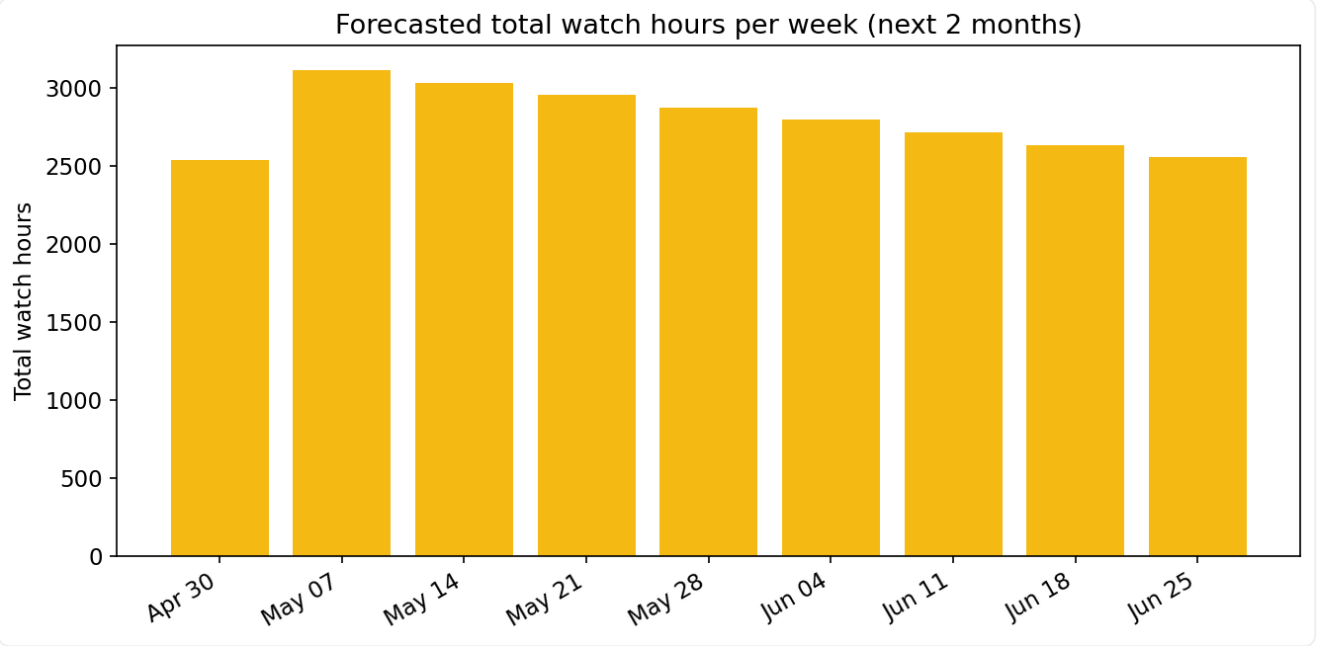
التغير عن آخر شهر فعلي	متوسط الشهر الثاني من التنبؤ	متوسط الشهر الأول من التنبؤ
-9.8%	**538.7**	**607.8**

1. أفضل 5 أيام متوقعة (أعلى مشاهدة - أوقات الذروة المحتملة)

Date	Weekday	Predicted watch hours
2018-05-04	Friday	650.4
2018-05-01	Tuesday	641.7
2018-05-07	Monday	638.7
2018-05-11	Friday	634.5
2018-05-08	Tuesday	625.8

2. ملخص أسبوعي للتنبؤ (الأداة الأسهل لقراءة المسؤولين)

Week starting	Avg/day	Total hours
2018-04-30	634.8	2539.4
2018-05-07	622.9	3114.3
2018-05-14	606.9	3034.6
2018-05-21	591.0	2954.9
2018-05-28	575.0	2875.1
2018-06-04	559.1	2795.4
2018-06-11	543.1	2715.7
2018-06-18	527.2	2636.0
2018-06-25	511.3	2556.3



نتائج بناء النموذج والتنبؤ

- **جودة النموذج**: على بيانات الاختبار، حقق النموذج خطأ متوسط مطلق (MAE) ≈43.57 ساعة فقط، بنسبة خطأ نسبية أقل من 6% من المتوسط اليومي.
- **الاتجاه العام**: اتجاه تنازلي واضح طوال فترة البيانات، يستمر النموذج بعكسه في تنبؤاته للشهرين القادمين.
- **التنبؤ للشهرين القادمين**: إجمالي متوقع **25,222 ساعة** مشاهدة خلال 44 يوم عمل قادم، بمتوسط **573.2 ساعة/يوم**، بانخفاض نسبته ≈9.8% عن متوسط آخر 30 يومًا، مع استمرار التراجع داخل فترة التنبؤ نفسها بنسبة إضافية ≈11.4% بين الشهر الأول والثاني.
- **أوقات الذروة المحتملة**: على مستوى الأسبوع، يوم **Friday** هو الأعلى تاريخيًا، بينما **Wednesday** هو الأدنى، ويُتوقع استمرار هذا النمط الأسبوعي. على مستوى الشهرين القادمين، تتركز الذروة المتوقعة في **الأسابيع الأولى من فترة التنبؤ** (الأقرب من الآن) بسبب

الاتجاه التنازلي العام، بينما تكون أدنى نقاط المشاهدة المتوقعة في نهاية فترة التنبؤ.

رابعًا: الخلاصة والتوصيات

- النموذج المستخدم (انحدار خطي بسيط يجمع الاتجاه الزمني وتأثير يوم الأسبوع) يوفر للمسؤولين أداة سهلة وسريعة لتقدير إجمالي ساعات المشاهدة المتوقعة لأي يوم أو أسبوع قادم، بدقة معقولة.
- الاتجاه التنازلي المستمر يستدعي تنبيه فرق المحتوى والتسويق لاتخاذ إجراءات لتحفيز المشاهدة (حملات تسويقية أو إصدارات محتوى جديد) قبل دخول فترة الانخفاض المتوقعة.
- صيانة الأنظمة أو إطلاق التحديثات التقنية يُفضّل أن تتم في أيام منخفضة الذروة (الأربعاء/الخميس)، بينما يجب تخصيص أكبر سعة لخوادم البث في أيام الذروة (الجمعة والإثنين)، خصوصًا في الأسابيع الأولى من فترة التنبؤ.
- يُنصح بإعادة تدريب النموذج بشكل دوري (شهريًا) كلما توفرت بيانات فعلية جديدة، لضمان مواكبة التنبؤ لأي تغيّر مستقبلي في سلوك المستخدمين لم يكن ظاهرًا في بيانات يناير-أبريل.

تم إعداد هذا التقرير بالاستعناد إلى تحليل بيانات stc TV Data Set T2 باستخدام Python (pandas, scikit-learn, matplotlib, plotly) ضمن بيئة Google Colab، وهو يلخّص الكود والنتائج الواردة في ملف stc_TV_T2_completed.ipynb.