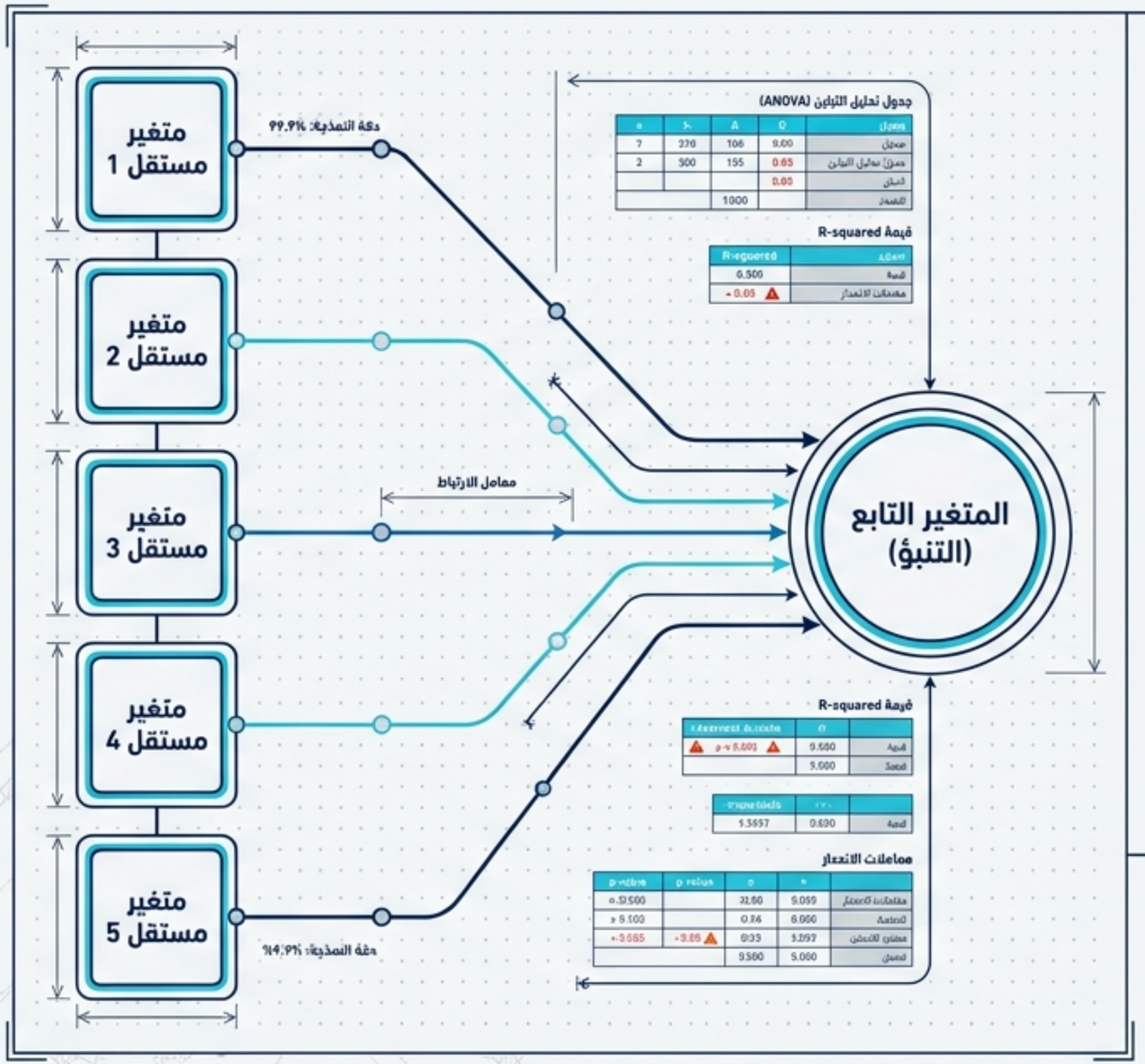


المخطط المعماري للعلاقات التنبؤية: الانحدار الخطي المتعدد

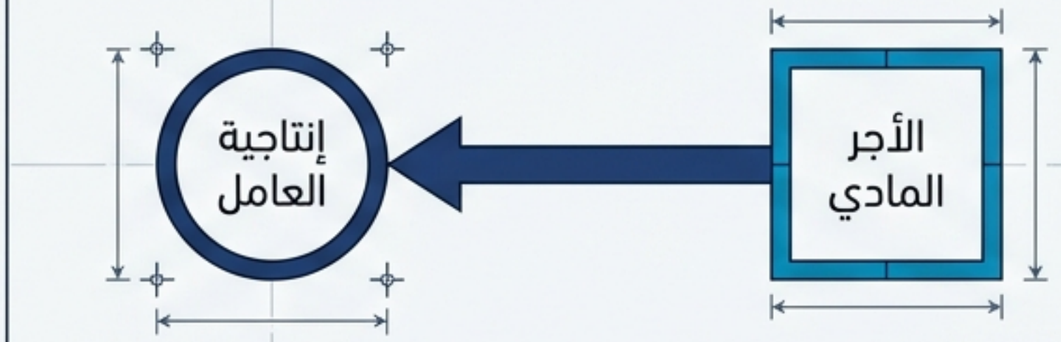
دليل تفصيلي ومرئي لبناء وتقييم النماذج الإحصائية

من إعداد الدكتورة سامية بوكحيل

NotebookLM

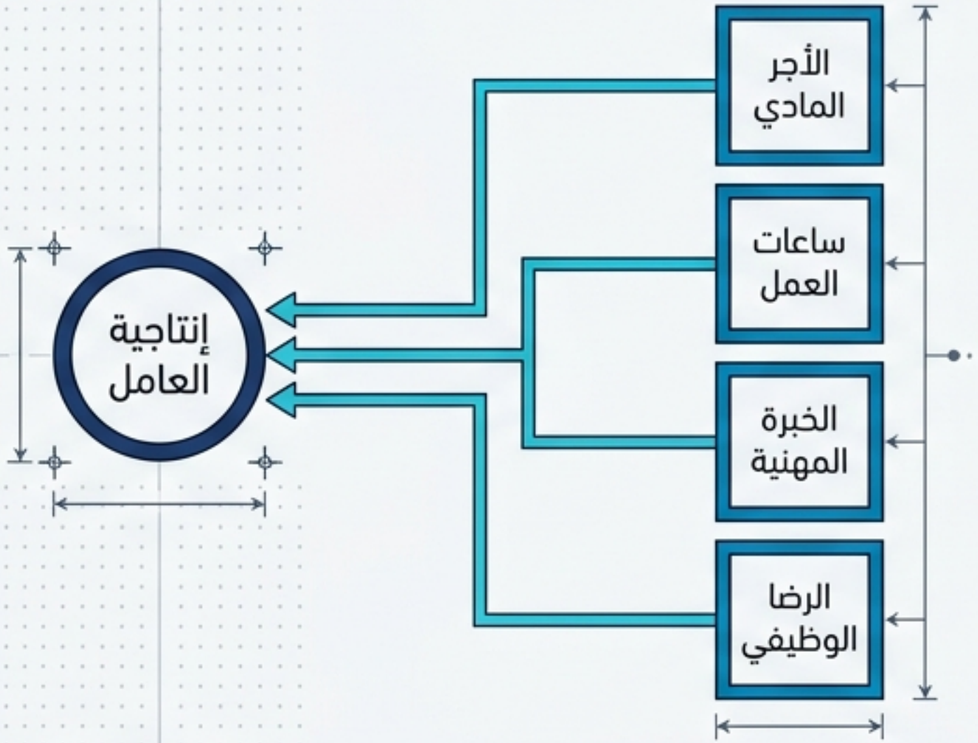
من البعد الواحد إلى النظام الشامل

الانحدار الخطي البسيط



الانحدار الخطي البسيط: دراسة تأثير متغير مستقل واحد على متغير تابع.

الانحدار الخطي المتعدد



الانحدار الخطي المتعدد: قياس التأثير المشترك لعدة متغيرات مستقلة (X_1, X_2, X_1) على ظاهرة واحدة (Y).

NotebookLM

الأهداف الأساسية للنموذج

شرح العلاقة السببية

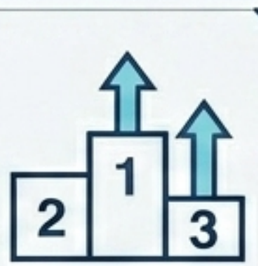
(إيجاد معادلة رياضية تعبر بدقة عن تأثير المتغيرات المستقلة مجتمعة على المتغير التابع).

التنبؤ المستقبلي

(استخدام البيانات الحالية لتقدير وتنبؤ قيم مستقبلية للظاهرة المدروسة).

ترتيب الأهمية

(معرفة أي المتغيرات المستقلة هو الأكثر تأثيراً والأقوى ارتباطاً بالمتغير التابع).



NotebookLM

مصفوفة التشخيص الأولى: متطلبات البيانات

✔ **حجم العينة (Sample Size)**	يجب أن يكون حجم العينة 4 أضعاف عدد المتغيرات المستقلة على الأقل ($N \geq 4K$).	
✔ **طبيعة المتغير التابع (Dependent Variable)**	يجب أن يكون متصلاً (كمي، نسبي، أو فتري). المتغيرات المستقلة يمكن أن تكون كمية أو رتبية/اسمية مرمزّة.	
✔ **العشوائية والاستقلالية (Randomization)**	اختيار العينة عشوائياً وضمان استقلالية درجات كل فرد عن الآخرين.	

NotebookLM

مصفوفة التشخيص الثانية: سلامة العلاقات

- **الارتباط الخطي (Linearity):**

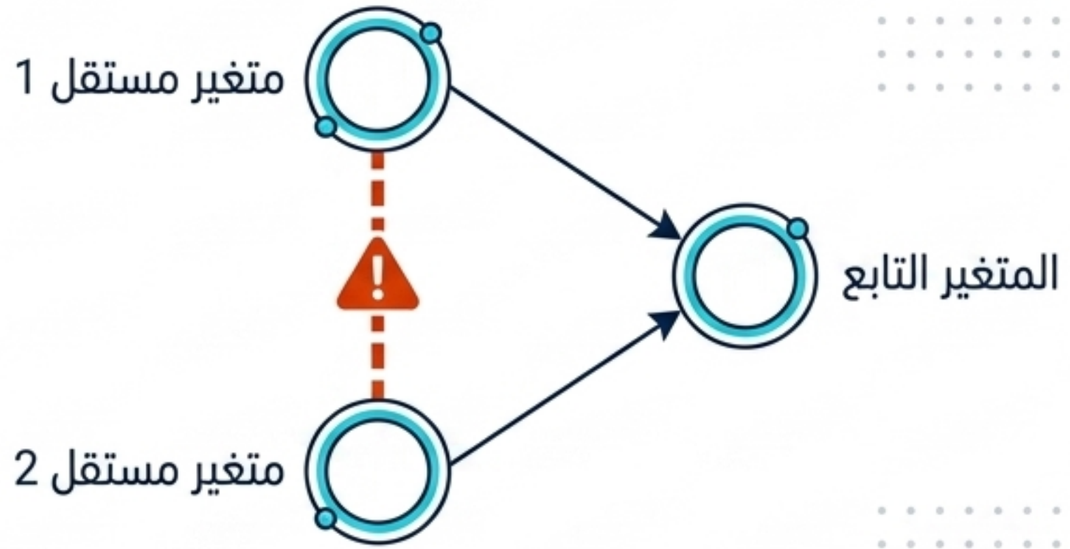
  ضرورة وجود علاقة خطية قوية بين كل متغير مستقل والمتغير التابع (Pearson > 0).

- **التباين (Variance)**

  تباين كل متغير مستقل يجب أن يكون أكبر من الصفر (ليس ثابتاً).

- **غياب التداخل الخطي (No Multicollinearity):**

  يجب ألا توجد علاقات ارتباطية قوية بين المتغيرات المستقلة نفسها.

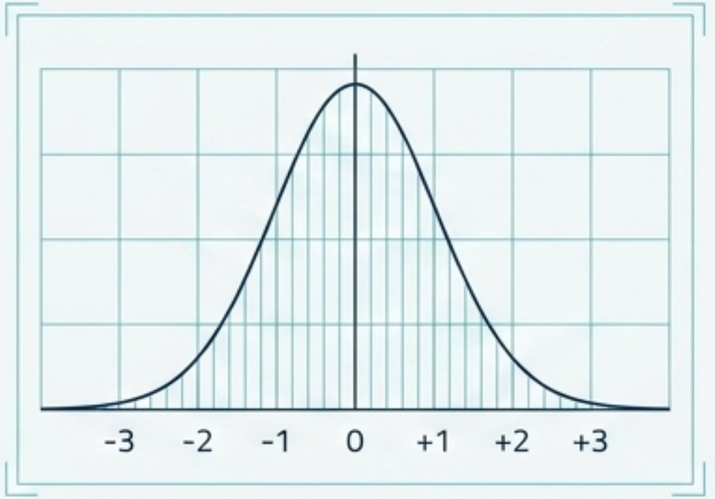


يتم قياس التداخل عبر **معامل تضخم التباين (VIF)**.
إذا كان $VIF \geq 10$، يوجد ازدواج خطي يهدد النموذج.
(الحل: حذف متغير، أو دمج الأبعاد، أو زيادة العينة).

NotebookLM

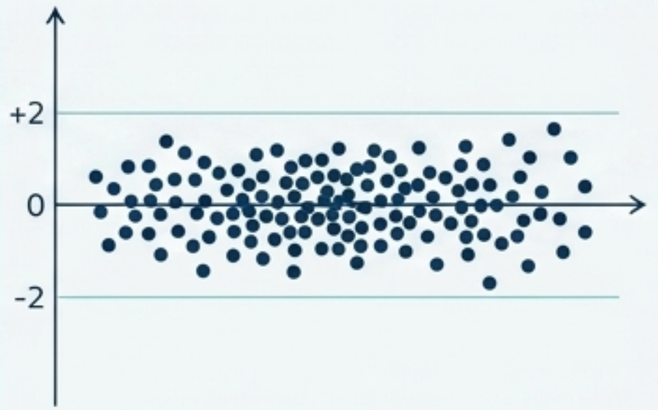
مصفوفة التشخيص الثالثة: فحص البواقي (الأخطاء العشوائية)

الاعتدالية (Normality)



تتوزع البواقي توزيعاً طبيعياً (متوسطها الحسابي = 0).

ثبات التباين (Homoscedasticity)



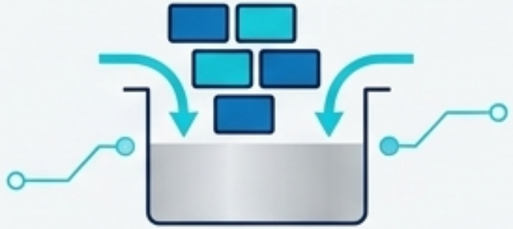
انتشار البواقي يكون متساوياً حول الصفر (95% منها تقع بنع بن -2- و +2).

غياب الارتباط الذاتي (No Autocorrelation)

استقلالية البواقي عن بعضها
(تختبر باستخدام مقياس ديربين -واتسون Durbin-Watson، حيث تقع القيمة الآمنة في المنطقة الوسطى بعيداً عن 0 و 4).

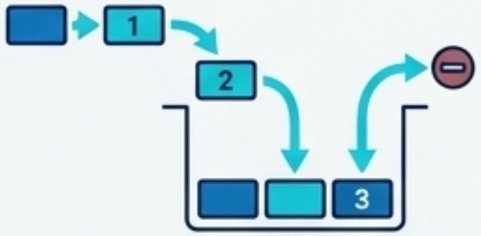
NotebookLM

هندسة النموذج: طرق إدخال المتغيرات المستقلة



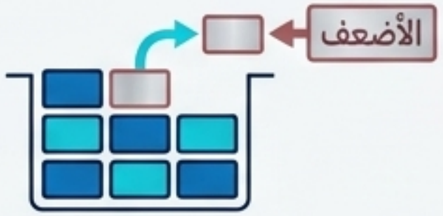
1. الطريقة المباشرة (Enter)

إدخال جميع المتغيرات دفعة واحدة (تستخدم حين لا يعرف الباحث أيهم الأهم).



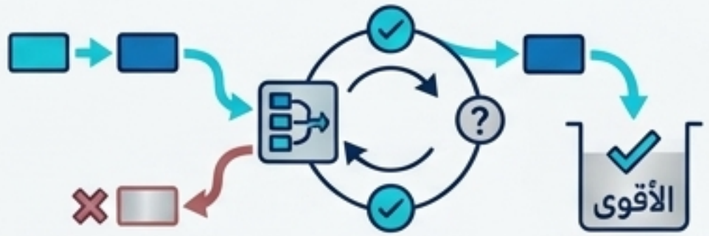
2. الإضافة المتتالية (Forward)

إدخال المتغير الأقوى ارتباطاً، ثم الذي يليه، والتوقف إذا لم يضف المتغير الجديد قوة للنموذج.



3. الحذف المتتالي (Backward)

إدخال الكل، ثم حذف المتغير الأضعف تأثيراً تدريجياً ليبقى النموذج الأمثل.

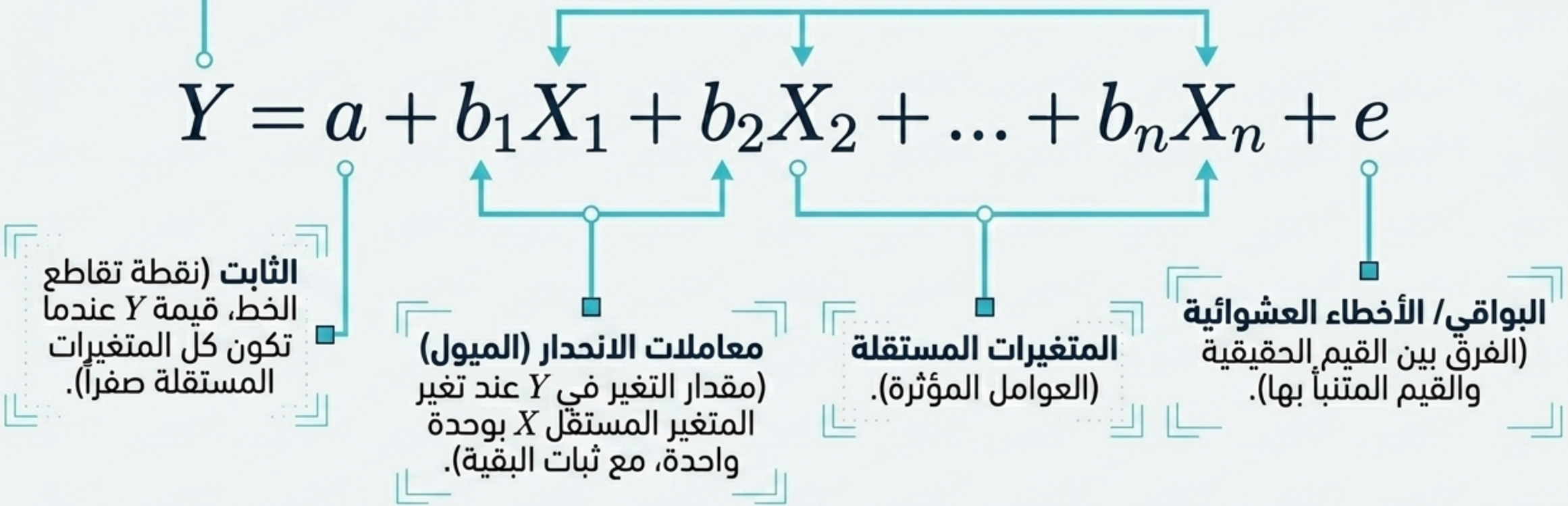


4. الخطوات المتتالية (Stepwise)

تجمع بين الإضافة والحذف، حيث يتم تقييم واختبار المتغيرات في كل خطوة للإبقاء على الأقوى فقط.

NotebookLM

المخطط الرياضي: تفكيك معادلة الانحدار



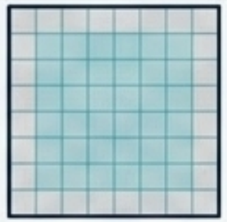
NotebookLM

قوة النموذج: تقييم جودة المطابقة

1.

معامل الارتباط المتعدد (R)

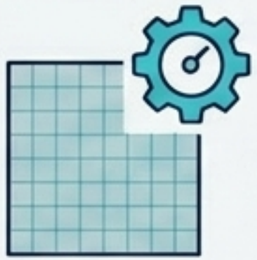
يقيس قوة واتجاه العلاقة بين المتغير التابع وجميع المتغيرات المستقلة مجتمعة (لا يمكن أن يكون سالباً).



2.

معامل التحديد (R^2)

نسبة التباين في المتغير التابع التي تفسرها المتغيرات المستقلة (مثلاً: 0.67 تعني أن النموذج يفسر 67% من الظاهرة).



3.

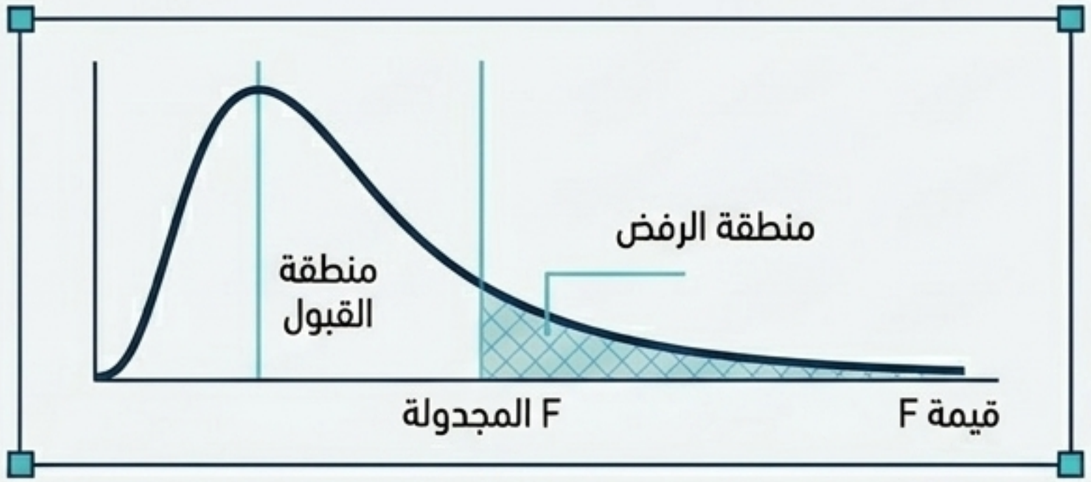
معامل التحديد المصحح (Adjusted R^2)

المعيار الأهم في الانحدار المتعدد. يصحح R^2 بناءً على عدد المتغيرات وحجم العينة، ليعاقب النموذج على إدخال متغيرات غير مفيدة لا تزيد من قوته التفسيرية.

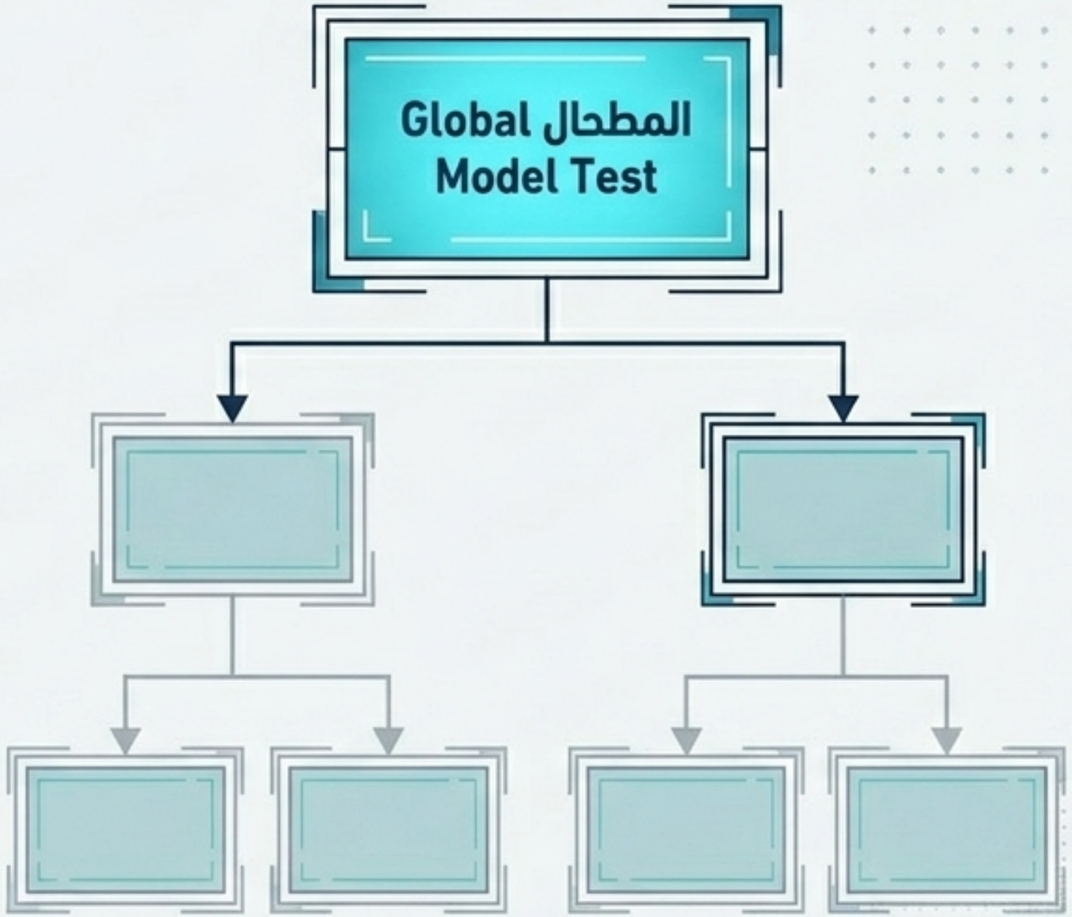
NotebookLM

التسلسل الهرمي للاختبارات (المرحلة 1): الدلالة الإحصائية للنموذج ككل

المحطة الأولى: اختبار فيشر (F-Test) / تحليل التباين (ANOVA)



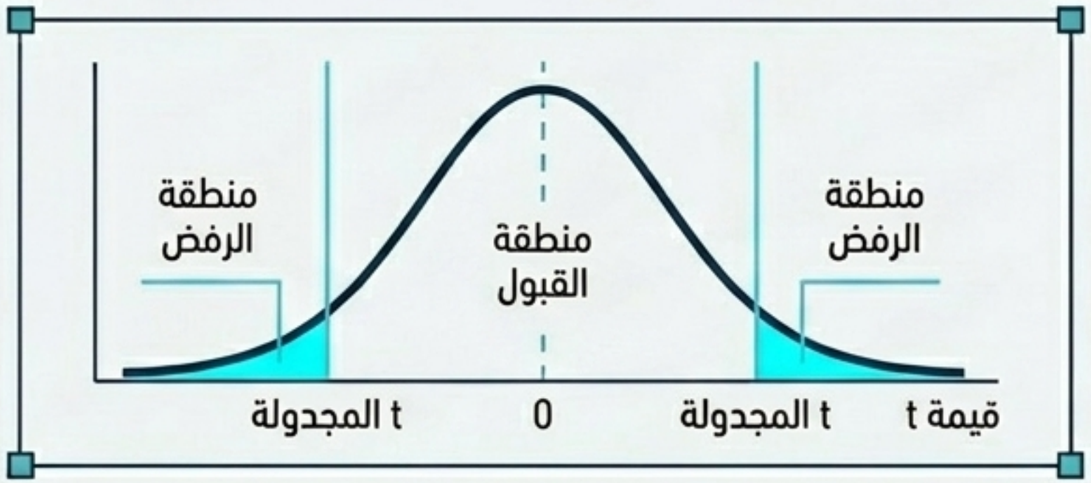
الفرضية الصفرية	1	جميع المعاملات مجتمعة ليس لها تأثير ($b_1 = b_2 = 0$) النموذج غير صالح للتنبؤ.
الفرضية البديلة	2	يوجد معامل واحد على الأقل له تأثير دال إحصائياً. النموذج صالح.
القرار	3	نقارن قيمة F المحسوبة بـ F المجدولة. إذا كانت المحسوبة ≥ **المجدولة، نرفض الفرضية الصفرية وننتقل للمرحلة الثانية.**



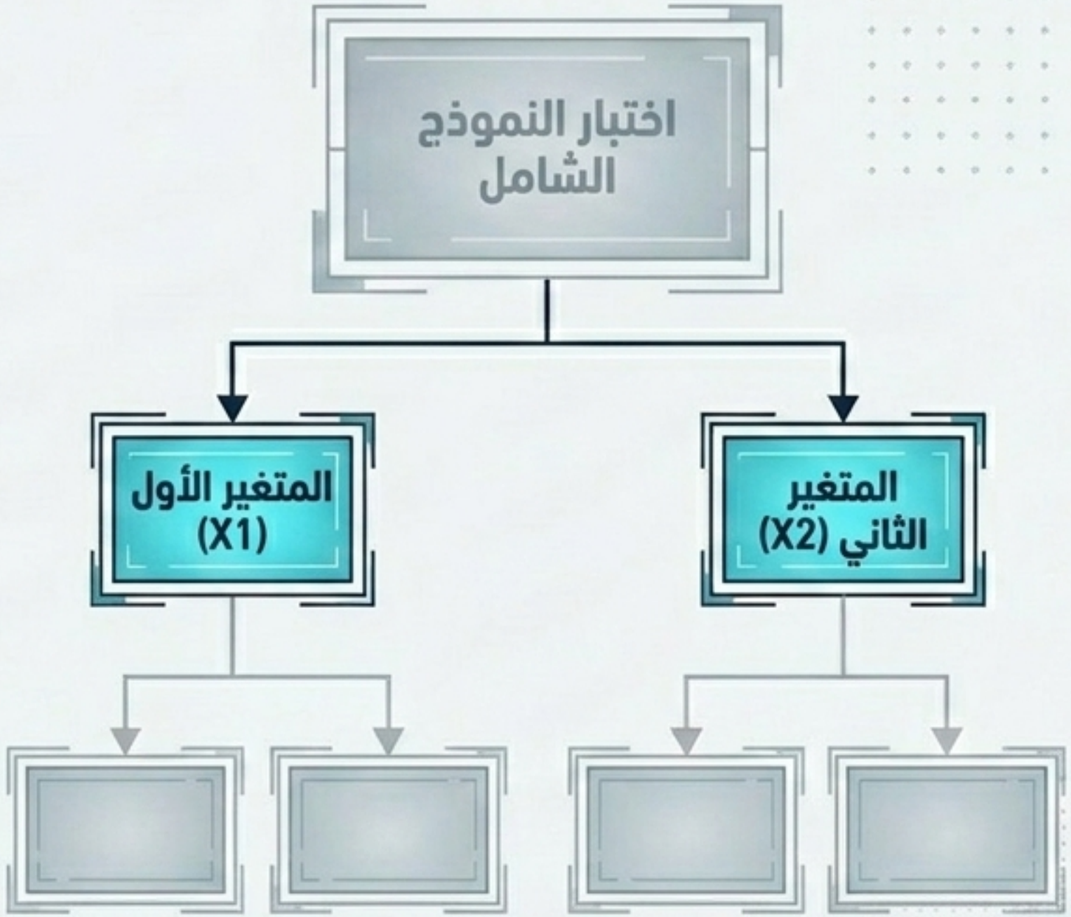
NotebookLM

التسلسل الهرمي للاختبارات (المرحلة 2): **الدلالة الإحصائية لكل متغير**

المحطة الثانية: اختبار ستودنت (t-Test)

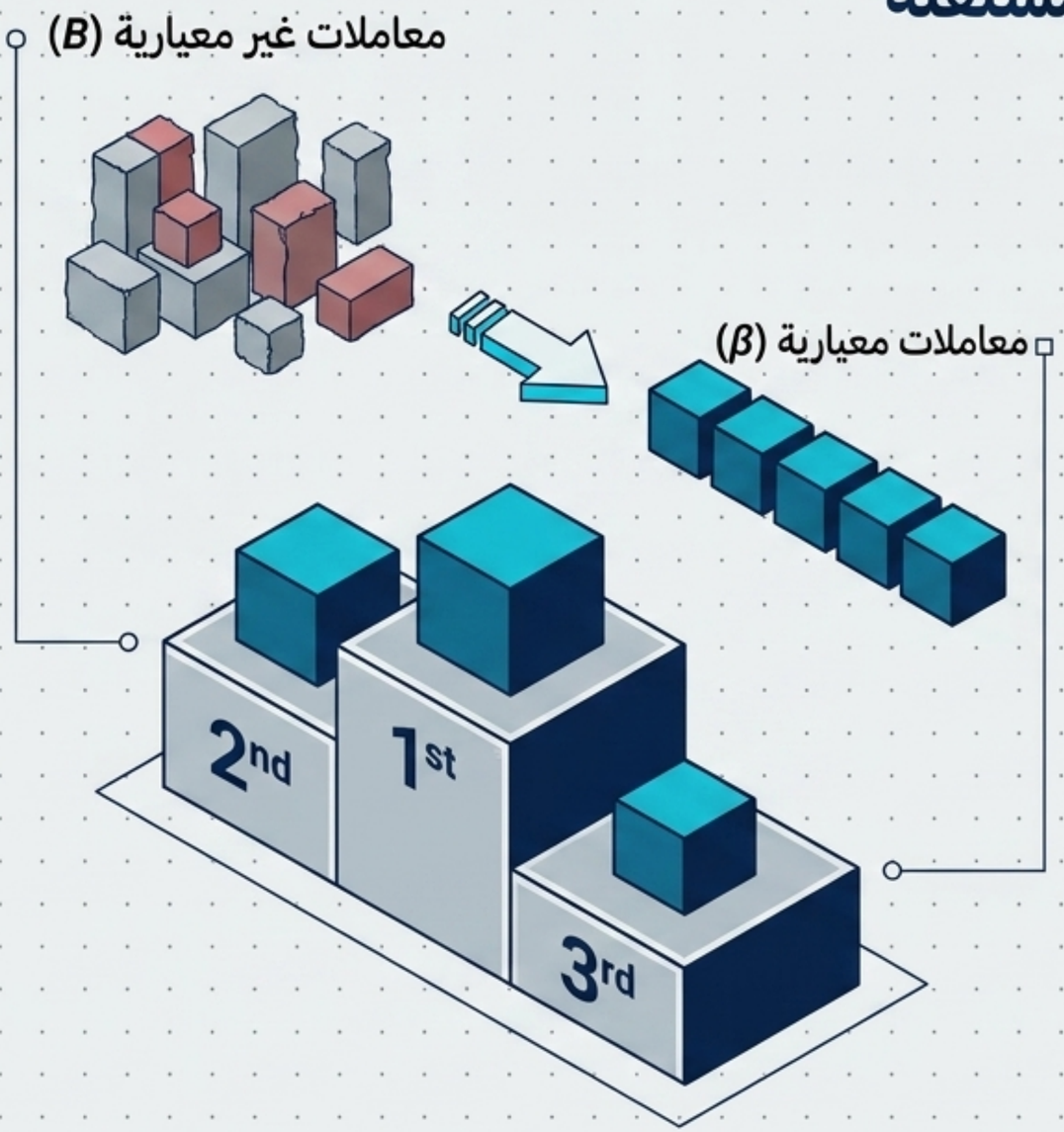


الهدف	إذا كان النموذج ككل دالاً، نختبر كل متغير مستقل على حدة لمعرفة أيها المؤثر الحقيقي.
الفرضية الصفرية	المتغير المستقل المعين ليس له تأثير ($b_i = 0$).
القرار	يتم قسمة معامل الانحدار على خطئه المعياري لحساب t. إذا كانت القيمة المطلقة لـ t المحسوبة $\geq t$ المجدولة، **يُعتبر المتغير دالاً إحصائيًا ويمكنه التنبؤ بالظاهرة.**



NotebookLM

استخلاص الرؤى: ترتيب أهمية المتغيرات المستقلة



لا يمكن مقارنة المعاملات غير المعيارية (*B*) لاختلاف وحدات القياس.
لتحديد المتغير الأقوى تأثيراً، نعتمد على:

1. المعاملات المعيارية (Standardized Beta - β)

المتغير ذو القيمة المطلقة الأكبر لـ β هو الأكثر تأثيراً.

2. قيمة اختبار *t*

القيمة المطلقة الأكبر تشير للتأثير الأقوى.

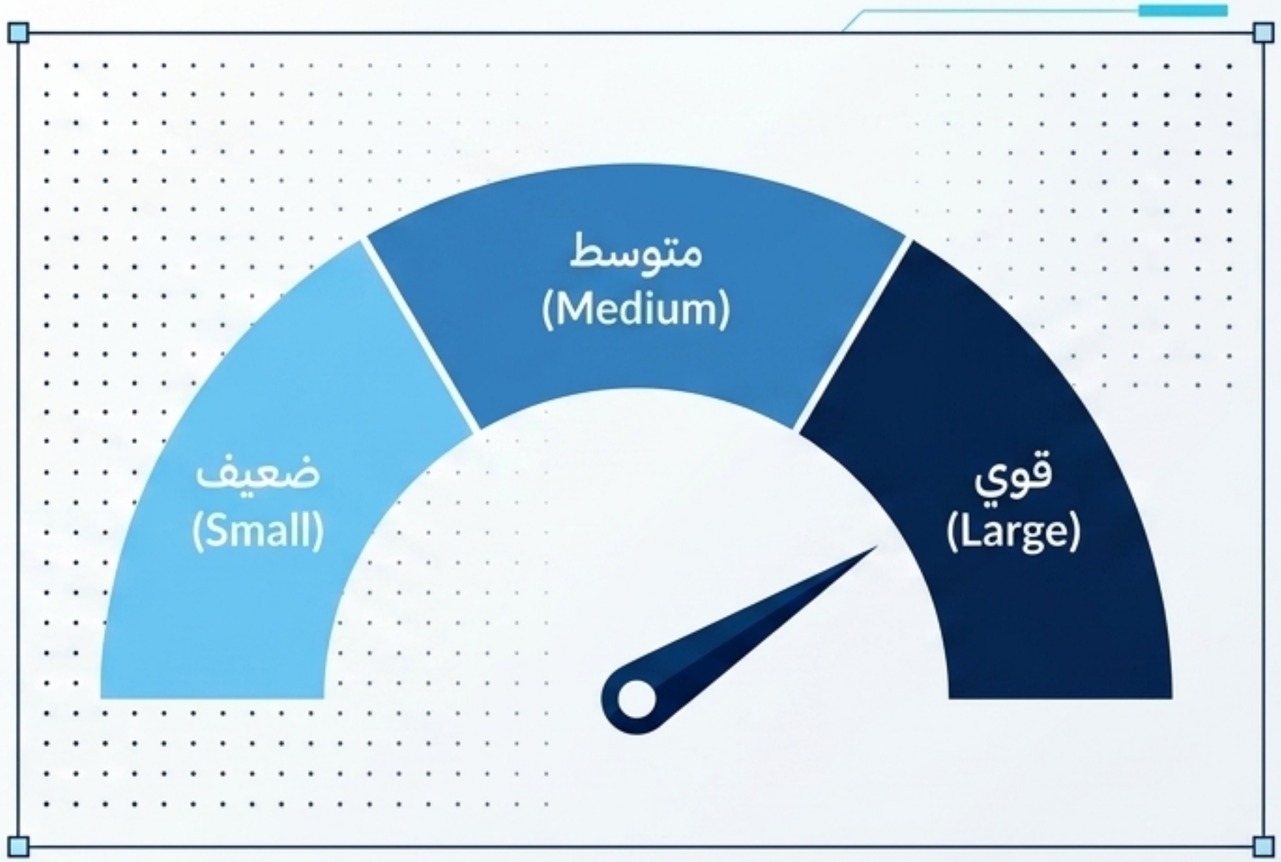
3. القيمة الاحتمالية (P-value)

في البرامج الإحصائية، الدلالة المعنوية الأصغر تعكس قوة تأثير أكبر (طريقة عكسية).

NotebookLM

الدلالة العملية: قياس حجم التأثير

الدلالة الإحصائية تثبت وجود التأثير، لكن الدلالة العملية تقيس حجمه. نعتمد على قيمة معامل التحديد (R^2):



تأثير ضعيف (Small)

$R^2 < 0.09$ (أقل من 9%).

تأثير متوسط (Medium)

$0.09 \leq R^2 < 0.25$ (من 9% إلى 24.9%).

تأثير قوي (Large)

$R^2 \geq 0.25$ (25% فما فوق - للنموذج قدرة تفسيرية وتنبؤية عالية للظاهرة).

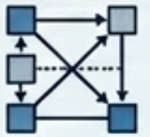
NotebookLM

الخلاصة: سير العمل التحليلي الشامل

1. تدقيق الشروط

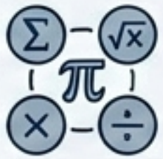
فحص العينة، التداخل الخطي، وتوزيع البواقي (Diagnostics).

2. بناء النموذج



اختيار طريقة إدخال المتغيرات (Enter/Stepwise).

3. حساب المعالم



استخراج R^2 المصحح وتشكيل المعادلة ($Y = a + bX$).

4. اختبار النموذج ككل

الدلالة الكلية عبر تحليل التباين. F-test.

5. تصفية المتغيرات

اختبار الأهمية الفردية عبر t-test.

6. استخراج الرؤى

ترتيب المتغيرات المؤثرة وتحديد حجم التأثير العملي.

NotebookLM