earthUI: واجهة تفاعلية لحزمة earth (شرائح الانحدار التكيّفية المتعددة المتغيرات)

ترجمة عربية أُعدّت بمساعدة الذكاء الاصطناعي وروجعت تحريريًا؛ النسخة الإنجليزية هي المُعتمدة.

William Bert Craytor

Vol. 1, Issue 1 · Summer 2026
نُشر في 1 يوليو 2026
DOI: 10.67330/7fj41b65
2026 William Bert Craytor. Published by Pacific Vista Net. ©
يُنشر هذا المقال وفق شروط رخصة المشاع الإبداعي: النسب إلى المُصنَّف 0.4.

الملخص. earthUI واجهة مستعمل رسومية لحزمة `earth` في لغة R، وهي تنفّذ شرائح الانحدار التكيّفية المتعددة المتغيرات (MARS). تتيح ثلاثة أنماط للاستعمال --- النمذجة التنبؤية العامة، وتقييم العقارات، وتحليل مناطق السوق --- وتقود المستعمل عبر استيراد البيانات، وضبط النموذج، والملاءمة، وتأويل تشخيصات earth ومخرجاته الرسومية. يوثّق هذا المقال متطلبات صيغة البيانات في earthUI، وسير عمل النمذجة، وأشكال المخرجات، والمرجع الكامل لوظائفه.

الكلمات المفتاحية: MARS، شرائح الانحدار التكيّفية المتعددة المتغيرات، earth، R، واجهة مستعمل رسومية، تقييم العقارات

ملاحظة الناشر (يوليو 2026): يصف هذا المقال earthUI الإصدار 0.10.0. وقُبيل النشر، استوعب earthUI الإصدار 0.11.0 التطبيقين المرافقين glmnetUI و mgcvUI (الموصوفين في المقالين التاليين)، فدمج سير عمل MARS والشبكة المرنة والنماذج الجمعية المعمّمة في تطبيقٍ واحد. أما الواجهة وسير العمل الموصوفان هنا فلم يتغيّرا بالدمج.

المحتويات

1 مقدمة **3**
1.1 ما هو earthUI؟ . . . 3
2.1 حزمةٌ واحدة، ثلاثة روتينات . . . 4
3.1 أنماط الغرض الثلاثة . . . 4
4.1 الخصائص الخاصة بالعقارات . . . 4
5.1 تنصيب earthUI وvProlog . . . 5
6.1 البدء . . . 5

2 متطلبات بيانات الإدخال من MLS **6**
1.2 صيغة الملف وبنيته . . . 6
2.2 الأعمدة المطلوبة في نمط التقييم . . . 6
3.2 أعراف تسمية الأعمدة الخاصة . . . 7
4.2 جودة البيانات واكتمالها . . . 7
5.2 موضع صف العقار موضوع التقييم . . . 7

3 النمط العام **8**
1.3 نظرة عامة . . . 8
2.3 سير العمل في الشريط الجانبي . . . 8
3.3 ألسنة اللوحة الرئيسة . . . 8
4.3 بقاء الإعدادات . . . 9
5.3 النمط الداكن . . . 9
6.3 ضبط اللغة والمنطقة . . . 9
1.6.3 البلدان المدعومة . . . 10
2.6.3 تجاوز الورق . . . 10
3.6.3 حفظ الافتراضات . . . 10

4 نمط التقييم **11**
1.4 معالجة صف العقار موضوع التقييم . . . 11
2.4 التاريخ الفعّال وعمر البيعة . . . 11
3.4 تعيينات الأعمدة الخاصة . . . 11
4.4 إسهام الباقي (تقدير الخصائص النادرة) . . . 12
5.4 نظرة عامة على تعديلات RCA . . . 12

5 نمط تحليل منطقة السوق **13**
1.5 الفروق عن نمط التقييم . . . 13
2.5 متى يُستعمل نمط السوق . . . 13

6 انتقاء المتغيرات **14**
1.6 المتغير الهدف أو المتغيرات الهدف . . . 14
2.6 جدول المتنبئات . . . 14
3.6 رصد أنواع البيانات وتجاوزه . . . 14
4.6 عَلَما العامل والخطية . . . 14
5.6 مرجع أنواع الأعمدة الخاصة . . . 15
6.6 الأهداف المتعددة . . . 15

7 انتقاء المعاملات **16**
1.7 المعاملات الأساسية . . . 16
2.7 المرور الأمامي . . . 16
3.7 التفاعلات المسموح بها . . . 16
4.7 التشذيب . . . 16
5.7 التحقق المتقاطع . . . 17
6.7 ترشيح المجموعة الفرعية . . . 17
7.7 قيم المعاملات المُوصى بها . . . 17
1.7.7 معاملات المرور الأمامي . . . 17
2.7.7 معاملات التشذيب . . . 18
3.7.7 معاملات التحقق المتقاطع . . . 19
4.7.7 نموذج التباين . . . 19
5.7.7 ملف مخرجات Earth . . . 19
8.7 افتراضات الإعدادات . . . 19

8 ملاءمة نموذج Earth **21**
1.8 زرّ الملاءمة . . . 21
2.8 نافذة الملاءمة ومخرجات التتبّع . . . 21
3.8 سجلّ الملاءمة . . . 21

9 تنزيل البيانات **22**
1.9 أعمدة المخرجات . . . 22
2.9 ترتيب الأعمدة وتنسيق Excel . . . 22
3.9 درجات CQA . . . 22
4.9 الترتيب وصف العقار موضوع التقييم . . . 23
5.9 التصدير التلقائي إلى mgcvUI . . . 23
6.9 الاستيراد إلى glmnetUI . . . 23
1.6.9 المزايا الرئيسة . . . 23
2.6.9 مهم: يجب أن يتطابق عمود الأوزان . . . 23
3.6.9 سير العمل . . . 24

10 حسابات RCA وتنزيلها **25**
1.10 فتح حوار RCA . . . 25
2.10 استكمال درجة CQA . . . 25
3.10 أعمدة المخرجات . . . 25
4.10 لسان تعديلات RCA . . . 25

11 شبكة المقارنة بالمبيعات **27**
1.11 نظرة عامة والغرض . . . 27
2.11 نافذة انتقاء المقارنات . . . 27
3.11 تخطيط الشبكة . . . 27
4.11 سعر البيع والتنازلات وصافي سعر البيع . . . 28
5.11 الصفوف المجمَّعة (الموقع والموقع والعمر) . . . 28
6.11 صفوف متغيرات النموذج . . . 28
7.11 صف CQA|الباقي وصفوف الخصائص الباقية . . . 28
8.11 صيغة سعر البيع المعدَّل . . . 28
9.11 حماية الأوراق والخلايا القابلة للتحرير . . . 28
10.11 العمل بالشبكة في Excel . . . 29

12 تنزيل التقارير **30**
1.12 توليد تقرير Quarto . . . 30
2.12 تحويل تقرير Quarto . . . 30
3.12 محتويات التقرير . . . 30

13 وحدتا المعالجة اللاحقة: الشبكة المرنة والنماذج الجمعية المعمّمة **32**
1.13 لماذا الملاءمات المُسنِدة؟ . . . 32
2.13 التشغيل والحالة المشتركة . . . 32
3.13 روتين الشبكة المرنة . . . 32
4.13 روتين النماذج الجمعية المعمّمة . . . 32
5.13 سير عملٍ عملي . . . 32

14 مجموعة البيانات التجريبية: Appraisal_1.csv **33**
1.14 الوصف . . . 33
2.14 الأعمدة . . . 33
3.14 بداية سريعة مقترحة . . . 34

15 تجريبي: استخراج النصّ بـProlog وقواعد الاشتقاق **35**
1.15 ماذا تفعل . . . 35
2.15 المتطلبات: vProlog (Trealla)، لا SWI-Prolog . . . 35
3.15 تفعيلها وتشغيلها . . . 35
4.15 قواعد الاشتقاق . . . 36
5.15 ملف البيانات النشط المُصدَّر . . . 36
6.15 الواجهة البرمجية والملفات المرفقة . . . 36
7.15 تجربةٌ ذات صلة: أوليّات الأقاليم المناخية . . . 36
1.7.15 ما وراء كاليفورنيا: حزم الولايات القضائية ووحدات مناطق السوق . . . 37

1 مقدمة

1.1 ما هو earthUI؟

earthUI واجهة مستعمل رسومية لحزمة `earth` في لغة R، وهي تنفّذ شرائح الانحدار التكيّفية المتعددة المتغيرات (MARS) (Friedman 1991). ومحرّك النمذجة هو حزمة `earth` لستيفن ميلبورو (Milborrow 2024)، الموثَّقة تفصيلًا لدى Milborrow

(n.d.[a]) وMilborrow (n.d.[b]). وهي تعمل تطبيقَ Shiny محليًا --- فلا تسجيل دخول، ولا خادم، ولا حسابات. تشغّلها من R، وتستورد مجموعة بيانات (CSV أو Excel)، وتضبط نموذجك، وتلائمه تفاعليًا.

ويوفّر التطبيق سير عملٍ كاملًا: استيراد البيانات، وضبط المتغيرات، وملاءمة النموذج، والرسوم التشخيصية، وأهمية المتغيرات، ومعادلات النموذج، والتقارير القابلة للتنزيل بصيغة HTML أو Word أو PDF.

2.1 حزمةٌ واحدة، ثلاثة روتينات

اعتبارًا من الإصدار الحالي، صارت حزمة earthUI **تحتوي** أيضًا وحدتَي المعالجة اللاحقة فيها، وقد كانتا في الإصدارات السابقة حزمتين منفصلتين هما glmnetUI وmgcvUI. وتنصيبُ earthUI ينصّب الروتينات الثلاثة؛ ويعمل كلٌّ منها تطبيقَ Shiny محليًا خاصًّا به:

- `earthUI::launch()` --- تطبيق earth (MARS) الموصوف في هذا المقال، على المنفذ 7878. و`earth()` هو محرّك التقييم المركزي.
- `earthUI::launch_glmnet()` --- وحدة المعالجة اللاحقة للشبكة المرنة (المنفذ 7879)، وهي تستورد دوال أساس نموذج earth وتعيد تقديرها تحت انتظام لاسو/الشبكة المرنة (انظر ``الاستيراد من glmnetUI" أدناه).
- `earthUI::launch_mgcv()` --- وحدة المعالجة اللاحقة للنماذج الجمعية المعمّمة (المنفذ 7880)، وهي تستعمل مواضع عُقد نموذج earth نقاطَ انطلاقٍ للحدود الملساء في `mgcv` (انظر ``التصدير التلقائي إلى mgcvUI" أدناه).

وتتشارك الروتينات الثلاثة شجرةَ مشاريع واحدة، وقاعدةَ إعداداتٍ واحدة، ونسخةً واحدة من البنية التحتية المساندة، فيكون المشروع المفتوح في أحد التطبيقات متاحًا فورًا للبقية. ووحدتا glmnet وmgcv بيّنةٌ مُسنِدة لنتيجة قيمة نموذج earth، لا سيرَي عملٍ منافسين: فالنمط المقصود هو ملاءمة نموذج earth وصقلُه أولًا، ثم تشغيل الوحدتين على المشروع نفسه متى أُريد التعضيد.

3.1 أنماط الغرض الثلاثة

ينظّم earthUI العمل في **مشاريع** --- وهي كياناتٌ مسمّاة تجمع ملفات إدخالك ومخرجاتك وإعداداتك في موضعٍ ثابت. ولكل مشروع **غرضٌ** يُختار في معالج المشروع الجديد عند إنشائه. والأنماط ثلاثة:

- **العام** --- انحدار Earth لأي نوعٍ من المجتمعات أو مجموعات البيانات. وهو النمط الافتراضي. ويوفّر سير عمل MARS كاملًا من غير أي إضافاتٍ خاصة بمجال.
- **للتقييم** --- انحدار Earth مفصَّلٌ لتقييم العقارات. ويضيف خصائصَ خاصة بتقييم العقار المفرد، منها معالجة العقار موضوع التقييم، وتعيينات الأعمدة الخاصة، ومنهج قيد البواقي (RCA)، وتوليد شبكة المقارنة بالمبيعات.
- **تحليل منطقة السوق** --- انحدار Earth مفصَّلٌ لدراسات مناطق السوق. ويضيف خصائصَ لتحليل مجموعاتٍ من العقارات في سوقٍ محدَّدة بدل تقييم عقارٍ مفرد.

ومحرّك النمذجة الجوهري واحدٌ في الأنماط الثلاثة كلها --- فأنت تلائم دائمًا نموذج Earth (MARS). وإنما يتحكّم ضبط الغرض في الأدوات وعناصر الواجهة الإضافية المتاحة.

والغرض ثابتٌ لكل مشروع. فغرضُ المشروع --- مع بلده وولايته ومقاطعته ومدينته --- يُقفل عند إنشائه. والانتقال إلى مشروعٍ آخر يعيد earthUI إلى حالةٍ افتراضية نظيفة (فتُمسح البيانات المستوردة، ونتائجُ النموذج، والألسنة، وضبطُ المتغيرات، ومعاملاتُ earth جميعًا) ثم يستعيد الضبط المحفوظ لذلك المشروع بحسب غرضه.

4.1 الخصائص الخاصة بالعقارات

حين يُختار **للتقييم** أو **تحليل منطقة السوق**، يفعّل earthUI عدة خصائصَ مصمَّمة لتحليل العقارات:

- **تعيينات الأعمدة الخاصة** --- يمكن وسم كل متنبئ بدورٍ خاص مثل `contract_date`، و`dom`، و`concessions`، و`latitude`، و`longitude`، و`living_area`، و`lot_size`، و`actual_age`، و`effective_age`، و`area`، و`site_dimensions`، أو `display_only`. وتتحكّم هذه التعيينات في كيفية معالجة العمود أثناء الملاءمة وفي المخرجات وفي توليد شبكة المبيعات. انظر الفصل 6 للمرجع الكامل.
- **تقريب خطَّي العرض والطول** --- تُقرَّب تلقائيًا إلى ثلاث منازل عشرية الأعمدةُ المعيَّنة خطَّ عرضٍ أو خطَّ طول، منعًا لفرط الملاءمة.
- **عمود عمر البيعة** --- حين يُعيَّن عمودٌ `contract_date` ويُقدَّم تاريخٌ فعّال، يحسب earthUI عمودًا باسم `sale_age` (الأيام بين تاريخ البيع والتاريخ الفعّال) ويستبدله متنبئًا.
- **حسابات RCA (في نمط التقييم فقط)** --- في نمط التقييم، يستطيع earthUI بعد ملاءمة النموذج أن يحسب مخرجات منهج قيد البواقي (RCA)، وهي تُنتج تعديلاتٍ لكل مقارِن، وملخّصاتٍ لصافي التعديلات وإجماليها، وسعر بيعٍ معدَّلًا للعقار موضوع التقييم.

- **شبكة المقارنة بالمبيعات (في نمط التقييم فقط)** --- تولّد مصنَّف Excel متعدّد الأوراق بمعادلاتٍ لتعديلات المقارنات، وتفكيكاتٍ للخصائص الباقية، وأسعارِ بيعٍ معدَّلة. انظر الفصل 11.

> **تنبيه:** يصلح النمط العام لأي مجموعة بيانات --- مالية أو علمية أو هندسية أو عقارية. وإنما يضيف نمطا التقييم والسوق خصائصَ تيسيرية لمحترفي العقارات.

5.1 تنصيب earthUI وvProlog

earthUI منشورةٌ على CRAN، فالتنصيب المعتاد سطرٌ واحد في R:

```
install.packages("earthUI")
```

ويصف هذا المقال الإصدار الحالي. فإن كان CRAN ما يزال يحمل إصدارًا أقدم حين تقرأ هذا، فالنسخة التطويرية تُنصَّب من GitHub (ولاحظ `subdir` --- فالحزمة تقيم في مجلد `pkg/` من المستودع):

```
remotes::install_github("wcraytor/earthUI", subdir = "pkg")
```

vProlog (اختياري). يستعمل مرفقُ معالجة الملاحظات التجريبي (انظر ``استخراج النصّ بـProlog" قرب نهاية هذا المقال) حزمةَ vProlog --- وهي محرّك Prolog مضمَّن ومكتفٍ بذاته. وهي عمدًا **ليست** اعتمادًا صلبًا: فـearthUI تُنصَّب وتعمل من دونها، وتبقى خصائص Prolog معطَّلةً ببساطة حتى تتوفّر. وتُوزَّع vProlog عبر مستودع r-universe الخاص بالمؤلّف، الذي يقدّم ثنائياتٍ مبنيّة مسبقًا لـWindows وmacOS (فلا حاجة إلى سلسلة أدوات ترجمة):

```
install.packages("vProlog",
                 repos = c("https://wcraytor.r-universe.dev",
                           getOption("repos")))
```

ولا يلزم تنصيبُ Prolog على مستوى النظام (SWI-Prolog أو غيره) --- فمحرّك Trealla مترجَمٌ داخل الحزمة نفسها. وكلُّ ما عدا ذلك في هذا المقال يعمل من دون vProlog.

6.1 البدء

لاستعمال earthUI:

1. **نصّب الحزمة** من CRAN كما وُصف أعلاه.
2. **شغّل التطبيق**: نفّذ `earthUI::launch()` في R، أو نفّذ من سطر الأوامر `Rscript -e 'earthUI::launch()'`. فيُفتح التطبيق في متصفّحك على المنفذ 7878. ويمكنك كذلك الوصول إليه مباشرةً بالانتقال إلى `http://localhost:7878` في متصفّحك. (وتبدأ وحدتا المعالجة اللاحقة على النحو نفسه: `launch_glmnet()` على المنفذ 7879، و`launch_mgcv()` على المنفذ 7880.)
3. **أنشئ مشروعًا أو افتحه** في القسم 1 من الشريط الجانبي. ويلتقط معالجُ المشروع الجديد اسمَ المشروع، والغرضَ (عام، أو للتقييم، أو تحليل منطقة السوق)، والموقعَ (البلد والولاية والمقاطعة والمدينة)، واختياريًا ملفَّ بياناتٍ أوّليًا يُنسخ إلى مجلد `in/` الخاص بالمشروع.
4. **استورد بياناتك** في القسم 2، الذي يُدرِج ملفات CSV وExcel في مجلد `in/` بالمشروع النشط. وتظهر الملفات التي تضيفها يدويًا عبر Finder أو Explorer بعد نقر ``تحديث".
5. **اضبط المتغيرات** --- اختر هدفك أو أهدافك ومتنبئاتك، واضبط أنواع البيانات، وأسنِد أي أدوارٍ خاصة للأعمدة.
6. **اضبط معاملات Earth** --- الدرجة، وnprune، ومرشّحات المجموعة الفرعية، وسائر وسائط `earth()`.
7. **لائم النموذج** --- انقر ``ملاءمة نموذج Earth" وراجع النتائج في اللوحة الرئيسة.
8. **صدّر** --- نزّل التنبؤات بصيغة Excel، أو ولّد حزم تقارير Quarto، أو (في نمط التقييم) احسب تعديلات RCA وشبكات المقارنة بالمبيعات.

ويُحفَظ الضبط في جانب الخادم في قاعدة بيانات المشروع، مفهرسًا بالمشروع والغرض، عبر زرَّي ``حفظ الحالي بوصفه الافتراضي" في القسمين 3 و4 (انظر الفصل 3، ``بقاء الإعدادات").

2 متطلبات بيانات الإدخال من MLS

في سير عمل تقييم العقارات وتحليل السوق، تأتي بيانات الإدخال عادةً من تصديرٍ لخدمة الإدراج المتعدد (MLS). ويصف هذا الفصل بنية الملف المتوقعة والأعمدة التي تستطيع earthUI استعمالها.

1.2 صيغة الملف وبنيته

يقبل earthUI ملفات **CSV** و**Excel** (`.xlsx`، `.xls`). وعند الاستيراد تُحوَّل أسماء الأعمدة تلقائيًا إلى صيغة `snake_case` --- فمثلًا يصير "Living SqFt" هو `living_sqft`، ويصير "Contract Date" هو `contract_date`، ويصير "Sale Price" هو `sale_price`. ويضمن هذا التوحيد اتساق الإحالات إلى الأعمدة في سير العمل كله. أما فاصل CSV وعلامة الكسر العشري المستعملان عند الاستيراد فيحدّدهما ضبط اللغة والمنطقة (انظر الفصل 3، "ضبط اللغة والمنطقة").

وينبغي أن يكون ملف بياناتك جدولًا مسطّحًا بصفٍّ واحد لكل عقار وعمودٍ واحد لكل سمة. ويجب أن يحتوي الصف الأول من الملف على عناوين الأعمدة.

2.2 الأعمدة المطلوبة في نمط التقييم

مع أن earthUI يعمل مع أي مجموعة من الأعمدة، فإن سير عمل التقييم الكامل (تعديلات RCA مع شبكة المقارنة بالمبيعات) يستفيد من وجود الأعمدة الآتية في تصدير MLS لديك:

العمود	النوع الخاص	الغرض
سعر البيع	(الهدف)	المتغير المستجيب للنموذج
تاريخ التعاقد	`contract_date`	يُستعمل لحساب `sale_age` (الأيام من التاريخ الفعّال)
تاريخ الإدراج	`listing_date`	يُستعمل مع تاريخ التعاقد لحساب أيام العرض في السوق إن لم يوجد عمودٌ لها
أيام العرض في السوق	`dom`	أيام العرض في السوق؛ تظهر في شبكة المبيعات
التنازلات	`concessions`	التنازلات البيعية؛ صافي سعر البيع = سعر البيع − التنازلات في شبكة المبيعات
مساحة المعيشة (قدم مربعة)	`living_area`	تتيح البواقي للقدم المربعة (`residual_sf`، `cqa_sf`)
مساحة الأرض	`lot_size`	يُجمَّع في صف "مساحة الموقع | الأبعاد" في شبكة المبيعات
أبعاد الموقع	`site_dimensions`	يُجمَّع مع مساحة الأرض في شبكة المبيعات
خط العرض	`latitude`	يُقرَّب إلى ثلاث منازل عشرية؛ ويُستعمل في حساب القرب وفي مجموعة الموقع
خط الطول	`longitude`	يُقرَّب إلى ثلاث منازل عشرية؛ ويُستعمل في حساب القرب وفي مجموعة الموقع
معرّف المنطقة	`area`	يُجمَّع في صف "الموقع: خط الطول | خط العرض | المنطقة" في شبكة المبيعات
العمر الفعلي	`actual_age`	يُجمَّع في صف "العمر الفعلي | العمر الفعّال" في شبكة المبيعات
العمر الفعّال	`effective_age`	يُجمَّع مع العمر الفعلي في شبكة المبيعات
رقم القطعة	(بلا)	يظهر في صف APN من شبكة المبيعات
معرّف الإدراج	(بلا)	يظهر في صف APN من شبكة المبيعات
العنوان	`display_only`	يظهر في صف العنوان من شبكة المبيعات؛ ويُستبعد من النموذج

ويمكن أن تكون أسماء أعمدة جدول البيانات بلغةٍ أجنبية --- أما الأسماء "الخاصة" فبالإنجليزية كي يستطيع برنامج R أن يعاملها معاملةً خاصة. وفيما عدا ذلك تظهر أسماء الأعمدة كما أعطيتها في نماذج الانحدار والرسوم، وفي تقارير المخرجات إن كنت تجري تقييمًا.

وليست كل الأعمدة مطلوبة. فـearthUI يتكيّف --- فإن غاب عمود، أُسقطت الخاصية المقابلة له ببساطة. غير أن نماذج تسعير العقارات يُنصح فيها بشدة ببعض الأعمدة لبلوغ ملاءمةٍ مقبولة:

1. **عمر البيعة** --- عدد الأيام بين تاريخ التعاقد على البيع والتاريخ الفعّال. وإذا استُعمل سجلّ مبيعاتٍ ممتدّ لسنوات، فكثيرًا ما يؤدي `sale_age` دورًا محوريًا.
2. **مساحة المعيشة** --- وتُعرف أيضًا بأسماء مثل "Living Sqft" و"GLA" (مساحة المعيشة الإجمالية).

3. **إجمالي عدد الحمّامات** --- مجموع الحمّامات الكاملة وحمّامات الربع والنصف وثلاثة الأرباع.
4. **مواقف المرآب** أو **مساحة المرآب** --- مواقف المرآب أو مساحته بالأقدام المربعة.
5. **مساحة الأرض** --- مساحة أرض العقار.
6. **خط الطول وخط العرض**، و**معرّف المنطقة** إن توفّر. وهذه المتغيرات الموقعية تعين النموذج على مراعاة التباين السعري الجغرافي.

3.2 أعراف تسمية الأعمدة الخاصة

يتعرّف earthUI الأعمدة بـ**تعيين نوعها الخاص**، لا باسم العمود. فيمكنك تسمية أعمدتك بما شئت في تصدير MLS --- والمهم أن تُسنِد النوع الخاص الصحيح في جدول ضبط المتغيرات (الفصل 6).

4.2 جودة البيانات واكتمالها

- **القيم المفقودة (NA)**: تُحذف تلقائيًا قبل الملاءمة الصفوفُ التي فيها قيمٌ مفقودة في أي عمودٍ متنبئ أو عمود الهدف. ويُضبط `na.action` داخليًا على `na.fail` دائمًا، مع معالجة حذف القيم المفقودة قبل استدعاء `earth()`.
- **أعمدة التواريخ**: ينبغي أن تكون التواريخ بصيغةٍ تستطيع R تحليلها (مثل `2025-06-15`، أو `06/15/2025`، أو `June 15, 2025`). ويرصد earthUI أعمدة التواريخ تلقائيًا حين تُحلَّل 80% من القيم على الأقل بنجاح. وتُقرأ السنواتُ ذات الرقمين (مثل `06/15/25`) على أنها أقربُ سنةٍ مطابقة، فتُحلَّل تصديرات MLS الحديثة تحليلًا صحيحًا؛ أما السنوات ذات الأربعة أرقام فتُؤخذ حرفيًا دائمًا.
- **الأعمدة العددية**: يجب أن تكون عددية حقولُ سعر البيع ومساحة المعيشة ومساحة الأرض والتنازلات وما شابهها. فإن كان نظام MLS لديك يصدّر الأسعار برموز العملات أو بفواصل الآلاف (مثل "$350,000" أو "350.000,00")، فقد تحتاج إلى تنظيفها قبل الاستيراد.
- **أعمدة العوامل**: ينبغي أن تحتوي المتغيراتُ الفئوية كمعرّف المنطقة أو الطراز أو الحالة على عددٍ قابل للإدارة من القيم المتمايزة. ويرصد earthUI العواملَ تلقائيًا، ويمكنك تجاوز هذا الرصد.

> **تنبيه:** راجع عمود *T/F/NA* في جدول المتنبئات بعد الاستيراد --- فهو يعرض أعداد الصواب والخطأ والقيم المفقودة لكل عمود. والأعمدة التي فيها أكثر من *30%* قيمٌ مفقودة تُعلَّم بالأحمر. وفكّر في استبعاد الأعمدة عالية القيم المفقودة أو في تنظيف البيانات قبل الاستيراد.

5.2 موضع صف العقار موضوع التقييم

في نمط **التقييم**، **يجب أن يكون الصف 1 هو العقار موضوع التقييم**. وسائر الصفوف مبيعاتٌ مقارِنة. ويمكن ترك سعر بيع العقار موضوع التقييم فارغًا (قيمةً مفقودة) أو ضبطه على أي قيمة --- إذ يعامله earthUI معاملة القيمة المفقودة أثناء الملاءمة على كل حال.

وفي نمط **تحليل منطقة السوق**، يكون وضع عقارٍ موضوع تقييم في الصف 1 اختياريًا، ولا يُستبعد أي صفٍّ تلقائيًا. ولحجب صف العقار موضوع التقييم (أو أي صفٍّ آخر) عن الملاءمة، أعطِه قيمةً صفرية في عمود أوزانٍ معيَّن أو استعمل مرشّح مجموعةٍ فرعية.

أما في النمط **العام** فلا معاملة خاصة للصفوف --- إذ تُعامَل الصفوف كلها معاملةً واحدة.

3 النمط العام

1.3 نظرة عامة

النمط العام هو النمط الافتراضي عند تشغيل earthUI. وهو يوفّر سير عمل نمذجة MARS كاملًا لأي مجموعة بيانات --- لا للعقارات وحدها. فيمكنك استعمال earthUI للبيانات العلمية، أو التحليل المالي، أو الدراسات الهندسية، أو أي مسألة انحدارٍ تريد فيها استكشاف العلاقات اللاخطية والتفاعلات بين المتغيرات.

وفي النمط العام تُسقِط الواجهةُ الخصائصَ الخاصة بالعقارات (الأعمدة الخاصة، وعمر البيعة، وتقريب الإحداثيات، وRCA). ويُبسَّط الشريط الجانبي ليتركّز على انتقاء المتغيرات، وضبط المعاملات، وملاءمة النموذج، والتصدير.

2.3 سير العمل في الشريط الجانبي

يُنظَّم الشريط الجانبي في أقسامٍ مرقّمة قابلة للطي تقودك من اختيار المشروع إلى التصدير:

1. المشروع --- اختر مشروعًا قائمًا أو أنشئ آخر عبر معالج المشروع الجديد، الذي يلتقط الغرض والبلد والموقع الإداري (وتُقفل كلها بعد الإنشاء). ويحدّد المشروعُ النشط موضعَ قراءة ملفات الإدخال وموضعَ كتابة كل المخرجات --- فلا يوجد حقلٌ منفصل لمجلد المخرجات.

2. استيراد البيانات --- اختر ملف CSV أو Excel من مجلد `in/` بالمشروع النشط؛ ويعيد رابطُ ``تحديث" مسحَ المجلد بحثًا عن ملفاتٍ أُضيفت يدويًا. وفي ملفات Excel متعددة الأوراق يظهر منتقي أوراق. وتُحوَّل أسماء الأعمدة تلقائيًا إلى snake_case.

3. ضبط المتغيرات --- منتقي المتغير الهدف (ويدعم أهدافًا متعددة)، والتاريخ الفعّال، وجدول المتنبئات بمربعات تأشير للإدراج والعَوملة والخطية والكمون. انظر الفصل 6 للتفاصيل الكاملة.

4. معاملات استدعاء Earth --- كل وسائط الدالة `earth()`: الدرجة، والجزاء، وnk، وطريقة التشذيب، والتحقق المتقاطع، وغيرها. انظر الفصل 7 للمرجع الكامل للمعاملات.

تنفيذ معالجة Prolog --- قسمٌ اختياري غير مرقّم بين القسمين 4 و5، ومعطَّلٌ ما لم تُفعَّل خاصية Prolog التجريبية في الإعدادات. انظر الفصل 15.

5. ملاءمة نموذج Earth --- حقلُ بذرة عشوائية والزرُّ الذي يشغّل النموذج لاتزامنيًا. انظر الفصل 8 لتفاصيل الملاءمة.

6. تنزيل المتغير الهدف المقدَّر والبواقي --- (وعنوانه ``تنزيل أسعار البيع المقدَّرة والبواقي" في نمطَي التقييم والسوق) يصدّر التنبؤات والبواقي ودرجات CQA وإسهامات كل دالة g في ملف Excel. انظر الفصل 9.

7–8. تعديلات RCA وشبكة المبيعات --- في نمط التقييم فقط. انظر الفصلين 10 و11.

9. توليد تقرير Quarto --- يكتب حزمة تقرير Quarto مكتفيةً بذاتها (مصدر `qmd` . مع الأصول) في مجلد مخرجات المشروع. ويحمل الرقم 7 خارج نمط التقييم. انظر الفصل 12.

10. تحويل تقرير Quarto --- يصيّر أي ملف `qmd` . إلى HTML أو Word أو PDF. انظر الفصل 12.

3.3 ألسنة اللوحة الرئيسة

بعد الملاءمة توفّر اللوحة الرئيسة عشرة ألسنة:

اللسان	المحتوى
البيانات	معاينةٌ للبيانات المستوردة. وفي نمط التقييم تُقسَم إلى جدول العقار موضوع التقييم وجدول المبيعات المقارِنة.
المعادلة	معادلة النموذج المُلائَم معروضةً بـMathJax/LaTeX، تبيّن كل دالة أساسٍ ومعاملها.
الملخّص	المقاييس الرئيسة (R^2، وGRSq، وGCV، وRSS، وR^2 للتحقق المتقاطع) وجدولُ معاملات.
أهمية المتغيرات	رسمٌ شريطي وجدولٌ مرتَّب لدرجات أهمية المتنبئات.
الإسهام	رسوم إسهام دوال g. اختر دالةً من القائمة المنسدلة؛ فالمتنبئات أحادية المتغير تعرض رسومًا خطية، والتفاعلات الثنائية (الدرجة ≥ 2) تعرض رسومًا منظورية ثلاثية الأبعاد ورسومَ كنتور.
الارتباط	خريطةٌ حرارية لارتباطات المتنبئات.
التشخيصات	ثلاثة رسوم: البواقي في مقابل القيم المُلائَمة، ورسم Q-Q الطبيعي، والفعلي في مقابل المتنبأ به.
تعديلات RCA	(في نمط التقييم فقط) مدرَّجاتٌ لنسبة تعديل الباقي وصافي التعديل والتعديل الإجمالي عبر المقارنات، تُملأ بعد إتمام الخطوة 7 (حساب تعديلات RCA).
إسهام الباقي	يقدّر إسهامات السوق للخصائص النادرة أو المستبعَدة بانحدارٍ مشترك لبواقي النموذج على الأعمدة التي لم يستعملها النموذج قط (انظر 4.4 subsection).
تحليل التباين	جدول تفكيك تحليل التباين يبيّن إسهام كل دالة أساسٍ في مجموع مربعات البواقي.
مخرجات Earth	المخرجات النصّية الخام من `earth::summary()`، بما فيها تفاصيل مرور التشذيب.

4.3 بقاء الإعدادات

يحفظ earthUI ضبطك في جانب الخادم في قاعدة بيانات المشروع (`projects.sqlite` في جذر regProj)، مفهرسًا بالمشروع والغرض. والحفظُ **مسوقٌ بالأزرار**: فزرّ ``حفظ الحالي بوصفه الافتراضي" في القسم 3 يحفظ الهدفَ والتاريخَ الفعّال وضبطَ المتنبئات، في حين يحفظ زرُّ القسم 4 معاملاتِ استدعاء earth ومصفوفةَ التفاعلات المسموح بها. ويحدّث كل زرٍّ نصفه هو فقط، فلا يطمس أحدهما الآخر أبدًا، و**الملاءمة لا تحفظ تلقائيًا** --- فيمكنك التجريب بحرّية من غير أن تطمس افتراضاتك المحفوظة. وملفاتُ الإدخال المتعددة للمشروع الواحد (مستخرَجُ اختبارٍ صغير، ومجموعةُ البيانات الكاملة) تتشارك ضبطًا واحدًا، ولأن الإعدادات تقيم في شجرة المشروع فإنها تسافر معها حين تزامن مجلّد regProj أو تنسخه احتياطيًا أو تشاركه.

5.3 النمط الداكن

انقر أيقونة القمر/الشمس في الركن الأعلى لتبديل السمة بين الفاتحة والداكنة. وتُحفظ السمة المفضّلة في التخزين المحلي وتبقى عبر الجلسات.

6.3 ضبط اللغة والمنطقة

يدعم earthUI أعرافًا دولية لتنسيق الأعداد والتواريخ وملفات CSV عبر نظام لغةٍ ومنطقة قائم على البلد. وتختار قائمةُ **البلد** المنسدلة في قائمة **الإعدادات** (الترس) من شريط القوائم العلوي إعدادًا مسبقًا لواحدٍ وثلاثين بلدًا مدعومًا. ويضبط كل إعدادٍ مسبق ما يأتي:

- **فاصل CSV** --- الفاصلة (,) في الولايات المتحدة والمملكة المتحدة واليابان، أو الفاصلة المنقوطة (;) في معظم أوروبا حيث تُستعمل الفاصلة علامةً عشرية.
- **العلامة العشرية** --- النقطة (.) أو الفاصلة (,).
- **فاصل الآلاف** --- الفاصلة (الولايات المتحدة والمملكة المتحدة واليابان)، أو النقطة (ألمانيا وإيطاليا وإسبانيا)، أو المسافة (فنلندا وفرنسا وبولندا ودول البلطيق وأوكرانيا وروسيا)، أو الفاصلة العليا (سويسرا).
- **صيغة التاريخ** --- MM/DD/YYYY (الولايات المتحدة)، أو DD/MM/YYYY (معظم أوروبا)، أو YYYY-MM-DD (السويد وليتوانيا واليابان وكندا).
- **مقاس الورق** --- Letter (الولايات المتحدة وكندا والمكسيك) أو A4 (في سائر البلدان).

1.6.3 البلدان المدعومة

البلد	فاصل CSV	العشرية	الآلاف	التاريخ	الورق	
الولايات المتحدة، كندا (csv)	,	.	,	MDY/YMD	Letter	
المملكة المتحدة، أيرلندا، أستراليا، نيوزيلندا	,	.	,	DMY	A4	
ألمانيا، النمسا، بلجيكا، هولندا، إيطاليا، إسبانيا، البرتغال، الدنمارك، البرازيل	;	,	.	DMY	A4	
فرنسا، بولندا، التشيك، النرويج، لاتفيا، إستونيا	;	,	(مسافة)	DMY	A4	
فنلندا، السويد	;	,	(مسافة)	DMY/YMD	A4	
سويسرا	;	,	’	DMY	A4	
أوكرانيا، روسيا	;	,	(مسافة)	DMY	A4	
تركيا	;	,	.	DMY	A4	
اليابان، كوريا الجنوبية	,	.	,	YMD	A4	
المكسيك	,	.	,	DMY	Letter	
ليتوانيا	;	,	(مسافة)	YMD	A4	

2.6.3 تجاوز الورق

أسفل منتقي البلد، تتيح لك قائمة **الورق** المنسدلة (Letter أو A4) تجاوزَ مقاس صفحة التقرير من غير تبديل البلد. أما فاصل CSV والعلامة العشرية وترتيب صيغة التاريخ فتتبع الإعدادَ المسبق للبلد المختار؛ وحين تغيّر البلد يُصفَّر تجاوزُ الورق إلى افتراض ذلك البلد.

3.6.3 حفظ الافتراضات

انقر **حفظ** في قائمة الإعدادات لتخزين تفضيلات اللغة والمنطقة حفظًا عامًّا. وتنطبق هذه الافتراضات على كل الجلسات المقبلة أيًّا كان المشروع الذي تفتحه. أما الإعدادات الخاصة بكل مشروع (الهدف، والمتنبئات، والمعاملات) فتُحفظ على حدة في قاعدة بيانات المشروع؛ في حين تبقى افتراضاتُ اللغة والمنطقة عبر كل المشاريع في قاعدة بيانات SQLite على مستوى المستعمل. وهذه المقاربة ذات المستويين مصمَّمةٌ لمنظّماتٍ كشركات التدقيق التي تعمل على بياناتٍ من بلدانٍ متعددة --- فاضبط بلدك الأشيع افتراضًا، ثم عدّله عند الحاجة.

> **تنبيه:** يتكيّف تنسيقُ الأعداد على محاور الرسوم وتسميات الميل تلقائيًا مع لغتك ومنطقتك. فالإعداد الألماني يستعمل النقاط للآلاف (*200.000*)، والفنلندي يستعمل المسافات (*200 000*)، والسويسري يستعمل الفواصل العليا (*200’000*). ولا تُعرض رموز العملات --- فـ*earthUI* محايدٌ إزاء العملة.

4 نمط التقييم

حين تختار **للتقييم** غرضًا، يهيّئ earthUI نفسه لتقييم عقارٍ مفرد. وتبقى كل الخصائص الموصوفة في الفصل 3 متاحة؛ ولا يتناول هذا الفصل إلا الإضافات الخاصة بالتقييم.

1.4 معالجة صف العقار موضوع التقييم

في نمط التقييم، **يكون الصف 1 من مجموعة بياناتك هو العقار موضوع التقييم**، وسائر الصفوف مبيعاتٌ مقارِنة. ويجب أن يكون ملف إدخالك منظَّمًا على هذا النحو (انظر الفصل 2). ويمكن ترك سعر بيع العقار موضوع التقييم فارغًا أو ضبطه على أي قيمة --- إذ يعامله earthUI تلقائيًا معاملة القيمة المفقودة أثناء الملاءمة.

وبعد الاستيراد ينقسم لسان البيانات إلى قسمين: **العقار موضوع التقييم** (الصف 1) و**المبيعات المقارِنة** (الصفوف 2 فما فوق). ويُستبعد الصف 1 دائمًا من ملاءمة النموذج --- ويؤكّد ذلك الإشعارُ ``تخطّي الصف 1 (العقار موضوع التقييم). الملاءمة على N صفًّا." ومع ذلك يولّد النموذج بعد الملاءمة تنبؤاتٍ لصف العقار موضوع التقييم، تظهر باسم `est_<target>` في المخرجات.

2.4 التاريخ الفعّال وعمر البيعة

يظهر حقل **التاريخ الفعّال** في قسم ضبط المتغيرات في كل أنماط الغرض (وافتراضه تاريخ اليوم) --- فللتحليلات العامة وتحليلات السوق تاريخٌ فعّال أيضًا، لا للتقييمات وحدها --- لكنه لا **يُستعمل** إلا في نمطَي التقييم والسوق. فهناك، إن عيّنت عمودًا `contract_date` في قائمة العمود الخاص المنسدلة، حسب earthUI عمودًا باسم `sale_age` --- وهو عدد الأيام الصحيحة بين تاريخ التعاقد على كل بيعة والتاريخ الفعّال. ويحلّ هذا العمود محلّ عمود التاريخ الأصلي متنبئًا.

وفي أول مرة تنقر فيها ``ملاءمة" بعد تعيين تاريخ التعاقد، ينشئ earthUI عمود `sale_age` ويُشعرك بأن تنقر ``ملاءمة" ثانيةً لإدراجه.

3.4 تعيينات الأعمدة الخاصة

في نمطَي التقييم والسوق تظهر قائمة **الخاص** المنسدلة لكل متنبئ في جدول ضبط المتغيرات. وهذه هي القائمة الكاملة للأنواع الخاصة وآثارها:

النوع الخاص	الأثر
`no`	الافتراض --- بلا معاملةٍ خاصة
`contract_date`	يستدعي الحساب التلقائي لـ`sale_age` من التاريخ الفعّال
`listing_date`	يُستعمل مع تاريخ التعاقد لحساب أيام العرض في السوق إن لم يوجد عمود `dom`
`dom`	يعرّف عمود أيام العرض في السوق؛ ويظهر في شبكة المبيعات
`concessions`	يعرّف التنازلات البيعية؛ ويُستعمل في صيغة صافي سعر البيع (سعر البيع − التنازلات) في شبكة المبيعات
`latitude`	تُقرَّب القيم إلى ثلاث منازل عشرية؛ وتُستعمل في قرب Haversine وفي مجموعة الموقع بشبكة المبيعات
`longitude`	تُقرَّب القيم إلى ثلاث منازل عشرية؛ وتُستعمل في قرب Haversine وفي مجموعة الموقع بشبكة المبيعات
`area`	يُجمَّع في صف ``الموقع: خط الطول | خط العرض | المنطقة" في شبكة المبيعات
`living_area`	يتيح البواقي للقدم المربعة (`residual_sf`، `cqa_sf`) في مخرجات التنزيل
`lot_size`	يُجمَّع في صف ``مساحة الموقع | الأبعاد" في شبكة المبيعات
`site_dimensions`	يُجمَّع مع مساحة الأرض في شبكة المبيعات
`actual_age`	يُجمَّع في صف ``العمر الفعلي | العمر الفعّال" في شبكة المبيعات
`effective_age`	يُجمَّع مع العمر الفعلي في شبكة المبيعات
`display_only`	يُدرَج العمود في التصديرات لكنه يُستبعد من ملاءمة النموذج (ويُسمح بأعمدةٍ متعددة)
`remarks`	(تجريبي) عمود ملاحظاتٍ نصّية حرّة تُحلّله خطوةُ Prolog الاختيارية إلى أعمدة خصائص (ويُسمح بأعمدةٍ متعددة)؛ انظر الفصل 15
`list`	(تجريبي) عمودٌ مفصول بفواصل تقسّمه خطوةُ Prolog الاختيارية إلى أعمدة FALSE/TRUE أحادية التنشيط (ويُسمح بأعمدةٍ متعددة)؛ انظر الفصل 15

ولا يُسمح إلا بعمودٍ واحد لكل نوعٍ خاص (عدا `display_only` و`remarks` و`list`). وإسنادُ نوعٍ خاص إلى عمودٍ ثانٍ يمسحه تلقائيًا من الأول. وتظهر شارةٌ زرقاء صغيرة بجوار اسم المتغير تبيّن نوعه الخاص المُسنَد.

4.4 إسهام الباقي (تقدير الخصائص النادرة)

الخصائصُ النادرةُ أكثرَ من أن تحتملها قناةُ الانحدار --- كورشةٍ موجودة في حفنةٍ من المبيعات --- تنتمي إلى طبقة الباقي في RCA. ولسانُ **إسهام الباقي** يقدّر إسهامها السوقي بانحدار بواقي النموذج المُلائَم على الأعمدة التي لم يستعملها النموذج قط، بطريقة المربعات الصغرى الاعتيادية. واختيارُ عدة أعمدة يقدّرها **مجتمعةً**، فتتقاسم المرافقُ المرتبطة (مثل `outbuilding_sqft` مع `pr_has_workshop`) الفضلَ بدل حسابه مرتين. وتعطي الأعلامُ الثنائية تقديراتٍ إجمالية وتُدرِج باقيَ كل بيعةٍ معلَّمة بيّنةً؛ وتعطي الأعمدةُ العددية إسهاماتٍ لكل وحدة؛ والخطأُ المعياري المُبلَّغ خطأٌ معياري واحد (و$t \geq 2$ هو المعيار المعتاد للدلالة). وتتلقّى الحدودُ المُسنَدة بأقل من عشر مشاهدات تنبيهًا تلقائيًا في ملاحظة ملف العمل المولَّدة، وهي توثّق المنهج والتقديرات والتحفّظات في فقرةٍ واحدة للتقرير. ولا يُستبعد من قائمة المرشَّحين إلا المتنبئات التي **استعملها** النموذج فعلًا --- فالعَلَم النادر المعروض على الملاءمة والذي لم ينتقه المرور الأمامي قط يبقى مؤهَّلًا، إذ لم يمتصّ النموذج شيئًا من قيمته.

5.4 نظرة عامة على تعديلات RCA

يحسب زرّ **حساب تعديلات RCA وتنزيلها** (القسم 7 من الشريط الجانبي، ولا يظهر إلا في نمط التقييم بعد الملاءمة) تعديلاتٍ مستمدّةً من السوق لكل مقارِنٍ نسبةً إلى العقار موضوع التقييم. وسير عمل RCA كاملًا موصوفٌ في الفصل 10.

وبعد حساب تعديلات RCA يصير زرّ **توليد شبكة المبيعات وتنزيلها** (القسم 8 من الشريط الجانبي) متاحًا. وسير عمل شبكة المبيعات موصوفٌ في الفصل 11.

> **تنبيه:** درجة *CQA* التي تُسنِدها إلى العقار موضوع التقييم تتحكّم في مقدار ما يُنسب إليه من توزيع البواقي. والدرجة *5.00* تضعه عند وسيط المقارنات.

5 نمط تحليل منطقة السوق

حين تختار **تحليل منطقة السوق** غرضًا، يوفّر earthUI الخصائصَ الخاصة بالعقارات نفسَها التي يوفّرها نمط التقييم (الأعمدة الخاصة، وعمر البيعة، وتقريب الإحداثيات)، لكنه موجَّهٌ نحو تحليل مجموعةٍ من العقارات لا نحو تقييم عقارٍ مفرد.

1.5 الفروق عن نمط التقييم

- **لا استبعاد تلقائي للعقار موضوع التقييم** --- فبخلاف نمط التقييم، لا يستبعد نمطُ السوق الصفَّ 1 تلقائيًا من الملاءمة. بل تُدرَج الصفوف كلها؛ ولحجب صف العقار موضوع التقييم عن الملاءمة، أعطِه قيمةً صفرية في عمود أوزانٍ معيَّن (كما في مجموعة البيانات التجريبية) أو استعمل مرشّح مجموعةٍ فرعية.
- **لا قسمَي RCA وشبكة المبيعات** --- فخطوتا ``حساب تعديلات RCA وتنزيلها" و``توليد شبكة المبيعات وتنزيلها" غير متاحتين. ويتركّز نمط السوق على ملاءمة النموذج والمخرجات، لا على تعديلات كل مقارِن.
- **ترقيم الشريط الجانبي** --- من دون خطوتَي RCA وشبكة المبيعات، يحمل قسمُ توليد تقرير Quarto الرقم 7 بدل 9 (ويبقى قسم تحويل التقرير برقم 10).

2.5 متى يُستعمل نمط السوق

يناسب نمط تحليل منطقة السوق حين تكون:

- تبني نموذج انحدارٍ لحيٍّ أو منطقة سوقٍ لفهم محرّكات القيمة
- تحلّل كيف تؤثّر متغيراتٌ كالمساحة بالأقدام المربعة والعمر ومساحة الأرض والموقع في أسعار البيع عبر مجموعةٍ من العقارات
- تعدّ إسنادًا لتحليل ظروف السوق أو لترسيم حدود الحي
- تعمل على مجموعة بياناتٍ تتضمّن عقارًا موضوع تقييم في الصف 1 لكنك تريد خيار إدراجه في الملاءمة أو استبعاده منها

> **تنبيه:** نمط السوق مفيدٌ أيضًا لانحدار العقارات العام حين تريد خصائص الأعمدة الخاصة (عمر البيعة، وتقريب الإحداثيات) من غير أن تحتاج إلى سير عمل تعديلات *RCA*.

6 انتقاء المتغيرات

القسم 3 من الشريط الجانبي --- **ضبط المتغيرات** --- هو حيث تختار أي الأعمدة يشارك في النموذج وكيف يُعامَل.

1.6 المتغير الهدف أو المتغيرات الهدف

تُدرِج قائمة **المتغير الهدف (المستجيب)** المنسدلة أعلى القسم 3 كلَّ عمودٍ في مجموعة بياناتك. فاختر عمودًا واحدًا لنموذجٍ معياري أحادي المستجيب، أو أعمدةً متعددة لنموذجٍ متعدد المستجيبات. وحين تُختار أهدافٌ متعددة يلائم النموذجُ كلَّ المستجيبات في آن باستعمال `earth(cbind(y1, y2, ...) ~ .)`.
وتُستبعد الأعمدة المختارة أهدافًا تلقائيًا من قائمة المتنبئات.

2.6 جدول المتنبئات

أسفل منتقي الهدف يُدرِج جدولٌ كلَّ عمودٍ متبقٍّ مع الحقول الآتية:

العمود	الوصف
المتغير	اسم العمود (ويظهر الاسم كاملًا في تلميحٍ منبثق إن كان مقتطعًا)، يسبقه **مبدِّل +/×** (انظر أدناه). وفي نمطَي التقييم والسوق تعرض شارةٌ زرقاء النوعَ الخاص المُسنَد.
النوع	قائمة نوع البيانات المنسدلة: numeric، integer، character، logical، factor، Date، POSIXct
الإدراج	مربع تأشير --- أدرِج هذا العمود متنبئًا في النموذج
العامل	مربع تأشير --- عامِل هذا العمود متغيرًا فئويًا
الخطية	مربع تأشير --- ألزِمه الدخول الخطي وحده (بلا دوال مفصلة)
الكمون	مربع تأشير --- احجب هذا العمود عن النموذج بالكامل (فيُستبعد من earth ومن البيانات المصدَّرة إلى glmnetUI أو mgcvUI)، حفظًا له لعمل المتغيرات الكامنة
الخاص	قائمة منسدلة (في نمطَي التقييم والسوق فقط) --- انظر مرجع أنواع الأعمدة الخاصة أدناه
T/F/NA	ثلاثة أعداد لكل عمود: في المنطقية، FALSE/مفقود؛ وفي العددية، 0 / $\neq 0$ / $=$/مفقود. وتظهر بالأحمر حين تتجاوز القيمُ المفقودة 30%

ويشرح سطرُ تلميحٍ فوق الجدول الاختصارات: ``النوع = نوع بيانات العمود، والإدراج = أدرِجه متنبئًا، والعامل = عامِله فئويًا، والخطية = خطيٌّ فقط (بلا مفصلات)، والخاص = دور العمود (مثل `contract_date`)."

التعطيل اللطيف بدل الحذف. يجعل مبدِّلُ × بجوار اسم كل متغير الصفَّ رماديًا ويحجب العمود عن النموذج من غير إزالته من البيانات؛ وينقلب إلى + أخضر يعيد تمكين العمود. فلا تُتلَف الأعمدة أبدًا.

3.6 رصد أنواع البيانات وتجاوزه

يرصد earthUI أنواع البيانات تلقائيًا عند الاستيراد. فيتعرّف الأعمدة العددية والصحيحة والمنطقية والعاملية وأعمدة التواريخ. أما الأعمدة النصّية التي تبدو تواريخ (بأن تُحلَّل 80% من القيم على الأقل في مقابل صيغ التواريخ الشائعة) فتُصنَّف `Date`.
ويمكنك تجاوز أي رصدٍ بتغيير قائمة **النوع** المنسدلة. وحين تغيّر عمودًا إلى `character` أو `factor`، يُؤشَّر مربعُ العامل تلقائيًا. وتغييرُ الأنواع يؤثّر في كيفية تمرير العمود إلى الدالة `earth()`.

4.6 عَلَما العامل والخطية

- **العامل** --- حين يكون مؤشَّرًا، يُعامَل العمود متغيرًا فئويًا. وهذا ملائمٌ لأعمدةٍ كـ`style` أو `area_id` أو `grade` التي تمثّل مجموعاتٍ منفصلة لا قياساتٍ متصلة. وتدخل أعمدةُ العوامل النموذجَ متغيراتٍ مؤشِّرة (صوريّة).
- **الخطية** --- حين يكون مؤشَّرًا، يُلزَم العمود بدخول النموذج خطيًا --- فلا تُنشأ له دوالُّ مفصلة (خطية مقطعيًا). وهذا مفيدٌ حين تعلم أن لمتغيرٍ ما علاقةً خطية محضة بالهدف.

5.6 مرجع أنواع الأعمدة الخاصة

في نمطَي التقييم والسوق توفّر قائمة **الخاص** المنسدلة الخيارات الآتية. ويمكن إسناد كل نوعٍ إلى عمودٍ واحد على الأكثر (عدا `display_only` و`remarks` و`list`، فهي تسمح بأعمدةٍ متعددة):

أنواع التاريخ والوقت:

- `contract_date` --- يستدعي الحساب التلقائي لـ`sale_age`. ويحلّ محلّ عمود التاريخ الأصلي عمودٌ صحيح يقيس الأيام بين تاريخ البيع والتاريخ الفعّال.
- `listing_date` --- يُستعمل احتياطيًا لحساب أيام العرض في السوق (أيام العرض = تاريخ التعاقد − تاريخ الإدراج) حين لا يُعيَّن عمود `dom` صراحةً.
- `dom` --- يعرّف عمود أيام العرض في السوق. ويظهر في صف APN وصفٌ تاريخ البيع من شبكة المبيعات.

الأنواع النقدية:

- `concessions` --- يعرّف التنازلات البيعية (ائتمانات البائع، وحوافز المشتري، وغيرها). ويُستعمل في شبكة المبيعات لحساب صافي سعر البيع: صافي سعر البيع = سعر البيع − التنازلات.

أنواع الحجم والموقع:

- `latitude` --- تُقرَّب القيم تلقائيًا إلى ثلاث منازل عشرية منعًا لفرط الملاءمة. وتُستعمل في حسابات قرب Haversine (المسافة من العقار موضوع التقييم إلى كل مقارِن) وتُجمَّع في صف الموقع من شبكة المبيعات.
- `longitude` --- المعاملة نفسها في التقريب كخط العرض. وتُستعمل في القرب وفي مجموعة الموقع.
- `area` --- عادةً معرّف حيٍّ أو منطقة. ويُجمَّع مع خطَّي العرض والطول في صف ``الموقع: خط الطول | خط العرض | المنطقة" من شبكة المبيعات.
- `living_area` --- يتيح حسابات البواقي للقدم المربعة (`residual_sf` و`cqa_sf`) في مخرجات التنزيل.
- `lot_size` --- يُجمَّع في صف ``مساحة الموقع | الأبعاد" من شبكة المبيعات.
- `site_dimensions` --- يُجمَّع مع مساحة الأرض في شبكة المبيعات (مثل "75x120").

أنواع العمر:

- `actual_age` --- يُجمَّع في صف ``العمر الفعلي | العمر الفعّال" من شبكة المبيعات.
- `effective_age` --- يُجمَّع مع العمر الفعلي في شبكة المبيعات.

أنواع العرض:

- `display_only` --- يُدرَج العمود في تصديرات Excel لكنه يُستبعد من ملاءمة النموذج بالكامل. استعمل هذا لحقول العناوين، أو أرقام MLS، أو معرّفات القطع، أو غيرها من البيانات المرجعية التي لا ينبغي أن تكون متنبئًا. ويمكن أن تحمل هذا التعيين أعمدةٌ متعددة.

أنواع استخراج النصّ (تجريبية):

- `remarks` --- عمود نصٍّ حرّ (مثل `public_remarks`) تحلّله خطوةُ Prolog الاختيارية إلى أعمدة خصائص. ولا يفعل شيئًا حتى تُفعَّل تلك الخاصية في الإعدادات. ويُسمح بأعمدةٍ متعددة. انظر الفصل 15.
- `list` --- عمودٌ مفصول بفواصل (مثل `existing_structures`) تقسّمه الخطوةُ نفسها إلى أعمدة FALSE/TRUE أحادية التنشيط على الصورة `<col>_<value>`. ويُسمح بأعمدةٍ متعددة. انظر الفصل 15.

6.6 الأهداف المتعددة

اختيارُ أكثر من متغيرٍ هدف يلائم نموذج Earth متعدد المستجيبات. وحين تُختار أهدافٌ متعددة:

- يتنبّأ النموذج بكل المستجيبات في آن، مع تشارك دوال الأساس عبر الأهداف
- يصير زرّ **wp (أوزان المستجيبات)** نشطًا، فيتيح لك إسناد وزنٍ عددي لكل هدف
- تُعطَّل نماذجُ التباين (`varmod.method`) (فهي غير مدعومة في تعدّد المستجيبات)
- تعرض ألسنةُ النتائج مقاييسَ ومعادلاتٍ ورسومًا تشخيصية لكل مستجيب

> **تنبيه:** الأعمدة المعيَّنة `display_only` تبقى في مجموعة بياناتك وتظهر في تصدير *Excel*، لكنها تُستبعد من جدول المتنبئات بالكامل. استعمل هذا لأعمدة المعرّفات، أو العناوين، أو غيرها من الحقول المرجعية.

7 انتقاء المعاملات

القسم 4 من الشريط الجانبي --- **معاملات استدعاء Earth** --- يتيح الوصول إلى كل الوسائط التي تقبلها الدالة `earth()`. ولكل معاملٍ أيقونة مساعدةٍ زرقاء (؟) تحمل شرحًا في تلميحٍ منبثق. والمعاملات منظَّمةٌ في أقسامٍ فرعية قابلة للطي.

1.7 المعاملات الأساسية

المعامل	الافتراض	الوصف
subset	(فارغ)	تعبير ترشيح الصفوف. انظر ``ترشيح المجموعة الفرعية" أدناه.
weights	NULL	منتقي عمودٍ لأوزان الحالات (الصفوف). ولا تُدرَج إلا الأعمدة العددية.
wp	NULL	أوزان المستجيبات في النماذج متعددة الأهداف. ويفتح الزرُّ حوارًا فيه مُدخَلٌ عددي لكل هدف (والافتراض 1.0 لكلٍّ منها).
keepxy	معطَّل	الاحتفاظ بـx وy والمجموعة الفرعية والأوزان في كائن النموذج.
trace	0	مستوى تتبّع مخرجات الملاءمة (من 0 إلى 5).
glm	بلا	عائلة GLM اختيارية: `gaussian` أو `binomial` أو `poisson`.
degree	1	أقصى رتبة تفاعل. وضبطُ الدرجة على $2 \geq$ يفعّل التحقق المتقاطع تلقائيًا ويُظهر مصفوفة التفاعلات المسموح بها.
penalty	2	جزاء GCV لكل عقدة. والقيم الأعلى تنتج نماذجَ أبسط.

2.7 المرور الأمامي

المعامل	الافتراض	الوصف
nk	تلقائي	أقصى عدد حدودٍ قبل التشذيب. ويُملأ تلقائيًا بالقيمة المُوصى بها للبيانات المحمَّلة (انظر ``قيم المعاملات المُوصى بها" أدناه).
thresh	0.001	عتبة الخطوة الأمامية. والقيم الأصغر تسمح بمزيدٍ من الحدود.
minspan	تلقائي	أدنى مدًى بين العُقد. ويُملأ تلقائيًا بالقيمة المُوصى بها. والقيم السالبة تضبط أقصى عدد عُقدٍ لكل متنبئ.
endspan	تلقائي	مدى الطرف --- أدنى مسافة من عقدةٍ إلى حافة البيانات. ويُملأ تلقائيًا بالقيمة المُوصى بها.
newvar.penalty	0	جزاءُ إدخال متغيرٍ جديد (يشجّع إعادة استعمال المتنبئات القائمة).
fast.k	20	عدد الحدود الأبوية التي تُنظَر في خوارزمية MARS السريعة.
fast.beta	1	يتحكّم في معامل التقادم في MARS السريعة.

3.7 التفاعلات المسموح بها

حين تُضبط **الدرجة** على 2 أو أكثر، تظهر مصفوفةُ تفاعلاتٍ أسفل المعاملات الأساسية. وهي شبكةٌ مثلثية علوية من مربعات التأشير، واحدٌ لكل زوج متنبئات. والمربع المؤشَّر يعني السماح للمتنبئين بالتفاعل؛ وإزالةُ التأشير تمنع ذلك التفاعل بعينه.

ونقرُ اسم متنبئ (في عنوان الصف أو العمود) يبدّل كل تفاعلات ذلك المتغير. ويوفّر مربعا **السماح بالكل** و**مسح الكل** في الأعلى تحكّمًا جماعيًا. وتستعمل المصفوفةُ عناوينَ لاصقة فتبقى تسميات الأعمدة والصفوف ظاهرةً عند التمرير.

ويذكّرك تنبيهٌ إعلامي بأن: ``حدود التفاعل تزيد خطر فرط الملاءمة. وقد فُعِّل التحقق المتقاطع (بعشر طيّات)."

4.7 التشذيب

المعامل	الافتراض	الوصف
pmethod	backward	طريقة التشذيب: `backward`، أو `none`، أو `exhaustive`، أو `forward`، أو `seqrep`، أو `cv`.
nprune	NULL	أقصى عدد حدودٍ بعد التشذيب. اتركه فارغًا للانتقاء التلقائي.

5.7 التحقق المتقاطع

المعامل	الافتراض	الوصف
nfold	10	عدد طيّات التحقق المتقاطع. واضبطه على 0 لتعطيله.
ncross	20	عدد تكرارات التحقق المتقاطع.
stratify	مفعَّل	طبِّق عيّنات التحقق المتقاطع بحيث يكون لكل طيّة توزيعٌ متقارب للمستجيب.

وحين تكون الدرجة $2 \geq$، يُفعَّل التحقق المتقاطع تلقائيًا (فتُضبط nfold على 10) للاحتراس من فرط الملاءمة الناشئ عن التفاعلات.

6.7 ترشيح المجموعة الفرعية

يقبل المُدخَل النصّي **subset** تعبيرًا بلغة R يرشّح الصفوف المستعملة في ملاءمة النموذج. ويمكنك كتابة التعبير مباشرةً (مثل `sale_age < 365 & area_id == 460`) أو استعمال زرّ **بناء المرشّح...** لبنائه بصريًا.

حوار بناء المرشّح --- انقر ``بناء المرشّح..." لفتح حوارٍ موجَّه. ولكل صفٍّ شرطٍ قائمةُ أعمدةٍ منسدلة، ومُعامِلٌ (<، >، =<، =>، ==، =!)، ومُدخَلُ قيمةٍ يتكيّف مع نوع العمود: مُدخَلٌ عددي للأعداد، ومنتقي تواريخ للتواريخ، وقائمةٌ منسدلة بالقيم المتمايزة للأعمدة النصّية والعاملية. وتُوصَل الشروط بروابط ``و" (&) أو ``أو" (|). وتعرض معاينةٌ في الأسفل التعبيرَ وعددَ الصفوف المطابقة. وانقر **تطبيق** لإدراج التعبير في المُدخَل النصّي.

و**أعمدة التواريخ** يجب أن تستعمل غلافَي `as.Date("...")` أو `as.POSIXct("...")` في التعبيرات اليدوية. أما حوار بناء المرشّح فيتولّى ذلك تلقائيًا.

و**ترشيح المجموعة الفرعية غير متلِف** --- إذ تبقى الصفوف المستبعَدة في مجموعة البيانات وتتلقّى تنبؤاتٍ في تصدير Excel.

> **تنبيه:** استعمل الأقواس لتجميع شروط ``أو" عند دمجها بـ``و". فمن غير أقواس، يُقيَّم `A & B | C` على أنه `(A & B) | C`، وقد لا يكون ذلك مرادك.

7.7 قيم المعاملات المُوصى بها

يعرض earthUI قيمةً مُوصى بها أسفل المعاملات الرئيسة في القسم 4. وتتحدّث هذه التوصيات تفاعليًا بحسب عدد صفوف الملاءمة (n) والمتنبئات المختارة (p). والصيغُ مستمدّةٌ من ورقة فريدمان عن MARS، ومن خوارزميات earth الداخلية، ومن اختبارٍ تجريبي. وتُملأ حقول **nk** و**minspan** و**endspan** مسبقًا بقيمها المُوصى بها مع تحميل البيانات؛ ويُحفظ أي تعديلٍ يدوي.

1.7.7 معاملات المرور الأمامي

nk (أقصى عدد حدودٍ قبل التشذيب):

$$\mathrm{nk} = \min\bigl(100,\ \max(21,\ 2p + 1,\ \lfloor n/10 \rfloor)\bigr)$$

p=10	p=5	n
21	21	30
21	21	50
21	21	100
21	21	200

n	p=5	p=10
500	50	50
1000	100	100
1500	100	100

أما افتراض earth، وهو $\min(200, \max(20, 2p)) + 1$، فلا يراعي حجم مجموعة البيانات وقد يكون أصغر ممّا ينبغي في المجموعات الكبيرة، فيقيّد المرورَ الأمامي قبل أن يستكشف كل المتنبئات استكشافًا كافيًا. والحدُّ $\lfloor n/10 \rfloor$ يتيح نحو حدٍّ واحد لكل عشر مشاهدات --- وهي قاعدةٌ تقريبية معتادة لتفادي فرط المعلمة. أما السقف عند 100 فيمنع تناقص العوائد.

minspan (أدنى عدد مشاهداتٍ بين العُقد):

$$\text{minspan} = \min\bigl(16,\ \lfloor 5 + n/50 \rfloor\bigr)$$

endspan (أدنى عدد مشاهداتٍ من حدود البيانات):

$$\text{endspan} = \min\bigl(16,\ \lfloor 5 + n/28 \rfloor\bigr)$$

n	minspan	endspan
30	5	6
50	6	6
100	7	8
200	9	12
300	11	15
500	15	16
1500	16	16

وتستعمل minspan المحسوبة تلقائيًا في earth معادلةَ فريدمان رقم 43: $\lfloor (-\ln(-\ln 0.95) + \ln(p \cdot n))/(2.5 \ln 2) \rfloor$، وهي تتدرّج كـ$\ln(np)$ وتنمو نموًّا أبطأ ممّا ينبغي في المجموعات الكبيرة. فعند n=1500 لا تعطي إلا $\approx$7، بما يتيح نحو 245 موضع عقدةٍ مرشَّح لكل متنبئ متصل --- وهو كثيرٌ يتيح للمرور الأمامي ملاءمةَ الضوضاء. أما الصيغة المُوصى بها فتستهدف نحو مئة مرشَّح لكل متنبئ عند n الكبيرة، مع ترك الأمر للحساب التلقائي في earth عند n الصغيرة (حيث هو مُختبَرٌ جيدًا). وتُضبط endspan أكبرَ قليلًا من minspan لتوفير حمايةٍ إضافية عند الحدود، وهو ما يفيد حين تكون للبيانات أذيالٌ رقيقة (وذلك شائعٌ في العقارات).

penalty (جزاء GCV لكل عقدة):

$$\text{penalty} = \begin{cases} 3 & \text{degree if} > 1 \\ 2 & \text{degree if} = 1 \end{cases}$$

وهذا افتراض earth نفسه --- 2 للنماذج الجمعية، و3 لنماذج التفاعل (فالجزاء الأعلى يعوّض اتساع فضاء البحث).

newvar.penalty (جزاء إدخال متنبئ جديد):

المُوصى به: **0.1** حين تكون المتنبئات مرتبطة (كمساحة المعيشة مع عدد غرف النوم وعدد الحمّامات ومساحة الأرض). فهذا يحيّز المرورَ الأمامي نحو إعادة استعمال المتنبئات الموجودة في النموذج بدل إدخال بدائلَ مرتبطة. والآلية: أن تحسّنَ مجموع مربعات البواقي لكل متنبئ جديد يُضرب في $1/(1 + \text{penalty})$. فمع newvar.penalty $= 0.1$، يجب أن يحسّن المتغيرُ الجديد مجموعَ مربعات البواقي بما لا يقلّ عن 10% فوق ما تحسّنه عقدةٌ أخرى على متغيرٍ قائم. وهذا ينتج نماذجَ أبسط بمتنبئاتٍ أقل من غير أن يؤثّر في تقديرات المعاملات النهائية (إذ يُزال الجزاء بعد الانتقاء). واتركه عند 0 إن لم تكن المتنبئات مرتبطة. ولاحظ أن earth لا يدعم `newvar.penalty` غير الصفري مع أوزان الحالات --- فـearthUI يعطّل المُدخَل (مُلزِمًا إياه بـ0) كلّما عُيِّن عمود أوزان.

2.7.7 معاملات التشذيب

pmethod: المُوصى به: **backward**. فالتشذيب الخلفي الافتراضي القائم على GCV حتميّ --- إذ تنتج البياناتُ نفسُها النموذجَ نفسَه دائمًا، بلا اعتمادٍ على بذور العشوائية. وهذا مهمٌّ لقابلية الاستنساخ، ولا سيما في عمل التقييم. أما طريقة `cv` فتستعمل التحقق المتقاطع للتشذيب، وهو ما يُدخل اعتمادًا على البذرة.

nprune: المُوصى به: **اتركه فارغًا** (NULL). ودع GCV ينتقي العدد الأمثل من الحدود. فضبطُ nprune يفرض سقفًا صارمًا يتجاوز حكمَ GCV.

3.7.7 معاملات التحقق المتقاطع

nfold (طيّات التحقق المتقاطع): المُوصى به بحسب نطاق حجم العيّنة:

n (صفوف الملاءمة)	nfold
0--299	5
300--399	6
400--499	7
500--599	8
600--699	9
+700	10

فزيادةُ الطيّات تدرّب كل طيّة على بياناتٍ أكثر (فيقلّ تحيّز التقدير) لكنها تصغّر طيّة الاختبار المحجوزة. ويتدرّج الجدولُ صعودًا إلى العشر طيّات المعيارية مع نموّ العيّنة، مع إبقاء كل طيّة اختبارٍ كبيرةً كفايةً (نحو خمسين صفًّا أو أكثر) لمقياسٍ مستقر لكل طيّة، ويرجع إلى خمس طيّات في العينات الصغيرة --- حيث يبقى تقدير CVR^2 مشوَّشًا، فارفع ncross (إلى 15–20 مثلًا) لتوسيط التباين. أما السقف عند عشرٍ فيعكس أن كل طيّةٍ إضافية ملاءمةُ earth كاملة مقابل مكسبٍ تشخيصي لا يُذكر. ومع `pmethod = "backward"` لا يؤثّر التحقق المتقاطع في النموذج --- بل يحسب CVR^2 التشخيصي فحسب ويوفّر بواقيَ لنموذج التباين.

ncross (تكرارات التحقق المتقاطع): المُوصى به بحسب نطاق حجم العيّنة:

n (صفوف الملاءمة)	ncross
0--199	20
200--399	16
400--699	12
700--1199	10
1200--2999	7
+3000	5

فالتكرارات توسِّط حظَّ تقسيم الطيّات لتثبيت CVR^2؛ والعيّنات الأكبر تحتاج إلى تكراراتٍ أقل لأن تحققًا متقاطعًا واحدًا يكون فيها أثبتَ أصلًا. كما تغذّي بواقي التحقق المتقاطع المكرَّر نموذجَ التباين.

4.7.7 نموذج التباين

varmod.method: المُوصى به: **lm** --- وهو افتراض ميلبورو: نموذج تباينٍ خطي حصين لفترات التنبؤ (إذ يُقدَّر التشتّت من بواقيَ مشوَّشة، فيعمّم نموذجٌ بسيط تعميمًا أفضل). فإن أخفق `lm` في التقارب على بياناتك، فجرّب `rlm` أو `const` (متجانس التباين، ويتقارب دائمًا)، أو ارجع إلى `earth` (وهو يعالج التشتّت اللاخطي أيضًا). ويقتضي `nfold > 1` و `ncross > 1`.

5.7.7 ملف مخرجات Earth

بعد كل ملاءمةٍ ناجحة يكتب earthUI ملفًّا باسم `<filename>_earth_output_<timestamp>.txt` في مجلد المخرجات. ويحتوي هذا الملف على حدود النموذج، والإحصاءات الملخِّصة (R^2، وGRSq، وCVRSq)، وتفاصيلِ نموذج التباين، وسجلِّ التتبّع. ويُنشأ ملفٌّ واحد لكل ملاءمة، فيوفّر سجلًّا تراكميًا لمقارنة تشكيلات المعاملات.

8.7 افتراضات الإعدادات

يتحكّم زرُّ اختيارٍ أعلى القسم 4 في الافتراضات التي تُحمَّل:

- **استعمال آخر إعدادات لملف الإدخال** --- يستعيد الإعدادات التي استعملتها آخر مرة مع هذا الملف بعينه
- **استعمال الإعدادات الافتراضية** --- يطبّق افتراضاتك المخصَّصة المحفوظة

- **افتراضات Earth** --- يعيد كل المعاملات إلى قيم المصنع

وانقر **حفظ الحالي بوصفه الافتراضي** أسفل القسم 4 لتخزين معاملات earth الحالية ومصفوفة التفاعلات المسموح بها؛ في حين يخزّن الزرُّ المقابل في القسم 3 الهدفَ والتاريخَ الفعّال وضبطَ المتنبئات. وكلاهما يُحفظ لكل مشروعٍ وغرض في قاعدة بيانات المشروع.

8 ملاءمة نموذج Earth

1.8 زرّ الملاءمة

يحتوي القسم 5 من الشريط الجانبي على زرّ **ملاءمة نموذج Earth** مع حقل **بذرة العشوائية**. والبذرةُ تجعل إسنادَ طيّات التحقق المتقاطع قابلًا للاستنساخ --- وتستدعي قائمةٌ منسدلة آخر عشر بذورٍ استُعملت، وتُسجَّل البذرة المستعملة في لسان مخرجات Earth بعد الملاءمة.
وقبل الملاءمة يتحقّق earthUI من صحة ضبطك --- إذ يجب أن يكون هدفٌ مختارًا ومتنبئٌ واحد على الأقل مُدرَجًا. وفي نمطَي التقييم والسوق تُقرَّب أعمدة خطَّي العرض والطول، ويُحسب عمر البيعة من التاريخ الفعّال وتاريخ التعاقد (إن عُيِّن).

2.8 نافذة الملاءمة ومخرجات التتبّع

حين تنقر ``ملاءمة" تظهر طبقةٌ داكنة مشروطة فيها:

- **ترويسة** --- ``ملاءمة نموذج Earth" مع مؤقّتٍ للزمن المنقضي يعدّ الثواني وزرّ **إجهاض** يلغي الملاءمة الخلفية
- **سجلّ التتبّع** --- منطقةُ نصٍّ ثابتة العرض قابلة للتمرير تعرض المخرجات الآنية من الدالة `earth()`: معلومات مجموعة البيانات، وتقدّمَ المرور الأمامي، وطيّات التحقق المتقاطع، ووقت الإتمام
- **الترميز اللوني** --- الأسطر الافتراضية بالأخضر، وأسطر التحقق المتقاطع بالأصفر، والأخطاء بالأحمر

ويلائم earthUI النماذجَ لاتزامنيًا باستعمال `callr::r_bg()`، الذي يشغّل الحساب في عملية R منفصلة. وهذا يُبقي التطبيق مستجيبًا أثناء الملاءمات الطويلة. ويقرأ مراقِبٌ يستطلع كل 300 جزءٍ من الألف من الثانية مخرجاتِ العملية الخلفية ويبثّها إلى سجلّ التتبّع.
وعند اكتمال الملاءمة تظهر رسالة ``تمّ في X.X ثانية"، ويُضاف زرّ إغلاق (X) إلى النافذة، وتُزال الخلفية فتُصيَّر ألسنةُ النتائج عند الطلب. ثم تُغلَق النافذة تلقائيًا بمجرد انتهاء تصيير المخرجات المرئية --- أو يمكنك إغلاقها قبل ذلك.
وعند النجاح تُلحَق علامةُ تأشيرٍ خضراء بزرّ الملاءمة. وإن وقع خطأ، عُرضت رسالتُه في سجلّ التتبّع وظهر إشعار.

3.8 سجلّ الملاءمة

تكتب كلُّ ملاءمة (ناجحةً كانت أو مخفقة) ملفَّ سجلٍّ في مجلد مخرجاتك:

- **اسم الملف**: `<datafile>_earth_log_<YYYYMMDD_HHMMSS>.txt`
- **المحتوى**: الختم الزمني، وكل مخرجات التتبّع من عملية الملاءمة، وأي رسائل خطأ
- **الموضع**: مجلد مخرجات المشروع النشط

> **تنبيه:** إن لم تكن *callr* منصَّبة، رجع *earthUI* إلى الملاءمة التزامنية بشريط تقدّمٍ بسيط. فنصّب *callr* للحصول على تجربة مخرجات التتبّع كاملة.

9 تنزيل البيانات

بعد الملاءمة، ينزّل القسم 6 من الشريط الجانبي --- وعنوانه ``تنزيل المتغير الهدف المقدَّر والبواقي" في النمط العام و``تنزيل أسعار البيع المقدَّرة والبواقي" في نمطَي التقييم والسوق --- ملفَّ Excel فيه التنبؤات والتشخيصات. وتُستعمل هذه المخرجات في الخطوة 7 (RCA) لإسناد تصنيف CQA (الحالة والجودة والجاذبية) إلى العقار موضوع التقييم. والمخرجاتُ مرتَّبةٌ بـ`residual_sf` و`cqa_sf` لتعينك على تقدير موضع العقار موضوع التقييم في الترتيب. فإن كان النموذج جيد الجودة، وجب أن تُرتَّب العقارات من الأقلّ جاذبيةً إلى الأكثر جاذبية بناءً على الخصائص الباقية التي لم تدخل الانحدار. وينبغي أن تقارب القيمةُ الوسطى الصفرَ، وأن يكون النصف الأدنى قيمًا سالبة والنصف الأعلى قيمًا موجبة. وستجد أردأ المنازل جودةً، أو ``ما يحتاج إلى إصلاح"، قرب أسفل الترتيب، وأجملَها في أعلاه. وثمة استثناءاتٌ عادةً لشذوذاتٍ كالحجوزات، والبيع بأقل من الدين، والبيع بحكم التركة، والبيع السريع اللازم لتغيير عملٍ أو لأسبابٍ أخرى. وتحرّي الشذوذات يكشف عادةً سببًا وجيهًا لشذوذ السعر.

وتتكيّف تسميةُ الزرّ مع نمط الغرض:

- العام/السوق: **تنزيل المخرجات (Excel)**
- التقييم: **تنزيل المخرجات الوسيطة (Excel)**

وصيغة اسم الملف هي `<datafile>_modified_<YYYYMMDD_HHMMSS>.xlsx`.

1.9 أعمدة المخرجات

تُلحَق الأعمدةُ الآتية بكل متغيرٍ هدف:

العمود	الوصف
`est_<target>`	تنبؤ النموذج، مثل `est_sale_price` (بمنزلةٍ عشرية واحدة)
`residual`	الفعلي ناقصًا المتنبأ به (بمنزلةٍ واحدة)
`cqa`	درجة الحالة والجودة والجاذبية (بمنزلتين، في المدى 0–10)
`residual_sf`	الباقي مقسومًا على مساحة المعيشة (إن عُيّنت، بمنزلةٍ واحدة)
`cqa_sf`	درجة CQA محسوبةً من الترتيب عبر `residual_sf` (بمنزلتين)
`<variable>_contribution`	قيمة إسهام كل دالة g (بمنزلةٍ واحدة، وعمودٌ لكل دالة)
`basis`	إسهام قيمة الحدّ الثابت، وهو واحدٌ لكل العقارات (بمنزلةٍ واحدة)
`calc_residual`	الفعلي ناقصًا (الأساس + كل الإسهامات) --- عمود تحقّق (بمنزلةٍ واحدة)

وفي النماذج متعددة الأهداف تُضاف لاحقةٌ على الصورة `_<i>` لتمييز أعمدة كل هدف.

2.9 ترتيب الأعمدة وتنسيق Excel

تُوضع أعمدة الترتيب في **أقصى اليسار** في ملف المخرجات، بهذا الترتيب: `residual_sf`، و`cqa_sf`، و`residual`، و`cqa`. وهذا يسهّل مسحَ القائمة المرتَّبة وتقديرَ موضع العقار موضوع التقييم في توزيع CQA.

وتُطبَّق على هذه الأعمدة تنسيقاتُ أعدادٍ في Excel:

العمود	التنسيق	مثال
`residual_sf`	عددي، بمنزلتين عشريتين	12.34
`cqa_sf`	عددي، بمنزلتين عشريتين	7.25
`residual`	عددي، بلا منازل عشرية	15,200
`cqa`	عددي، بمنزلتين عشريتين	6.80

3.9 درجات CQA

ترتّب درجةُ CQA (الحالة والجودة والجاذبية) باقيَ كل صفٍّ في مقابل سائر البواقي. فـCQA لصفٍّ ما هي نسبةُ الصفوف ذات الباقي المُشار الأصغر، مضروبةً في عشرة. وهذا يعطي مقياسًا من 0 إلى 10 حيث:

- **CQA مرتفعة** ($\approx$ 9–10) --- بيع العقار بأكثر بكثير ممّا تنبّأ به النموذج (باقٍ موجب كبير)

- **CQA منخفضة** ($\approx$ 0–1) --- بيع بأقل بكثير ممّا تنبّئ به النموذج (باقٍ سالب كبير)
- **CQA** $\approx$ **5** --- الباقي قريبٌ من وسيط كل البواقي

وحين يُعيَّن عمود `living_area`، توفّر `cqa_sf` الترتيبَ نفسه بناءً على البواقي للقدم المربعة.

4.9 الترتيب وصف العقار موضوع التقييم

في نمطَي التقييم والسوق تُرتَّب الصفوف المقارِنة تنازليًا بـ`residual_sf` (أو بـ`residual` إن لم تُعيَّن مساحة معيشة). ويبقى صف العقار موضوع التقييم (الصف 1) في موضعه الأول ولا يُرتَّب. وفي النمط العام تُصدَّر الصفوف بترتيبها الأصلي.

وهذا الترتيب، مقترنًا بأعمدة الترتيب في أقصى اليسار، يتيح للمُقيّم أن يمسح المقارنات سريعًا من الأكثر مغالاةً في التنبؤ إلى الأكثر بخسًا فيه، وأن يقدّر موضع درجة CQA المُسنَدة إلى العقار موضوع التقييم في ذلك التوزيع.

وتُحاذى مستوياتُ العوامل في بيانات التنبؤ مع بيانات التدريب قبل استدعاء `predict()`. والصفوفُ ذات مستويات العوامل غير المرئية سابقًا تُنتج تنبؤاتٍ مفقودة.

> **تنبيه:** ينبغي أن يساوي عمود *`calc_residual`* عمودَ *`residual`* دائمًا (في حدود التقريب). فإن اختلفا اختلافًا كبيرًا، دلّ ذلك على مشكلةٍ حسابية جديرة بالتحرّي.

5.9 التصدير التلقائي إلى mgcvUI

في كل ملاءمةٍ ناجحة بدرجةٍ $2 \geq$، يحفظ earthUI تلقائيًا كائنَ النتيجة كاملًا ملفَّ `rds`. في مجلد مخرجات المشروع. ويتبع اسم الملف النمط `<datafile>_earthUI_result_<YYYYMMDD_HHMMSS>.rds`.

ويمكن تحميل هذا الملف في وحدة المعالجة اللاحقة لـmgcv --- المضمَّنة في حزمة earthUI والمشغَّلة بـ`earthUI::launch_mgcv()` --- باستعمال `readRDS()`. ويستعمل روتين النماذج الجمعية المعمّمة مواضعَ عُقد نموذج earth ودوالَّ أساسه نقاطَ انطلاقٍ للحدود الملساء، فيتيح انتقالًا سلسًا من نمذجة MARS إلى نمذجة النماذج الجمعية المعمّمة.

وتُتخطّى النماذج ذات الدرجة الأكبر من 2 لأن mgcvUI لا يدعم إلا التفاعلات الثنائية. وحين يُشغَّل earthUI في نمط الثلاثية (من واجهة الثلاثية)، يتيح لك قسمٌ إضافي باسم **قفل مخرجات النموذج** أن تقفل ملفَّ `rds` لملاءمةٍ بعينها --- مع أساس متنبئاتها وإعدادات CQA المشتركة --- بوصفه نموذج earth المعتمَد الذي يستورده glmnetUI وmgcvUI.

6.9 الاستيراد إلى glmnetUI

ويمكن كذلك استيراد ملف `rds`. نفسِه المحفوظ لروتين النماذج الجمعية المعمّمة إلى **وحدة المعالجة اللاحقة للشبكة المرنة** --- وهي كذلك مضمَّنةٌ في حزمة earthUI وتُشغَّل بـ`earthUI::launch_glmnet()`. وفي تطبيق glmnet، انتقل إلى **القسم 2: الاستيراد من earthUI** واستعمل زرّ الاستعراض لاختيار ملف `.rds`.

ويستعمل glmnetUI المقاربة المعيارية `earth::model.matrix()`: إذ تصير دوالُّ أساس نموذج earth (المفصلات والتفاعلات والحدود الخطية) أعمدةَ مصفوفة التصميم في glmnet. ثم يجري glmnet انحدارًا مُنتظَمًا على أعمدة الأساس هذه، فينتقيها ويقلّصها عبر لاسو/الشبكة المرنة. وهذا يجمع بين بناء earth التكيّفي للأساس وبين انتظام glmnet.

1.6.9 المزايا الرئيسة

- **تفاعلاتٌ قابلة للتفسير**: لتفاعلات earth القائمة على المفصلات (بما فيها الحدود الثلاثية) معنًى هندسي واضح، بخلاف تفاعلات glmnet الأصلية القائمة على الجداءات التي يصعب شرحها في المحاكم أو في سياقات التدقيق.
- **لاخطيةٌ تلقائية**: تلتقط دوالُّ المفصلة في earth علاقاتٍ لاخطية كان نموذج glmnet المعياري بالمتنبئات الخام سيفوّتها.
- **انتقاء المتغيرات**: يستطيع glmnet تشذيبَ أساس earth أكثر، فيُسقط حدود المفصلة التي لا تُسهم.

2.6.9 مهم: يجب أن يتطابق عمود الأوزان

> **تحذير**
>
> إن استعمل earthUI عمودَ أوزان (لاستبعاد صف العقار موضوع التقييم بوزنٍ صفري مثلًا)، وجب تعيينُ عمود الأوزان نفسِه في ضبط المتغيرات لدى glmnetUI. فإن اختلف ضبط الأوزان، لم تتحاذَ أعدادُ الصفوف ولم يمكن إلحاقُ أساس earth بمصفوفة نموذج glmnet.

3.6.9 سير العمل

1. لائم نموذج earth في earthUI (بأي درجة).
2. يُحفظ ملف نتيجة `.rds` تلقائيًا في مجلد مخرجات المشروع.
3. وفي glmnetUI، افتح **القسم 2: الاستيراد من earthUI** واستعرض إلى ملف `..rds`.
4. تحقّق من استعمال **ملف البيانات نفسه** و**عمود الأوزان نفسه** في التطبيقين.
5. اضبط معاملات glmnet في القسم 5 (وتُعطَّل مصفوفةُ التفاعلات تلقائيًا عند وجود استيرادٍ نشط من earthUI).
6. انقر **ملاءمة النموذج**. فتظهر أعمدةُ أساس earth في ألسنة المعادلة والمعاملات والإسهامات.

10 حسابات RCA وتنزيلها

يظهر زرّ **حساب تعديلات RCA وتنزيلها** في القسم 7 من الشريط الجانبي، ولا يظهر إلا في نمط التقييم بعد ملاءمة نموذج. ويُنتج منهج قيد البواقي (RCA) تعديلاتٍ مستمدّةً من السوق لكل مقارِن نسبةً إلى العقار موضوع التقييم.

1.10 فتح حوار RCA

نقرُ الزرّ يفتح حوارًا مشروطًا صغيرًا فيه:

- **نوع الدرجة** --- أزرارُ اختيارٍ لانتقاء **CQA** أو **CQA للقدم المربعة** (ولا يتاح خيار القدم المربعة إلا حين يُعيَّن عمود `living_area`). فإن اخترت ``CQA للقدم المربعة''، كانت درجةُ باقي العقار موضوع التقييم، بناءً على درجة CQA التي تُسنِدها إليه، هي مساحةَ معيشته مضروبةً في `residual_sf` المطابقة لدرجة CQA_SF المعطاة.
- **قيمة CQA** --- مُدخَلٌ لدرجة CQA للعقار موضوع التقييم (0.00–10.00). وتمثّل هذه الدرجة موضعَه الذي تعتقده في توزيع جودة المقارنات.
- زرّ **توليد** --- يحسب مخرجات RCA وينزّلها بصيغة Excel. ويبقى الزرّ معطَّلًا حتى تُدخَل درجةٌ صحيحة.

2.10 استكمال درجة CQA

حين تنقر ``توليد'' يستكمل earthUI باقيَ العقار موضوع التقييم من أزواج CQA/الباقي للمقارنات:

1. تُرتَّب درجات CQA للمقارنات وبواقيها تصاعديًا بـCQA
2. ويطبّق استكمالٌ خطي (`stats::approx()`) قيمةَ CQA التي أدخلتها على باقٍ
3. وإن استعملتَ CQA للقدم المربعة، حُوِّل الباقي للقدم المربعة إلى باقٍ إجمالي بضربه في مساحة معيشة العقار موضوع التقييم
4. وتُحسب القيمة المقدَّرة للعقار موضوع التقييم على النحو: تنبؤ النموذج + الباقي المستكمَل

3.10 أعمدة المخرجات

يتضمّن ملف RCA بصيغة Excel (`<datafile>_adjusted_<YYYYMMDD_HHMMSS>.xlsx`) كلَّ أعمدة المخرجات الوسيطة، مضافًا إليها:

العمود	الوصف
`subject_value`	تنبؤ النموذج + الباقي المستكمَل (الصف 1 فقط)
`subject_cqa`	درجة CQA التي أدخلتها (الصف 1 فقط)
`<variable>_adjustment`	إسهام العقار موضوع التقييم ناقصًا إسهام المقارِن (لكل دالة g)
`residual_adjustment`	باقي العقار موضوع التقييم ناقصًا باقي المقارِن
`net_adjustments`	مجموع كل التعديلات (الإسهام + الباقي)
`gross_adjustments`	مجموع القيم المطلقة لكل التعديلات
`residual_pct`	تعديل الباقي بوصفه كسرًا من سعر بيع المقارِن
`net_adj_pct`	صافي التعديلات بوصفه كسرًا من سعر بيع المقارِن
`gross_adj_pct`	التعديلات الإجمالية بوصفها كسرًا من سعر بيع المقارِن
`adjusted_sale_price`	سعر بيع المقارِن + صافي التعديلات

وتخبرك أعمدةُ التعديل، لكل مقارِن، بمقدار ما ينسبه النموذجُ من الفرق إلى كل متغير. و`adjusted_sale_price` هو سعرُ بيع المقارِن بعد تطبيق كل التعديلات المستمدّة من النموذج --- أي مجموعةٌ من أسعار البيع المعدَّلة ينبغي أن تتجمّع حول القيمة المقدَّرة للعقار موضوع التقييم.

4.10 لسان تعديلات RCA

بمجرد حساب مخرجات RCA، يعرض لسانُ **تعديلات RCA** في اللوحة الرئيسة مدرَّجاتٍ لتوزيعات نسبة تعديل الباقي وصافي التعديل والتعديل الإجمالي عبر المقارنات (مجموعةً لكل هدف في النماذج متعددة الأهداف)، بما يعطي فحصًا بصريًا سريعًا لمدى تجمّع التعديلات.

تنبيه: درجة *CQA* البالغة *5.00* تضع العقار موضوع التقييم عند وسيط المقارنات. والدرجاتُ فوق *5* تدلّ على أنه فوق متوسط الجودة؛ ودونها تدلّ على أنه دون المتوسط.

11 شبكة المقارنة بالمبيعات

1.11 نظرة عامة والغرض

شبكة المقارنة بالمبيعات مصنَّف Excel يُولَّد في نمط التقييم (القسم 8 من الشريط الجانبي) بعد حساب تعديلات RCA. وهي تعرض العقار موضوع التقييم إلى جانب المقارنات المختارة في صيغة شبكةٍ مبنينة، بمعادلات Excel تحسب تلقائيًا التعديلات وسعرَ بيعٍ معدَّلًا لكل مقارِن.

والشبكةُ مصمَّمةٌ لملف عمل المُقيّم --- فهي تجمع بين التعديلات المستمدّة من انحدار نموذج Earth وبين خلايا قابلة للتحرير يستطيع فيها المُقيّم توزيعَ باقي CQA على خصائص عقارية بعينها. واسم ملف المخرجات هو `SalesGrid_<YYYYMMDD_HHMMSS>.xlsx`.

2.11 نافذة انتقاء المقارنات

نقرُ ``توليد شبكة المبيعات وتنزيلها" يفتح حوارًا مشروطًا تختار فيه المقارنات المُدرَجة:

- **المقارنات المُوصى بها** مؤشَّرةٌ مسبقًا. وهي المقارنات ذات نسبة تعديلٍ إجمالي دون 25% من سعر البيع، مرتَّبةً تصاعديًا بعمر البيعة (فالمبيعات الأحدث أولًا) ومحدودةً بثلاثين.
- **المقارنات الإضافية** مُدرَجةٌ أدناها، بلا تأشيرٍ افتراضًا. وتعديلاتها الإجمالية 25% فأكثر لكنها تبقى متاحةً للاختيار (ويُعرض منها حتى خمسون، مرتَّبةً تصاعديًا بنسبة التعديل الإجمالي).
- ويمكن اختيار **ثلاثين مقارِنًا** كحدٍّ أقصى، بما يُنتج حتى **عشر أوراق** (ثلاثة مقارنات لكل ورقة).

وبعد تأكيد اختيارك، تُرتَّب المقارنات المختارة تصاعديًا بنسبة التعديل الإجمالي، فتظهر أشبهُ المقارنات في الورقة 1.

3.11 تخطيط الشبكة

لكل ورقةٍ تخطيطٌ من عشرين عمودًا يتّسع للعقار موضوع التقييم مع ثلاثة مقارنات. وأعمدةُ كل كيان هي:

الكيان	الأعمدة (خمسة لكل كيان)
العقار موضوع التقييم	التسمية، والقيمة الواقعية 1، والقيمة الواقعية 2، والقيمة الواقعية 3، وإسهام القيمة (VC)
كل مقارِن	القيمة الواقعية 1، والقيمة الواقعية 2، والقيمة الواقعية 3، وإسهام القيمة (VC)، والتعديل

والصفوف من الأعلى إلى الأسفل هي:

1. **العنوان** --- اسم الورقة (مثل ``شبكة المقارنة بالمبيعات الوسيطة --- الورقة 1 من 3")
2. **الترويسات** --- ``العقار موضوع التقييم" وعناوين أعمدة المقارنات مع العنوان
3. **العنوان** --- عنوان العقار كاملًا
4. **APN | #MLS | أيام العرض | القرب من العقار موضوع التقييم** --- رقم القطعة، ومعرّف الإدراج، وأيام العرض في السوق، ومسافة Haversine (بالأميال) من العقار موضوع التقييم
5. **سعر البيع | التنازلات | صافي سعر البيع** --- سعر البيع، والتنازلات، وصيغة صافي سعر البيع
6. صف ترويسة **خصائص الانحدار**
7. **القيمة الأساس** --- الحدّ الثابت للنموذج
8. **تاريخ البيع | خارج السوق | داخل السوق** --- تاريخ التعاقد، وعمر البيعة، وأيام العرض في السوق
9. **الصفوف المجمَّعة** (شرطية) --- مجموعات الموقع ومساحة الموقع والعمر
10. **صفوف متغيرات النموذج** --- صفٌّ لكل متنبئ (عدا المتغيرات المجمَّعة)
11. **صف فاصل فارغ**
12. صف ترويسة **الخصائص الباقية**
13. **CQA|الباقي** --- درجة CQA مع صيغة الباقي المتبقّي
14. **صفوف الخصائص الباقية** --- صفوفٌ مسمّاة وأخرى فارغة لإدخال المُقيّم
15. **إجمالي إسهام القيمة / صافي التعديل**
16. **نسبة صافي التعديل**
17. **نسبة التعديل الإجمالي**
18. **سعر البيع المعدَّل** --- صفُّ صيغة
19. تذييل **حقوق النشر**

4.11 سعر البيع والتنازلات وصافي سعر البيع

يعرض الصف 5 ثلاث قيمٍ لكل مقارِن:

- **سعر البيع** --- سعر بيع المقارِن من البيانات
- **التنازلات** --- القيمة من العمود المعيَّن `concessions` (وصفرٌ إن لم يُعيَّن)
- **صافي سعر البيع** --- صيغة Excel: سعر البيع − التنازلات

ويعرض عمودُ العقار موضوع التقييم ``غير متاح" لسعر البيع (إذ هو مجهول) وأي قيمة تنازلاتٍ إن توفّرت.

5.11 الصفوف المجمَّعة (الموقع والموقع والعمر)

حين تُعيَّن أنواعٌ خاصة بعينها وتظهر تلك المتغيرات في النموذج، ينشئ earthUI صفوفًا مجمَّعة تضمّ المتغيرات المتصلة:
الموقع: خط الطول | خط العرض | المنطقة --- يظهر حين يكون أيٌّ من `longitude` أو `latitude` أو `area` في النموذج. ويعرض القيمَ الواقعية لكل متغيرٍ مكوِّن، وإسهامَ قيمةٍ مجمَّعًا (مجموع الإسهامات المفردة)، وللمقارنات تعديلًا (إسهام العقار موضوع التقييم المجمَّع − إسهام المقارِن المجمَّع). ويُنسَّق بخلفيةٍ زرقاء فاتحة على خلايا إسهام القيمة.
مساحة الموقع | الأبعاد --- يظهر حين يكون `lot_size` أو `site_dimensions` في النموذج. وببنيةٍ مماثلة لمجموعة الموقع.
العمر الفعلي | العمر الفعّال --- يظهر حين يكون `actual_age` أو `effective_age` في النموذج. وببنيةٍ مماثلة.
والمتغيراتُ التي تستهلكها الصفوفُ المجمَّعة **تُستبعد من صفوف متغيرات النموذج المفردة** أدناه، منعًا للحساب المزدوج في مجاميع التعديلات.

6.11 صفوف متغيرات النموذج

أسفل الصفوف المجمَّعة (أو مباشرةً أسفل القيمة الأساس إن لم توجد صفوفٌ مجمَّعة)، يعرض صفٌّ لكل متنبئ نموذجٍ متبقٍّ:

- **القيم الواقعية** للعقار موضوع التقييم ولكل مقارِن (أي قيم البيانات الفعلية)
- **إسهام القيمة (VC)** --- إسهام دالة g في القيمة المتنبأ بها لذلك الصف
- **التعديل** (في المقارنات فقط) --- إسهام العقار موضوع التقييم ناقصًا إسهام المقارِن

7.11 صف CQA|الباقي وصفوف الخصائص الباقية

يعرض صف **CQA|الباقي** درجةَ CQA لكل عقار ويحتوي على صيغةٍ لـ**الباقي المتبقّي** --- أي الجزء من الباقي الذي لم يُوزَّع بعدُ على خصائصَ بعينها. والصيغة هي:

$$\text{الباقي المتبقّي} = \text{الباقي الإجمالي} - \sum(\text{إسهامات الخصائص الباقية أدناه})$$

وأسفل هذا **صفوف الخصائص الباقية** --- خمسةُ صفوفٍ مسمّاة (الموقع، والإطلالة/الجاذبية، والحالة، والجودة، وغيرها) مع ستة صفوفٍ فارغة. وهذه خلايا إدخالٍ يُدخل فيها المُقيّم إسهامات القيمة للخصائص التي لم يلتقطها النموذج. وكلما ملأ المُقيّم قيمًا، تناقصت صيغةُ الباقي المتبقّي تلقائيًا.
ولكل مقارِن يحتوي عمود التعديل على صيغة: إسهام خاصية العقار موضوع التقييم − إسهام خاصية المقارِن.

8.11 صيغة سعر البيع المعدَّل

يحتوي صف **سعر البيع المعدَّل** على صيغ Excel:

- **العقار موضوع التقييم**: مجموع كل خلايا إسهام القيمة من القيمة الأساس حتى آخر صفٍّ خاصيةٍ باقية. ويمثّل هذا تنبؤَ النموذج الإجمالي للعقار موضوع التقييم مضافًا إليه أي خصائصَ باقية وزّعها المُقيّم.
- **المقارنات**: صافي سعر البيع + مجموع كل خلايا التعديل من أول صفٍّ مجمَّع أو صفٍّ متغير حتى آخر صفٍّ خاصيةٍ باقية. وهذا يعطي سعر بيع المقارِن بعد كل التعديلات المستمدّة من الانحدار والمُدخَلة من المُقيّم.

9.11 حماية الأوراق والخلايا القابلة للتحرير

كل ورقةٍ **محميّة** منعًا للتعديل العرَضي على الصيغ والبيانات. وتقفل الحمايةُ:

- كل خلايا الصيغ (صافي سعر البيع، والباقي المتبقّي، والتعديلات، وسعر البيع المعدَّل، وغيرها)

- كل خلايا البيانات (العناوين، وأسعار البيع، والقيم الواقعية، وإسهامات القيمة من النموذج)
- التسميات والترويسات

والخلايا **غير المقفلة** (القابلة للتحرير) وحدها هي مُدخَلات إسهام القيمة للخصائص الباقية --- أي الخلايا أسفل صف CQA|الباقي حيث يُدخل المُقيّم التفكيكات. وهي مُنسَّقةٌ بخلفيةٍ صفراء فاتحة للدلالة على أنها قابلة للتحرير.

10.11 العمل بالشبكة في Excel

1. **افتح الملف** في Excel أو في تطبيق جداول بياناتٍ متوافق.
2. **راجع تعديلات الانحدار** --- فالتعديلات المستمدّة من النموذج مملوءةٌ مسبقًا ومقفلة.
3. **وزّع الباقي** --- في صفوف الخصائص الباقية (الخلايا الصفراء)، أدخِل تقديرك لمقدار ما يُنسب من الباقي المتبقّي إلى كل خاصية (الموقع، والإطلالة، والحالة، والجودة، وغيرها). وستتناقص صيغةُ الباقي المتبقّي كلما وزّعت.
4. **تحقّق من سعر البيع المعدَّل** --- فهذه الصيغة تتحدّث تلقائيًا كلما أدخلت قيم الخصائص الباقية.
5. **الأوراق المتعددة** --- إن اخترت أكثر من ثلاثة مقارنات، تنقّل بين الأوراق. ولكل ورقةٍ العقارُ موضوع التقييم نفسه في العمود الأيسر مع ثلاثة مقارنات مختلفة.

تنبيه: الغاية هي توزيع باقي *CQA* على صفوف الخصائص حتى يقارب الباقي المتبقّي الصفرَ. وهذا يبرهن على قدرتك على تفسير الفرق بين تقدير الانحدار لكل مقارِن وسعرِ بيعه الفعلي.

12 تنزيل التقارير

بعد ملاءمة نموذج، يكون توليدُ التقرير سيرَ عمل Quarto من خطوتين: **توليد تقرير Quarto** (القسم 9 من الشريط الجانبي في نمط التقييم، و7 في غيره) يكتب حزمةَ تقريرٍ مكتفية بذاتها، و**تحويل تقرير Quarto** (القسم 10) يصيّرها --- أو أي ملف Quarto آخر --- إلى الصيغ النهائية.

1.12 توليد تقرير Quarto

نقرُ **توليد تقرير Quarto** يكتب حزمة Quarto مكتفيةً بذاتها تحت مجلد مخرجات المشروع في `<base>_qmd/`، تحتوي على مصدر `qmd`. المملوء، وكلِّ أصول الرسوم المولَّدة مسبقًا (PNG وPDF)، وبياناتِ التقرير، وقالبِ Word المرجعي. ولا يجري تصييرٌ في هذه الخطوة --- فيمكن تحرير الحزمة يدويًا، أو دمجها مع مصادر Quarto أخرى عبر تضمينات Quarto، أو معاينتها مستقلةً بـ`quarto preview`.

2.12 تحويل تقرير Quarto

يوفّر القسم 10 منتقي ملفات (يشير افتراضيًا إلى آخر حزمةٍ مولَّدة في المشروع الحالي --- لكن يمكنك استعراض أي ملف `qmd`.) ومربعاتِ تأشير الصيغ:

- **HTML** --- ملف `html`. بتصيير الرياضيات بـKaTeX، وصورٍ مضمَّنة، وفهرس محتويات. ويستعمل سمة Flatly Bootstrap وخط Roboto Condensed.
- **Word** --- ملف `docx`. مناسبٌ للتحرير والتوزيع. ويتضمّن فهرس محتويات وأرقام صفحات، ويستعمل قالبًا مرجعيًا مخصَّصًا لتنسيق متّسق.
- **PDF** --- مُنضَّدٌ بـLuaLaTeX لمخرجاتٍ بجودةٍ مهنية. ويتبع مقاسُ الورق ضبطَ اللغة والمنطقة (Letter أو A4). ويتضمّن صفحاتٍ أفقية لمصفوفات التفاعل الكبيرة، ويستعمل Roboto Condensed مع Latin Modern Math للمعادلات. ولا يُعرض خيار PDF إلا حين تُرصد توزيعةُ LaTeX.

وانقر **تحويل** لتصيير الصيغ المختارة. ويجري التصيير في الخلفية --- إذ يعرض حوارٌ مشروط الزمنَ المنقضي وتقدّمَ Quarto بينما يبقى التطبيق مستجيبًا. ويؤكّد إشعارٌ موضعَ الملف عند الإتمام.

وتُسجَّل كل العمليات (ملاءمة النموذج، وتنزيلات البيانات، وحسابات RCA، وتوليد شبكة المبيعات، وتصيير التقارير) بأختامٍ زمنية للبدء والانتهاء وبالأزمنة المنقضية في `<datafile>_earthui_log.txt` في مجلد المخرجات، لاستكشاف الأعطال ومراقبة الأداء.

3.12 محتويات التقرير

يتضمّن كل تقريرٍ مولَّد الأقسامَ الآتية:

1. **وصف مجموعة البيانات** --- عدد المشاهدات، والمتغير الهدف أو المتغيرات الهدف، والمتنبئات المستعملة، والمتنبئات الفئوية
2. **تحديد النموذج** --- الدرجة، وحالة التحقق المتقاطع، وعدد الحدود، والمتنبئات في النموذج
3. **التفاعلات المسموح بها** (في الدرجة $2 \geq$ فقط) --- مصفوفة التفاعلات بعلامات تأشير، مُنسَّقةً بحسب نوع المخرجات
4. **ملخّص النتائج** --- R^2، وGRSq، وGCV، وRSS، وR^2 للتحقق المتقاطع (إن كان مفعَّلًا)
5. **معادلة النموذج** --- معادلة نموذج Earth كاملةً بترميز LaTeX
6. **المعاملات ودوال الأساس** --- جدولٌ بكل الحدود ومعاملاتها ودوال الأساس التي تعرّفها
7. **أهمية المتغيرات** --- رسمٌ شريطي وجدولٌ مرتَّب لدرجات أهمية المتنبئات
8. **إسهامات دوال g** --- رسومٌ لكل دالة: رسومٌ خطية للحدود أحادية المتغير، ورسومٌ منظورية ثلاثية الأبعاد ورسومُ كنتور للتفاعلات الثنائية
9. **مصفوفة الارتباط** --- خريطةٌ حرارية لارتباطات المتنبئات
10. **التشخيصات** --- رسوم البواقي في مقابل القيم المُلائَمة، ورسم Q-Q الطبيعي، والفعلي في مقابل المتنبأ به
11. **تفكيك تحليل التباين** --- جدولٌ يبيّن إسهام كل دالة أساسٍ في مجموع مربعات البواقي
12. **مخرجات Earth** --- المخرجات النصّية الخام من `earth::summary()`، بما فيها تفاصيل مرور التشذيب

وفي النماذج متعددة المستجيبات تتكرّر الأقسام 8–4 و10 لكل متغيرٍ هدف.

تنبيه: تقتضي التقاريرُ حزمةَ `quarto` في *R* وتنصيبًا لـ*Quarto*. ولمخرجات *PDF* يجب أن تكون توزيعةُ *LaTeX* منصَّبةً كذلك (مثلًا عبر `tinytex::install_tinytex()`) --- وإلا أُخفي مربعُ تأشير *PDF*.

13 وحدتا المعالجة اللاحقة: الشبكة المرنة والنماذج الجمعية المعمّمة

ترفق حزمة earthUI وحدتَي معالجةٍ لاحقة --- روتينَ الشبكة المرنة المبني على `glmnet`، وروتينَ النماذج الجمعية المعمّمة المبني على `mgcv`. ولكلٍّ منهما مقالٌ خاص به في هذا العدد يتناوله بعمق؛ ويشرح هذا القسم **الغرض** منهما في سير عمل earthUI، وكيف تتعاون الروتينات الثلاثة.

1.13 لماذا الملاءمات المُسنِدة؟

نموذج earth هو الذي يُنتج نتيجة القيمة. ووحدتا المعالجة اللاحقة موجودتان لـ**تعضيدها**: بتقديرين إضافيين للعقار نفسِه، ينتجهما محرّكان مختلفان منهجيًا، ملائَمين على بيانات المشروع نفسها. فحين تصل إعادةُ تقدير أساس earth بالشبكة المرنة وملساءُ نموذجٍ جمعي معمّم مستقلة إلى القيمة نفسها جوهريًا، لم تعد نتيجة المُقيّم قائمةً على افتراضات نموذجٍ واحد --- وتلك ميزةٌ ذات وزن في سياقات المراجعة والتقاضي. وحين **تختلفان**، كان الاختلافُ نفسه تشخيصيًا، وهو يشير عادةً إلى بياناتٍ رقيقة، أو حدٍّ مفرط المرونة، أو مقارِنٍ شاذّ يستحقّ إعادة النظر. ولا تعلو أيٌّ من الوحدتين على نموذج earth؛ فهما بيّنةٌ مُسنِدة، تُعرض عمدًا من غير ترجيح.

2.13 التشغيل والحالة المشتركة

كلُّ روتينٍ تطبيقُ Shiny محلي منفصل:

- `earthUI::launch_glmnet()` --- تطبيق الشبكة المرنة، على المنفذ 7879.
- `earthUI::launch_mgcv()` --- تطبيق النماذج الجمعية المعمّمة، على المنفذ 7880.

وتقرأ التطبيقات الثلاثة كلها شجرةَ المشاريع نفسَها وقاعدةَ الإعدادات نفسَها، فيكون المشروع المفتوح لديك في earthUI متاحًا فورًا في الوحدتين --- مع الحقول المنقولة التي يحفظها earthUI له: المستجيب (الهدف)، والتاريخ الفعّال، ونوعَ درجة CQA للعقار موضوع التقييم وقيمتَها في RCA، وتعيينَ مساحة المعيشة. فلا تعيد إدخال هذه؛ بل تقرؤها الوحدتان لتسري القيمُ نفسها عبر الملاءمات الثلاث.

3.13 روتين الشبكة المرنة

يستورد روتين glmnet دوالَّ أساس نموذج earth --- مفصلاتِه وتفاعلاتِه وحدودَه الخطية --- أعمدةً لمصفوفة تصميم، ويعيد تقديرها تحت انتظام لاسو/الشبكة المرنة (انظر ``الاستيراد إلى glmnetUI" أعلاه للآلية ولتحفّظ عمود الأوزان). والنتيجة تُبقي حدودَ earth القابلة للتفسير وذات المعنى الهندسي، مع ترك المُنتظِم يقلّص أو يُسقط ما لا يستحقّ منها. انظر مقال glmnetUI في هذا العدد للمعالجة الكاملة.

4.13 روتين النماذج الجمعية المعمّمة

ويمضي روتين mgcv في الاتجاه الآخر: فبدل إعادة استعمال أساس earth، يلائم حدودًا ملساء مُجازاة جديدة، مستعملًا مواضعَ عُقد نموذج earth نقاطَ انطلاق (انظر ``التصدير التلقائي إلى mgcvUI" أعلاه). ولأن النموذج الجمعي المعمّم يعيد اشتقاق الصور الدالّية تحت نموذجٍ تنعيمي مختلف، كان اتفاقُه مع نموذج earth تعضيدًا مستقلًا حقًّا لا صدًى. وتُدعم نماذج earth ذات الدرجة $2 \leq$. انظر مقال mgcvUI في هذا العدد للمعالجة الكاملة.

5.13 سير عملٍ عملي

1. لائم **نموذج earth أولًا** وصقلْه، في earthUI، حتى ترضى عنه. وكل ملاءمةٍ ناجحة (بدرجةٍ $2 \leq$) تصدّر تلقائيًا ملفَّ `.rds` الذي تستهلكه الوحدتان.
2. افتح روتينَي **glmnet** و**mgcv** على المشروع نفسه، واستورد نتيجة earth، ولائم. واعمل في روتينٍ واحد في المرة --- فالملاءمات طويلة والروتينات تتنسّق عبر المشروع المشترك، فالإيقاع المعقول تتابعيٌّ لا متواز.
3. وكرِّر بحرّية. فكل تطبيقٍ يحتفظ بنتائجه وأنت تتنقّل بينها، ولا شيء يُلزمك بتوليد تقارير في الأثناء.
4. ولا تعد إلى كل روتينٍ لتوليد تقرير Quarto الخاص به إلا حين ترضى عن النتائج الثلاث كلها. ويسجّل كل منهجٍ ملاءمتَه ونتيجةَ قيمته في سجلّ الثلاثية بالمشروع، فتبقى التقارير الثلاثة قابلةً للتتبّع إلى تشغيلٍ مقارِنٍ واحد.

تنبيه: الوحدتان اختياريتان على كل مستوى: فسير عمل *earthUI* نفسِه --- الملاءمة، و*RCA*، وشبكات المبيعات، والتقارير --- كاملٌ من دونهما. فاستعن بهما حين تستوجب النتيجةُ تعضيدًا: عقارٌ صعب، أو قطاعُ سوقٍ رقيق، أو مهمّةٌ يُرجَّح أن تواجه مراجعة.

14 مجموعة البيانات التجريبية: Appraisal_1.csv

يتضمّن earthUI مجموعةَ بيانات MLS تجريبية لاستكشاف سير عمل التقييم. وحمّلها برمجيًا بـ:

```
demo_file <- system.file("extdata", "Appraisal_1.csv", package = "earthUI")
df <- import_data(demo_file)
```

أو، في تطبيق Shiny، انسخها إلى مجلد `in/` بمشروعٍ ما (أو أرفقها ملفَّ بياناتٍ أوّليًا عند إنشاء المشروع) واخترها في القسم 2 (استيراد البيانات).

1.14 الوصف

يحتوي الملف على 1,501 بيعةً سكنية مع عقارٍ واحد موضوع تقييم في الصف 1 (أي 1,502 صفًّا في المجمل) من تصدير MLS محاكى. وتمثّل البيانات مبيعات منازل أحادية الأسرة في سوقٍ متعددة المناطق، بمدًى من أحجام العقارات وأعمارها ومواقعها.

وهذه ليست بياناتٍ حقيقية، لكنها مبنيّةٌ على حيٍّ واقعي في شمال كاليفورنيا. وقد غُيِّرت كل معلومات التعريف أو أُزيلت.

2.14 الأعمدة

العمود	النوع	النوع الخاص	الوصف
weight	numeric	---	وزن المشاهدة (0 = استبعاد من الملاءمة)
id	numeric	display_only	معرّف السجل الداخلي
property_id	numeric	display_only	معرّف العقار في MLS
listing_id	character	display_only	رقم الإدراج في MLS
parcel_number	character	display_only	رقم القطعة لدى مُقيّم المقاطعة (APN)
street_address	character	display_only	عنوان العقار
city_name	character	display_only	المدينة
postal_code	character	display_only	الرمز البريدي
county_name	character	display_only	المقاطعة
contract_date	Date	contract_date	تاريخ التعاقد على البيع (يحسب sale_age)
sale_age	numeric	---	الأيام من تاريخ التعاقد إلى التاريخ الفعّال (محسوبةً مسبقًا)
coe_date	Date	display_only	تاريخ إغلاق الضمان
listing_status	character	display_only	حالة الإدراج (مثل ``مباع")
sale_price	numeric	(الهدف)	سعر البيع --- المتغير المستجيب
rent	numeric	---	الإيجار الشهري (للنماذج متعددة الأهداف)
list_price	numeric	display_only	سعر الإدراج
original_list_price	numeric	display_only	سعر الإدراج الأصلي
living_sqft	numeric	living_area	مساحة المعيشة الإجمالية بالأقدام المربعة
beds_total	integer	---	عدد غرف النوم
baths_total	numeric	---	إجمالي عدد الحمّامات (مثل 2.5 = كاملان + نصفي)
lot_size	numeric	lot_size	مساحة الأرض بالأقدام المربعة
area_id	integer	area	معرّف المنطقة في MLS
area_text	character	display_only	اسم المنطقة
age	numeric	actual_age	عمر العقار بالسنوات
year_built	integer	display_only	سنة البناء
latitude	numeric	latitude	خط العرض (مُقرَّبًا إلى ثلاث منازل عشرية للنموذج)
longitude	numeric	longitude	خط الطول (مُقرَّبًا إلى ثلاث منازل عشرية للنموذج)
latitude6	numeric	display_only	خط العرض بالدقة الكاملة
longitude6	numeric	display_only	خط الطول بالدقة الكاملة
garage_spaces	integer	---	عدد مواقف المرآب
fp_count	integer	---	عدد المواقد
no_of_stories	numeric	---	عدد الطوابق
style	character	---	الطراز المعماري
view	character	---	نوع الإطلالة (مثل ``الحي" أو ``التلال")
dom	integer	display_only	أيام العرض في السوق
listing_date	Date	listing_date	تاريخ الإدراج
sale_concessions	numeric	concessions	تنازلات البائع

3.14 بداية سريعة مقترحة

1. شغّل earthUI: `earthUI::launch()`
2. أنشئ مشروعًا بغرض **للتقييم** في معالج المشروع الجديد، مرفقًا به `Appraisal_1.csv` ملفَّ بياناتٍ أوّليًا
3. اختر `Appraisal_1.csv` في القسم 2 (استيراد البيانات)
4. اختر `sale_price` هدفًا
5. أسنِد الأنواع الخاصة كما في الجدول أعلاه
6. أدرِج المتنبئات: `sale_age`، و`living_sqft`، و`baths_total`، و`lot_size`، و`area_id` (عاملًا)، و`age`، و`latitude`، و`longitude`، و`garage_spaces`
7. اضبط الدرجة على 1، وانقر **ملاءمة نموذج Earth**
8. نزّل المخرجات الوسيطة (الخطوة 6)، وراجع ترتيب CQA
9. احسب تعديلات RCA (الخطوة 7) بدرجة CQA تقارب 5.00
10. ولّد شبكة المقارنة بالمبيعات (الخطوة 8)

تنبيه: مجموعة البيانات التجريبية طريقةٌ جيدة لاستكشاف كل خصائص *earthUI* قبل استيراد بيانات *MLS* الخاصة بك. والمتنبئاتُ المقترحة أعلاه تُنتج عادةً نموذجًا بـR^2 فوق *0.85*.

15 تجريبي: استخراج النصّ بـProlog وقواعد الاشتقاق

يُدخل الإصدار 0.10.0 مرفقًا **اختياريًا وتجريبيًا حقًّا** لاستخراج خصائصَ جاهزة للنمذجة من أعمدة النصّ الوصفي في تصدير MLS --- وأولها عمود ``الملاحظات العامة"، وكذلك ملاحظات الوكيل وقوائم الخصائص المفصولة بفواصل كخصائص الغرف أو المنشآت القائمة. ومحرّك الاستخراج هو Prolog، وعلى وجه الخصوص **النحو ذو الجمل المحدَّدة** (DCGs): إذ تُكتب القواعدُ التي تتعرّف عباراتٍ مثل ”move in ready“ أو ”3 car garage“ أو ”owned solar“ في نصّ الملاحظات قواعدَ نحوٍ بلغة Prolog.

وهذه الخاصية ما تزال قيد التطوير. ويمكن استعمالها **جزئيًا** في الاستخراج اليوم --- فالنحو المرفق يتعرّف مجموعةً مفيدة من الخصائص --- لكن تكييفها على مفردات نظام MLS لديك يقتضي منك قراءة Prolog وكتابته، ولا سيما DCGs. فإن لم يكن ذلك شأنك، فتخطَّ هذا الفصل بالكامل: فالخاصية معطَّلةٌ افتراضًا، وسائرُ earthUI يعمل من دونها تمامًا.

> **تحذير**
>
> هذه خاصيةٌ تجريبية قيد التطوير النشط. ونحوُ استخراجها ليس كاملًا ولا مُتحقَّقًا منه لعمل التقييم الإنتاجي --- فيجب مراجعة الأعمدة المستخرَجة في مقابل النصّ المصدر قبل الاعتماد عليها في نموذج. واستعمالُها استعمالًا فعّالًا يقتضي معرفةً عملية بـProlog وDCGs. وكلُّ ما تفعله إضافيٌّ وقابل للرجوع عنه؛ فإن أساء شيءٌ التصرّف، أطفئ مبدِّلَ الإعدادات وواصل عملك طبيعيًا.

1.15 ماذا تفعل

تحمل إدراجاتُ MLS قدرًا كبيرًا من المعلومات ذات الصلة بالقيمة في نصٍّ حرّ وفي حقول قوائمَ مفصولة. وتحوّل هذه الخطوةُ الاختيارية ذلك النصّ إلى أعمدةٍ عادية يستطيع انحدارك استعمالها، عبر دورَي عمودٍ خاصين (الفصل 6):

- **أعمدة `remarks`** يحلّلها نحوُ Prolog DCG إلى عمود خاصيةٍ واحد لكل سمةٍ مستخرَجة. وتُبنى بادئةُ العمود من الحروف الأولى لاسم العمود المصدر: فـ`public_remarks` يعطي أعمدة `pr_*`، و`agent_remark` يعطي `ar_*`، وهكذا.
- **أعمدة `list`** (مثل `existing_structures = "Shed, Workshop"`) تُقسَم إلى أعمدة FALSE/TRUE أحادية التنشيط على الصورة `<col>_<value>`. وهذا الجزء R محض ولا يحتاج إلى Prolog البتة.

والنحو المرفق (`inst/prolog/mls_remarks.pl`) يستخرج حاليًا، من جملة ما يستخرج: طبقةَ حالةٍ متدرّجة بدرجة حالةٍ من 1 إلى 5؛ وأعلامَ الإطلالة وأنواعها؛ وأعلامَ تجديد المطبخ والحمّام؛ والأرضيات الخشبية وغيرها؛ وقائمةَ التحسينات وعددها؛ ومواقفَ المرآب ومساحته؛ والمسبحَ والمنتجع وسائر الخصائص الخارجية؛ وعددَ المواقد؛ والملحقاتِ (سقيفة، حظيرة، ورشة، بيت صغير، بيت ضيافة، بيت كلاب، وغيرها) بمقاساتٍ محلَّلة من أبعادٍ مثل ”40x50“؛ ومواقفَ المركبات الترفيهية والتخزين؛ وتملّكَ الطاقة الشمسية في مقابل استئجارها، وتخزينَ البطاريات؛ ومساحةَ الأرض (بما فيها تحويل الأفدنة إلى أقدامٍ مربعة)؛ ومبالغَ الأسعار المتصلة بسياق البيع والمرتبطة بموضوعٍ ما (تخفيض السعر، ورسوم اتحاد الملاك، وإيجار الطاقة الشمسية أو سدادها، وضريبة Mello-Roos)؛ وإشاراتِ الموقع؛ ودرجةَ انطباعٍ للإيجابيات والسلبيات في المدى $[-1, 1]$.

2.15 المتطلبات: vProlog (Trealla)، لا SWI-Prolog

يجري تحليلُ الملاحظات على حزمة **vProlog** في R --- وهي محرّك Trealla Prolog مكتفٍ بذاته. وهي ليست على CRAN؛ فنصّبها من مستودع r-universe الخاص بالمؤلّف كما هو مبيَّن في ``تنصيب earthUI وvProlog" في مطلع هذا المقال (ثنائياتٌ مبنيّة مسبقًا لـWindows وmacOS، والمصدر في `github.com/wcraytor/vProlog`). ولا يلزم تنصيبُ Prolog على مستوى النظام (SWI-Prolog أو غيره) ولا يُستعمل. وإن غابت vProlog، تُخطّى تحليلاتُ الملاحظات مع ملاحظة، مع بقاء تقسيم `list` عاملًا. ويُبلِغ المساعدُ `vProlog_available()` عمّا إذا كان المحرّك والنحو المرفق موجودَين معًا.

3.15 تفعيلها وتشغيلها

1. افتح **الإعدادات** (الترس) وأشّر **تفعيل معالجة Prolog / أعمدة القوائم**. والمبدِّل معطَّلٌ افتراضًا، ويحمل ملاحظة ``خاصية متقدمة --- للمستعملين ذوي الخبرة بـProlog"، ويُتذكَّر لكل مشروع، ويظهر في حوار معلومات المشروع.
2. في جدول المتنبئات، اضبط دورَ **الخاص** للأعمدة المعنية على `remarks` أو `list` (ويجوز أن تحمل أعمدةٌ متعددة كلَّ دور).
3. استعمل قسم **تنفيذ معالجة Prolog** الواقع بين القسمين 4 و5 من الشريط الجانبي. وهو يقدّم خطوتين:

1. توسيع القوائم والملاحظات --- يشغّل نحوَ DCG على كل عمود `remarks`، ويقسّم كل عمود `list` تقسيمًا أحادي التنشيط. والتوسيعُ **تراكمي**: فلا تُعالَج إلا الأعمدة المعيَّنة منذ آخر تشغيل، ويُعطَّل الزرّ بمجرد توسيع كل ما عُيِّن. ويتتبّع شريطُ تقدّمٍ عمليةَ التحليل.

2. تنفيذ قواعد الاشتقاق --- يطبّق قواعد `derive/2` الخاصة بك على الأعمدة المُوسَّعة (انظر أدناه). وتُفحص القواعد نحويًا ويُتحقَّق منها بتشغيلٍ تجريبي قبل تنفيذها.

4.15 قواعد الاشتقاق

لكل صفٍّ من البيانات، يُثبت earthUI أعمدةَ الصف وقائعَ على الصورة `cell(Name, Value)` ويجمع كل حلول قواعدك `derive(Col, Val)` --- فيصير كلُّ `Col` متمايز عمودًا جديدًا، وتفوز **أولُ** قيمةٍ تُوجَد لكل عمود، فيجوز أن تتبعها جملةٌ احتياطية. وتدعم المحمولاتُ المساعدة (`count_true/2`، و`count_false/2`، و`count_true_prefix/2`، و`cell_or/3`، و`is_na/1`) التجميعَ ومعالجةَ القيم المفقودة، وتُسقط وقائعُ `drop/1` أعمدةً بعد تشغيل القواعد. مثال:

```
% Merge a collinear cluster of solar flags into one boolean.
derive(solar_any, B) :-
    ( cell(pr_has_solar, true)
    ; cell(pr_solar_owned, true)
    ; cell(pr_has_battery, true) ) -> B = true ; B = false.

% Count amenities instead of OR-ing them (numeric column).
derive(amenity_count, N) :- count_true_prefix(pr_has_, N).

% Threshold -> boolean, with a fallback clause.
derive(big_outbuilding, true)  :- cell(pr_outbuilding_sqft, S), S >= 1000.
derive(big_outbuilding, false).

% Discard a raw granular column once consolidated.
drop(pr_price_solar_lease).
```

وتقيم القواعد في ملفٍّ خاص بكل مشروع باسم `<project>_rules.pl`، مبذَّرٍ من قالبٍ معلَّق تعليقًا كاملًا، وتُحرَّر عبر **تحرير قواعد الاشتقاق...** في محرّر شفرةٍ عائم قابل للسحب يمكن كذلك إرساؤه في الشريط الجانبي. ويتحقّق زرّ **فحص** من صحة القواعد (مسحٌ نحوي، وفحصُ تحميل، وتشغيلٌ تجريبي على الصف 1 من بياناتك) من غير تنفيذها.

5.15 ملف البيانات النشط المُصدَّر

بعد أيٍّ من الخطوتين، يُكتب إطارُ البيانات المُعزَّز في مجلد إدخال المشروع باسم `<original_name>_<YYMMDD_HHMMSS>.xlsx` ويصير ملفَّ البيانات **النشط** للمشروع --- فإعادةُ فتح المشروع تحمّل البيانات المُعزَّزة، ولا تُعاد حسابات التوسيع أبدًا، والملفُّ نفسُه هو ما يستهلكه glmnetUI وmgcvUI. وتُحجب أعمدةُ الملاحظات المصدر تلقائيًا عن النموذج؛ أما الأعمدة المولَّدة فلا تُدرَج تلقائيًا متنبئات. وتحليلُ النصّ قد يُنتج أعمدةً كثيرة، فعلّم ما لا تريده (بعدُ) في الانحدار بوصفه **كامنًا**، أو عطّله بـ`x` (الفصل 6).

6.15 الواجهة البرمجية والملفات المرفقة

الخطُّ كلُّه مصدَّرٌ للاستعمال الدفعي: `execute_prolog_processing(df, remarks_cols, list_cols, rules_text)` ينسّق المرورات، وهو مبنيٌّ على `remarks_to_columns()` و`split_list_column()`، و`apply_user_rules()`، و`validate_rules()`. ومصادرُ Prolog المرفقة تحت `inst/prolog/` هي `mls_remarks.pl` (نحو DCG وتجميع الخصائص)، و`mls_analyze.pl` (مشغّل إدخالٍ وإخراج للتحليل الدفعي)، و`mls_rules.pl` (المحمولات المساعدة المتاحة لقواعد الاشتقاق)، و`climate_regions_ca.pl` (أدناه).

7.15 تجربةٌ ذات صلة: أوليّات الأقاليم المناخية

ثمة مرفقٌ تجريبي مرافق يعالج الخصائص **رقيقة البيانات** --- أي التي لا يستطيع الانحدارُ تقديرها لأنها لا تظهر إلا في حفنةٍ من المقارنات (كمبرّدٍ تبخيري مثلًا). فثلاثةُ مساعداتٍ في R --- `climate_region_for()`، و`market_area_profile()`، و`climate_feature_priors()` --- تصنّف موقعَ المشروع (المقاطعة والمدينة) في واحدٍ من سبعة أقاليم مناخية كاليفورنية يتعرّفها المشترون (وهي اختزالٌ عملي لمناطق البناء المناخية الستّ عشرة لدى لجنة الطاقة في كاليفورنيا)، وتعيد أوليّاتٍ ترتيبية للقيمة الإسهامية لخصائص التدفئة والتهوية والتكييف والتبريد (التدفئة، والتكييف، والمبرّد التبخيري، ومروحة المنزل الكاملة، والترشيح) بحسب الإقليم. وتحاكي وحدةُ Prolog `climate_regions_ca.pl` التصنيفَ نفسَه للاستعمال داخل قواعد الاشتقاق؛ والجانبان

مولَّدان من مصدر R معتمَد واحد فلا يمكن أن ينحرفا. والأوليّاتُ يُقصد بها أن **تُنير** تقييمَ الخصائص وأن تصلح فحصَ معقولية للمعاملات المقدَّرة --- وينبغي أن يراجع المُقيّم إسناداتِ الأقاليم (ولا سيما في المقاطعات الممتدّة على أقاليم عدة).

1.7.15 ما وراء كاليفورنيا: حزم الولايات القضائية ووحدات مناطق السوق

كاليفورنيا أولُ ولايةٍ قضائية لا الحدّ. فـearthUI مستعمَلٌ دوليًا، ويُعاد الآن هيكلةُ معرفة الأوليّات المناخية في صورة **حزم ولاياتٍ قضائية**: جداولِ وقائعَ مكتفية بذاتها على هيئة بيانات (تعريفات المناطق، وإسنادات الوحدات الإدارية، وأوليّات الخصائص) مفهرسةٍ برموز البلد والمستوى الإداري في نظام المشاريع، بحيث تستطيع الحزمةُ أن تصف ولايةً أمريكية، أو ولايةً ألمانية، أو بلدًا صغيرًا بأكمله. وثمة طبقةٌ مرافقة من **وحدات مناطق السوق** تلتقط المفردات المحلية وأوليّاتٍ مصقولة لمنطقة سوقٍ مسمّاة.

وكلتاهما مصمَّمتان ليؤلّفهما **في الميدان** مهندسو التقييم والمُقيّمون، بلغاتهم --- وواقعيًا بمساعدة النماذج اللغوية الكبيرة: فالعالم لا يملك كثيرًا من مبرمجي Prolog، ولذلك جُعلت مصنوعاتُ الميدان جداولَ لا برامج عمدًا. وصيغةُ الحزمة مخطَّطٌ موثَّق بمثالٍ تطبيقي كامل (كاليفورنيا نفسها)، وذلك صراحةً كي يستطيع مساعدٌ مثل Claude صياغةَ حزمةٍ صحيحة من وصفٍ بلغةٍ عادية لسوقٍ ما؛ ويولّد earthUI Prolog الأساسي من الحزمة، ويجب أن يجتاز مُتحقِّقٌ حاجب (فحوصُ المخطَّط والتغطية والتكامل الإسنادي والتشغيل التجريبي) قبل تفعيل أي حزمة. وتحمل كلُّ حزمةٍ مؤلَّفة في الميدان إسنادًا إلزاميًا --- المؤلّف المسجَّل، والتاريخ، والمرجع المستشهد به --- كي يقف وراء المعرفة المهنيُّ المسؤول لا أداةُ الصياغة، وكي تستطيع التقارير الاستشهادَ بالحزمة ونسختها بعينهما.

> **تنبيه:** ثمة نمطٌ عملي: وسّع الملاحظات، ثم ادمج أعلام `pr_*` الدقيقة في حفنةٍ من الأعداد أو القيم المنطقية المشتقّة بقواعد `derive/2`، ثم أسقِط بـ`drop/1` الأعلامَ الخام التي دمجتها، وأبقِ الأعمدةَ المشتقّة وحدها مرشّحةً للتنبؤ. وهذا يُبقي قائمة المتنبئات قصيرةً والتعدّدَ الخطي قابلًا للإدارة.

References

Friedman, Jerome H. (1991). **?**Multivariate Adaptive Regression Splines.**?** In: *The Annals of Statistics* 19.1, pp. 1–141. DOI: 10.1214/aos/1176347963.

Milborrow, Stephen (Oct. 1, 2024). *earth: Multivariate Adaptive Regression Splines*. R package. URL: https://CRAN.R-project.org/package=earth.

— (n.d.[a]). *Notes on the earth Package*. URL: http://www.milbo.org/doc/earth-notes.pdf.

— (n.d.[b]). *Variance Models in earth*. URL: http://www.milbo.org/doc/earth-varmod.pdf.