

glmnetUI: una interfaz interactiva para el paquete glmnet (Lasso y Elastic Net)

Traducción al español, elaborada con asistencia de IA y revisada por la redacción; la versión en inglés es la autorizada.

William Bert Craytor

Volumen 1, Número 1 · Verano 2026

Publicado el 1 de julio de 2026

DOI: [10.67330/ja8s2k39](https://doi.org/10.67330/ja8s2k39)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

Este artículo de acceso abierto se distribuye bajo los términos de la [licencia Creative Commons Atribución 4.0](#).

Resumen. glmnetUI es una interfaz gráfica de usuario para el paquete de R glmnet, que ajusta modelos lineales generalizados regularizados mediante las penalizaciones Lasso, ridge y elastic-net. Ofrece tres modos de uso — modelado predictivo general, tasación inmobiliaria y análisis de zona de mercado — y guía al usuario a través de la importación de datos, la configuración del modelo, el ajuste con validación cruzada, las trayectorias de coeficientes y los informes descargables. Este artículo documenta los requisitos de formato de datos de glmnetUI, el flujo de trabajo de modelado, las pantallas de salida y la referencia completa de funciones.

Palabras clave: Lasso, regresión ridge, elastic net, regresión regularizada, glmnet, R, interfaz gráfica de usuario, valuación inmobiliaria

Nota del editor (julio de 2026): Poco antes de la publicación, glmnetUI se consolidó dentro de earthUI (versión 0.11.0 y posteriores), que ahora ofrece el flujo de trabajo de elastic net aquí descrito junto con los flujos de trabajo de MARS y GAM. glmnetUI sigue disponible en su forma final publicada (0.5.0) en <https://github.com/wcraytor/glmnetUI>; el desarrollo futuro continúa en earthUI. Este artículo se refiere a glmnetUI tal como fue publicado.

Índice

1. Introducción	3
1.1. Historia y antecedentes	3
1.1.1. El Lasso	3
1.1.2. Regresión ridge y elastic net	4
1.1.3. El paquete glmnet	4
1.1.4. Comparación con MARS (earth) y GAM (mgcv)	4
1.1.5. Flujo de trabajo recomendado	5
1.2. ¿Qué es glmnetUI?	6
1.3. Tres modos de uso	6
1.4. Funciones específicas del sector inmobiliario	7
1.5. Primeros pasos	7
2. Requisitos de los datos de entrada del MLS	9
2.1. Formato y estructura del archivo	9
2.2. Columnas requeridas en el modo de tasación	9
2.3. Convenciones de nomenclatura de columnas especiales	10
2.4. Calidad e integridad de los datos	10
2.5. Colocación de la fila del sujeto	11

3. Modo de uso general	12
3.1. Panorama general	12
3.1.1. Skip First Row (omitir la primera fila)	12
3.1.2. Ajustes por propósito	12
3.2. El flujo de trabajo de la barra lateral	12
3.3. Proyectos	13
3.4. Importación desde earthUI (opcional)	13
3.5. Pestañas del panel principal	14
3.6. Persistencia de los ajustes	14
3.7. Modo oscuro	15
3.8. Configuración regional y de idioma	15
3.8.1. Guardar los valores predeterminados	15
4. Modo de tasación	16
4.1. Tratamiento de la fila del sujeto	16
4.2. Fecha efectiva y edad de la venta	16
4.3. Designaciones especiales de columnas	16
4.4. Panorama de los ajustes RCA	16
5. Modo de análisis de zona de mercado	17
5.1. Diferencias respecto al modo de tasación	17
5.2. Cuándo usar el modo de mercado	17
6. Selección de variables	18
6.1. Variable objetivo	18
6.2. La tabla de predictores	18
6.3. Detección de tipos de datos y anulaciones	18
6.4. Signos esperados y su aplicación	18
6.5. Referencia de los tipos de columna especiales	19
7. Selección de parámetros	20
7.1. Alpha (parámetro de mezcla)	20
7.2. Selección de lambda	20
7.3. Family (familia)	20
7.4. Standardize (estandarizar)	21
7.5. Restricciones de signo	21
7.6. Relaxed lasso	21
7.7. Semilla aleatoria	21
7.8. Matriz de interacciones	22
7.9. Bloquear el efecto principal (solo interacción)	22
7.10. Parámetros avanzados	22
7.11. Valores predeterminados de los ajustes	23
8. Ajuste del modelo	24
8.1. El botón de ajuste	24
8.2. Qué ocurre durante el ajuste	24
9. Pestañas de resultados	25
9.1. Data Preview	25
9.2. Equation	25
9.3. Correlation	25
9.4. Summary	25

9.5. Coefficients	26
9.6. Variable Importance	26
9.7. Contributions	26
9.8. ANOVA	26
9.9. Diagnostics	26
9.10. Glmnet Output	27
9.11. Report	27
10.Descarga de datos	28
10.1. Columnas de salida	28
10.2. Puntuaciones CQA	28
11.Cálculos RCA y descarga	29
11.1. Apertura del diálogo RCA	29
11.2. Interpolación de la puntuación CQA	29
11.3. Columnas de salida	29
12.Sales Comparison Grid	30
12.1. Diálogo de selección de comparables	30
12.2. Estructura del libro	30
12.3. Protección de la hoja y celdas residuales	30
13.Descarga de informes	31
13.1. El flujo de trabajo Quarto de dos pasos	31
13.2. Formatos de informe	31
13.3. Contenido del informe	31
14.Comparación con earthUI	33
14.1. Diferencias principales	33
14.2. Funciones compartidas	33
14.3. Cuál usar en cada caso	33
15.Conjunto de datos de demostración: Appraisal_1.csv	35
15.1. Descripción	35
15.2. Columnas	35
15.3. Inicio rápido sugerido	36
16.Requisitos del sistema y resolución de problemas	37
16.1. Plataformas admitidas	37
16.2. Dependencias opcionales	37
16.3. Notas por plataforma	37
16.4. Degradación elegante	37
16.5. Resolución de problemas	37

1 Introducción

1.1 Historia y antecedentes

1.1.1. El Lasso

El **Lasso** (**L**east **A**bsolute **S**hrinkage and **S**election **O**perator) fue introducido por Robert Tibshirani en 1996. El nombre es un juego de palabras deliberado con el lazo del vaquero — una

cuerda que restringe y selecciona. Tibshirani lo eligió porque el método literalmente “enlaza” los coeficientes, restringiendo sus valores absolutos y llevando algunos exactamente a cero, con lo cual selecciona qué variables permanecen en el modelo.

La intuición de Tibshirani consistió en añadir una penalización L_1 (suma de los valores absolutos de los coeficientes) al objetivo de mínimos cuadrados. A diferencia de la penalización L_2 de la regresión ridge (Hoerl y Kennard, 1970), que contrae los coeficientes hacia cero pero nunca lo alcanza, la penalización L_1 produce soluciones *dispersas* en las que muchos coeficientes son exactamente cero. Esto convierte al Lasso a la vez en un método de regularización y en un método de selección de variables.

1.1.2. Regresión ridge y elastic net

La **regresión ridge** fue propuesta por Arthur Hoerl y Robert Kennard en 1970. Añade una penalización L_2 (suma de los cuadrados de los coeficientes) que contrae todos los coeficientes proporcionalmente pero mantiene todas las variables en el modelo. Ridge es eficaz cuando los predictores están correlacionados (multicolinealidad), pero no simplifica el modelo.

El **elastic net**, introducido por Hui Zou y Trevor Hastie en 2005, combina ambas penalizaciones. El parámetro de mezcla α controla la combinación: $\alpha = 1$ es Lasso puro, $\alpha = 0$ es ridge puro, y los valores intermedios dan un compromiso. El elastic net supera una limitación del Lasso: cuando los predictores están muy correlacionados, el Lasso tiende a seleccionar uno e ignorar el resto, mientras que el elastic net agrupa los predictores correlacionados.

1.1.3. El paquete glmnet

El paquete de R **glmnet** fue desarrollado por Jerome Friedman, Trevor Hastie y Robert Tibshirani en la Universidad de Stanford. Implementa la regularización elastic net para modelos lineales generalizados mediante un algoritmo de descenso por coordenadas extremadamente eficiente (Friedman, Hastie y Tibshirani 2010), extendido después a los modelos de riesgos proporcionales de Cox (Simon et al. 2011) y a todas las familias de modelos lineales generalizados (Tay, Narasimhan y Hastie 2023). El paquete admite las familias gaussian (lineal), binomial (logística), poisson y otras familias de GLM.

Entre las características principales de **glmnet** se incluyen:

- **Trayectoria de regularización:** ajusta el modelo sobre una malla de valores de λ en una sola pasada, mucho más rápido que ajustarlo por separado en cada λ .
- **Validación cruzada:** `cv.glmnet()` selecciona automáticamente la intensidad óptima de regularización mediante validación cruzada de k pliegues.
- **Cotas de los coeficientes:** admite límites superiores e inferiores para coeficientes individuales (restricciones de signo).
- **Relaxed lasso:** reajusta el modelo sobre las variables seleccionadas con penalización reducida o nula, lo que disminuye el sesgo.

1.1.4. Comparación con MARS (earth) y GAM (mgcv)

La aplicación **earth** y sus dos Post Processing Modules (módulos de posprocesamiento; **glmnet** y **mgcv**) utilizan motores de modelado distintos. La tabla siguiente resume las diferencias principales:

Característica	glmnet (Elastic Net)	earth (MARS)	mgcv (GAM)
Tipo de modelo	Lineal (con regularización)	Lineal por tramos (splines adaptativos)	No lineal suave (splines penalizados)
No linealidad	Solo mediante interacciones o expansión de base	Automática mediante funciones bisagra con nodos guiados por los datos	Automática mediante funciones suaves (thin plate, cúbicas, etc.)
Selección de variables	Incorporada mediante la penalización L_1 (Lasso)	Incorporada mediante pasos hacia adelante/atrás con poda por GCV	Opcional mediante contracción de doble penalización (<code>select=TRUE</code>)
Interacciones	Productos cruzados por pares (difíciles de interpretar)	Productos basados en bisagras (interpretables, hasta grado 3)	Suavizados de producto tensorial (<code>te()</code> , <code>ti()</code>)
Interpretabilidad	Los coeficientes son pesos lineales; fáciles de explicar individualmente	Las funciones g muestran efectos parciales lineales por tramos para cada variable	Curvas suaves de efecto parcial; muy intuitivas
Más indicado para	Datos de alta dimensión, selección de variables, muestras pequeñas	Detección automática de no linealidad e interacciones	Relaciones no lineales suaves, modelado flexible
Ajustes RCA	Ajustes lineales por variable	Ajustes lineales por tramos mediante funciones g	Ajustes suaves mediante efectos parciales
Control del sobreajuste	Lambda validado por validación cruzada	Poda basada en GCV	Estimación del parámetro de suavizado por REML/GCV

MARS (Multivariate Adaptive Regression Splines) fue introducido por Jerome Friedman en 1991. Construye modelos lineales por tramos seleccionando de manera adaptativa funciones bisagra $\max(0, x-k)$ y las posiciones de sus nodos a partir de los datos. La implementación en R es el paquete **earth** de Stephen Milborrow. MARS destaca por descubrir automáticamente relaciones no lineales e interacciones sin que el usuario tenga que especificarlas. Sus funciones g (términos de base agrupados por variable) proporcionan curvas de efecto parcial muy interpretables.

Los **GAM** (Generalized Additive Models, modelos aditivos generalizados) fueron formalizados por Hastie y Tibshirani (1990). La implementación en R **mgcv**, de Simon Wood, ofrece splines de regresión penalizados con selección automática del suavizado mediante REML o GCV. Los GAM producen curvas suaves de efecto parcial que resultan intuitivas de visualizar y de explicar. Cuando **earthUI** exporta las posiciones de los nodos a **mgcvUI**, los nodos derivados de **earth** sirven de punto de partida para los términos suaves del GAM, combinando la colocación adaptativa de nodos de MARS con la estimación suave del GAM.

1.1.5. Flujo de trabajo recomendado

Para la tasación inmobiliaria y aplicaciones similares que requieren modelos interpretables y defendibles:

1. **Comience con earthUI** para descubrir las variables importantes, las relaciones no lineales y las interacciones presentes en sus datos. La selección automática de base y las funciones `g` de `earth` proporcionan excelentes intuiciones iniciales.
2. **Refine con mgcvUI** si desea efectos parciales más suaves. Importe los nodos de `earthUI` en `mgcvUI` para una transición fluida de modelos lineales por tramos a modelos suaves.
3. **Utilice glmnetUI** cuando necesite una selección de variables agresiva (Lasso), cuando el número de predictores se acerque o supere el número de observaciones, o cuando desee un modelo puramente lineal con regularización por motivos de defendibilidad.

1.2 ¿Qué es glmnetUI?

`glmnetUI` es una interfaz gráfica de usuario para el paquete `glmnet` de R. Se ejecuta como una aplicación Shiny local — no hay inicio de sesión, ni servidor, ni cuentas. Se inicia desde R, se importa un conjunto de datos (CSV o Excel), se configura el modelo y se ajusta de manera interactiva.

A partir de la versión actual, `glmnetUI` ya no es un paquete que se instale por separado: se distribuye *dentro* del paquete `earthUI` como su Post Processing Module de elastic net, y se inicia con `earthUI::launch_glmnet()`. La aplicación en sí — y todo lo descrito en este artículo — no ha cambiado; solo difieren la instalación y el comando de inicio. Las tres rutinas (`earth`, `glmnet`, `mgcv`) comparten un mismo árbol de proyectos, una misma base de datos de configuración y una misma copia de la infraestructura de soporte.

La aplicación proporciona un flujo de trabajo completo: importación de datos, configuración de variables, ajuste del modelo, gráficos de diagnóstico, importancia de las variables, ecuaciones del modelo e informes descargables en formato Word, PDF o HTML.

La **regresión elastic net** combina dos técnicas de regularización:

- **Regresión ridge** ($\alpha = 0$): contrae proporcionalmente todos los coeficientes hacia cero. Adecuada cuando se considera que todas las variables son relevantes.
- **Regresión Lasso** ($\alpha = 1$): puede contraer coeficientes exactamente a cero, realizando una selección automática de variables. Produce modelos más sencillos e interpretables.
- **Elastic net** ($0 < \alpha < 1$): una combinación de ambos enfoques. Resulta útil cuando los predictores están correlacionados.

1.3 Tres modos de uso

Todo proyecto de `glmnetUI` tiene un **purpose** (propósito o modo de uso), que se elige al crearlo (Sección 1, *Project*). El propósito del proyecto activo determina qué herramientas y elementos de la interfaz aparecen; para trabajar en otro modo, cree o abra un proyecto con ese propósito. Los tres modos son:

- **General** — regresión elastic net para cualquier tipo de población o conjunto de datos. Este es el modo predeterminado. Ofrece el flujo de trabajo completo de regresión regularizada sin añadidos específicos de ningún dominio.
- **For Appraisal** (para tasación) — regresión elastic net adaptada a la tasación inmobiliaria. Añade funciones específicas para la valuación de una sola propiedad, incluidos el tratamiento de la propiedad sujeto, las designaciones especiales de columnas y el Residual Constraint Approach (RCA).
- **Market Area Analysis** (análisis de zona de mercado) — regresión elastic net adaptada a los estudios de zona de mercado. Añade funciones para analizar grupos de propiedades dentro de un mercado definido, con exclusión opcional de la fila del sujeto.

En los tres modos el motor de modelado subyacente es idéntico — siempre se ajusta un modelo elastic net (`glmnet`). El ajuste del propósito controla qué herramientas y elementos de interfaz adicionales están disponibles.

1.4 Funciones específicas del sector inmobiliario

Cuando se selecciona **For Appraisal** o **Market Area Analysis**, glmnetUI activa varias funciones diseñadas para el análisis inmobiliario:

- **Designaciones especiales de columnas** — cada predictor puede etiquetarse con un papel especial como `contract_date`, `dom`, `concessions`, `latitude`, `longitude`, `living_area`, `lot_size`, `actual_age`, `effective_age`, `area`, `site_dimensions` o `display_only`. Estas designaciones controlan cómo se trata la columna durante el ajuste y en las salidas.
- **Redondeo de latitud y longitud** — las columnas designadas como latitud o longitud se redondean automáticamente a 3 decimales para evitar el sobreajuste.
- **Columna Sale Age** — cuando una columna se designa como `contract_date` y se proporciona una fecha efectiva, glmnetUI calcula una columna `sale_age` (días entre la fecha de venta y la fecha efectiva) y la sustituye como predictor.
- **Cálculos RCA (solo en tasación)** — en el modo de tasación, después de ajustar el modelo glmnetUI puede calcular la salida del Residual Constraint Approach (RCA), que produce ajustes por comparable, resúmenes de ajuste neto/bruto y un precio de venta ajustado para la propiedad sujeto.

***Consejo:** El modo General funciona con cualquier conjunto de datos — financiero, científico, de ingeniería o inmobiliario. Los modos de tasación y de mercado simplemente añaden funciones de conveniencia para los profesionales del sector inmobiliario.*

1.5 Primeros pasos

Para usar glmnetUI:

1. **Instale earthUI** en R: `install.packages("earthUI")` — la rutina de glmnet está incluida. (Véase “Instalación de earthUI y vProlog” en el artículo sobre earthUI de este número para la versión de desarrollo y los componentes opcionales).
2. **Inicie la aplicación:** ejecute `earthUI::launch_glmnet()` en R. La aplicación se abre en su navegador web en el puerto 7879. También puede acceder directamente a la aplicación navegando a `http://localhost:7879` en su navegador. La aplicación recuerda el último modo de uso empleado y lo restaura automáticamente.
3. **Cree o abra un proyecto** en la Sección 1 de la barra lateral. Un proyecto fija el propósito y la ubicación de trabajo (país, estado, condado, ciudad) y sus carpetas de entrada y salida. Los proyectos se almacenan bajo una raíz `regProj` compartida y se comparten con las aplicaciones hermanas earthUI y mgcvUI. Abrir un proyecto con un propósito distinto es la forma de cambiar de modo.
4. **Importe sus datos** en la Sección 2 seleccionando un archivo CSV o Excel de la carpeta de entrada del proyecto activo. Añada archivos a esa carpeta y haga clic en **Refresh** (actualizar) para listarlos.
5. **Importe desde earthUI (opcional)** — la Sección 3 permite importar un archivo `.rds` de resultados de earthUI para utilizar las funciones de base bisagra de earth con la regularización de glmnet. Omita este paso si no dispone de un resultado de earthUI.
6. **Configure las variables** — elija su variable objetivo y sus predictores, fije los tipos de datos, los signos esperados y asigne los papeles especiales de columna.
7. **Fije los parámetros de glmnet** — `alpha`, `lambda`, `family`, restricciones de signo, interacciones y otras opciones.
8. **Ajuste el modelo** — haga clic en “Fit Glmnet Model” (ajustar el modelo glmnet) y revise los resultados en el panel principal.
9. **Exporte** — descargue las predicciones en Excel, genere y convierta un informe Quarto o (en modo de tasación) calcule los ajustes RCA y una Sales Comparison Grid (cuadrícula

de comparación de ventas).

Los ajustes se guardan automáticamente en el almacenamiento local de su navegador y se restauran cuando vuelve a cargar el mismo archivo de entrada.

2 Requisitos de los datos de entrada del MLS

Para los flujos de trabajo de tasación inmobiliaria y análisis de mercado, sus datos de entrada provienen normalmente de una exportación de un Multiple Listing Service (MLS, servicio de listado múltiple). Este capítulo describe la estructura de archivo esperada y las columnas que glmnetUI puede utilizar.

2.1 Formato y estructura del archivo

glmnetUI acepta archivos **CSV** y **Excel** (.xlsx, .xls). Al importar, los nombres de las columnas se convierten automáticamente a **snake_case** — por ejemplo, “Living SqFt” pasa a ser **living_sqft**, “Contract Date” pasa a ser **contract_date** y “Sale Price” pasa a ser **sale_price**. Esta normalización garantiza referencias de columna coherentes en todo el flujo de trabajo. El separador de CSV y el signo decimal utilizados durante la importación vienen determinados por la configuración regional (véase el capítulo 3, “Locale & Regional Settings”).

Su archivo de datos debe ser una tabla plana con una fila por propiedad y una columna por atributo. La primera fila del archivo debe contener los encabezados de columna.

2.2 Columnas requeridas en el modo de tasación

Aunque glmnetUI funciona con cualquier conjunto de columnas, el flujo de trabajo completo de tasación se beneficia de disponer de las columnas siguientes:

Columna	Tipo especial	Finalidad
Sale Price	(objetivo)	Variable respuesta del modelo
Contract Date	contract_date	Se usa para calcular sale_age (días desde la fecha efectiva)
Listing Date	listing_date	Se usa con la fecha de contrato para calcular el DOM si no existe columna DOM
Days on Market	dom	Días en el mercado; se muestra en las exportaciones
Concessions	concessions	Concesiones de venta; Net SP = Sale Price – Concessions
Living Area (SF)	living_area	Permite residuos por pie cuadrado (residual_sf , cqa_sf)
Lot Size	lot_size	Columna de tamaño del terreno
Site Dimensions	site_dimensions	Agrupada con el tamaño del terreno
Latitude	latitude	Redondeada a 3 decimales
Longitude	longitude	Redondeada a 3 decimales
Area ID	area	Identificador de zona de mercado / vecindario
Actual Age	actual_age	Edad de la propiedad
Effective Age	effective_age	Edad efectiva de la propiedad
Address	display_only	Se muestra en las exportaciones; se excluye del modelo

Los nombres de columna de la hoja de cálculo pueden estar en otro idioma — los nombres “especiales” están en inglés para que el programa de R pueda darles un tratamiento especial. Por lo demás, los nombres de columna suministrados aparecen en los modelos de regresión, en los gráficos y, si se hacen tasaciones, en los informes de salida.

No todas las columnas son obligatorias. glmnetUI se adapta — si falta una columna, la función correspondiente simplemente se omite. No obstante, para los modelos de precios inmobiliarios se recomiendan encarecidamente ciertas columnas para lograr un ajuste aceptable:

1. **Sale Age** — el número de días entre la fecha de venta según contrato y la fecha efectiva de la tasación o del análisis. Si se emplea un historial de ventas de varios años, especialmente para períodos superiores a 5 años, **sale_age** suele desempeñar un papel central en la estimación del precio de venta.
2. **Living Area** — también aparece con nombres como “Living Sqft”, “GLA” (gross living area, superficie habitable bruta), etc. Es otro determinante principal del precio de venta.
3. **Total Bath Count** — el número total de baños completos, de un cuarto, de medio baño y de tres cuartos. Por ejemplo, dos baños completos y un medio baño darían un valor de 2,5.
4. **Garage Bays** o **Garage Area** — el número de plazas de garaje o la superficie del garaje.
5. **Lot Size** — la superficie de terreno de la propiedad, normalmente en pies cuadrados o acres.
6. **Longitude**, **Latitude** y, si está disponible, **Area ID**. Estas variables de localización ayudan al modelo a tener en cuenta la variación geográfica de los precios.

2.3 Convenciones de nomenclatura de columnas especiales

glmnetUI identifica las columnas por su **designación de tipo especial**, no por su nombre de columna. Puede dar a sus columnas el nombre que desee en la exportación del MLS — lo que importa es que asigne el tipo especial correcto en la tabla de configuración de variables (capítulo 6).

Por ejemplo, su MLS podría exportar la superficie habitable como “GLA”, “Living SqFt”, “liv_area” o “gross_living_area”. Tras la importación (donde pasa a snake_case), basta con designarla como **living_area** en el desplegable Special. glmnetUI la utilizará entonces para los cálculos de residuos por pie cuadrado con independencia de su nombre original.

2.4 Calidad e integridad de los datos

- **Valores faltantes (NA)**: las filas con valores NA en cualquier columna predictora u objetivo se eliminan automáticamente antes del ajuste.
- **Columnas de fecha**: las columnas de texto que coinciden con formatos de fecha habituales se analizan automáticamente al importar. Los formatos con año de 2 dígitos tienen prioridad cuando no se detecta un año de 4 dígitos.
- **Columnas numéricas**: el precio de venta, la superficie habitable, el tamaño del terreno, las concesiones y campos similares deben ser numéricos. Si su MLS exporta los precios con símbolos de moneda o separadores de miles (por ejemplo, “\$350,000” o “350.000,00”), es posible que deba limpiarlos antes de la importación.
- **Columnas de factor**: las columnas de texto, TRUE/FALSE y de tipo factor de R se tratan automáticamente como categóricas — una regresión solo puede manejarlas como factores. Una columna **numérica** (por ejemplo, un código numérico de zona o de estilo) se convierte en factor solo cuando se marca su casilla **Factor**; glmnetUI nunca infiere el carácter categórico a partir del rango de valores de una columna numérica, de modo que predictores numéricos discretos como **bath_count** permanecen continuos. El papel Special es independiente de la decisión sobre el factor.

***Consejo:** Revise la columna NA en la tabla de predictores después de la importación. Las columnas con muchos valores faltantes pueden provocar la eliminación de filas. Considere excluir las columnas con muchos NA o limpiar los datos antes de importarlos.*

2.5 Colocación de la fila del sujeto

En el modo **Appraisal**, la fila 1 debe ser la propiedad sujeto. Todas las filas restantes son ventas comparables. La fila del sujeto se excluye del ajuste del modelo. Tras el ajuste, el modelo genera igualmente predicciones para la fila del sujeto.

En el modo **Market Area Analysis**, colocar el sujeto en la fila 1 es opcional.

En el modo **General** no hay tratamiento especial de filas — todas las filas se tratan por igual.

3 Modo de uso general

3.1 Panorama general

El modo General Purpose (uso general) es el predeterminado al iniciar glmnetUI. Ofrece el flujo de trabajo completo de regresión elastic net para cualquier conjunto de datos — no solo inmobiliarios. Puede usar glmnetUI para datos científicos, análisis financiero, estudios de ingeniería o cualquier problema de regresión.

En el modo General, la interfaz omite las funciones específicas del sector inmobiliario (columnas especiales, edad de la venta, redondeo de coordenadas, RCA). La barra lateral se simplifica para centrarse en la selección de variables, la configuración de parámetros, el ajuste del modelo y la exportación.

3.1.1. Skip First Row (omitir la primera fila)

Una casilla **Skip first row** (omitir la primera fila) aparece en la parte superior de la barra lateral una vez abierto un proyecto General o Market. Cuando está marcada, la fila 1 se excluye del ajuste del modelo. Esto resulta útil cuando la fila 1 contiene una observación objetivo o de referencia que no debe influir en el modelo. En el modo Appraisal, la fila 1 (la propiedad sujeto) siempre se excluye automáticamente.

3.1.2. Ajustes por propósito

Los ajustes (selección de predictores, parámetros del modelo, interacciones) se guardan por separado para cada combinación de archivo de entrada y modo de uso. Cambiar entre los modos General, Appraisal y Market conserva los ajustes de cada modo de forma independiente.

3.2 El flujo de trabajo de la barra lateral

La barra lateral se organiza en secciones numeradas y plegables que le guían desde la selección del proyecto hasta la exportación del informe. La mayoría de las secciones aparecen solo una vez abierto un proyecto activo, y varias son específicas del propósito del proyecto. La carpeta de salida se deriva del proyecto activo — no hay un campo aparte para la carpeta de salida.

1. Project — crear o abrir un proyecto. Un proyecto fija el propósito y la ubicación de trabajo (país, estado, condado, ciudad) y sus carpetas de entrada (`<os>_in/`) y de salida. Los proyectos se almacenan bajo una raíz `regProj` compartida (por omisión `~/regProj`, modificable con la variable de entorno `REGPROJ_ROOT`) y se comparten con las aplicaciones hermanas `earthUI` y `mgcvUI`, de modo que los modelos de las tres conviven lado a lado. Haga clic en **Close Project** (cerrar proyecto) para cambiar.

2. Import Data — seleccionar un archivo CSV o Excel de la carpeta de entrada del proyecto activo. Añada archivos a la carpeta mediante el Finder o el Explorador, y haga clic en **Refresh** para listarlos. Para los archivos de Excel con varias hojas aparece un selector de hoja. Los nombres de columna se convierten automáticamente a `snake_case`.

3. Import from earthUI (opcional) — importar un archivo `.rds` de resultados de `earthUI` para utilizar las funciones de base bisagra de `earth` con la regularización elastic net de `glmnet`. El botón de exploración tiene el mismo estilo que el selector de archivos de la Sección 2. Este paso es opcional — omítalo si no dispone de un resultado de `earthUI` que importar.

4. Variable Configuration — selector de la variable objetivo, tabla de predictores con casillas Include, Factor y Force, además de los desplegables Sign. La columna Special (para designar fechas de contrato, coordenadas, etc.) aparece solo en los modos Appraisal y Market. Véase el capítulo 6 para todos los detalles.

5. glmnet Call Parameters — toda la configuración del modelo: alpha, lambda, family, standardize, restricciones de signo, relaxed lasso, semilla aleatoria, matriz de interacciones y parámetros avanzados (lambda.min.ratio, nlambd, métrica de pérdida de CV, umbral de convergencia, iteraciones máximas, intercepto). Véase el capítulo 7 para la referencia completa de parámetros.

6. Fit Glmnet Model — el botón que ejecuta el modelo. Los resultados aparecen en las pestañas de la derecha.

7. Download Output — exporta las predicciones, los residuos, las puntuaciones CQA y las contribuciones por variable como archivo de Excel a la carpeta de salida del proyecto. Disponible en todos los modos de uso una vez ajustado el modelo.

8. Calculate RCA Adjustments & Download — calcula los ajustes RCA por variable para cada comparable e interpola el residuo del sujeto mediante CQA. Aparece solo en los modos Appraisal y Market, después del ajuste. Véase el capítulo 11.

9. Generate Sales Grid & Download — genera la Sales Comparison Grid intermedia, ordenando los comparables por porcentaje de ajuste bruto. Aparece solo en el modo Appraisal, después de calcular los ajustes RCA. Véase el capítulo 12.

10. Generate Quarto Report — escribe un paquete de informe Quarto (.qmd) autocontenido — con todos los gráficos y tablas — en la carpeta de salida del proyecto. Disponible una vez ajustado el modelo.

11. Convert Quarto Report — renderiza el paquete .qmd generado a HTML, Word o PDF. El PDF requiere una instalación de LaTeX; si no se detecta ninguna, la opción de PDF queda oculta. Véase el capítulo 13.

En los modos Market y General se omiten los pasos exclusivos de tasación y las dos secciones de informe se renumeran en consecuencia.

3.3 Proyectos

glmnetUI organiza el trabajo en **proyectos** bajo una carpeta raíz **regProj** compartida (el mismo árbol que usan earthUI y mgcvUI). Cada proyecto posee una carpeta de entrada (<os>_in) y carpetas de salida por motor (por ejemplo, <os>_out_glmnet).

Para crear un proyecto, haga clic en **+ New...** en la Sección 1 y proporcione:

- **Purpose** — General, Appraisal o Market Area.
- **Location** — la jerarquía política/administrativa del proyecto. El conjunto de niveles que se rellenan depende del país que se elija, de modo que las convenciones de denominación extranjeras se admiten automáticamente — un proyecto estadounidense usa estado/condado/ciudad, uno francés usa región/département/commune, uno japonés usa prefectura/ciudad-o-barrio/distrito, y así sucesivamente.
- **Project Name** — una etiqueta corta, limitada a 8 caracteres (letras, dígitos, _ y -).

La aplicación no usa literalmente el nombre que usted teclea: construye un nombre de carpeta único con marca de tiempo anteponiendo los códigos de ubicación y la fecha y hora de creación, con la forma <country>_<admin level codes>_<YYYYMMDD-HHMMSS>_<your name> — por ejemplo, un proyecto estadounidense llamado **bur1** pasa a ser **us_ca_081_bur1in_20260604-131500_bur1**. El límite de 8 caracteres se aplica solo a la parte que usted teclea. Solo los proyectos recién creados reciben el prefijo de marca de tiempo; los proyectos ya existentes en disco conservan sus nombres originales y siguen siendo plenamente accesibles.

3.4 Importación desde earthUI (opcional)

La Sección 3 de la barra lateral — **Import from earthUI** — permite importar un archivo .rds de resultados de earthUI. Esto permite a glmnetUI usar las funciones de base bisagra de earth (términos lineales por tramos) con la regularización elastic net de glmnet, combinando la

detección automática de no linealidad de earth con la selección de variables y las restricciones de coeficientes de glmnet.

Cuando se importa un resultado de earthUI, son los predictores del modelo earth — y no las casillas Include — los que determinan el ajuste de glmnet, porque glmnet se ajusta sobre la expansión de base de earth. La tabla de configuración de variables lo deja explícito: **Include** y **Type** quedan bloqueados, las filas de los predictores que el modelo earth no utilizó se deshabilitan y una nota explica que el conjunto de predictores procede del modelo earth. **Force** y **Sign** siguen siendo editables para los predictores de earth, ya que estos aún se aplican a la base de earth en el momento del ajuste.

El botón de exploración tiene el mismo estilo que el selector de archivos de la Sección 2. Este paso es totalmente opcional — si no dispone de un archivo de resultados de earthUI, simplemente omita la Sección 3 y pase a la Sección 4.

3.5 Pestañas del panel principal

Tras el ajuste, el panel principal ofrece las pestañas siguientes:

Pestaña	Contenido
Data Preview	Vista previa de los datos importados. En modo de tasación, se divide en las tablas Subject Property y Comparable Sales.
Equation	La ecuación del modelo ajustado, mostrada en LaTeX/MathJax, con cada coeficiente no nulo y los términos de interacción.
Correlation	Mapa de calor de las correlaciones entre predictores numéricos (disponible antes del ajuste).
Summary	Métricas clave (R^2 , R^2 ajustado, GR^2 , R^2 de CV, RMSE, MAE) y métricas de tasación (COD, PRD, razón mediana).
Coefficients	Tabla de coeficientes no nulos con avisos de violación de signo.
Variable Importance	Diagrama de barras y tabla ordenada de la importancia de los predictores ($ \beta \times sd(x)$). Interacción con plotly al pasar el cursor, cuando está disponible.
Contributions	Gráficos de efecto parcial por variable con etiquetas de pendiente. Dispersión + línea de ajuste para predictores numéricos, diagramas de caja para factores, superficie/dispersión/mapa de calor 3D para interacciones.
ANOVA	Tabla de descomposición de la varianza: SS por variable, % de la SS del modelo, coeficiente.
Diagnostics	Cuatro gráficos: trayectoria de coeficientes, error de CV, valores reales frente a predichos y residuos frente a valores ajustados.
Glmnet Output	Salida bruta del modelo: semilla aleatoria empleada, impresión del modelo, lambda seleccionado, lambda.min/1se, gamma (si es relaxed) y vector completo de coeficientes.
Report	Campos de configuración y exportación del informe.
RCA Adjustments	Resultados del análisis RCA (modo de tasación, tras el cálculo).

3.6 Persistencia de los ajustes

glmnetUI guarda automáticamente su configuración en el almacenamiento local del navegador, indexada por el nombre del archivo de entrada. Cuando vuelve a cargar el mismo archivo, se restauran todos los ajustes: selección del objetivo, casillas de los predictores, tipos de datos, signos esperados, parámetros de glmnet y matriz de interacciones. El último modo de uso empleado

también se guarda globalmente y se restaura al volver a iniciar la aplicación. Hay tres opciones disponibles:

- **Use last settings for input file** — restaurar los ajustes guardados para este archivo concreto
- **Use default settings** — aplicar los valores predeterminados globales guardados
- **glmnet defaults** — restablecer los valores de fábrica ($\alpha = 1$, CV lambda.1se, 10 pliegues, gaussian, standardize activado)

Un botón **Save current as default** guarda todos los ajustes actuales como valor predeterminado global.

3.7 Modo oscuro

Haga clic en el icono de luna/sol de la esquina superior derecha para alternar entre los temas claro y oscuro. La preferencia de tema se guarda en el almacenamiento local y persiste entre sesiones. Todos los elementos de la interfaz (tablas, gráficos, tarjetas, botones) se adaptan al tema seleccionado mediante variables CSS.

3.8 Configuración regional y de idioma

glmnetUI admite convenciones internacionales de formato numérico y de CSV mediante un sistema de configuración regional basado en el país. El desplegable **Settings** de la barra de título ofrece selectores **Country** y **Paper** para 31 países admitidos. Cada preajuste configura:

- **Separador de CSV** — coma (,) para EE. UU., Reino Unido y Japón, o punto y coma (;) para gran parte de Europa, donde la coma se usa como signo decimal.
- **Signo decimal** — punto (.) o coma (,).
- **Separador de millares** — coma (EE. UU./Reino Unido/Japón), punto (Alemania/Italia/España), espacio (Finlandia/Francia/Polonia/Bálticos/Ucrania/Rusia) o apóstrofo (Suiza).
- **Tamaño de papel** — Letter (EE. UU./Canadá/México) o A4 (en el resto de los casos).

3.8.1. Guardar los valores predeterminados

Haga clic en **Save as my default** para almacenar sus preferencias regionales globalmente. Estos valores predeterminados se aplican a todas las sesiones futuras con independencia del archivo de datos que cargue.

***Consejo:** El formato numérico de los ejes de los gráficos y de las etiquetas de pendiente se adapta automáticamente a su configuración regional. La configuración alemana usa puntos para los millares (200.000), la finlandesa usa espacios (200 000) y la suiza usa apóstrofes (200'000). No se muestran símbolos de moneda — glmnetUI es agnóstico respecto de la moneda.*

4 Modo de tasación

Cuando selecciona **For Appraisal** como propósito, glmnetUI se configura para la valuación de una sola propiedad. Todas las funciones descritas en el capítulo 3 siguen disponibles; este capítulo cubre únicamente las adiciones específicas de la tasación.

4.1 Tratamiento de la fila del sujeto

En el modo de tasación, **la fila 1 de su conjunto de datos es la propiedad sujeto** y todas las filas restantes son ventas comparables. Su archivo de entrada debe organizarse en consecuencia (véase el capítulo 2). El precio de venta del sujeto puede dejarse en blanco o fijarse en cualquier valor — glmnetUI lo trata automáticamente como NA durante el ajuste.

Tras la importación, la pestaña Data Preview se divide en dos secciones: **Subject Property** (fila 1) y **Comparable Sales** (filas 2 y siguientes). La fila 1 siempre se excluye del ajuste del modelo. Tras el ajuste, el modelo genera igualmente predicciones para la fila del sujeto.

4.2 Fecha efectiva y edad de la venta

En los modos de tasación y de mercado aparece un campo **Effective Date** (fecha efectiva) en la sección de configuración de variables (con la fecha de hoy por omisión). Si designa una columna como **contract_date** en el desplegable de la columna Special, glmnetUI calcula una columna **sale_age** — el número entero de días entre la fecha de contrato de cada venta y la fecha efectiva. Esta columna se añade como predictor.

Cuando cambia la fecha efectiva, **sale_age** se recalcula automáticamente.

4.3 Designaciones especiales de columnas

En los modos de tasación y de mercado aparece un desplegable **Special** para cada predictor en la tabla de configuración de variables. Véase el capítulo 6 para la referencia completa de los tipos especiales y sus efectos.

4.4 Panorama de los ajustes RCA

El botón **Calculate RCA Adjustments & Download** (sección 8 de la barra lateral, visible en los modos Appraisal y Market tras el ajuste) calcula ajustes derivados del mercado para cada comparable en relación con el sujeto. El flujo de trabajo RCA completo se describe en el capítulo 11.

***Consejo:** La puntuación CQA que asigne al sujeto controla qué parte de la distribución de residuos se atribuye al sujeto. Una puntuación de 5,00 sitúa al sujeto en la mediana de los comparables.*

5 Modo de análisis de zona de mercado

Cuando selecciona **Market Area Analysis** como propósito, glmnetUI ofrece las mismas funciones específicas del sector inmobiliario que el modo de tasación (columnas especiales, edad de la venta, redondeo de coordenadas), pero está orientado a analizar un grupo de propiedades en lugar de valorar un único sujeto.

5.1 Diferencias respecto al modo de tasación

- **El tratamiento de la fila del sujeto es flexible** — en el modo de mercado, la fila 1 no se trata automáticamente como propiedad sujeto.
- **No hay Sales Comparison Grid** — el paso “Calculate RCA Adjustments & Download” (Sección 8) está disponible en el modo de mercado, pero el paso de Sales Grid (Sección 9) es exclusivo de la tasación.

5.2 Cuándo usar el modo de mercado

El modo de análisis de zona de mercado es apropiado cuando usted está:

- Construyendo un modelo de regresión para un vecindario o zona de mercado con el fin de entender los factores de valor
- Analizando cómo variables como la superficie, la antigüedad, el tamaño del terreno y la ubicación afectan a los precios de venta en un grupo de propiedades
- Preparando el respaldo de un análisis de condiciones de mercado o de una delimitación de vecindario
- Trabajando con un conjunto de datos en el que desea las funciones de columna especial (edad de la venta, redondeo de coordenadas) pero no necesita el flujo de trabajo de ajustes RCA

***Consejo:** El modo de mercado también es útil para la regresión inmobiliaria general cuando se desean las funciones de columna especial pero no se necesita el flujo de trabajo de ajustes RCA.*

6 Selección de variables

La Sección 4 de la barra lateral — **Variable Configuration** — es donde se elige qué columnas participan en el modelo y cómo se tratan.

6.1 Variable objetivo

El desplegable **Target (response) variable** situado en la parte superior de la Sección 4 enumera todas las columnas de su conjunto de datos. Seleccione una columna como variable respuesta (por ejemplo, el precio de venta). La columna objetivo se excluye automáticamente de la lista de predictores.

6.2 La tabla de predictores

Debajo del selector de objetivo, una tabla enumera todas las columnas restantes con los campos siguientes:

Columna	Descripción
Variable	Nombre de la columna
Type	Desplegable de tipo de dato: numeric , integer , character , Date , POSIXct
Include	Casilla — incluir esta columna como predictor en el modelo
Factor	Casilla — tratar la variable como factor categórico (crea columnas indicadoras en la matriz del modelo)
Force	Casilla — forzar su entrada en el modelo (factor de penalización = 0, nunca descartada por el lasso)
Special	Desplegable (solo en tasación/mercado) — véase la referencia de tipos de columna especiales más abajo
Sign	Signo esperado del coeficiente: positivo, negativo o cualquiera
NAs	Recuento de valores faltantes

6.3 Detección de tipos de datos y anulaciones

glmnetUI detecta automáticamente los tipos de datos al importar. Se reconocen las columnas numéricas, enteras y de fecha. Las columnas de texto que parecen fechas (patrones de formato de fecha habituales) se clasifican como **Date**.

Puede anular cualquier detección cambiando el desplegable **Type**. Cambiar los tipos afecta a cómo se codifica la columna en la matriz del modelo.

Las columnas de texto, TRUE/FALSE y de tipo factor de R se tratan automáticamente como categóricas. Una columna numérica se convierte en factor solo cuando se marca su casilla **Factor** — glmnetUI nunca infiere el carácter categórico a partir del rango de valores de una columna numérica, de modo que predictores numéricos discretos como un recuento de baños permanecen continuos salvo indicación en contrario. El ajuste Factor persiste mediante localStorage.

6.4 Signos esperados y su aplicación

Para cada predictor, el desplegable **Sign** especifica si se espera que su coeficiente sea positivo, negativo o de cualquier signo. Tras el ajuste:

- Los coeficientes con signos inesperados se **marcan en la tabla Coefficients**

- Los avisos de signo aparecen en la pestaña Summary

Cuando la casilla **Enforce Sign Constraints** está activada en los parámetros de glmnet (capítulo 7), se obliga al modelo a producir coeficientes que coincidan con los signos esperados. Esto se implementa mediante `upper.limits` y `lower.limits` en glmnet.

6.5 Referencia de los tipos de columna especiales

En los modos de tasación y de mercado, el desplegable **Special** ofrece las opciones siguientes:

Ponderación:

- **weight** — columna de peso de la observación (disponible en todos los modos; solo se permite una; las filas con peso = 0 se excluyen del ajuste)

Tipos de fecha y hora:

- **contract_date** — activa el cálculo automático de **sale_age** a partir de la fecha efectiva
- **listing_date** — se usa como alternativa para calcular los días en el mercado
- **dom** — identifica la columna de días en el mercado
- **sale_age** — designa una columna de edad de venta ya existente, de modo que cualquier columna pueda servir como edad de la venta en lugar de que glmnetUI calcule una a partir de **contract_date**

Tipos monetarios:

- **concessions** — identifica las concesiones de venta (créditos del vendedor, incentivos al comprador, etc.)

Tipos de transacción:

- **sale_type** — identifica la columna de tipo de venta o transacción

Tipos de tamaño y localización:

- **latitude** — los valores se redondean automáticamente a 3 decimales
- **longitude** — mismo tratamiento de redondeo que la latitud
- **area** — identificador de zona de mercado o de vecindario (por ejemplo, **area_id**, **neighborhood**, **market_area**)
- **living_area** — permite los cálculos de residuos por pie cuadrado (**residual_sf** y **cqa_sf**)
- **lot_size** — columna de tamaño del terreno
- **site_dimensions** — agrupada con el tamaño del terreno (por ejemplo, “75x120”)

Tipos de antigüedad:

- **actual_age** — columna de edad de la propiedad
- **effective_age** — edad efectiva de la propiedad

Tipos de visualización:

- **display_only** — la columna se incluye en las exportaciones a Excel pero queda totalmente excluida del ajuste del modelo. Úsela para campos de dirección, números de MLS, identificadores de parcela u otros datos de referencia. Varias columnas pueden llevar esta designación.

***Consejo:** Las columnas designadas como **display_only** permanecen en su conjunto de datos y aparecen en la exportación a Excel, pero quedan totalmente excluidas de la tabla de predictores. Úselas para columnas de identificador, direcciones u otros campos de referencia.*

7 Selección de parámetros

La Sección 5 de la barra lateral — **glmnet Call Parameters** — da acceso a todas las opciones de configuración del modelo elastic net. Cada parámetro tiene un icono azul de ayuda (?) con una explicación emergente.

7.1 Alpha (parámetro de mezcla)

El parámetro alpha controla el tipo de regularización:

Valor de alpha	Tipo	Comportamiento
0	Ridge	Contrae todos los coeficientes proporcionalmente; nunca fija ninguno en cero
1	Lasso	Puede fijar coeficientes exactamente en cero (selección de variables)
0 a 1	Elastic Net	Combinación de ridge y lasso

Hay dos métodos de selección de alpha disponibles:

- **Fixed** — fijar un único valor de alpha con el deslizador (por omisión: 1,0)
- **Grid Search** — probar varios valores de alpha y seleccionar el que tenga el menor error de CV. Se configuran el rango y el número de valores.

7.2 Selección de lambda

Lambda controla la intensidad de la regularización:

Método	Descripción
Cross-validation	Recomendado. Prueba muchos valores de lambda mediante validación cruzada de k pliegues.
Manual	Introducir un valor concreto de lambda si tiene motivos para ello.

Al usar la validación cruzada hay dos opciones de lambda disponibles:

- **lambda.1se** (por omisión) — el mayor lambda dentro de 1 error estándar del mínimo. Produce modelos más sencillos y más regularizados. Recomendado para el trabajo de tasación.
- **lambda.min** — el lambda que minimiza el error de validación cruzada. Produce el modelo con mejor ajuste, pero potencialmente más complejo.

El parámetro **Number of CV Folds** controla cómo se divide el conjunto de datos para la validación cruzada (por omisión: 10 pliegues).

7.3 Family (familia)

Elija la familia de distribución de su variable respuesta:

Familia	Caso de uso
gaussian	Respuestas continuas (p. ej., el precio de venta). La más habitual.
binomial	Resultados binarios (p. ej., vendido/no vendido).
poisson	Datos de recuento (p. ej., número de ventas).

7.4 Standardize (estandarizar)

Cuando está marcada (por omisión), todos los predictores se escalan para tener media 0 y desviación típica 1 antes del ajuste. Esto garantiza que la penalización trate a todos los predictores por igual, con independencia de su escala original. Los coeficientes se devuelven en la escala original. Normalmente conviene dejarla activada.

7.5 Restricciones de signo

Cuando **Enforce Sign Constraints** está activado, los coeficientes se restringen para que coincidan con los signos esperados fijados en la tabla de variables (capítulo 6):

- Un signo “positive” obliga a que el coeficiente sea ≥ 0
- Un signo “negative” obliga a que el coeficiente sea ≤ 0
- Las variables fijadas en “either” quedan sin restricción

Esto se implementa mediante los parámetros `upper.limits` y `lower.limits` de `glmnet()`.

7.6 Relaxed lasso

El lasso ordinario usa la misma penalización para la selección de variables Y para la estimación de los coeficientes, lo que puede contraer en exceso hacia cero coeficientes importantes. El **relaxed lasso** separa estos pasos:

1. Primero, seleccionar las variables mediante el lasso (la penalización determina qué coeficientes sobreviven)
2. Después, reajustar las variables supervivientes con menos penalización o sin ella para obtener estimaciones menos sesgadas

El deslizador **Gamma** controla el grado de relajación:

Gamma	Efecto
0	Totalmente relajado (reajuste por MCO, sin contracción alguna)
1	Sin relajación (igual que el glmnet ordinario)
Intermedio	Combina los ajustes relajado y penalizado

Con validación cruzada, gamma se elige automáticamente para obtener un rendimiento óptimo.

7.7 Semilla aleatoria

Un campo de texto **Random seed**, precargado con un entero aleatorio, hace reproducible la asignación de pliegues de la validación cruzada — se llama a `set.seed()` antes del ajuste, y la semilla empleada se muestra en la pestaña Glmnet Output y en el mensaje de estado del ajuste. Después de cada ajuste se genera automáticamente una nueva semilla (usando la hora del sistema, no el generador de números aleatorios de R). Las últimas 5 semillas empleadas para el archivo actual se muestran como enlaces en los que se puede hacer clic para recuperarlas fácilmente; el historial de semillas se reinicia al cargar un nuevo archivo.

7.8 Matriz de interacciones

Una matriz triangular superior interactiva permite controlar qué pares de variables pueden interactuar. Cada celda contiene una casilla — marcada significa que el término de interacción $x_1:x_2$ se incluye en la matriz del modelo.

- **Allow All** — marcar todos los pares de interacción
- **Clear All** — desmarcar todos los pares (el valor por omisión para un archivo nuevo es todos desmarcados)
- **Hacer clic en el nombre de una variable** — alternar todas las interacciones de esa variable

Las selecciones de interacción se guardan en el localStorage del navegador para cada archivo de entrada y se restauran al volver a cargar el mismo archivo.

Cuando hay alguna interacción activada, el informe exportado incluye una matriz **Enabled Interactions** (los predictores en ambos ejes y una marca de verificación donde se incluyó el término producto $x_1:x_2$), renderizada en apaisado y análoga a la matriz Allowed Interactions de earthUI. Tenga en cuenta que el lasso puede haber contraído a cero el coeficiente de un término activado.

***Consejo:** Con muchos predictores, el lasso anulará automáticamente los términos de interacción innecesarios. Empezar con "Allow All" y dejar que el modelo seleccione es un enfoque razonable.*

Advertencia

Activar las interacciones de glmnet producirá ajustes que pueden resultar difíciles de interpretar y explicar ante un tribunal o en una auditoría. Use las interacciones con cautela y verifique que los ajustes resultantes sean razonables.

7.9 Bloquear el efecto principal (solo interacción)

Haga **clic con el botón derecho** sobre el nombre de una variable en la matriz de interacciones para impedir que entre en el modelo como efecto principal. Aparece un indicador “1” en negrita después del nombre de la variable. Cuando está bloqueada:

- La variable se **elimina** de los términos principales de la fórmula
- La variable **sí puede** aparecer en términos de interacción ($x_1:x_2$)
- Vuelva a hacer clic con el botón derecho para desbloquearla
- El estado persiste mediante localStorage para cada archivo de entrada

Un caso de uso típico: al modelar ajustes temporales que deben escalar con el tamaño de la propiedad, bloquee **sale_age** como efecto principal y permítalo solo en una interacción con **living_area**. Esto crea un término de interacción **sale_age:living_area** en lugar de un coeficiente **sale_age** plano — la tasa de apreciación temporal varía según el tamaño de la vivienda.

7.10 Parámetros avanzados

Una sección plegable “Advanced” al final de glmnet Call Parameters expone ajustes adicionales. Estos usan valores por omisión razonables, pero son visibles para que toda la configuración del modelo pueda documentarse ante un tribunal o en una auditoría.

- **Lambda Min Ratio** — razón entre el lambda más pequeño y el más grande de la trayectoria de regularización. Por omisión: 0,00001. Para conjuntos de datos pequeños (<50 observaciones), un valor mayor (p. ej., 0,01) puede mejorar el rendimiento.

- **Number of Lambda Values** — número de valores de lambda en la trayectoria. Por omisión: 100.
- **CV Loss Metric** — función de pérdida para la validación cruzada: MSE (por omisión para gaussian), MAE (robusta frente a valores atípicos) o desviación.
- **Convergence Threshold** — umbral de convergencia del descenso por coordenadas. Por omisión: 1e-07.
- **Max Iterations** — número máximo de iteraciones del descenso por coordenadas. Por omisión: 100.000. Auméntelo si aparecen avisos de convergencia.
- **Fit Intercept** — incluir un término de intercepto (constante). Por omisión: activado. El intercepto es el “valor base” en el RCA de tasación.

7.11 Valores predeterminados de los ajustes

Tres opciones controlan cómo se inicializan los parámetros al cargar un archivo:

- **Use last settings for input file** — restaurar exactamente los ajustes empleados la última vez con este archivo
- **Use default settings** — aplicar los valores predeterminados globales guardados
- **glmnet defaults** — restablecer los valores de fábrica (alpha = 1, CV lambda.1se, 10 pliegues, gaussian, standardize activado)

Haga clic en **Save current as default** para almacenar todos los ajustes actuales como valor predeterminado global para archivos futuros.

8 Ajuste del modelo

8.1 El botón de ajuste

La Sección 6 de la barra lateral contiene el botón **Fit Glmnet Model**. Al hacer clic en él se ejecuta el modelo con su configuración actual.

8.2 Qué ocurre durante el ajuste

Cuando hace clic en Fit, glmnetUI:

1. **Prepara los datos** — elimina las filas con NA, aplica los pesos (si se han designado), excluye la fila del sujeto (en modo de tasación) y codifica las variables de factor como columnas indicadoras.
2. **Construye la matriz del modelo** — crea la matriz de diseño incluyendo los términos de interacción permitidos por la matriz de interacciones.
3. **Aplica las restricciones** — si la imposición de signos está activada, fija `upper.limits` y `lower.limits` sobre los coeficientes.
4. **Ejecuta la validación cruzada** — si se ha seleccionado CV, llama a `cv.glmnet()` para hallar el λ óptimo. Si el relaxed lasso está activado, llama a `cv.glmnet(..., relax = TRUE)`.
5. **Ajusta el modelo final** — almacena el modelo ajustado, los coeficientes, las predicciones y los residuos.
6. **Actualiza todas las pestañas de resultados** — se rellenan Summary, Coefficients, Equation, Variable Importance, Contributions, ANOVA y Diagnostics.

Un mensaje de estado bajo el botón Fit muestra el número de observaciones, las filas excluidas y el λ seleccionado.

9 Pestañas de resultados

9.1 Data Preview

Muestra los datos importados como una DataTable interactiva. En los modos de tasación y de mercado, la vista previa se divide en dos tablas: **Subject Property** (fila 1) y **Comparable Sales** (filas 2 y siguientes).

Cada celda se mantiene en una sola línea y se trunca al ancho de la columna, de modo que los campos anchos de texto libre (como las observaciones sobre la propiedad) no estiren una fila a lo largo de la página. **Haga doble clic en cualquier celda** para abrir una ventana emergente que muestre su contenido completo, sin truncar.

9.2 Equation

Muestra la ecuación del modelo ajustado renderizada en LaTeX mediante MathJax. Muestra todos los coeficientes no nulos con un formato matemático adecuado:

- Los coeficientes positivos se muestran con signo +
- Los coeficientes negativos se muestran con signo −
- Los términos de interacción se muestran como $x_1 \times x_2$
- Los números se formatean con separadores y con los decimales adecuados

9.3 Correlation

Una matriz en forma de mapa de calor que muestra las correlaciones de Pearson entre todos los predictores numéricos y la variable respuesta. Disponible inmediatamente después de la importación de los datos — no se requiere ajustar el modelo.

- Gradiente de color: azul (−1) → blanco (0) → rojo (+1)
- Los valores de correlación se imprimen en cada celda
- El tamaño del texto y de los ejes se adapta al número de variables
- Útil para identificar multicolinealidad antes del ajuste

9.4 Summary

Estadísticos de ajuste del modelo mostrados como tarjetas:

- **R^2** — proporción de varianza explicada
- **R^2 ajustado** — ajustado por el número de predictores
- **GR^2** — R^2 generalizado a partir de la razón de desviación de glmnet en el lambda seleccionado
- **R^2 de CV** — R^2 de validación cruzada (cuando se usa CV)
- **RMSE** — raíz del error cuadrático medio
- **MAE** — error absoluto medio

En los modos de tasación y de mercado, tarjetas adicionales muestran:

- **Median Ratio** — mediana de (predicho / real)
- **COD** — coeficiente de dispersión (cuanto menor, mejor)
- **PRD** — diferencial relacionado con el precio

Aparece un **aviso de sobreajuste** cuando el R^2 de entrenamiento supera al R^2 de CV en más de 0,1.

Debajo de las tarjetas, una tabla de coeficientes con avisos de signo duplica la pestaña Coefficients por comodidad.

9.5 Coefficients

Una tabla de todos los coeficientes no nulos que muestra:

- Nombre de la variable
- Valor del coeficiente
- Signo esperado (de la configuración de variables)
- Marca de aviso de signo si la dirección del coeficiente no coincide con lo esperado

9.6 Variable Importance

Magnitud del coeficiente estandarizado: $|\beta| \times \text{sd}(x)$ para cada predictor, agregada entre las columnas indicadoras en el caso de las variables de factor. Dos modos de visualización:

- **plotly interactivo** (cuando el paquete **plotly** está instalado): diagrama de barras horizontales con información emergente que muestra la importancia exacta, el coeficiente y el porcentaje relativo.
- **ggplot2 estático** (alternativa): diagrama de barras horizontales con etiquetas de eje formateadas con separadores de millares.

Debajo del gráfico, una DataTable muestra Variable, Importance, Coefficient y Relative %.

9.7 Contributions

Gráficos de efecto parcial por variable que muestran la contribución de cada predictor a las predicciones del modelo. Seleccione una variable en el desplegable.

Para predictores numéricos: un diagrama de dispersión de los valores x de la variable frente a su contribución ($\beta \times x$), con un segmento de línea rojo y una etiqueta de pendiente. La etiqueta de pendiente muestra el efecto marginal por unidad (por ejemplo, $+1,234.56/\text{unit}$), con unidades adaptativas para variables de rango pequeño como la latitud y la longitud (por ejemplo, $+0.12/0.001$). El subtítulo muestra el valor del intercepto (base).

Para variables de factor: un diagrama de caja que muestra la distribución de los valores de contribución por nivel del factor.

Para interacciones de dos variables: tres visualizaciones apiladas — un gráfico de superficie 3D (plotly interactivo cuando está disponible, o **persp()** estático en caso contrario) de la contribución conjunta, un diagrama de dispersión con los puntos coloreados según la contribución y un mapa de calor de la contribución media sobre los rangos de ambas variables.

9.8 ANOVA

Tabla de descomposición de la varianza que muestra, para cada predictor:

- **SS** — suma de cuadrados: $\sum(\text{contribution}_j - \overline{\text{contribution}_j})^2$
- **% de la SS del modelo** — porcentaje de la suma de cuadrados total del modelo
- **Coefficient** — el coeficiente dominante de ese predictor

Incluye una fila de intercepto y una fila TOTAL MODEL. Los términos de interacción (que contienen “:”) se muestran como grupos separados.

9.9 Diagnostics

Cuatro gráficos de diagnóstico con fuentes grandes (tamaño base 16 pt), 15 marcas de eje y etiquetas con separadores de millares (sin notación científica):

1. **Coefficient Path** (trayectoria de coeficientes) — muestra cómo cambian todos los coeficientes a medida que aumenta $\log(\lambda)$. Las líneas se colorean por variable. Ayuda a visualizar la trayectoria de regularización y la selección de variables.

2. **CV Error** — error de validación cruzada en función de $\log(\lambda)$. Los puntos muestran el error medio de CV; las barras de error muestran ± 1 error estándar. Las líneas verticales discontinuas marcan $\lambda_{\min}$ (azul) y λ_{1se} (verde).
3. **Actual vs Predicted** — diagrama de dispersión de los valores observados frente a los predichos. Los puntos deberían agruparse en torno a la línea de referencia discontinua a 45 grados. Una dispersión mayor indica más error de predicción.
4. **Residuals vs Fitted** — diagrama de dispersión de los residuos frente a los valores ajustados. Debería mostrar una dispersión aleatoria en torno a la línea del cero (discontinua). La presencia de patrones sugiere una especificación incorrecta del modelo (p. ej., no linealidad, heterocedasticidad).

Cada gráfico tiene un botón **Download PNG** para guardarlo a 150 DPI.

9.10 Glmnet Output

Salida bruta del modelo que muestra: la semilla aleatoria empleada, la impresión del modelo, el λ seleccionado, $\lambda_{\min}/\lambda_{1se}$, γ (si el relaxed lasso está activado) y el vector completo de coeficientes.

9.11 Report

Permite configurar los campos del informe (nombre del tasador, dirección de la propiedad, fecha del informe, número de expediente) y exportarlo a Word (.docx) o PDF directamente desde la pestaña. Para el flujo de trabajo completo de informes basado en Quarto (HTML, Word o PDF), véase el capítulo 13.

10 Descarga de datos

Tras el ajuste, el botón **Download Output (Excel)** (sección 7 de la barra lateral, disponible en todos los modos de uso) exporta un archivo de Excel con las predicciones y los diagnósticos a la carpeta de salida del proyecto. El encabezado de la sección dice “Download Estimated Sale Prices & Residuals” en los modos de tasación y de mercado, y “Download Estimated Target Variable(s) & Residuals” en el modo General.

10.1 Columnas de salida

Columna	Descripción
<code>est_<target></code>	Predicción del modelo (p. ej., <code>est_sale_price</code>)
<code>residual</code>	Real – predicho
<code>cqa</code>	Puntuación de condición-calidad-atractivo (escala 0–10)
<code>residual_sf</code>	Residuo / superficie habitable (si se ha designado <code>living_area</code>)
<code>cqa_sf</code>	CQA calculada a partir del orden según <code>residual_sf</code>
<code><var>_contribution</code>	Contribución de cada variable a la predicción
<code>basis</code>	Valor del intercepto (el mismo para todas las filas)

Una marca de verificación blanca aparece en el botón de descarga tras completarse correctamente.

10.2 Puntuaciones CQA

La CQA ordena el residuo de cada fila frente a todos los demás en una escala de 0 a 10:

- **CQA alta (~9–10):** se vendió por mucho más de lo predicho — probablemente condición, calidad o atractivo superiores
- **CQA baja (~0–1):** se vendió por mucho menos de lo predicho — probablemente condición inferior o venta forzada
- **CQA ~5:** cerca de la mediana

En los modos de tasación y de mercado, las filas se ordenan por `residual_sf` de forma descendente cuando se ha designado una columna `living_area`.

11 Cálculos RCA y descarga

El flujo de trabajo RCA (Residual Constraint Approach) está disponible en los modos **Appraisal** y **Market Area Analysis**, después de ajustar el modelo. La Sales Comparison Grid subsiguiente (capítulo 12) es exclusiva de la tasación.

11.1 Apertura del diálogo RCA

Haga clic en el botón **Calculate RCA Adjustments & Download** de la sección 8 de la barra lateral. Elija:

- El tipo de puntuación **CQA** o **CQA per SF**. Si se selecciona “CQA per SF”, el residuo del sujeto se escala por la superficie habitable.
- Introduzca la puntuación CQA del sujeto (0,00–9,99; por omisión 5,00)

11.2 Interpolación de la puntuación CQA

1. Se ordenan las puntuaciones CQA y los residuos de los comparables
2. Una interpolación lineal asocia su valor de CQA a un residuo
3. Valor del sujeto = predicción del modelo + residuo interpolado

11.3 Columnas de salida

Columna	Descripción
subject_value	Predicción del modelo + residuo interpolado
<var>_adjustment	Contribución del sujeto – contribución del comparable
residual_adjustment	Residuo del sujeto – residuo del comparable
net_adjustments	Suma de todos los ajustes
gross_adjustments	Suma de los valores absolutos de los ajustes
adjusted_sale_price	Precio de venta del comparable + ajustes netos

Una marca de verificación blanca aparece en el botón tras completarse correctamente el cálculo.

12 Sales Comparison Grid

La Sales Comparison Grid (cuadrícula de comparación de ventas) está disponible únicamente en el modo **Appraisal**, después de calcular los ajustes RCA (paso 8). Genera un libro de Excel con formato apto para incluirse en los informes de tasación.

12.1 Diálogo de selección de comparables

Al hacer clic en el botón **Generate Sales Grid & Download** de la sección 9 de la barra lateral se abre un diálogo modal que enumera todas las ventas comparables de la salida RCA. Los comparables se dividen en dos grupos:

- **Recommended comps** (marcados de antemano): ajuste bruto < 25 % del precio de venta, ordenados por edad de la venta (los más recientes primero).
- **Other comps** (desmarcados por omisión): los comparables restantes, ordenados por edad de la venta.

Seleccione hasta 30 comparables y haga clic en **Generate Sales Grid** para crear el libro.

12.2 Estructura del libro

El libro generado contiene hasta 10 hojas, cada una con 3 comparables uno al lado del otro:

- **Filas de encabezado:** direcciones del sujeto y de los comparables, precios de venta, fechas de venta y días en el mercado
- **Filas agrupadas:** localización (latitud, longitud, zona), tamaño del terreno (superficie del lote, dimensiones del sitio) y antigüedad (edad real, edad efectiva) con encabezados de grupo y subelementos
- **Filas de variables del modelo:** una fila por cada coeficiente no nulo del modelo que muestra la contribución del sujeto, las contribuciones de los comparables y los porcentajes de ajuste
- **Filas de características residuales:** celdas vacías y desbloqueadas para que el tasador introduzca características no captadas por el modelo (p. ej., condición, calidad, vistas)
- **Filas de resumen:** % de ajuste neto, % de ajuste bruto y precio de venta ajustado, con fórmulas activas de Excel

12.3 Protección de la hoja y celdas residuales

Cada hoja está protegida con una contraseña para evitar ediciones accidentales, pero las celdas de valor de las características residuales están explícitamente desbloqueadas. Esto permite a los tasadores introducir valores para características ausentes del modelo sin dejar de preservar los cálculos de ajuste basados en fórmulas.

13 Descarga de informes

13.1 El flujo de trabajo Quarto de dos pasos

La exportación de informes desde la barra lateral es un flujo de trabajo Quarto de dos pasos:

1. **Generate Quarto Report** (sección 10 de la barra lateral) — escribe un paquete Quarto autocontenido (el código fuente `.qmd` más los gráficos, `report_data.rds` y `reference.docx`) en la carpeta de salida de glmnet del proyecto.
2. **Convert Quarto Report** (sección 11 de la barra lateral) — renderiza un archivo `.qmd` a los formatos de salida elegidos. La ruta apunta por omisión al paquete generado más recientemente, pero puede navegar hasta cualquier `.qmd`.

13.2 Formatos de informe

En el paso Convert hay tres formatos disponibles:

- **HTML** — documento HTML autocontenido. Funciona en todas las plataformas sin más software que pandoc. Opción recomendada por omisión.
- **Word (.docx)** — documento con formato. Adecuado para su edición y distribución. Funciona en todas las plataformas.
- **PDF** — generado mediante Quarto con LaTeX. El tamaño de papel sigue la configuración regional (Letter o A4). **Requiere una instalación de LaTeX** (véase el apéndice de requisitos del sistema). Si no se detecta LaTeX, la opción de PDF se oculta automáticamente de las casillas de formato.

Los informes se generan con Quarto cuando está disponible, con rmarkdown como alternativa. La pestaña Report ofrece además botones directos de exportación a Word y a PDF.

13.3 Contenido del informe

Los informes incluyen el contenido de todas las pestañas:

- Descripción del conjunto de datos (nombre del archivo, observaciones, respuesta, predictores, propósito)
- Especificación del modelo (alpha, lambda, family, standardize, relaxed, coeficientes no nulos)
- Resumen de resultados (R^2 , R^2 ajustado, R^2 generalizado, R^2 de CV, RMSE, MAE)
- Ecuación del modelo (renderizada en LaTeX)
- Tabla de coeficientes
- Importancia de las variables (diagrama de barras y tabla)
- Gráficos de contribución por variable equivalentes a los de la pestaña Contributions:
 - **Predictores numéricos:** diagrama de dispersión con línea de ajuste lineal
 - **Variables de factor:** diagrama de caja por nivel
 - **Interacciones:** dispersión coloreada según la contribución, mapa de calor de la contribución media e instantánea estática de la superficie 3D (persp)
- Matriz Enabled Interactions (cuando hay alguna interacción activada) — los predictores en ambos ejes, con una marca de verificación donde se incluyó el término producto $x_1:x_2$, renderizada en apaisado
- Mapa de calor de la matriz de correlaciones
- Tabla de descomposición ANOVA
- Gráficos de diagnóstico:
 - Trayectoria de coeficientes

- Error de validación cruzada
 - Valores reales frente a predichos
 - Residuos frente a valores ajustados
 - Gráfico Q-Q normal
- Salida del modelo

Todas las etiquetas de eje y las leyendas de color usan números con separador de millares (sin notación científica).

Una marca de verificación blanca aparece en el botón del informe tras generarse correctamente, y se muestra un cronómetro mientras se renderiza el informe.

14 Comparación con earthUI

La aplicación earth y su Post Processing Module de glmnet forman un único conjunto de herramientas para el modelado de regresión. Comparten el mismo formato de datos, los mismos tipos de columna especiales, el mismo flujo de trabajo RCA y los mismos conjuntos de datos de demostración, pero usan motores de modelado distintos.

14.1 Diferencias principales

Característica	glmnetUI	earthUI
Método	Elastic net (glmnet)	MARS/Earth (earth)
Selección de variables	Automática mediante la penalización lasso	Automática mediante poda hacia adelante/atrás
Relaciones de los coeficientes	Solo efectos lineales	Lineales por tramos (funciones bisagra)
Interacciones	Por pares mediante la matriz de interacciones	Hasta grado 3
Signos de los coeficientes	Imponibles mediante restricciones	No imponibles directamente
Ajuste relajado	Relaxed lasso disponible	No aplicable
Ajuste de alpha	Búsqueda en malla disponible	No aplicable (no hay parámetro de mezcla)
Sales Comparison Grid	Disponible (paso 9)	Disponible (paso 8)
Gráficos de funciones g	No aplicable	Visualización de términos agrupados
Formatos de informe	HTML, Word, PDF	HTML, Word, PDF

14.2 Funciones compartidas

Ambas herramientas ofrecen:

- Importación de CSV y Excel con conversión de columnas a snake_case
- Compatibilidad con configuraciones regionales de 31 países, con separador de CSV, signo decimal y tamaño de papel
- Designaciones especiales de columnas (contract_date, living_area, area, latitude, longitude, etc.)
- Cálculo automático de sale_age a partir de la fecha efectiva
- Alternancia entre modo claro y oscuro
- Persistencia de los ajustes mediante localStorage
- Mapa de calor de correlaciones, importancia de las variables, contribuciones, ANOVA, diagnósticos
- Puntuación CQA y flujo de trabajo de ajustes RCA
- Sales Comparison Grid con selección de comparables y salida a Excel con formato
- Compatibilidad con los mismos conjuntos de datos de demostración (Appraisal_1.csv)

14.3 Cuál usar en cada caso

- Use **glmnetUI** cuando desee un modelo lineal sencillo e interpretable con selección automática de variables. El elastic net resulta especialmente adecuado cuando se dispone de muchos predictores y se desea que el modelo seleccione los más importantes.

- **Use earthUI** cuando espere relaciones no lineales (por ejemplo, que el efecto de la superficie habitable sobre el precio cambie según el tamaño) o necesite funciones bisagra para captar efectos de umbral. Los modelos earth también pueden captar interacciones de orden superior (grados 2 y 3). earthUI ofrece además una función experimental de extracción de observaciones basada en Prolog/DCG que no tiene equivalente en glmnetUI.
- **Use ambos** para comparar modelos lineales y no lineales sobre los mismos datos. Si ambos producen valores de R^2 similares, puede preferirse el modelo glmnet, más sencillo, por su interpretabilidad.

15 Conjunto de datos de demostración: Appraisal_1.csv

15.1 Descripción

El conjunto de datos de demostración se distribuye con glmnetUI (y se comparte con earthUI). Localícelo con:

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

El archivo contiene 1.502 ventas residenciales (más 1 propiedad sujeto en la fila 1) procedentes de una exportación simulada de un MLS. Los datos representan ventas de viviendas unifamiliares en un mercado con varias zonas, con una variedad de tamaños, antigüedades y ubicaciones.

No se trata de datos reales, sino de datos basados en un vecindario realista del norte de California. Toda la información identificativa ha sido alterada o eliminada.

15.2 Columnas

Columna	Tipo	Tipo especial	Descripción
weight	numeric	weight	Peso de la observación (0 = excluir del ajuste)
id	numeric	display_only	Identificador interno del registro
property_id	numeric	display_only	Identificador de propiedad del MLS
listing_id	character	display_only	Número de listado del MLS
parcel_number	character	display_only	Número de parcela del tasador del condado (APN)
street_address	character	display_only	Dirección de la propiedad
city_name	character	display_only	Ciudad
postal_code	character	display_only	Código postal
county_name	character	display_only	Condado
contract_date	Date	contract_date	Fecha del contrato de venta (calcula sale_age)
sale_age	numeric	—	Días desde la fecha de contrato hasta la fecha efectiva
sale_price	numeric	(objetivo)	Precio de venta — variable respuesta
living_sqft	numeric	living_area	Superficie habitable bruta en pies cuadrados
beds_total	integer	—	Número de dormitorios
baths_total	numeric	—	Recuento total de baños (p. ej., 2,5 = 2 completos + 1 medio)
lot_size	numeric	lot_size	Superficie del lote en pies cuadrados
area_id	integer	area	Identificador de zona del MLS
age	numeric	actual_age	Edad de la propiedad en años
latitude	numeric	latitude	Latitud (redondeada a 3 decimales para el modelo)
longitude	numeric	longitude	Longitud (redondeada a 3 decimales para el modelo)
garage_spaces	integer	—	Número de plazas de garaje
days_on_market	integer	dom	Días en el mercado
listing_date	Date	listing_date	Fecha de listado
sale_concessions	numeric	concessions	Concesiones del vendedor

15.3 Inicio rápido sugerido

1. Inicie la rutina de glmnet: `earthUI::launch_glmnet()`
2. Cree un proyecto nuevo (Sección 1) con el propósito **For Appraisal**
3. Coloque **Appraisal_1.csv** en la carpeta de entrada del proyecto y selecciónelo en la Sección 2 (haga clic en **Refresh** si es necesario)
4. Seleccione **sale_price** como objetivo
5. Asigne los tipos especiales como se muestra en la tabla anterior
6. Incluya los predictores: **sale_age**, **living_sqft**, **baths_total**, **lot_size**, **area_id** (como factor), **age**, **latitude**, **longitude**, **garage_spaces**
7. Mantenga $\alpha = 1$ (Lasso), CV con `lambda.1se`
8. Haga clic en **Fit Glmnet Model**
9. Revise las pestañas Summary, Coefficients y Diagnostics
10. Descargue la salida (paso 7) y revise el orden por CQA
11. Calcule los ajustes RCA (paso 8) con una puntuación CQA de ~5,00
12. Genere una Sales Comparison Grid (paso 9) con los comparables recomendados

16 Requisitos del sistema y resolución de problemas

16.1 Plataformas admitidas

glmnetUI funciona en **macOS**, **Windows** y **Linux**. Se recomienda encarecidamente RStudio Desktop (2023.06 o posterior) — incluye **pandoc**, necesario para los informes.

16.2 Dependencias opcionales

- **LaTeX** (TinyTeX, MiKTeX o MacTeX): solo para la generación de informes en PDF. Si no se detecta, la opción de PDF se oculta automáticamente. Instálelo con: `tinytex::install_tinytex()`
- **sysfonts + showtext**: fuente Roboto Condensed para los gráficos. Recurre automáticamente a la fuente sans-serif del sistema si no está disponible o no hay conexión.
- **earth** ($\geq 5.3.0$): solo necesario para la canalización de importación desde earthUI.

16.3 Notas por plataforma

macOS: la que menos problemas presenta. Instale TinyTeX para el PDF: `tinytex::install_tinytex()`

Windows: funciona bien con RStudio. Los equipos corporativos o con restricciones pueden tener limitaciones en el directorio temporal — la aplicación avisará en la consola pero continuará sin persistencia de los ajustes.

Linux: puede necesitar bibliotecas del sistema: `sudo apt install libcurl4-openssl-dev libssl-dev libxml2-dev libsqlite3-dev libfontconfig1-dev`

16.4 Degradación elegante

Componente ausente	Comportamiento
Sin LaTeX	La opción de PDF se oculta. HTML y Word siguen disponibles.
Sin internet / fallan las fuentes	Se usa la sans-serif del sistema. Se registra un mensaje en la consola.
Sin RSQLite / sistema de archivos de solo lectura	Los ajustes no persisten. La aplicación funciona con normalidad.
Directorio temporal sin permiso de escritura	La generación del informe falla con un error claro. Fije TMPDIR en una ubicación con permiso de escritura.

16.5 Resolución de problemas

“**PDF option not available**”: ejecute `tinytex::install_tinytex()` en R y reinicie la aplicación.

“**Settings will not persist**”: compruebe los permisos del directorio de datos del usuario (macOS: `~/Library/Application Support/R/glmnetUI/`, Linux: `~/.local/share/R/glmnetUI/`, Windows: `%APPDATA%/R/data/glmnetUI/`).

El puerto 7879 ya está en uso: ejecute `lsof -ti:7879 | xargs kill` (macOS/Linux) o `Stop-Process -Id (Get-NetTCPConnection -LocalPort 7879).OwningProcess` (Windows PowerShell).

Referencias

- Friedman, Jerome, Trevor Hastie y Robert Tibshirani (2010). «Regularization Paths for Generalized Linear Models via Coordinate Descent». En: *Journal of Statistical Software* 33.1, págs. 1-22. DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Simon, Noah et al. (2011). «Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent». En: *Journal of Statistical Software* 39.5, págs. 1-13. DOI: [10.18637/jss.v039.i05](https://doi.org/10.18637/jss.v039.i05).
- Tay, J. Kenneth, Balasubramanian Narasimhan y Trevor Hastie (2023). «Elastic Net Regularization Paths for All Generalized Linear Models». En: *Journal of Statistical Software* 106.1, págs. 1-31. DOI: [10.18637/jss.v106.i01](https://doi.org/10.18637/jss.v106.i01).