



Universidad de  
**San Andrés**

DEPARTAMENTO DE ECONOMÍA

MAESTRÍA EN ECONOMÍA

**Herramientas computacionales - Gráficos en R y principios de Schwabish**

*Matias Harari y Mariana Santi*

2021

## Gráficos sobre resultados de elecciones electorales

La Figura 1 incumple con los principios 1 y 2 de Schwabish ya que el formato puede ajustarse para que el mensaje sea más claro y contiene elementos visuales innecesarios. En primer lugar, el mapa especifica las coordenadas de latitud y longitud en los ejes, lo cual no aporta información relevante y vuelve más compleja la lectura. En segundo lugar, presenta colores estridentes que no resultan agradables a la vista. En tercer lugar, el título de la leyenda resulta innecesario ya que dicha información puede presentarse en el título. Por último, para simplificar la lectura de los datos, es conveniente colocar porcentajes en lugar de proporciones. En la Figura 2 se muestra el mismo gráfico corrigiendo los errores comentados. Como se puede observar, el mensaje es más claro y la escala de colores, además de ser más agradable, refleja de mejor manera el mensaje ya que establece con el mismo color un tono más fuerte para los estados con mayor cantidad de votos en relación a su población y un tono más claro para el caso contrario. Esta opción resulta superadora ya que al tratarse de una sola variable, es más natural utilizar un único color.

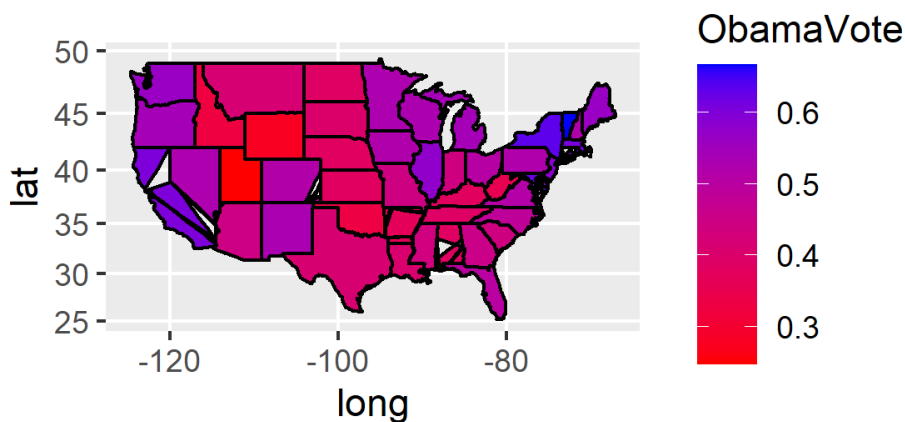


Figura 1: Mapas sobre resultados de las elecciones de 2012 con errores

### Votos a Obama por estados En porcentaje

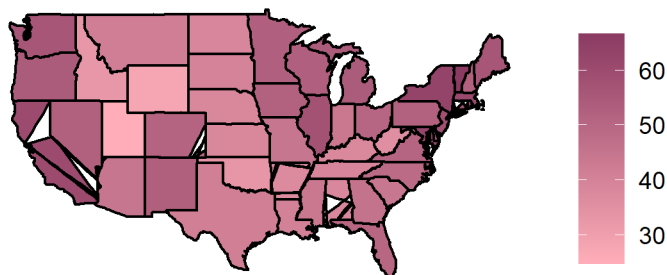


Figura 2: Mapas sobre resultados de las elecciones de 2012 corregido

## Scatter plots con datos de préstamos

La Figura 3 también incumple con los principios 1 y 2. En primer lugar, los títulos de los ejes no son claros. Las bases de datos suelen contener nombres de variables reducidos y deben ser corregidos en los gráficos o se debe aclarar a qué refieren las abreviaciones en el documento. En segundo lugar, dado que el objetivo es observar la correlación entre los pagos recibidos y otras dos variables, resulta conveniente colocar a la primera variable en el eje *y* para colocar sólo una vez el título y así evitar repetir información. En tercer lugar, los valores que indican los ejes deberían expresarse en otra unidad (en este caso, en miles) para simplificar la lectura. Por último, la leyenda podría incluirse en el título para evitar duplicar información. En la Figura 4, se corrigen dichos errores.

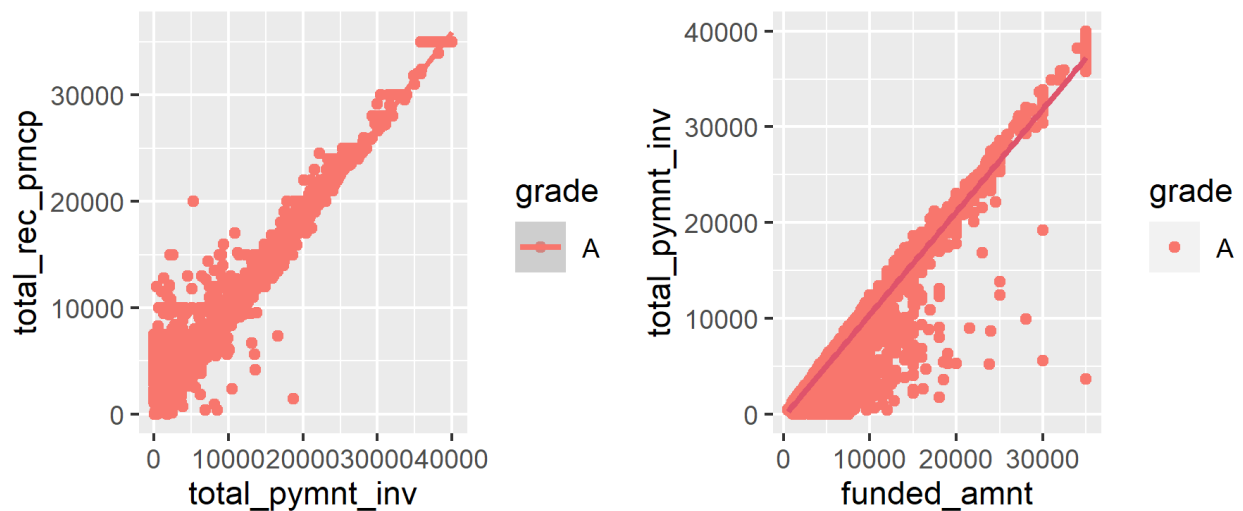


Figura 3: Scatter plots sobre clientes con calificación crediticia A con errores

Variables relacionadas con los pagos de clientes de calificación crediticia A  
En miles de USD

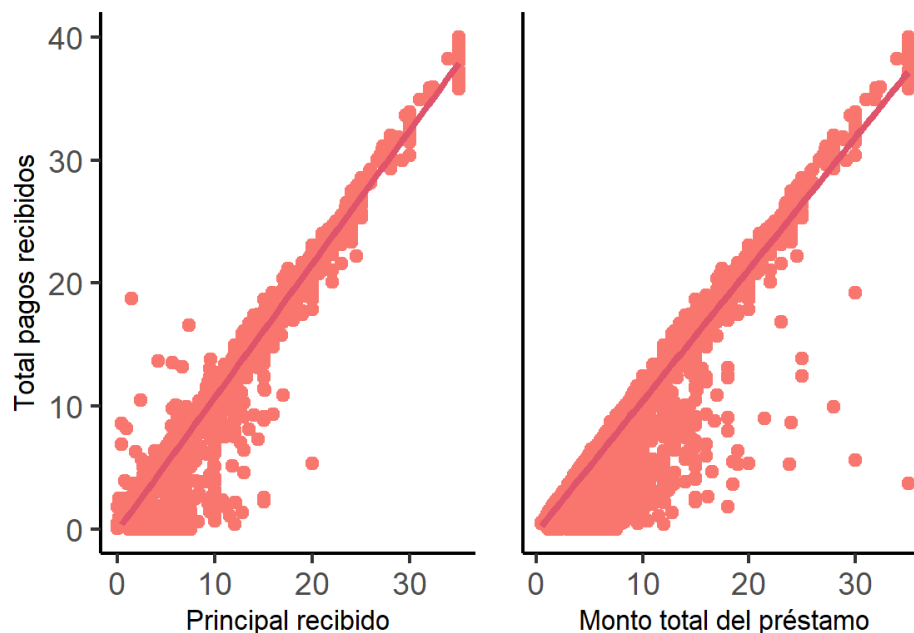


Figura 4: Scatter plots sobre clientes con calificación crediticia A corregido

## Boxplots con datos de préstamos

La Figura 5 incumple con los 3 criterios. En primer lugar, al igual que en el caso anterior se presentan los títulos de los ejes con abreviaciones de los nombres de las variables que no son claros, se utilizan números de varias cifras que dificultan la lectura y además los valores del eje  $x$  se superponen. En segundo lugar, siguiendo el principio 3, sería conveniente incluir el texto de la leyenda dentro del mismo gráfico. Si bien no es deseable que se repitan las calificaciones crediticias, teniendo en cuenta que el objetivo es comparar las distribuciones de acuerdo a la calificación crediticia y la propiedad de la vivienda del cliente, resulta más comprensible para poder comparar entre categorías armar sub-gráficos a partir de la titularidad de la vivienda como se muestran en la Figura 5. En este caso, se tiene en el eje  $x$  a las distintas calificaciones y de esta forma es más sencillo comparar entre ellas.

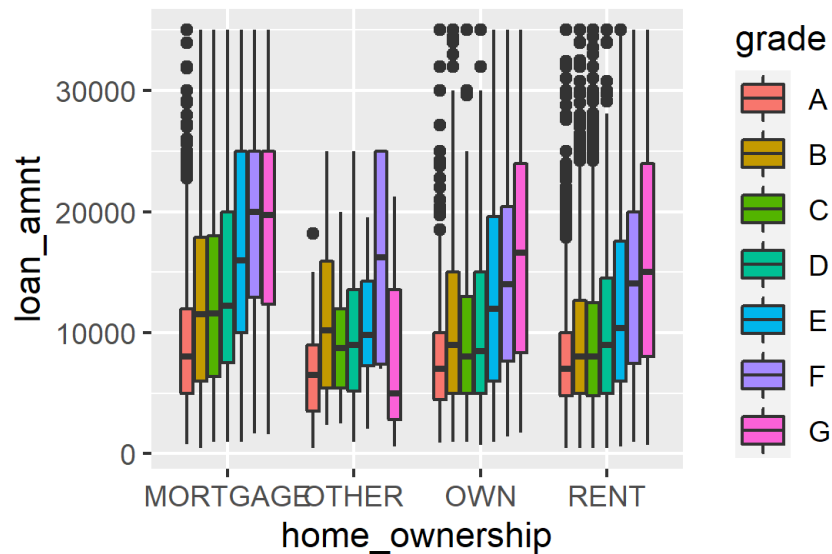


Figura 5: Boxplot con errores

### Loans according to credit grade and home ownership

Thousand dollars

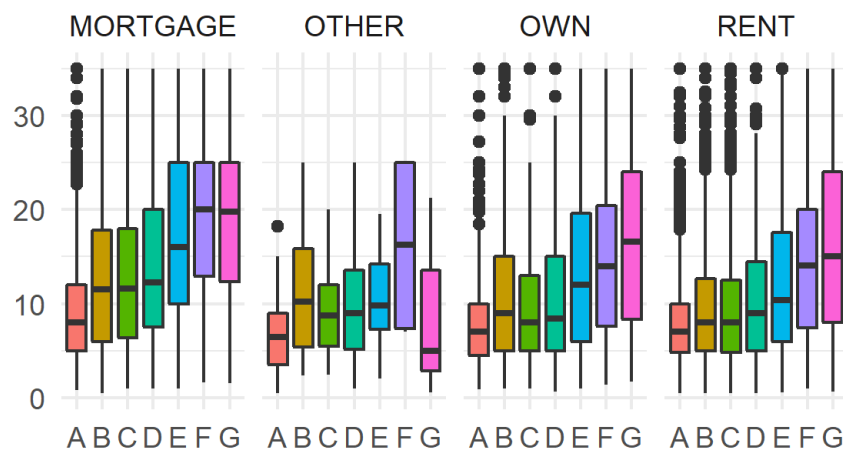


Figura 6: Boxplot corregido

Por último, se presenta la Figura 7 como nueva propuesta para comparar los niveles de préstamos según la calificación crediti-

cia y la titularidad de la propiedad. Los gráficos no muestran exactamente lo mismo, ya que los boxplot muestran distribuciones de los préstamos condicionadas en las variables mencionadas mientras que la Figura 7 contiene el promedio. Sin embargo, si el objetivo es analizar los montos de préstamos otorgados para ver si se encuentran diferencias según la calificación crediticia y titularidad de la propiedad, puede resultar de interés analizar los promedios ya que aportan un mensaje más claro sobre las relaciones.

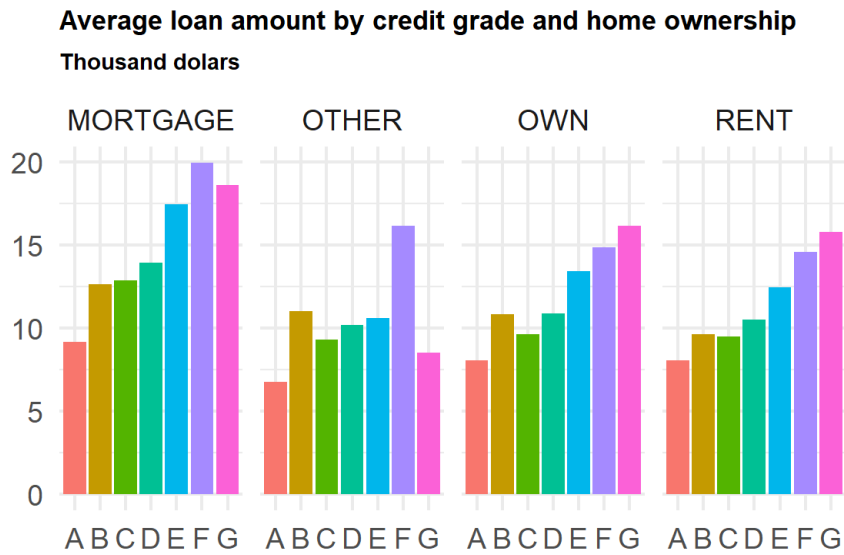


Figura 7: Nueva propuesta