

Cómo incorporar variables socio-económicas en el análisis del voto en las elecciones y/o el Referendum 2022 en Montevideo

julio/2022

Caso Edad e Ingreso del núcleo familiar en la votación del Referendum 2022

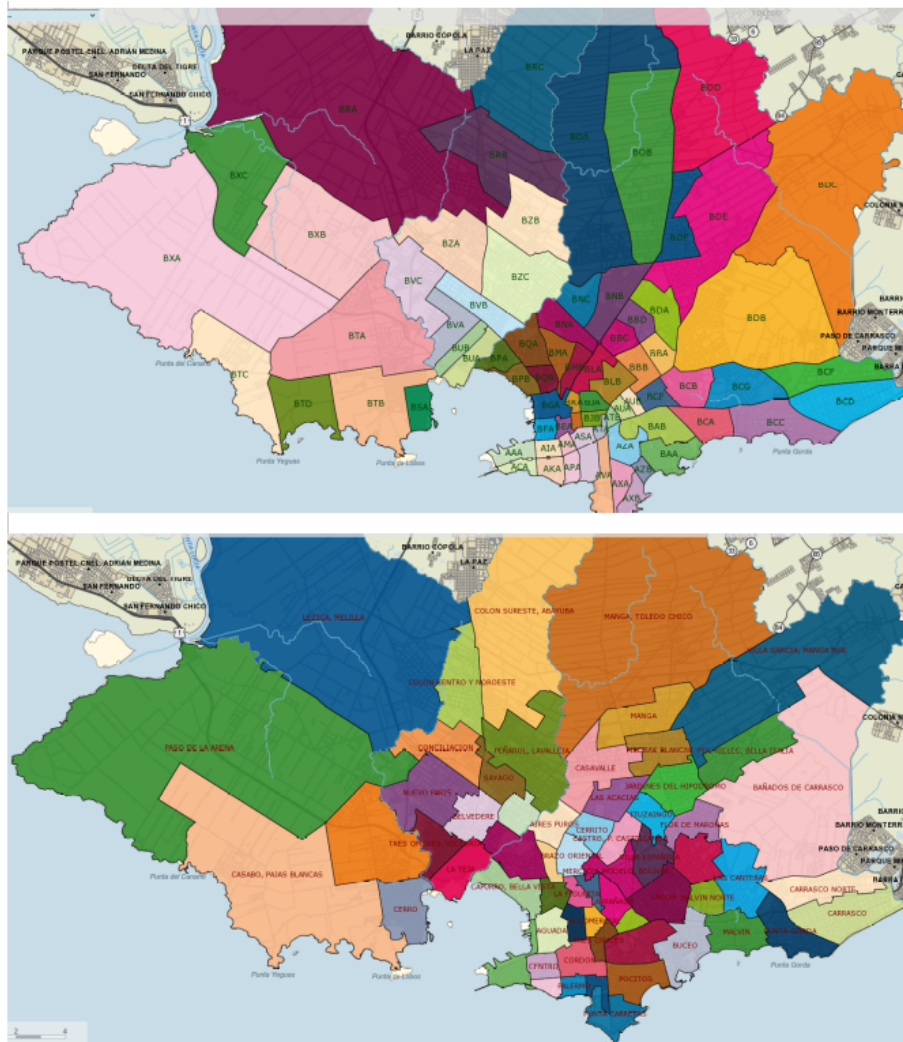
1 Justificación

El Padrón de la Corte Electoral contiene la fecha de nacimiento de los ciudadanos y es posible deducir la edad media de los habilitados y/o votantes de cada circuito de Montevideo. A partir de los resultados en Montevideo del Referendum de marzo/2022 hice una nota sobre el análisis de la dependencia del voto *SÍ* en el Referendum en función de la edad, para cada municipio de Montevideo. Ver [Porcentaje de votos SÍ vs edad en el Referendum](#). Ahí se construyen modelos para cada uno de los ocho municipios, y se comparan los modelos entre sí. En todos los municipios se observa la dependencia del voto en función de la edad: la probabilidad del voto *SÍ* decrece monótonamente con la edad.

Es razonable pensar que variables socio-económicas (ingreso per cápita, nivel educativo, etc.) están correlacionadas con el voto *SÍ* y que deberían ser incluidas junto a la edad en un modelo del voto. Pero esas variables no están disponibles para la unidad medible que es el circuito electoral.

En esta nota se propone un método para incorporar información socio-económica a nivel de los circuitos electorales en un modelo de la probabilidad del voto, ampliando el modelo de la nota anterior basado en la edad como única variable explicativa del voto *SÍ*. Este método se ejemplifica con la votación *SÍ* en el Referendum del 27/marzo/2022, con resultados igualmente válidos para el voto *NO* dado que las probabilidades son complementarias.

NOTA: *Agradezco a la Corte Electoral que me facilitó la edad media de los ciudadanos habilitados de cada circuito en el Referendum 27/03/2022.*



2 Integración de información del INE

La Corte Electoral divide Montevideo en 75 Distritos Electorales con límites geográficos asociados a las Series. El Instituto Nacional de Estadística (en lo que sigue el INE), divide Montevideo en 62 Barrios con límites precisos, pero estos límites no coinciden con los electorales. En la figura 1 parte superior se incluye la división de Montevideo en Distritos Electorales, y en la parte inferior la división de Montevideo en Barrios según el IN. (Esta información ha sido tomada del Servicio de Información Geográfica (SIG) de la Intendencia Municipal).

No hay una relación de inclusión de los Distritos Electorales en los Barrios: se observa que Distritos y Barrios se solapan.

Cómo integrar información del INE a los Distritos Electorales

Para cada Barrio y Distrito Electoral se han tomado las coordenadas de un punto central de los mismos. Las coordenadas corresponden al mismo sistema UTM 21 sur que usa el SIG.

La figura 2 presenta los centros de los Barrios con un triángulo rojo y los centros de los Distritos Electorales con un cuadrado azul. Cada centro tiene asociada una latitud y una longitud medidas en cientos de metros a partir de un origen ubicado en el punto UTM (500kmts, 6000 kmts) que corresponde a [-57 longitud, -36.1447 latitud] dados en grados decimales.

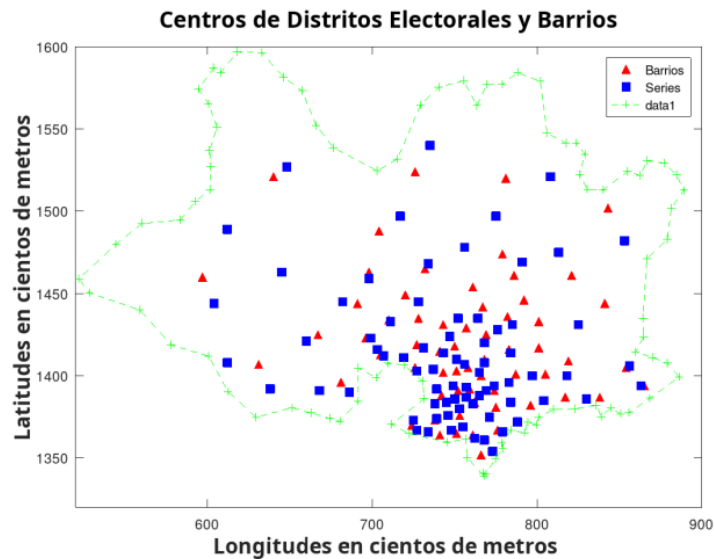


Fig. 2: Centros de Barrios y Distritos Electorales

El INE realiza una encuesta continua de hogares que contiene información sobre diversas variables para cada uno de los 62 barrios de Montevideo. Una de ellas es el Ingreso mensual del Núcleo Familiar, (a partir de aquí en adelante simplemente Ingreso). A partir de la información de diciembre de 2019 del INE se actualizaron los valores del Ingreso para la fecha 27/03/2022 usando la relación de las Unidades indexadas, UI, del INE (factor 1.215).

Interpolación desde los Barrios hacia los Distritos Electorales

Es posible estimar el Ingreso mensual del Núcleo Familiar en los Distritos Electorales mediante una interpolación de la data del INE en los Barrios. Aquí se usó la función `griddata`¹ del software OCTAVE que se usó tanto para los algoritmos como para los gráficos de esta nota. Se asume que el valor del Ingreso en los circuitos es el asociado a su Distrito Electoral.

¹ Biharmonic Spline Interpolation

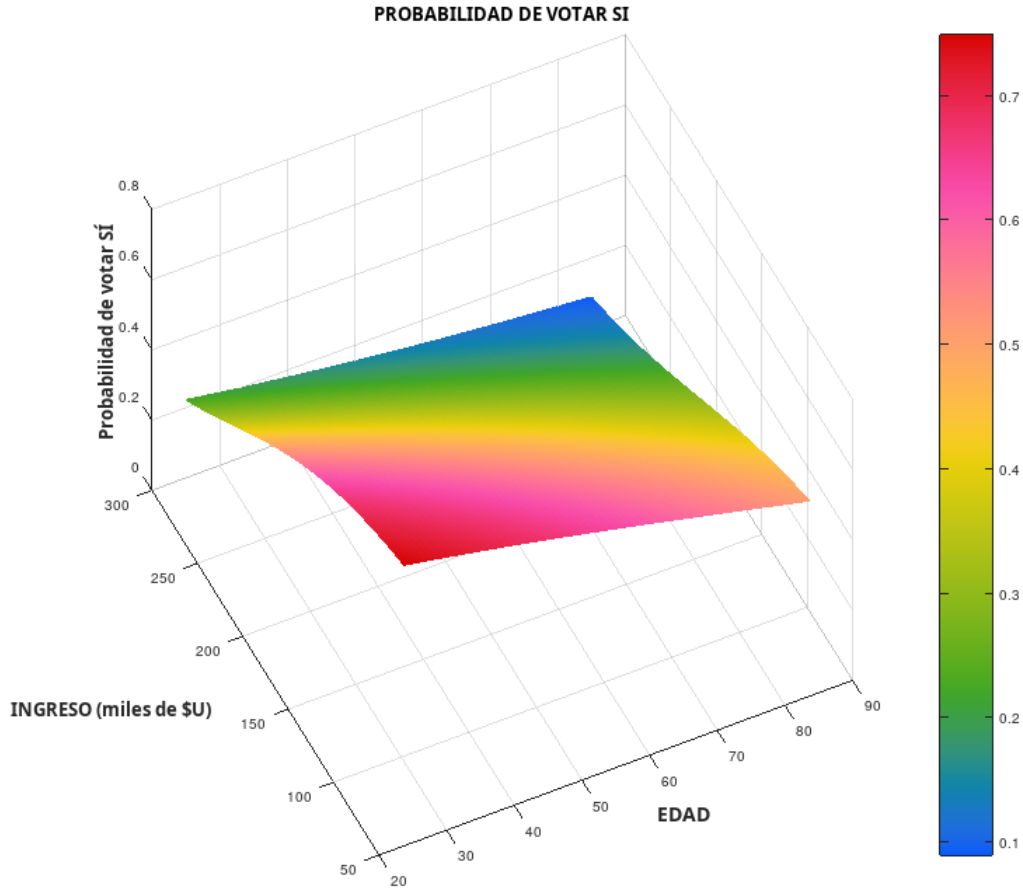


Fig. 3: Probabilidad de voto SÍ en función de Edad e Ingreso

3 Construcción de un modelo de probabilidad del voto *SÍ* en función de la Edad y el Ingreso

Para cada uno de los 2565 circuitos se tiene, además del número de votos *SÍ* y *NO*, la edad media de los votantes y el Ingreso como variables "causales". Puesto que el voto en el Referendum es dicotómico *SÍ* – *NO*, (codificados como 1 y 0), es apropiado un modelo de Regresión Logística donde la probabilidad del voto *SÍ* es función de las variables explicativas Edad e Ingreso. Analizados varios modelos posibles² se opta por el siguiente:

$$Prob(SÍ) = \frac{1}{1 + e^{-(a + b \times Edad + c \times Ingreso + d \times Ingreso^2)}} \quad (1)$$

donde la probabilidad de votar *SÍ* se expresa en función de la edad y el Ingreso del votante. En la función 1 e es

²Se consideraron modelos hasta el orden 2 en Edad e Ingreso

la constante matemática $e \cong 2,71828$ y los valores de los parámetros hallados son: $a = 1.76405$; $b = -0.01787$; $c = -0.00256$; $d = -0.00002$.³

La superficie de respuesta de la probabilidad del voto *SÍ* se presenta en la figura 3. La escala de colores indica el valor de esta probabilidad. Se observa que la probabilidad del *SÍ* disminuye cuando la Edad aumenta y cuando el Ingreso aumenta. El máximo se presenta en la esquina roja, (menor Edad e Ingreso) y el mínimo en la azul (mayor Edad e Ingreso).

Superficies de nivel

Para visualizar mejor la relación entre la probabilidad del voto *SÍ* con las variables Edad e Ingreso la figura 4 presenta 5 curvas de probabilidad constante (30, 40, 50, 60 y 70%). Debajo de cada curva existe mayor probabilidad del voto *SÍ* (por ejemplo todos los puntos por debajo de la curva azul tienen probabilidad de votar *SÍ* superior al 60%).

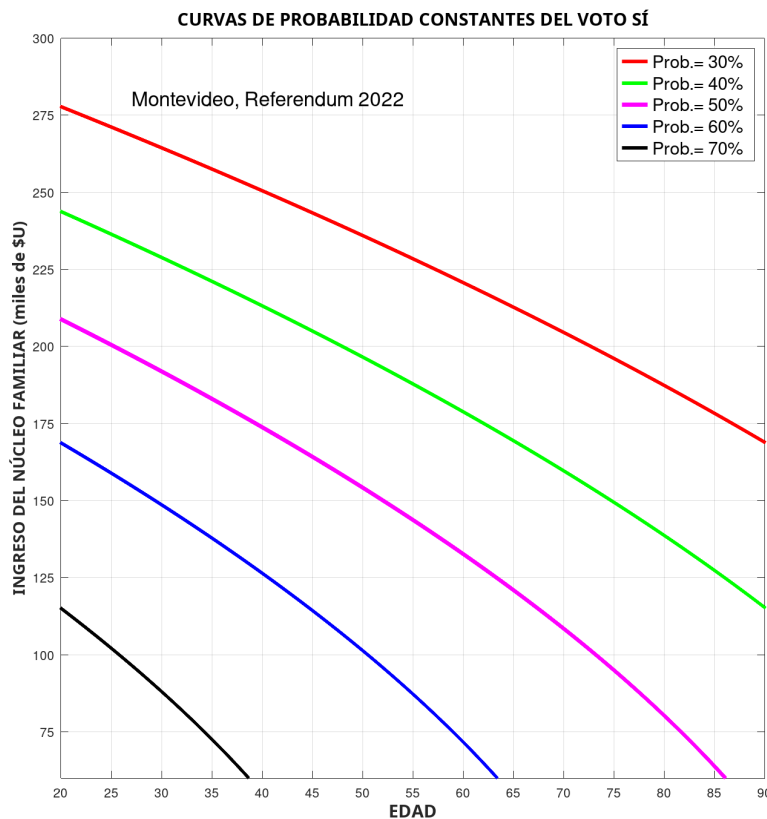


Fig. 4: Curvas de Nivel de la Probabilidad del voto *SÍ*

³Como los coeficiente b, c, y d son negativos la probabilidad decrece cuando Edad y/o Ingreso aumenta.

Cómo ver e interpretar estas curvas:

- Si nos movemos por una misma curva, un joven de 25 años y un Ingreso de 200.000 \$U tiene una probabilidad del 50% de votar *SÍ* (curva de color magenta); la misma que tiene una persona de 63 años con mucho menor Ingreso (125.000 \$U).
- Si comparamos valores sobre dos curvas, ese mismo joven aumenta su probabilidad de votar *SÍ* a 70% (curva de color negro) si su Ingreso se reduce a 105.000 \$U.

Por último, el método es válido también para el análisis y construcción de modelos del voto en elecciones de respuesta binaria (por ejemplo, segundas vueltas de elecciones futuras).

Salvador Pintos

C.I. 1585831/3

<https://salvador-pintos.github.io>