

Métodos Multicriterio Orientados al Tratamiento Automático de Datos

Trabajo N°1

Integrantes:

- Ezequiel Acosta
- Candelaria Bosch
- Gabriela Gastelu
- Agustina Maini
- Maricel Santos



Docentes:




Juan B. Cabral (FaMAF-UNC-CONICET)

Nadia Luzcywo (FCE-UNC)

Octubre 2023

Resumen

En el presente informe , se examinaron los posibles sesgos presentes en tres conjuntos de datos específicos . Estos conjuntos de datos abordan los siguientes temas de manera general:

1. Indicador de residuos peligrosos municipales reciclados por país en una serie de tiempo determinada .
2. Intensidad hídrica del valor agregado agrícola .
3. Capacidades Financieras en Argentina - BCRA .

El análisis se centró en identificar cualquier sesgo cognitivo que estuvo presente en la extracción de conclusiones o en la organización de la información en las investigaciones.

Caso 1

Residuos municipales reciclados (toneladas) EN_MWT_RCYV

Descripción del indicador

Los residuos municipales reciclados son uno de los indicadores oficiales mundiales que se tienen en cuenta para los Objetivos de Desarrollo Sostenible (ODS). En particular, este indicador compone al objetivo 12 al cual citamos a continuación 📌

Garantizar modalidades de consumo y producción sostenibles”, meta 12.4: “De aquí a 2020, lograr la gestión ecológicamente racional de los productos químicos y de todos los desechos a lo largo de su ciclo de vida, de conformidad con los marcos internacionales convenidos, y reducir significativamente su liberación a la atmósfera, el agua y el suelo a fin de minimizar sus efectos adversos en la salud humana y el medio ambiente.

Los datos provienen de la base de datos de los ODS de Naciones Unidas, junto a indicadores proxis y complementarios regionales disponibles en [CEPALSTAT](#).

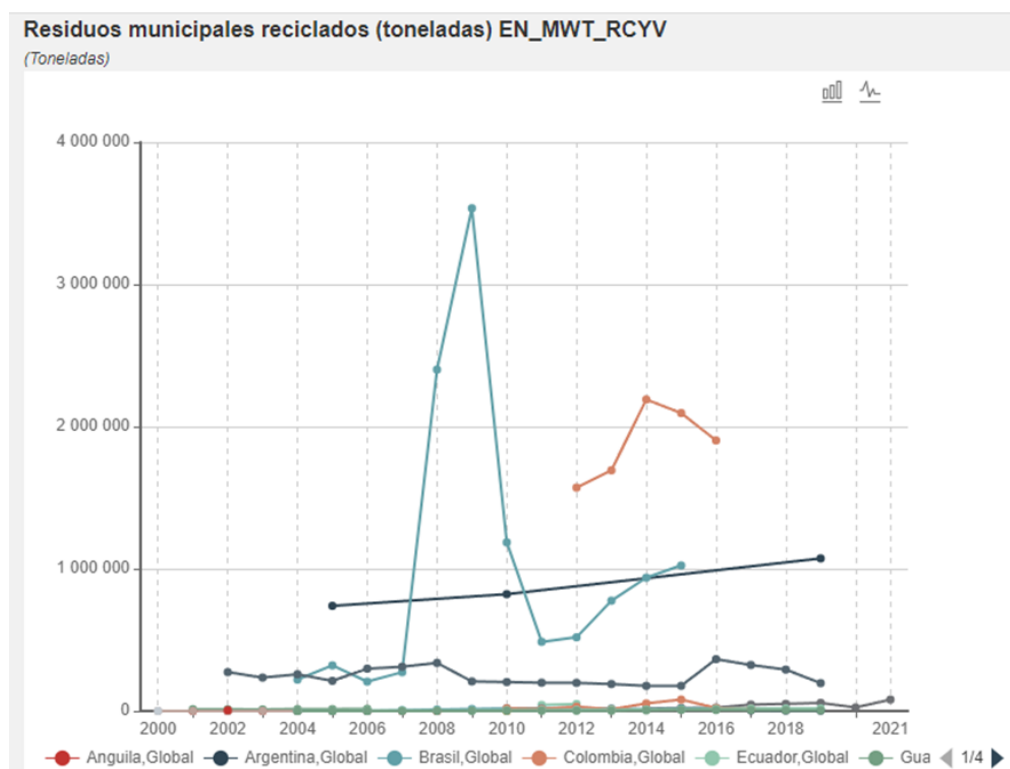
La Comisión Económica para América Latina (CEPAL), que es quién informa este estadístico, es una de las cinco comisiones regionales de las Naciones Unidas. Se fundó en 1948 para contribuir al desarrollo económico de América Latina, coordinar las acciones encaminadas a su promoción y reforzar las relaciones económicas de los países entre sí y con las demás naciones del mundo. Posteriormente, su labor se amplió a los países del Caribe y se incorporó el objetivo de promover el desarrollo social.

La ficha técnica del indicador “12.4.2 a) Desechos peligrosos generados per cápita y b) proporción de desechos peligrosos tratados, desglosados por tipo de tratamiento, Residuos municipales reciclados (toneladas) EN_MWT_RCYV” nos indica el documento donde se establecen las líneas generales donde figura la metodología completa.

En primer lugar, se puede mencionar que este indicador está **desagregado**, por lo cual se entiende que sólo indica los residuos peligrosos generados y reciclados por las municipalidades de los países. Los países que figuran en esta base son:

Anguila, Argentina, Brasil, Colombia, Cuba, Curaçao, Ecuador, Guadalupe, Guatemala, Guayana Francesa, Islas Vírgenes Británicas, Martinica, Montserrat, Perú, San Vicente y las Granadinas, Suriname, Uruguay

El periodo medido es de los años 2000 a 2021. En la siguiente gráfica se puede observar todos los datos ofrecidos 📌



Este indicador surge de la siguiente suma:

Residuos peligrosos generados = residuos peligrosos recogidos a través de servicios municipales o empresas privadas + residuos peligrosos entregados por el generador a instalaciones de tratamiento o eliminación + estimación de residuos peligrosos no contabilizados.



El último término, es el que hace que este indicador sea “nivel II”, según el documento titulado "Global Chemicals and Waste Indicator Review Document (UNEP, forthcoming)" donde figura la metodología completa. “..Para todos los indicadores en este documento, se propone un enfoque progresivo de monitoreo.(...) Nivel II: indicadores centrales que se recomienda recopilar a nivel nacional en todos los países y que se recopilarán a nivel mundial para informar sobre los Objetivos de Desarrollo Sostenible (ODS).” Hace esta categorización ya que reconoce que el nivel de residuos peligrosos generados depende no sólo del tamaño demográfico del país si no también del tipo de actividades que allí se realicen, así como de su desarrollo económico y social.

Por otro lado, podría preguntarse por qué es necesario tener en cuenta el último término de estimación, a lo que el documento indica: “Los residuos peligrosos que no están contabilizados pueden provenir, entre otras fuentes, de hogares, pequeñas empresas y explotaciones agrícolas. Debido a las pequeñas cantidades de residuos peligrosos generadas por entidad y a la falta de acceso o dificultad para acceder a servicios de recolección de residuos peligrosos, así como a la debilidad o escasa aplicación de normativas de protección ambiental, los residuos peligrosos pueden mezclarse con residuos municipales y/o ser eliminados de manera no controlada (por ejemplo, vertidos ilegales, quema a cielo abierto). En el caso de algunos países de bajos ingresos, debido a la insuficiencia de normativas de protección ambiental o de supervisión de la implementación, puede ocurrir la eliminación o quema de residuos peligrosos en lugar de su tratamiento, incluso en el caso de productores de cantidades significativas, lo que aumenta aún más la cantidad de residuos peligrosos no contabilizados. Una cantidad significativa de residuos

peligrosos no contabilizados también puede desviarse hacia el sector informal, ya sea antes o después de ingresar al sistema de recolección.”

Para poder hacer esta estimación se sugieren varios métodos: como tener en cuenta el tipo de industria y en base a esto hacer estimaciones usando el principio del balance de masas, utilizando el Análisis de Flujo de Materiales. Por otro lado, también propone analizar un ratio de residuos peligrosos [kg/cápita/día] informado para ciertos países como base comparativa para los de la misma región o semejantes.

Posibles Sesgos

A partir del análisis del indicador y los datos que lo conforman, se detectaron los siguientes posibles sesgos en el mismo:

✓ Igualación: los países no generan la misma cantidad de residuos peligrosos. No tienen el mismo tamaño de población. En el gráfico pareciera que los países más chicos no reciclan, pero quizás en proporción con la cantidad que generan, el reciclado puede ser alto. A la inversa, quizás los países más grandes generan muchos residuos y aún no tienen la capacidad suficiente para aumentar su porcentaje de reciclado como deberían.

✓ Ganancia o pérdida: no sería lo mismo mostrar la cantidad de residuos reciclados que si el indicador fuese la cantidad de residuos que quedan sin reciclar. De esta manera, sería más evidente que países como Brasil, Argentina y Colombia son los que más generan y los que más cantidad de residuos peligrosos dejan sin reciclar; mientras que países chicos, de menor crecimiento económico, generan poca cantidad y, por eso, no reciclan grandes volúmenes de residuos.

✓ Datos faltantes: en los países más chicos se muestran datos igualados a cero mientras que en países de más desarrollo hay gran cantidad de datos faltantes, aunque los valores donde sí hay mediciones son mayores. Esto podría estar indicando que los métodos de medición no son los más adecuados en ninguno de los dos casos ya que no se puede hacer un seguimiento o comparación asertiva.

[SDG indicator metadata](#)

[Global Chemicals and Waste Indicator Review Document](#)

¿Cómo se podrían abordar estos sesgos?

A partir de los posibles sesgos detectados, se proponen las siguientes soluciones que disminuirían la probabilidad de que aparezcan:

- Recopilación de datos adicionales sobre la cantidad total de residuos peligrosos generados
- Relativización de los datos por tamaño de país (según número de habitantes o por km²)
- Inclusión de información sobre residuos no reciclados
- Acompañar la información con algún tipo de indicador sobre las principales industrias y/o consumos de cada país.

Caso 2

Intensidad hídrica del valor agregado agrícola

Descripción del indicador

Este indicador es informado también por CEPAL. A diferencia del anterior, incluye más países: *Antigua y Barbuda, Argentina, Barbados, Belice, Bolivia (Estado Plurinacional de), Brasil, Chile, Colombia, Costa Rica, Cuba, Dominica, Ecuador, El Salvador, Granada, Guatemala, Guyana, Haití, Honduras, Jamaica, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana, Saint Kitts y Nevis, San Vicente y las Granadinas, Santa Lucía, Suriname, Trinidad y Tabago, Uruguay, Venezuela (República Bolivariana de)*. El periodo también es mayor siendo 1990-2019.

Es un parámetro de la eficiencia en el uso de agua para riego relacionado a la producción económica que resulta de su uso. Se define como la intensidad del uso del agua para el sector de la agricultura (extracción de agua para uso agrícola / valor agregado de la agrícola a precios constantes en millones de dólares del 2010).

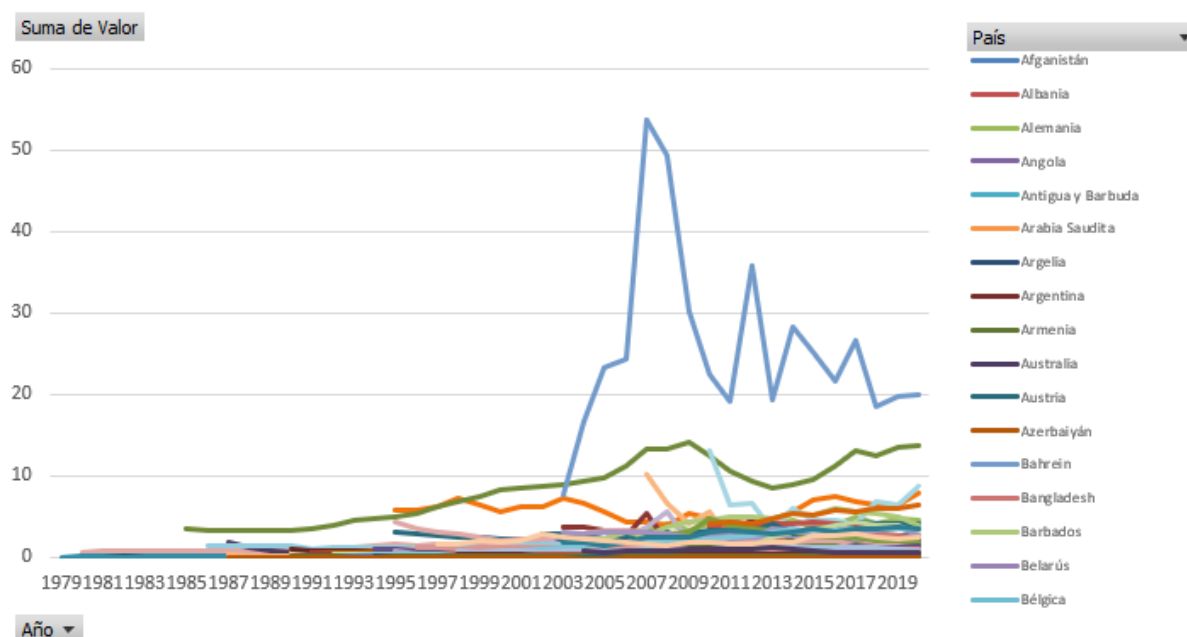


Tabla del indicador por país para los años 1979 - 2020

[FAO AQUASTAT Dissemination System](#)

Se busca que el número de este indicador sea el mínimo posible. Un bajo valor del indicador muestra que una menor cantidad de agua es usada para producir una unidad de valor agregado agrícola.

Posibles Sesgos

A partir del análisis del indicador y los datos que lo conforman, se detectan los siguientes posibles sesgos en el mismo:

✓ Igualación: Además que naturalmente los países tienen distintas extensiones geográficas, existen variaciones en las características del suelo y del clima que afectan a la fertilidad o cultivabilidad de la tierra que termina por incidir en la cantidad de agua necesaria para obtener una unidad de valor agregado agrícola.

✓ Sesgo de omisión de variables importantes: Tal como se menciona en el apartado anterior, la omisión de factores como la fertilidad de la tierra propia de la región en la que se hace el estudio, el tipo de cultivo, el tipo de clima, la cantidad de precipitaciones promedio puede estar sesgando la visión de la problemática.

¿Cómo se podrían abordar estos sesgos?

A partir de los posibles sesgos detectados, se proponen las siguientes soluciones que disminuirían la probabilidad de que aparezcan:

- Relativización de los datos por tamaño de país (según número de habitantes o por km²).
- Acompañar la información con algún tipo de indicador sobre el tipo de clima y/o cultivo que predomina.

Caso 3

Encuesta de Medición de Capacidades Financieras en Argentina - BCRA-CAF (2017)

Descripción del indicador

Esta encuesta de medición de las capacidades financieras de la población argentina se llevó a cabo con el objetivo de realizar un diagnóstico que permita identificar los conocimientos, habilidades, actitudes y comportamientos de los ciudadanos con relación a los temas financieros. La misma fue financiada por CAF -Banco de Desarrollo de América Latina- en colaboración con el Banco Central de la República Argentina (BCRA) en el año 2017.

Según se indica en la página del BCRA, para esta investigación fueron encuestadas cara a cara 1.224 personas mayores a 18 años de todos los niveles socioeconómicos, tanto residentes en las zonas urbanas como también en zonas rurales de 25 regiones del país. Tiene representatividad nacional y tiene un margen de error de +/- 2,8%. Las tareas de campo se desarrollaron del 14 de marzo al 19 de abril de 2017.

Detección de sesgos

Para el primer análisis, se prestará atención al apartado de **Resultados de la encuesta**. Allí, se realiza una evaluación de conceptos financieros entre los participantes de la muestra. Para el caso de la evaluación de la división se registra **un sesgo de anclaje**.

Viendo en detalle el párrafo en cuestión 📌

3.5.1 División

Las dificultades en el campo de las matemáticas en Argentina quedaron expuestas en los resultados de las pruebas Aprender, que consisten en la realización de una evaluación a nivel nacional en las escuelas públicas y privadas realizadas desde el año 2016 por parte del Ministerio de Educación de la Nación. Uno de los resultados más sorprendentes es que, de los estudiantes del último año de las escuelas secundarias, el 40,9% de los estudiantes está por debajo del nivel básico y el 29,3% en el nivel básico en matemáticas. Solo un 5,2% de los alumnos alcanzó el nivel avanzado en dicha disciplina.⁹

Esos resultados destacados anteriormente, son consistentes con los observados entre los adultos en el marco de la presente encuesta. A nivel nacional en Argentina solo el 59% puede realizar correctamente una división. Comparativamente estos niveles son superiores en la región y alcanzan al 85% en Colombia, 80% en Bolivia, 78% en Ecuador, 73% en Perú y 70% en Chile.

En esta investigación, se estudia sólo a la **población adulta** (18 o más edad). No obstante, como se muestra en el primer párrafo de la imagen anterior, se recuperaron los resultados de otra investigación sobre **niños y jóvenes** (resultados de las pruebas Aprender) generando así un efecto anclaje en los resultados de la investigación en adultos. El documento indica: “**sólo** el 59% puede realizar correctamente una división”, esto enfatiza dicho sesgo porque hace una extrapolación de la **información negativa** obtenida en la investigación con niños y la lleva al terreno de adultos. Se sabe que el estudio *Aprender* fue ampliamente difundido en los medios en el año en que los resultados se publicaron y generó controversia pública. Una noticia relacionada se puede consultar aquí [👉 Polémica: las preguntas de las Pruebas Aprender que generan controversia | El Destape](#)

Por otro lado, se decidió hacer foco en la **metodología** empleada para llevar a cabo el presente estudio que data de 2017. En particular, sobre la muestra total.

Tabla 1 | Características de la muestra.

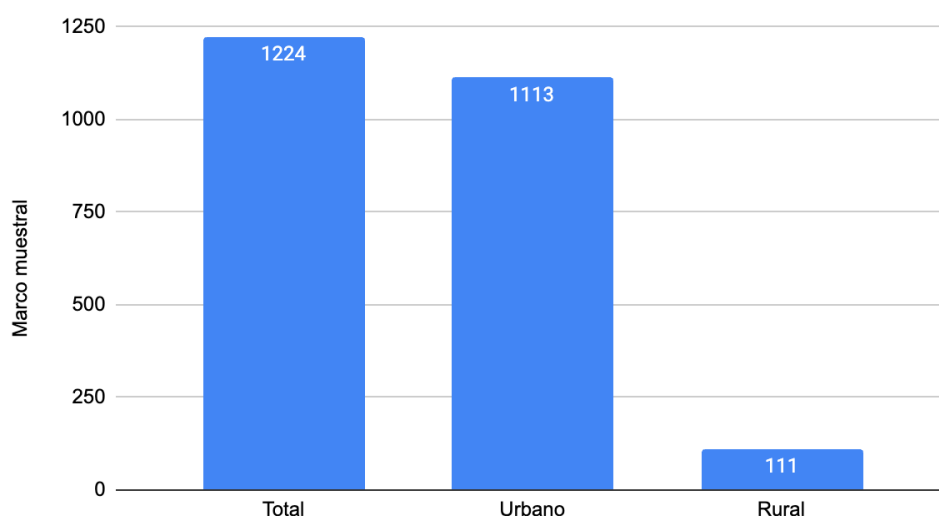
Marco Muestral	Total	Urbano	Rural	Hombres	Mujeres	18-29 años	30-49 años	50 años o más
Muestra	1224	1113	111	623	601	368	477	379
Margen de error (%)	2.8	2.9	9.3	3.9	4	5.1	4.5	5

Fuente: IPSOS Argentina.

La tabla muestra datos discretos sin una correlación que acompañe a la comprensión global que responda a la pregunta: ¿quiénes fueron las personas encuestadas?

En su lugar, se presentan gráficos que aclaren en mayor medida la relación entre las clases encuestadas, graficando solamente la proporción entre población rural y urbana.

Rural vs. urbano respecto del total



En esta nueva manera de representar la información, se observa que los registros de personas rurales es **significativamente menor** que las personas encuestadas de zonas urbanas lo cual hace que hagamos foco en dicha diferencia. En particular, ahora si se presta atención al **porcentaje de error** aceptado para cada grupo, ese porcentaje, es **dispar** entre ambas clases.

Urbano	Rural
1113	111
2.9	9.3

Se considera que hay un problema con las muestras al momento de sacar **conclusiones generales** que se pueden extraer.

En este estudio, se afirman conclusiones sobre **toda** la población argentina: “los resultados de esta encuesta ponen en evidencia un bajo de nivel de conocimientos financieros entre los adultos en Argentina” (p. 124). Pero por la diferencia entre los márgenes de error de las muestras *pareciera* que la conclusión “para todos” los argentinos no es acertada.

El índice de conocimientos financieros permite conocer el nivel de comprensión de conceptos financieros y de cálculos de matemática básicos. Los resultados de esta encuesta ponen en evidencia un bajo nivel de conocimientos financieros entre los adultos en Argentina. Es importante reparar en que solo un 60% puede realizar correctamente una división.

En principio, se comete una falacia de la generalización apresurada en la **conclusión** porque extiende los estudios a **toda la población argentina** en general asumiendo que la población es **homogénea**.

Para analizar este punto con más detalle, se vuelve a hacer el cálculo de tamaño muestral de las encuestas e indica que se deberían encuestar 1225 personas como mínimo (en el estudio se usan 1224).

Usando datos de Radiografía del trabajo argentino para hacer nuevamente el cálculo muestral tomando a 28.000.000 de argentinos económicamente activos en 2017.

Fuente: [Radiografía del trabajo argentino - Febrero 2018 \(final para pdf\)](#)

Calcula el tamaño de tu muestra

Tamaño de la población ⓘ
28000000

Nivel de confianza (%) ⓘ
95 ▼

Margen de error (%) ⓘ
2.8

Tamaño de la muestra
1,225

No obstante, esto asume que cualquier individuo seleccionado es **indistinto** para los resultados. La conclusión parecería asumir que el acceso a la información y a la tecnología es igual en las 25 regiones analizadas.

Siguiendo este argumento, si se analiza la proporción de personas en zonas urbanas y zonas rurales en 2017 entonces se encuentra que equivale a 3.634.137

2017	3,634,137
2018	3,617,403
2019	3,599,141
2020	3,579,773

Fuente: [Argentina - Población rural](#)

Si se vuelve a realizar el cálculo de tamaño muestral, y se encuentra que si se quisiera conservar el mismo nivel de error bajo (2.8) la muestra debería incluir a más representantes de la población rural.

Calcula el tamaño de tu muestra

Tamaño de la población ⓘ

3634137

Nivel de confianza (%) ⓘ

95 ▼

Margen de error (%) ⓘ

2.8

Tamaño de la muestra

1,225

El cálculo de la muestra para la población rural tiene otro cálculo asociado.

Calcula el tamaño de tu muestra

Tamaño de la población ⓘ

3.634.137

Nivel de confianza (%) ⓘ

95 ▼

Margen de error (%) ⓘ

9.3

Tamaño de la muestra

112

Continuando con el análisis de la metodología y perfil del entrevistado, en el punto 2.1 Metodología y muestra del estudio, se observan aspectos a considerar sobre las unidades muestrales.

Tabla 2 | Selección de unidades muestrales

Etapa	Unidad de muestreo	Marco muestral	Método de muestreo
1	ZONA	<ul style="list-style-type: none"> • Zonas urbanas: Cartografía del CENSO 2010 - INDEC. 	Probabilístico. Sistemático con inicio aleatorio, asegurando la dispersión dentro de cada localidad. La selección se realiza mediante un muestreo sistemático, con inicio aleatorio, asegurando la dispersión dentro de cada estrato y con probabilidad de selección proporcional al tamaño en función al número de viviendas que contiene. En las localidades que no tienen cartografía se realiza un muestreo aleatorio simple de las zonas en que se ha dividido el centro poblado.

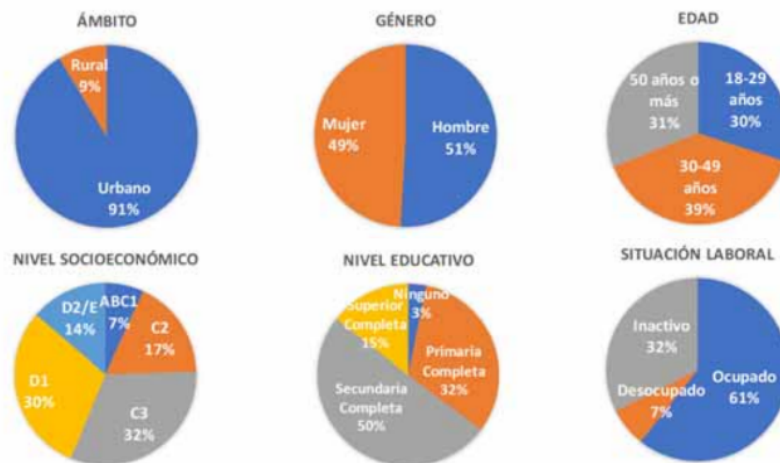
En la etapa de selección de zonas y del método de muestreo, los investigadores indican que el estudio se realizó de manera probabilística, sistemática y con inicio aleatorio, asegurando la dispersión dentro de cada estrato y con probabilidad de selección proporcional al tamaño en función al número de viviendas. No obstante, en ningún momento se muestra la fórmula según la cual se calculó esto y no es claro de entender. Este se podría corregir mostrando la fórmula según la cual se determina el método de muestreo y ejemplificando.

4	PERSONA	Realizado por el encuestador, como parte del método de selección aleatoria a realizar.	Dentro del hogar se selecciona a la persona miembro del hogar que cumpla con las condiciones requeridas. Cada encuestador debe completar una cuota de sexo y edad, de tal manera que la muestra final tenga la misma distribución del universo en estas variables.
---	----------------	--	---

En la etapa de determinar el método de muestreo de cada persona encuestada se indica que “Esto va a ser realizado por cada encuestador de manera de completar una cuota de sexo y edad” pero esto puede introducir error ya que no sabemos según qué criterio cada encuestador va a decidir si en cada casa va a ir entrevistando a un hombre o una mujer o de qué edad. Por ejemplo, quizás entrevista a todas mujeres de distintas edades y después entrevista a todos hombres de distintas edades y después, otro encuestador va entrevistando uno y uno alternativamente, y eso va generando error. Esto se podría corregir indicando un método de selección de persona miembro del hogar que sea igual para todos los encuestadores.

Cuando se muestra la distribución ponderada de la población encuestada se emplea el uso de gráficos de torta que no están ordenados de menor a mayor lo cual dificulta la interpretación de los mismos y podría mejorarse ordenándose de mayor a menor.

Gráfico 1 | Distribución ponderada de la población encuestada (*)



(*) Utiliza factores de expansión de la muestra.
Fuente: BCRA y CAF con base en Ipsos- Argentina (2017).

En la sección 3. Resultados de la encuesta, 3.1 Planificación financiera

“Los ingresos se consideran bastante estables (66% del total nacional) a través de todas las poblaciones sociodemográficas con la excepción del nivel socioeconómico D2/E (43%) y quienes se encuentran desempleados (48%).”

En la sección de resultados de la encuesta se establece que los que los ingresos se consideran bastante estables pero no se aclara A qué se refiere con que sean estables y esto puede generar errores en la interpretación del informe

3.1.3 Gastos inesperados

En la subsección de gastos inesperados indica que solo el 31% de los encuestados podría enfrentar un gasto inesperado equivalente a su ingreso mensual sin pedir un crédito o ayuda a sus familiares o amigos pero esto no es comparable de igual forma a personas con distintos niveles de ingresos. No es lo mismo enfrentar un gasto inesperado de medio millón de pesos si ese es el ingreso mensual, que enfrentar un gasto inesperado de \$180,000 si es ese el ingreso mensual, analizándolo desde el punto de vista del monto. Analizándolo desde el punto de vista del poder adquisitivo y la capacidad de ahorro quizás sea distinto entonces hay que buscar la manera de que se pueda comparar mejor esta capacidad de enfrentar gastos inesperados para todos los encuestados.

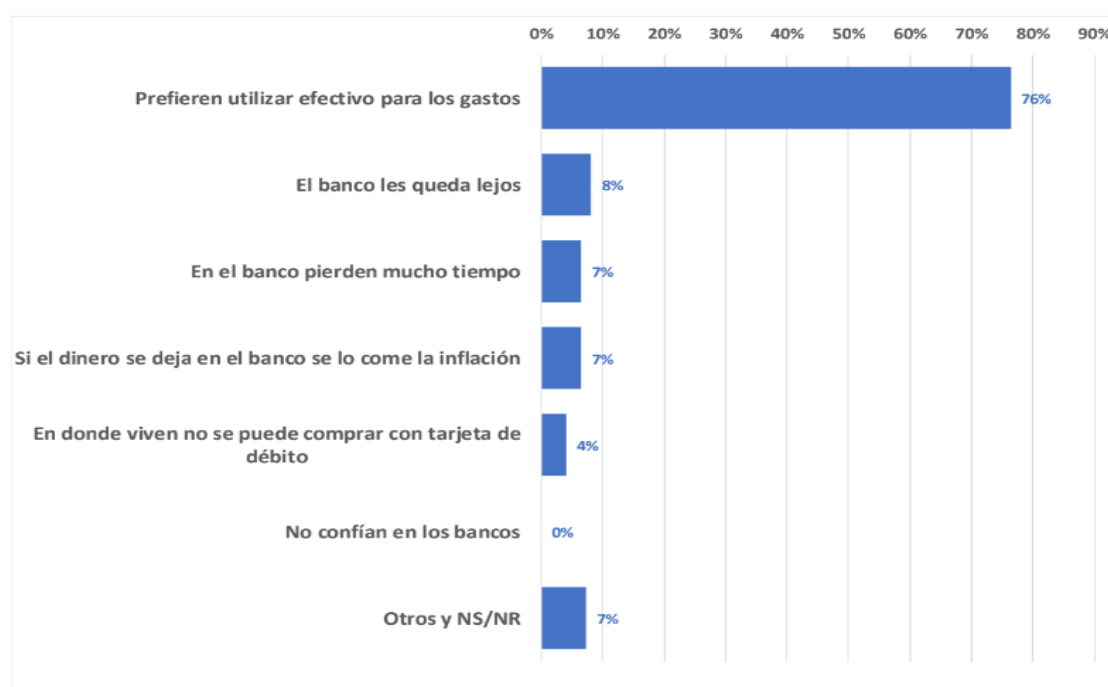
3.1.6 ¿Bancarizados o semi bancarizados?

“Dentro del grupo de quienes podríamos considerar “semi bancarizados”, se les ha consultado acerca de por qué extraen de una sola vez todo el dinero que reciben en su cuenta bancaria. Aquí encontramos que tres de cada cuatro personas han contestado que prefieren utilizar el efectivo para sus pagos (76%). Por su parte, la lejanía de las sucursales bancarias (8%), la pérdida de tiempo en las sucursales (7%), la pérdida de poder adquisitivo de los saldos por la inflación (7%), entre otros, son motivos argumentados para explicar la preferencia por las extracciones de todo el efectivo de una sola vez en el mes.

Es preciso destacar que la desconfianza en los bancos, no ha sido destacada como un factor relevante para la extracción de efectivo de una sola vez.”

En la sección de bancarizados o semi bancarizados indica que se les consultó a aquellas personas dentro del grupo de semi bancarizados acerca de por qué extraen de una sola vez todo el dinero que reciben en su cuenta bancaria. No obstante, no es transparente si las opciones que han indicado eran proporcionadas por el encuestador o si era una pregunta abierta, porque luego se concluye que no es un factor relevante para la extracción de efectivo en una única ocasión la desconfianza en los bancos, ni tampoco es claro cómo se llega a esa conclusión desde esa pregunta con estas opciones.

de una sola vez.



Fuente: BCRA y CAF con base en Ipsos- Argentina (2017).

¿Cómo se podrían abordar estos sesgos?

- Mejorando la manera de ordenar la información en los gráficos
- Aclarando los cálculos realizados
- Haciendo conclusiones que sean particulares de cada grupo de estudio y mostrar cómo se llega a conclusiones generales.

¡Muchas gracias!