

Reporte Técnico: Diferencias en el Consumo de Agua en la Ciudad de México entre Alcaldías

Julio Lezama Amastalli

Daniela Pinto Veizaga

Resumen—Usando fuentes de información pública, la técnica de Aprendizaje de Máquina k vecinos más cercanos (KNN, por sus siglas en inglés) y la construcción de dos indicadores, estudiamos el patrón del consumo de agua en la Ciudad de México (CDMX) y su relación con el Índice de Desarrollo Social (IDS) en dos dimensiones:

1. Diferencia promedio *per cápita* en el consumo de agua por alcaldía, respecto al consumo recomendado por la Organización Mundial de la Salud (OMS). [1].
2. Diferencia promedio *per cápita* en el consumo de agua entre escuelas de nivel preescolar entre alcaldías.

Nuestros hallazgos principales son:

- Hay alcaldías cuyo consumo promedio *per cápita* está por debajo de los 100 litros diarios que sugiere la OMS.
- El consumo de agua está positivamente correlacionado con el Índice de Desarrollo Social (IDS).

Encontramos un fenómeno análogo para las escuelas a nivel preescolar: las escuelas en alcaldías con mayores niveles de IDS presentan mayores promedios *per cápita* en el consumo de agua. Nuestras propuestas van en dos sentidos:

- Es necesario homologar claves, catálogos, preguntas, metodologías, etc. en los procesos de captación de información de la Ciudad de México, por ello se sugiere que se adopten los estándares del Instituto Nacional de Estadística y Geografía.
- Propuestas específicas de política pública que tengan por objetivo: 1) disminuir el costo adicional que enfrentan las personas en alcaldías con IDS bajos (por no recibir la cantidad de agua mínima necesaria recomendada por la OMS) y 2) campañas de cuidado del agua, orientadas a las alcaldías con IDS altos.

Este documento representa un reporte técnico, un reporte a nivel difusión, junto con el código de implementación en Python y R, se encuentra en el siguiente repositorio: https://dapivei.github.io/agua-cdmx/CONSUMO_AGUA_CDMX.html.

I. INTRODUCCIÓN

Al considerar el Objetivo 6 de los Objetivos de Desarrollo Sostenible [2] es: Garantizar la disponibilidad de agua y su gestión sostenible y el saneamiento para todos. Que conforme a [1] el derecho al agua es una condición necesaria para la realización de otros derechos humanos. Que conforme a [3] el porcentaje que representa el agua empleada en usos consuntivos respecto al agua renovable (grado de presión que se ejerce sobre el recurso hídrico) en la Ciudad de México (CDMX) está estimado en 148.2 por ciento para 2030 (muy alto). Y que en general se sabe de la poca disponibilidad de agua en CDMX [4]. Surgen dos preguntas: 1) ¿hay desigualdades en el consumo de agua entre los habitantes de distintas alcaldías de la CDMX?; y, 2) ¿el consumo de agua es equitativo entre personas de distintas alcaldías para ejercer su derecho a la educación?

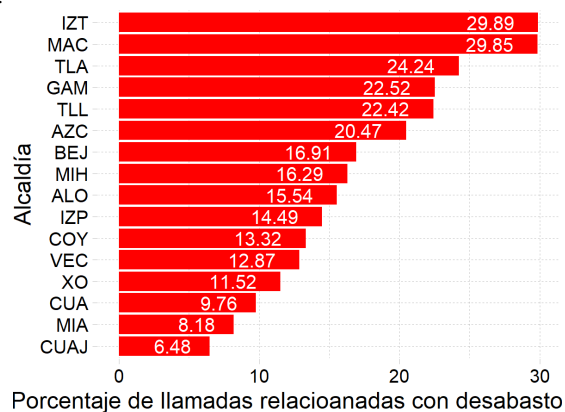
En este trabajo intentamos responder las preguntas anteriores y para ello presentamos la información de la siguiente manera. En la Sección II utilizando registros de llamadas al Sistema Unificado de Atención Ciudadana (SUAC) de la CDMX, relacionadas con desabasto de agua, mostramos que alcaldías como Iztacalco y la Magdalena Contreras casi alcanzan 30 por ciento del total de las llamadas en este rubro. Esto nos hace pensar que puede existir una relación entre niveles de desarrollo social y el acceso al agua. Por ello, en la Sección III explicamos cómo, haciendo uso de varias tablas provenientes de fuentes públicas con información incompleta y con la técnica de Aprendizaje de Máquina k vecinos más cercanos, es posible construir una base de datos con información que estima el consumo de agua a nivel manzana en la CDMX y que identifica a las escuelas. Con base en esta información, en la Sección IV construimos dos indicadores: 1) diferencia entre alcaldías en el promedio *per cápita* en el consumo de agua respecto a los 100 litros recomendados por Organización Mundial de la Salud (OMS) [1]; y, 2) Diferencia en el consumo promedio de agua *per cápita* entre escuelas a nivel preescolar entre alcaldías. En las Secciones V y VII mostramos nuestros resultados. Y finalmente en la Sección VI mostramos las propuestas de Política Pública con base en nuestros hallazgos y respondemos una de las preguntas guía del Datatón tu Dinero tus Datos.

II. JUSTIFICACIÓN

En la Figura 1 se muestra para todas las alcaldías de la CDMX el porcentaje de llamadas al SUAC que tienen que ver con desabasto de agua. Se observa que Iztacalco y la Magdalena Contreras alcanzan casi 30 por ciento del total de sus llamadas en este rubro. Esto indica el gran problema de agua que hay en estas dos alcaldías. Incluso los porcentajes más pequeños alcanzados en Cuajimalpa, Milpa Alta y Cuauhtémoc son altos con registros de 6, 8 y 9 por ciento, respectivamente. Además, si se compara el Índice de Desarrollo Social (IDS) de Iztacalco (0.81972) de consumo bajo de agua contra el de Cuauhtémoc (0.85880) de consumo alto de agua, se observa una posible correlación entre estas dos medidas. Por ello resulta importante conocer la relación entre la diferencia entre alcaldías en el promedio *per cápita* en el consumo de agua respecto a los 100 litros recomendados por Organización Mundial de la Salud (OMS) [1] y el IDS. Por otra parte, en [1] se indica el derecho al agua como una condición previa para ejercer otros derechos como la educación. Es decir, la falta de agua puede obstaculizar el derecho

a la educación. Por ello, resulta importante preguntarse si hay diferencias en el consumo de agua entre escuelas de distintas alcaldías con distintos niveles de IDS.

Figura 1. Porcentaje de llamadas al SUAC relacionadas con desabasto de agua.



III. METODOLOGÍA

A continuación, se describe la metodología empleada. Con el fin de hacer totalmente reproducible este trabajo, se documentan cada uno de los pasos ejecutados y se pone a disposición la implementación en Python y R en el siguiente repositorio https://dapivei.github.io/agua-cdmx/CONSUMO_AGUA_CDMX.html, donde además se muestra una versión de este documento en formato difusión.

III-A. Tablas y clasificaciones consultadas

A continuación, mostramos las fuentes consultadas para la realización de este trabajo y comentamos brevemente algunos de los retos con los que nos encontramos.

- La tabla `consumo-agua.csv` (en adelante `tabla agua`) contiene información del consumo de agua en metros cúbicos por manzana en la CDMX. Encontramos varios retos en esta tabla, como que los metadatos no son suficientes para documentar sus campos (por ejemplo, no es posible identificar claramente las diferencias entre los distintos tipos de consumo de agua). Y que solo contiene información de dos bimestres para aproximadamente 30 por ciento del total de las manzanas en CDMX. Por ello, es necesario imputar el consumo de agua de las manzanas faltantes. Esto se realiza mediante la técnica k vecinos más cercanos. Es importante señalar que la imputación se realiza por la falta de datos completos, de contarse con la `tabla agua` completa, y gracias a la reproducibilidad de este trabajo, solo se tendría que sustituir la tabla y se generarían en automático el resto de los Cuadros y Gráficas empleados.
- Base de datos 911 (en adelante `base 911`) de la Secretaría de Educación Pública (SEP). Provee información para todos los niveles de educación en México tanto público como privada. Contiene datos como número de

estudiantes por plantel, personal disponible, número de salones, dirección del inmueble, entre otros.

El reto que se encontró en esta base es que la dirección de las escuelas no está completa, en algunos casos se hace referencia solo a una calle o colonia.

- Tabla Índice de Desarrollo Social (en adelante `tabla ids`). Contiene el Índice de Desarrollo Social para cada alcaldía en CDMX.

El reto encontrado aquí es la no actualización del índice desde hace algunos años.

- Tabla Directorio Estadístico Nacional de Unidades Económicas (en adelante `DENUE`), aquí se encuentra información relacionada con las Unidades Económicas en México. Podemos encontrar coordenadas, clasificación del sector productivo conforme al Sistema de Clasificación Industrial de América del Norte (SCIAN) entre otros.

Esta base de datos es la más completa que encontramos, ya que estandariza la información con catálogos internacionales. Por ejemplo, las actividades productivas están clasificadas conforme al Sistema de Clasificación de América del Norte (SCIAN), lo que facilita mucho el tratamiento de los datos.

- Datos del Sistema Unificado de Atención Ciudadana (en adelante `tabla SUAC`, contiene las solicitudes de información, dudas, requerimientos, entre otros.

Esta tabla fue adecuada para los propósitos específicos de este trabajo, pero se podría homogeneizar con el resto de las tablas en el sistema de datos abiertos de CDMX.

- Tabla Espacio y Datos de México (en adelante `tabla mapas`). Contiene información geográfica a nivel entidad federativa, municipio, Área Geoestadística Básica (AGEB), localidad y manzana. La información es presentada en polígonos geográficos.

Esta base de datos, así como el `DENUE`, son elaborados por el Instituto Nacional de Geografía y Estadística (INEGI) y son totalmente interoperables. Si los gobiernos locales adoptan las buenas prácticas que sigue el INEGI, el sistema de datos abiertos en México puede alcanzar un potencial sobresaliente a nivel mundial.

III-B. Unión de Tablas

Para poder hacer interoperables las tablas mencionadas fue necesario generar un algoritmo para unir las. La idea general es que en la `tabla agua` se cuenta con coordenadas geográficas y en `tabla mapas`, `DENUE` y `base 911` se cuenta con polígonos geográficos, por lo que la unión considerada está basada en una relación de contención. Es decir, se busca el polígono al que cada manzana de la `tabla agua` pertenece en el resto de las tablas.

- Unión `tabla agua`, `tabla mapas`, `DENUE` y `base 911`. Se tomaron las coordenadas geográficas de la `tabla agua` para cada manzana y se buscó la contención de estas en los polígonos a nivel manzana en `tabla mapas`, `DENUE` y `base 911`.

Como era de esperarse, hay algunas manzanas de la tabla agua que no aparecen en tabla mapas, DENUe y base 911, por lo que una recomendación es actualizar esta tabla. Además, la unión podría no hacerse con la relación de contención mencionada, sino directamente con el identificador de manzana con el que cuentan las tablas de INEGI, esto actualmente aún no es posible debido a que la tabla agua no emplea las claves de INEGI.

III-C. Imputación por k vecinos más cercanos

Ante la falta de datos completos de la tabla agua es necesario imputar los datos faltantes con base en la información existente. Esto se hace mediante el algoritmo de k vecinos más cercanos, en adelante KNN por sus siglas en inglés. La idea general es asignar el consumo de agua, a las manzanas que no cuentan con este dato, el promedio de las 5 manzanas más parecidas a ella en términos de variables geográficas, de población y de acceso a agua. A continuación, establecemos formalmente nuestro problema:

- Datos completos: $\{u_n\}$, con $u_n \in \mathbb{R}^8$, una colección de manzanas caracterizadas por *latitud*, *longitud*, *localidad*, *alcaldía*, *población total*, *número de viviendas habitadas*, *si tiene acceso a agua entubada* y *consumo de agua*.
- Datos a imputar: $\{u_m\}$, con $u_m \in \mathbb{R}^7$, una colección de manzanas caracterizadas por *latitud*, *longitud*, *localidad*, *alcaldía*, *población total*, *número de viviendas habitadas* y *si tiene acceso a agua entubada*.
- Queremos asignar los consumos de agua de las manzanas de la colección de datos a imputar con base en la información de la colección de datos completos.

Para resolver este problema mediante KNN empleamos la distancia euclidiana, dada por $d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}$, con $x, y \in \mathbb{R}^n$. Y mediante la paquetería `scikit-learn` de Python resolvemos el problema de encontrar los $k = 5$ vecinos más cercanos de cada manzana en la colección de datos a imputar, dentro de las manzanas de la colección de datos completos. Y asignamos el consumo de agua como el promedio de consumo de los 5 vecinos más cercanos.

Una vez realizada la imputación, contamos con información completa (estimada) y es posible llevar a cabo la unión con el resto de las tablas. Con toda esta información es posible calcular indicadores muy interesantes, mismos que se muestran a continuación.

IV. INDICADORES

A continuación, enlistamos los indicadores empleados y damos una descripción de lo que miden. En el sitio https://dapivei.github.io/agua-cdmx/CONSUMO_AGUA_CDMX.html se muestran las fichas con el detalle de estos, así como el método de cálculo.

1. Diferencia promedio per cápita en el consumo de agua por alcaldía (diferencia en consumo de agua).
Mide la brecha que existe entre el consumo promedio per cápita por alcaldía en el consumo de agua por día en litros, comparado con los 100 litros recomendados

la OMS [1]. Un resultado negativo indica un déficit en el consumo, un resultado positivo indica un superávit y un resultado igual a cero indica que se consume lo que sugiere la OMS.

2. Consumo de agua per cápita promedio escuelas a nivel maternal y preescolar.

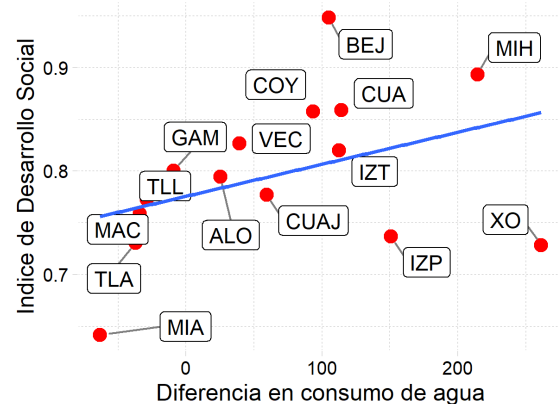
Mide el consumo de agua promedio per cápita por alcaldía en las escuelas a nivel maternal y preescolar. Considera el consumo de agua en las escuelas a nivel maternal y preescolar dividido por la población escolar que incluye estudiantes y empleados.

V. DIFERENCIAS EN CONSUMO DE AGUA PER CÁPITA ENTRE ALCALDÍAS Y SU RELACIÓN CON MEDIDAS SOCIO-ECONÓMICAS

En la Figura 2 se muestra el indicador diferencia en consumo de agua y el IDS por alcaldía. Se observa una correlación positiva entre estas dos medidas. Esta relación es la esperada y genera las siguientes hipótesis:

- En las alcaldías con IDS bajo, las personas consumen, vía la distribución entubada, menos agua que en las alcaldías con índices altos, lo que puede suponer una carga adicional para las personas en estas alcaldías ya que en promedio su acceso a la red de distribución de agua no satisface sus necesidades básicas, por lo que buscarán disminuir su déficit mediante el abastecimiento vía carros tanque (pipas) [5], lo que tiene asociado un costo periódico.
- Por otra parte, en las alcaldías con superávit e IDS altos, el consumo mayor no necesariamente tendría que ver con consumos personales altos o desperdicio. Es posible que se cuente con mayores actividades recreativas relacionadas con consumo de agua, como gimnasios, albercas, deportivos, centros comerciales, etc. Que provocan que el consumo promedio aumente.
- Los dos puntos anteriores, sugieren que además de falta de agua en las alcaldías con niveles bajos de IDS, también puede darse la situación que en las alcaldías con niveles altos de IDS las personas puedan acceder a otro tipo de servicios que aumentan su consumo promedio per cápita.

Figura 2. Relación en la Diferencia en el consumo de agua e IDS



VI. DIFERENCIAS EN CONSUMO DE AGUA PER CÁPITA ENTRE ESCUELAS EN DISTINTAS ALCALDÍAS

En la Figura 3 se muestra el indicador consumo de agua per cápita en escuelas a nivel maternal y preescolar y el IDS por alcaldía. Se observa una correlación positiva entre estas dos medidas, como en el caso de la Sección anterior. De esta manera, confirmamos que las escuelas que las alcaldías con niveles bajos de IDS consumen menos agua que las escuelas en alcaldías con niveles altos de IDS. Esto puede sugerir un problema para la educación, dada su naturaleza de derecho esencial para ejercer otros derechos.

Lamentablemente no es posible contar con calificaciones de los estudiantes por alcaldía, de contar con esos datos, podríamos intentar estimar un modelo en que la variable explicada sea el rendimiento académico y dentro de las variables explicativas se encuentre consumo de agua promedio, IDS y alcaldía. Sería interesante ver la significancia de cada una de estas variables en el rendimiento académico o si hay algún grado de colinealidad entre ellas.

VII. CONCLUSIONES Y RECOMENDACIONES

A continuación, mostramos nuestras conclusiones y recomendaciones con base en los hallazgos de este trabajo.

Nuestras recomendaciones respecto a las bases de datos son:

1. Homogeneizar uso de catálogos, metodologías, formas de preguntar en los instrumentos de captación de información en la CDMX respecto a INEGI.
2. Construir metadatos completos para todas las tablas de la CDMX.
3. Documentar de manera exhaustiva las tablas.
4. Presentar tablas con información completa.
5. Asignar un responsable para responder dudas acerca de las bases de datos, hacer públicos datos de contacto, e implementar un chat como medio de comunicación.
6. Continuar con la promoción del Datatón en futuras ediciones.

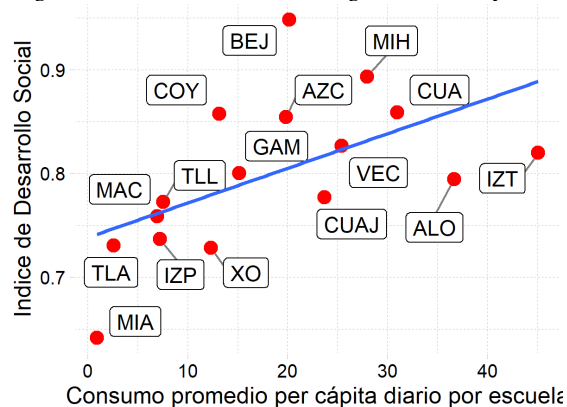
Nuestras recomendaciones de política pública respecto al problema del consumo de agua son:

1. Realizar un diagnóstico de la relación que existe entre el consumo de agua y
 - Rendimiento en todos los niveles de educación en CDMX y diferencias entre alcaldías y escuelas público y privadas y
 - Eficiencia o ineficiencia del uso de agua en los sectores productivos respecto al consumo de agua a nivel mundial en la industria.

Esto con el fin de diagnosticar problemas relacionados con el consumo de agua basados en la gran cantidad de datos que ofrece la CDMX. Y proponer soluciones.

2. Generar campañas de cuidado del agua, sobre todo en alcaldías con IDS alto.
3. Identificar a las industrias intensivas en uso de agua y concientizar a sus trabajadores, directivos y fundadores a cuidarla. Si bien es cierto que en CDMX la mayor parte del consumo se da en hogares, el objetivo de esta

Figura 3. Relación del consumo de agua en escuelas y el IDS



recomendación es facilitar el hábito en todos los rubros de la vida.

4. Atender eficientemente las solicitudes de pipas de agua en todas las alcaldías. Y dar prioridad a las alcaldías con niveles de IDS bajos.
5. Generar apoyos en alcaldías con IDS bajo para que las personas puedan cubrir sus necesidades de agua. De manera inicial mediante baños y lavaderos públicos de calidad y posiblemente de manera monetaria. Pero de manera permanente, asegurar que la red de abastecimiento de agua alcance a todos los habitantes de la CDMX.
6. Hacer consciente a toda la sociedad que el agua es un bien escaso y que es un derecho fundamental que debemos preservar para futuras generaciones.
7. Atender de manera prioritaria las fugas reportadas con independencia de las alcaldías en donde tengan lugar.

Nuestra respuesta a una de las preguntas guía del Datatón tu Dinero tus Datos es la siguiente:

1. A partir del análisis de información ¿Consideras que tu Ciudad es incluyente?
Consideramos que las diferencias en el consumo de agua promedio per cápita entre alcaldías con distintos niveles de IDS puede aumentar la desigualdad ya existente. Por ello, proponemos varias iniciativas de política pública encaminadas a aminorar esta brecha.

REFERENCIAS

- [1] Naciones Unidas (2002), "El derecho al agua (artículos 11 y 12 del Pacto Internacional de Derechos Económicos, Sociales y Culturales)" https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=E%2fC.12%2f2002%2f11&Lang=en.
- [2] Naciones Unidas (2018), "La Agenda 2030 y los Objetivos de Desarrollo Sostenible Una oportunidad para América Latina y el Caribe", disponible en https://repositorio.cepal.org/bitstream/handle/11362/40155/24/S1801141_es.pdf.
- [3] Comisión Nacional del Agua (2019), "Estadísticas del Agua en México 2018", Edición 2018, disponible en https://www.un.org/spanish/waterforlifedecade/human_right_to_water.shtml.
- [4] Instituto de Investigaciones Parlamentarias (2017), "La gestión del agua potable en la Ciudad de México. Los retos hídricos de la CDMX: gobernanza y Sustentabilidad. Primera edición: febrero de 2017".

- [5] Mora Efraín et al (2018), “Comercialización De agua por pipas en el oriente del valle de México”, Revista Mexicana de Ciencias Agrícolas, vol. 9, núm. 3, 2018.
- [6] Hugo Maguey (2018), “Más de 80 % del agua se va en uso agrícola y de la industria”, Gaceta UNAM, <https://www.gaceta.unam.mx/crisis-agua-industria/>
- [7] Instituto Nacional de Estadística y Geografía, “Cuéntame de México” <http://cuentame.inegi.org.mx/territorio/agua/ usos.aspx?tema=T>.
- [8] Instituto Nacional de Estadística y Geografía (2018) “Sistema de Clasificación Industrial de América del Norte, México SCIAN 2018” https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825099695.pdf.
- [9] Secretaría Inclusión y Bienestar Social, “Índice de Desarrollo Social de las Unidades Territoriales del Ciudad de México. Delegación, Colonia y Manzana”. <http://www.sideso.cdmx.gob.mx/index.php?id=551>.
- [10] Naciones Unidas (2014) Decenio Internacional para la Acción El agua fuente de vida 2005 - 2015 https://www.un.org/spanish/waterforlifedecade/human_right_to_water.shtml.

VIII. NOTAS

No olvides acceder a nuestro repositorio, te va a encantar:

https://dapivei.github.io/agua-cdmx/CONSUMO_AGUA_CDMX.html.