

Fundamentos de la Ciencia de Datos

Calidad de agua del Río de La Plata



Francisco Horianzki

Ulises Matteoda

Martino Masson Holzmann

Índice

Índice.....	2
Introducción.....	3
Variables Analizadas.....	4
Variables.....	4
Hipótesis.....	7
¿La interacción humana es un gran causante de la contaminación?.....	7
¿El olor en el río indica mayor contaminación por heces?.....	9
¿Las bacterias de las heces fecales son un factor significativo de la contaminación?.....	11
¿Existe una relación entre el oxígeno o la calidad del agua?.....	13
¿Existe el oxígeno suficiente en el agua?.....	14
¿Las colonias de coliformes impactan en la calidad del agua?.....	15
Conclusión.....	17
Bibliografía.....	18

Introducción

En el año 2021, se realizaron diversas mediciones para evaluar el estado del agua en el Río de la Plata a lo largo de las distintas estaciones del año. Este análisis incluye parámetros fundamentales como la temperatura del agua, la temperatura del aire, el color del agua, la presencia de espuma y otras observaciones relevantes. A través de estos datos, se busca comprender no solo la calidad del agua, sino también los posibles factores que afectan su salud ecológica. Este informe pretende brindar una visión integral del estado del agua en el río, identificando tendencias y áreas de preocupación.



Este mapa muestra la región del Río de la Plata en el área de Buenos Aires y La Plata. Los puntos marcados en azul representan los sitios específicos donde se realizaron evaluaciones de la calidad del agua. Cada uno de estos puntos corresponde a una ubicación de muestreo en la que se tomaron mediciones de diversos parámetros ambientales, como temperatura del agua, pH, niveles de coliformes fecales y otras variables químicas y biológicas, para analizar el estado de la calidad del agua en esta área. Estos sitios están distribuidos a lo largo de la costa, desde la Ciudad Autónoma de Buenos Aires hasta La Plata, lo que permite un monitoreo detallado de cómo se comporta la calidad del agua en distintos puntos del río.

Variables Analizadas

A continuación, se describen las variables consideradas en el análisis de la calidad del agua en el Río de la Plata durante 2021, así como su clasificación

Variables

Variable	Descripción	Tipo
Sitios	Localización específica donde se realizó el muestreo del agua	Cualitativa Nominal
Campaña	estación del año donde fue la medición	Cualitativa Nominal
Tem agua	La temperatura del agua influye en la cantidad de oxígeno que puede disolver el agua, mayor temperatura mas sensibilidad a desechos tóxicos	Cuantitativo Continuo
Tem aire	Mayor temperatura es posible el crecimientos de algas que disminuye el oxígeno y afecta la vida acuática	Cuantitativo Continuo
Od	Oxígeno disuelto, medido en miligramos por litro (mg/L), esencial para la vida acuática.	Cuantitativo Continuo
Ph	Medida de la acidez o alcalinidad del agua, en una escala de 0 a 14	Cuantitativo Continuo
Olores	Presencia de olores en el agua, que puede indicar contaminación	Cualitativa Nominal
Color	Color del agua, que puede ser un indicador de la calidad del agua	Cualitativa Nominal
Espumas	Presencia de espumas en la superficie del agua, que puede ser un signo de contaminación	Cualitativa Nominal
Mat Susp	Materia suspendida, que se refiere a partículas sólidas que flotan en el agua	Cualitativa Nominal

Colif fecales ufc 100ml	Unidades formadoras de colonias de coliformes fecales en 100 ml de agua, un indicador de contaminación fecal.	Cuantitativo Discreto
Escher coli ufc 100ml	Unidades formadoras de colonias de Escherichia coli en 100 ml de agua, otro indicador de contaminación fecal.	Cuantitativo Discreto
enteroc ufc 100ml	Unidades formadoras de colonias de enterococos en 100 ml de agua, que también indican contaminación fecal	Cuantitativo Discreto
nitrato mg l	Concentración de nitratos en miligramos por litro (mg/L), que puede indicar contaminación por fertilizantes	Cuantitativo Continuo
nh4 mg l	Concentración de amonio en miligramos por litro (mg/L), que puede ser un indicador de contaminación orgánica	Cuantitativo Continuo
p total l mg l	Fósforo total en miligramos por litro (mg/L), que incluye todas las formas de fósforo en el agua.	Cuantitativo Continuo
fosf ortofos mg l	Concentración de ortofosfatos en miligramos por litro (mg/L), que es un nutriente importante.	Cuantitativo Continuo
dbo mg l	Demanda biológica de oxígeno en miligramos por litro (mg/L), que mide la cantidad de oxígeno requerido por microorganismos para descomponer materia orgánica.	Cuantitativo Continuo
dqo mg l	Demanda química de oxígeno en miligramos por litro (mg/L), que mide la cantidad total de oxígeno requerido para oxidar materia orgánica e inorgánicas	Cuantitativo Continuo
turbiedad ntu	Turbidez del agua medida en unidades NTU (Nephelometric Turbidity Units), que indica la claridad del agua.	Cuantitativo Continuo
hidr deriv petr ug l	Hidrocarburos derivados del petróleo en microgramos por litro (µg/L), que indican contaminación por productos petroleros.	Cuantitativo Continuo
cr total mg l	Concentración total de cromo en miligramos por litro (mg/L), un metal pesado que puede ser tóxico.	Cuantitativo Continuo

cd total mg l	Concentración total de cadmio en miligramos por litro (mg/L), otro metal pesado que es tóxico en altas concentraciones.	Cuantitativo Continuo
clorofila a ug l	Concentración de clorofila en microgramos por litro (µg/L), que indica la cantidad de fitoplancton en el agua.	Cuantitativo Continuo
microcistina ug l	Concentración de microcistinas en microgramos por litro (µg/L), que son toxinas producidas por ciertas algas	Cuantitativo Continuo
ica	Índice de calidad del agua, que puede ser un valor calculado para evaluar la calidad general del agua.	Cuantitativo Continuo
calidad de agua	Clasificación general de la calidad del agua basada en los parámetros medidos	Cualitativo Ordinal

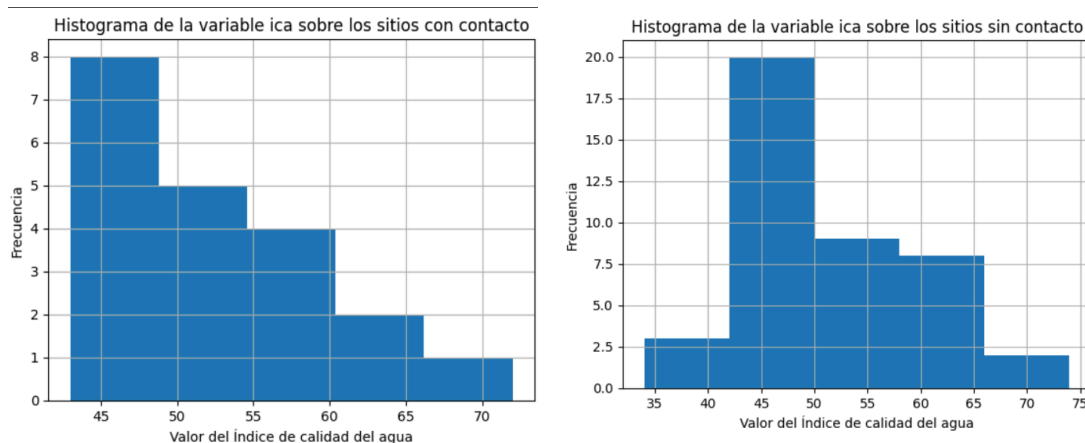
Hipótesis

¿La interacción humana es un gran causante de la contaminación?

El objetivo de este análisis es determinar si la actividad humana tiene un impacto significativo en la calidad del agua en diferentes sitios, diferenciando entre aquellos con y sin contacto humano.

Se recopilaron datos de varios sitios, incluyendo aquellos con alta probabilidad de contacto humano (como playas, balnearios y áreas de pesca) y aquellos sin contacto humano. Posteriormente, se dividieron en dos grupos para comparar la calidad del agua entre ellos.

una vez separados los datos analizamos los datos a partir de histogramas

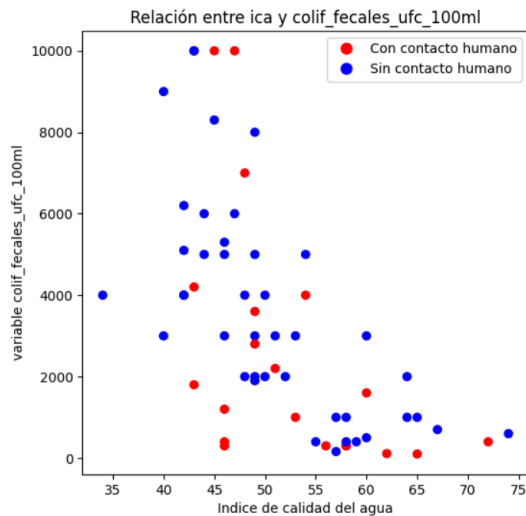


En estos gráficos, se observa que en los sitios con contacto directo, la calidad del agua parece más afectada o deteriorada.

Para evaluar la significancia estadística de esta relación, se realizaron pruebas de normalidad mediante el test de Shapiro-Wilk. Los resultados indicaron un p-valor de 0.024 para los sitios con contacto humano, lo cual sugiere que estos datos no siguen una distribución normal. Para los sitios sin contacto, el p-valor fue de 0.389, indicando que presentan normalidad.

Dado estos resultados, procedimos a verificar la homocedasticidad de los datos con el test de Levene, obteniendo un p-valor de 0., lo que confirma que los datos son homocedásticos. Al determinar que los datos son homocedásticos pero no normales, utilizamos el test de

Mann-Whitney U para evaluar la significancia de la relación. Este análisis arrojó un p-valor de 0.48, indicando que no hay suficiente evidencia para rechazar la hipótesis nula. Por lo tanto, no se observan diferencias significativas entre los sitios con y sin contacto humano.



En la presente imagen, no se observa una diferencia clara entre los puntos rojos (sitios con contacto humano) y los puntos azules (sitios sin contacto humano) en cuanto a la concentración de coliformes fecales. Ambos grupos muestran valores altos de coliformes fecales cuando el ICA es bajo y valores menores de coliformes a medida que el ICA aumenta. Este patrón es coherente con los resultados estadísticos previos, que indicaron que no existen diferencias significativas en la calidad del agua entre los sitios con y sin contacto humano.

Es importante considerar que el estudio puede estar limitado por factores como el tamaño de la muestra y la representatividad de los sitios seleccionados. Además, otros factores ambientales, como la cercanía a fuentes de contaminación industrial o agrícola, podrían estar influyendo en los resultados

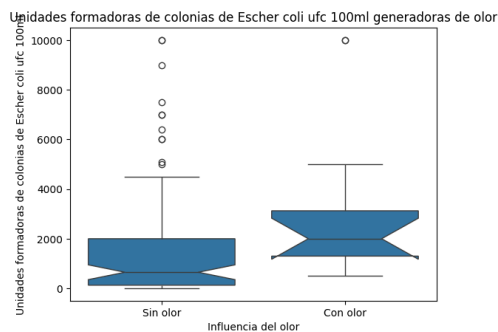
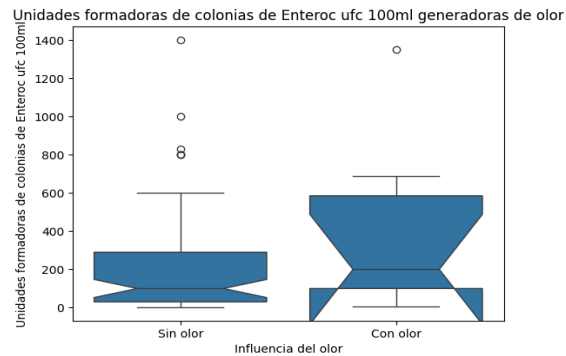
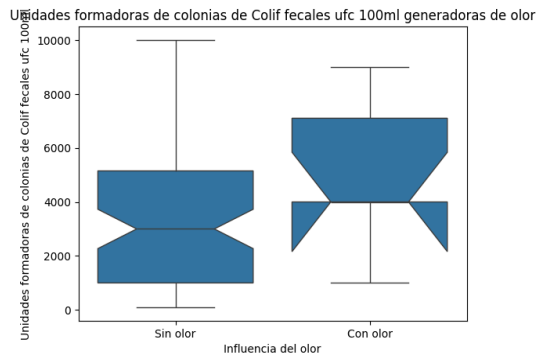
Conclusión

Los resultados sugieren que es necesario realizar un análisis más exhaustivo, considerando variables adicionales y ampliando el número de sitios observados. Asimismo, sería útil llevar a cabo estudios longitudinales para observar cómo evoluciona la calidad del agua en sitios con y sin contacto humano a lo largo del tiempo. Esto podría ayudar a comprender mejor la dinámica entre la actividad humana y la calidad del agua.

¿El olor en el río indica mayor contaminación por heces?

Si hay más olor, entonces la contaminación por heces fecales es mayor. Los olores fuertes indican la presencia de desechos orgánicos, lo que sugiere un mayor nivel de contaminación.

Para comenzar realizamos las siguientes imágenes



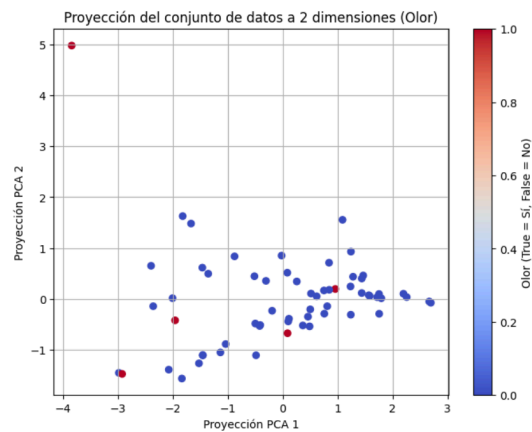
En los gráficos presentados se observan tres variables asociadas a bacterias indicadoras de contaminación fecal: *Escherichia coli*, coliformes fecales y enterococos. A primera vista, en todos los casos se aprecia una mayor cantidad de bacterias en las ubicaciones donde se percibe la presencia de olores, lo cual sugiere una posible relación entre la emisión de olores y el nivel de contaminación bacteriana.

Para evaluar la significancia estadística de esta relación, se realizaron pruebas para determinar si los datos se distribuían normalmente. Debido a que no se ajustan a una distribución normal, se decidió evaluar la homocedasticidad a partir del test de Levene.

En el caso de *Escherichia coli* y coliformes fecales, se aplicó la prueba de Mann-Whitney, ya que estas variables presentaban homocedasticidad. De estos, solo *Escherichia coli* mostró una diferencia estadísticamente significativa, con un p-valor de 0,029, lo cual indica una mayor población de esta bacteria en sitios con presencia de olores.

Para los enterococos, se utilizó la prueba de Kruskal-Wallis debido a que los datos no presentaban homocedasticidad. Los resultados mostraron un p-valor mayor a 0,05, por lo que no se puede

rechazar la hipótesis nula, indicando que no hay una diferencia significativa en la concentración de enterococos entre sitios con y sin olor.



La imagen presenta una proyección de datos en dos dimensiones utilizando Análisis de Componentes Principales (PCA). En el gráfico, los puntos están coloreados según la presencia de olor: los puntos rojos indican "Sí" y los puntos azules indican "No". El eje horizontal corresponde a la primera componente principal (Proyección PCA 1), mientras que el eje vertical representa la segunda componente principal (Proyección PCA 2)

Conclusión

Los resultados sugieren que, aunque hay una tendencia general a encontrar mayores concentraciones bacterianas en áreas con olor, la hipótesis de una asociación significativa entre el olor y la presencia de todas las bacterias indicadoras de contaminación fecal no se confirma con los datos obtenidos. De las tres variables analizadas, solo *Escherichia coli* mostró una diferencia significativa entre los sitios con y sin olor. Así que, en resumen, no se puede validar de manera concluyente la hipótesis, ya que no se observa una diferencia significativa en las tres variables bacterianas evaluadas.

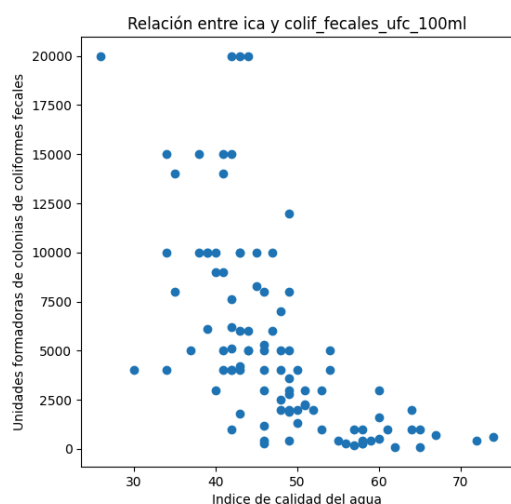
¿Las bacterias de las heces fecales son un factor significativo de la contaminación?

La contaminación por heces fecales es un factor significativo en el deterioro de la calidad del agua. Los microbios presentes en estos desechos pueden causar diversos problemas de salud, como diarrea, cólicos, náuseas, dolores de cabeza, entre otros síntomas. Esta contaminación representa un grave peligro para la salud de bebés, niños pequeños y personas con sistemas inmunológicos debilitado

Análisis de la Relación entre la Contaminación Bacteriana y el Índice de Calidad del Agua (ICA)

En el presente análisis, se observa la relación entre la contaminación fecal, medida en términos de Unidades Formadoras de Colonias de coliformes fecales (UFC) y enterococos (UFE), y el Índice de Calidad del Agua (ICA). Se presentan dos gráficos de dispersión (scatter-plots) que muestran estas relaciones, permitiéndonos evaluar el impacto de cada tipo de bacteria en la calidad del agua.

Figura 1: Relación entre UFC de Coliformes Fecales y el ICA

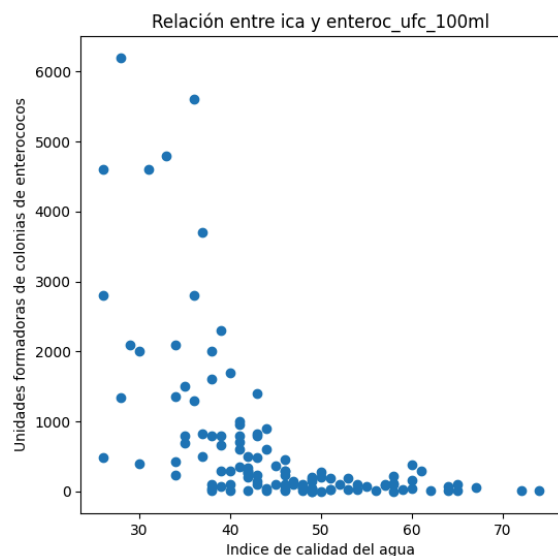


En el primer gráfico, se analiza la relación entre las UFC de coliformes fecales y el ICA. Los datos muestran una tendencia inversa clara: a medida que aumenta la cantidad de coliformes fecales, el ICA disminuye. Esto implica que una mayor presencia de coliformes fecales está asociada a una calidad de agua más baja, lo cual respalda la hipótesis planteada. Los coliformes fecales son indicadores directos de contaminación orgánica y su presencia en altos niveles sugiere un deterioro significativo en la calidad del agua.

Para validar la hipótesis, se realizó el test de Kruskal-Wallis para comparar dos grupos: uno correspondiente a aquellos con un Índice de Calidad del Agua (ICA) menor o igual a 50, y otro a aquellos con un ICA superior a 50. El análisis reveló un valor p de 0,000, lo que indica

una diferencia estadísticamente significativa entre ambos grupos en relación con la cantidad de unidades formadoras de coliformes fecales.

Figura 2: Relación entre UFE de Enterococos y el ICA



El segundo gráfico explora la relación entre las Unidades Formadoras de Colonias de Enterococos y el ICA. Aunque la tendencia inversa entre UFE y el ICA no es tan marcada como en el caso de los coliformes fecales, se puede observar que los sitios con ICA más bajo tienden a tener mayores niveles de enterococos. Los enterococos son otro indicador de contaminación fecal, y su presencia en altos valores también sugiere un impacto negativo en la calidad del agua.

En el caso de esta variable, también se realizó el test de Kruskal-Wallis para comparar los mismos grupos mencionados anteriormente, basados en los valores del Índice de Calidad del Agua (ICA), pero en relación con las unidades formadoras de enterococos. En este análisis, el p-valor obtenido fue de 0,002, lo que también indica una diferencia estadísticamente significativa entre los grupos comparados.

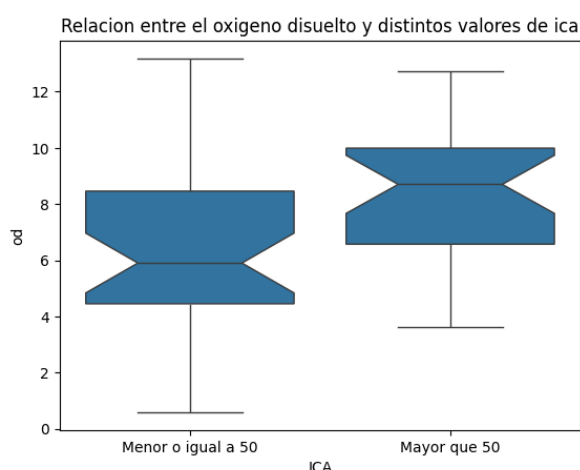
Conclusión

Los resultados de estos análisis sugieren que la contaminación bacteriana tiene un impacto considerable en la calidad del agua. La correlación negativa entre el ICA y los niveles de coliformes fecales y enterococos reafirma la importancia de controlar estos indicadores bacterianos para mejorar y preservar la calidad de los ecosistemas acuáticos. Mantener bajo control la carga de coliformes fecales y enterococos es una medida esencial para proteger la salud pública y asegurar la sostenibilidad de los recursos hídricos.

Estos hallazgos subrayan la necesidad de políticas de control y monitoreo regular de contaminantes fecales como parte de la gestión integral de los cuerpos de agua.

¿Existe una relación entre el oxígeno o la calidad del agua?

Existe una clara relación entre los niveles de oxígeno disuelto en el agua y el Índice de Calidad del Agua (ICA). A medida que aumentan los niveles de oxígeno disuelto, también se observa un incremento en el ICA.



El gráfico muestra una tendencia evidente: las ubicaciones con un ICA superior a 50 presentan, en general, niveles más altos de oxígeno disuelto (OD). Para confirmar esta observación, realizamos una prueba, dado que los datos cumplen con los supuestos de normalidad y homocedasticidad. Los resultados arrojaron un valor p de 0.007, lo que nos permite rechazar la hipótesis nula y afirmar que existe una diferencia significativa

en los niveles de OD entre sitios con ICA alto y aquellos con ICA más bajo.

Estos hallazgos destacan lo crucial que es el oxígeno disuelto (OD) para combatir la contaminación, ya que ayuda a mejorar la calidad del agua. Por lo tanto, es esencial que las entidades responsables de la gestión de recursos hídricos implementen estrategias para aumentar los niveles de OD. Esto podría tener un impacto muy positivo en la salud de los ecosistemas acuáticos y, en consecuencia, mejorar significativamente la calidad del agua.

¿Existe el oxígeno suficiente en el agua?

La Demanda Biológica de Oxígeno (DBO) representa la cantidad de oxígeno que los microorganismos necesitan para descomponer la materia orgánica en el agua. Cuando la DBO es alta, la competencia por el oxígeno aumenta, lo que reduce los niveles de oxígeno disuelto (OD) disponibles para otros organismos. Estudios en gestión de aguas residuales y calidad del agua, como los realizados por la Agencia de Protección Ambiental (EPA), muestran que una DBO elevada es un indicador clave de contaminación orgánica. Esto significa que en cuerpos de agua con alta DBO, los niveles de oxígeno suelen ser bajos, lo que compromete la salud de los ecosistemas acuáticos.

Para mejorar esta situación, es esencial encontrar formas de reducir esta demanda de oxígeno o incrementar el oxígeno disuelto en el agua, de modo que los procesos naturales de descomposición puedan ocurrir sin agotar el oxígeno necesario para la vida acuática.

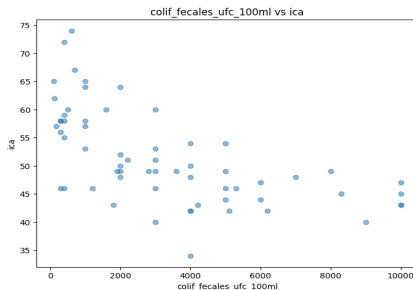
Las plantas acuáticas, a través de la fotosíntesis, son una fuente natural de oxígeno. La clorofila-a, un pigmento clave en la fotosíntesis, indica la cantidad de fitoplancton y algas en el agua. A niveles controlados, el fitoplancton es beneficioso porque produce oxígeno, lo que contribuye a mantener niveles saludables de OD. Sin embargo, un crecimiento excesivo de algas puede resultar contraproducente, ya que su posterior descomposición aumenta la DBO y consume más oxígeno.

Conclusión

Si logramos un equilibrio donde la carga orgánica en el agua esté controlada y los niveles de fitoplancton y algas se mantengan en rangos moderados, se podría aprovechar la fotosíntesis de estas plantas para aportar oxígeno de manera natural. Una baja DBO aseguraría que este oxígeno esté disponible para toda la vida acuática en el río. De esta manera, no solo se mantendrían los niveles de oxígeno disuelto necesarios para los organismos acuáticos, sino que también se mejoraría la calidad del agua, promoviendo un ambiente más limpio y saludable en el Río de la Plata.

¿Las colonias de coliformes impactan en la calidad del agua?

Las unidades formadoras de colonias de coliformes fecales (colif_fecales_ufc_100ml) son un indicador comúnmente utilizado para evaluar la calidad del agua. Se sospecha que las altas concentraciones de coliformes fecales pueden deteriorar significativamente la calidad del agua, afectando tanto a los ecosistemas acuáticos como a la salud pública.



En esta imagen se puede notar claramente que a medida que aumentan las unidades formadoras de colonias de coliformes fecales, la calidad del agua disminuye.

Validación del Modelo de Regresión Lineal

Para asegurar la validez del modelo de regresión lineal aplicado, se evaluó la normalidad de los residuos a través de dos pruebas estadísticas: el test de Omnibus y el test de Jarque-Bera. Los resultados obtenidos son los siguientes:

- **Test de Omnibus:** p-valor = 0.597.
- **Test de Jarque-Bera:** p-valor = 0.777.

Dado que ambos p-valores superan el nivel de significancia común (0.05), no existe evidencia suficiente para rechazar la hipótesis nula de normalidad de los residuos, lo que indica que estos se distribuyen normalmente.

Conclusión

Los resultados del análisis de regresión lineal sugieren que las colonias de coliformes fecales tienen un impacto significativo en la calidad del agua. Se observó que, al incrementarse la concentración de coliformes fecales, el Índice de Calidad del Agua (ICA) disminuye, lo cual sugiere un deterioro de la calidad del agua.

Conclusión

El análisis de la calidad del agua en el Río de la Plata evidencia la influencia de múltiples factores ambientales y de actividad humana en la salud del ecosistema acuático. Las mediciones de coliformes fecales, enterococos y oxígeno disuelto sugieren que la contaminación orgánica y bacteriana impactan significativamente en la calidad del agua, reduciendo el Índice de Calidad del Agua (ICA) y amenazando tanto el entorno natural como la salud pública. La actividad humana cercana a las zonas de muestreo, aunque no mostró una diferencia estadística concluyente entre áreas con y sin contacto humano, resalta la necesidad de monitoreo constante para identificar fuentes específicas de contaminación.

Es importante destacar que la base de datos utilizada para este informe tiene limitaciones significativas, ya que el número de muestras obtenidas fue reducido. Esto restringe la solidez de las conclusiones y dificulta una interpretación integral de los resultados. Para fortalecer futuros estudios, sería esencial ampliar el tamaño de la muestra y diversificar los sitios de muestreo, permitiendo así obtener un panorama más representativo de la calidad del agua en distintas condiciones y estaciones del año.

Además, el estudio subraya la importancia de mantener niveles adecuados de oxígeno disuelto y controlar los niveles de bacterias fecales, ya que estos indicadores reflejan tanto la carga orgánica como la capacidad del ecosistema para sostener vida acuática saludable. La implementación de políticas de control y reducción de contaminantes, así como el uso de medidas preventivas en áreas de alta actividad humana, son esenciales para preservar la calidad del agua y la sostenibilidad de este recurso hídrico vital.

En conjunto, los hallazgos recalcan la urgencia de una gestión integral de la calidad del agua, promoviendo acciones que disminuyan la carga de contaminantes y fomenten prácticas sostenibles, garantizando así la conservación del Río de la Plata y su biodiversidad.

Bibliografía

- ¿Qué son los coliformes fecales?
<https://www.microlabindustrial.com/parametros/patogenos/182/coliformes-fecales>
- ¿Qué te pueden hacer?.
<https://www.deq.nc.gov/tier1-spanishfecalecolipdf/open#:~:text=Coliformes%20fecales%20o%20E.de%20cabeza%20u%20otros%20s%C3%ADntomas.>
- ¿Cómo ayuda el oxígeno a la contaminación?
[Oxígeno Disuelto en Aguas Residuales: Importancia y Métodos de Medición en el Tratamiento de Aguas | Instituto del Agua](#)
- dataset del río de la plata
<https://www.fororiodelaplata.com.ar/la-contaminacion-en-el-rio-de-la-plata-continua-en-aumento/#:~:text=Esta%20es%20la%20verdad%20asumida,en%20armon%C3%ADa%20con%20la%20naturaleza.>