Propuesta de Proyecto

Seminario de Analítica y Ciencia de Datos

# Presentación del Proyecto

Según un estudio de la cámara de comercio del oriente antioqueño del 2023, está región está experimentando un crecimiento exponencial de la población, lo que puede estar provocado un aumento significativo en los agentes contaminantes atmosféricos de esta subregión. Sumado a sus condiciones climáticas, esto podría derivar en escenarios críticos en la calidad del aire. Analizar estos datos resulta fundamental para apoyar a las autoridades ambientales en su preparación y respuesta ante estos fenómenos, brindándoles herramientas tecnológicas que faciliten la implementación de mecanismos de acción y prevención eficaces

# Objetivo del Proyecto

Desarrollar herramientas de análisis de datos que permitan a Autoridades Ambientales y/o entes que les compete, identificar patrones en las variables atmosféricas relacionadas con la calidad del aire, contribuyendo de manera efectiva a la prevención de problemas de salud y la mitigación de los efectos negativos de los contaminantes en el Oriente Antioqueño, apoyando la toma de decisiones y la implementación de estrategias de control ambiental.

# Contexto del Problema

Según la Organización Mundial de la Salud (OMS), la contaminación atmosférica se cobra trece vidas en el mundo cada minuto (OMS, 2022). Más del 32 % de las muertes se deben a la contaminación atmosférica, que afecta negativamente al bienestar general, al clima y a la economía (Wu et al., 2021). La Agencia de Protección del Medio Ambiente de Estados Unidos (EPA) define la contaminación atmosférica como sustancias, agentes físicos o regulares que degradan el aire y dañan sus propiedades (EPA, 2023). Entre los contaminantes atmosféricos criterio se encuentran el Material Particulado (MP), el Ozono (O3) troposférico, el dióxido de nitrógeno (NO2) y el dióxido de azufre (SO2) proceden principalmente de actividades humanas y procesos naturales (OMS, 2021).

Los factores meteorológicos como la temperatura, la precipitación o la humedad, tienen un gran efecto el cual puede ser negativo o positivo, en las proporciones de los contaminantes presentes en el aire. Los cambios en las condiciones meteorológicas pueden amplificar el impacto del cambio climático al influir en la formación, el transporte y la dispersión de los contaminantes (Ramanathan, 2020). Un estudio anterior ha determinado que los casos de olas de calor e inversiones de temperatura contribuyeron a un aumento de las concentraciones de O3 en un 28 % y un 14 %, respectivamente, sobre la región (Joshi et al., 2020). Un estudio realizado en China reveló que un aumento de 10 μg/m3 en el O3 provocaba un incremento del 0,55 % en la mortalidad asociada a enfermedades respiratorias (Y. Zhang et al., 2021).

La calidad y la contaminación ha sido un problema cada vez más preocupante debido a los efectos pueden variar desde enfermedades respiratorias hasta problemas cardiacos y derrames cerebrales, en donde la población vulnerable, como niños y personas de tercera edad, son los más propensos a sufrir las consecuencias de una mala calidad el aire en la zona donde viven. (OMS,2021). El registro, estudio y análisis de la información nos permite ser capaces de tomar acciones en el presenten y ser proyectado.

En Colombia mediante la Resolución 2254 de 2017 expedida por el Ministerio de Medio ambiente, estableció los niveles máximos permisibles de contaminantes criterio mencionados anteriormente, estos y datos meteorológicos son monitoreados por las Autoridades Ambientales (Corporaciones Autónomas Regionales, Corporaciones de Desarrollo Sostenible y Autoridades Ambientales Urbanas) mediante Sistemas de Vigilancia de la Calidad del Aire – SVCA, diseñados y operados conforme a los criterios que el Ministerio de Ambiente y Desarrollo Sostenible ha definido. La información de los contaminantes atmosféricos es reportada por las Corporaciones Autónomas Regionales y las Autoridades Ambientales de los grandes centros urbanos en el Subsistema de Información sobre Calidad del Aire – SISAIRE (Minambiente, 2021).

El Oriente Antioqueño es una de las 9 subregiones que componen el departamento de Antioquia. La subregión aportó en el 2022 el 10.4% del PIB de Antioquia, siendo la segunda subregión de mayor importancia económica en el departamento después del Valle de Aburrá. Según la Cámara de Comercio del Oriente Antioqueño, el producto interno bruto (PIB) de la subregión se duplicó en los últimos 10 años, pasando de COP$5.75 billones en el 2009 a COP$11.06 billones en el 2018. A medida que las capacidades territoriales del área metropolitana del Valle de Aburrá se agotan, las empresas y los habitantes han buscado nuevas posibilidades de expansión en el Oriente Antioqueño. Lo anterior ha llevado a un crecimiento del 59,5% en la cantidad total de empresas ubicadas en la subregión del 2015 al 2022, y un crecimiento de la población del 11% en estos mismos años (Cámara de Comercio, 2023). Este crecimiento exponencial llevó a que en el 2021 el oriente contara con una población de 706.477 habitantes, que representan el 10.4% de Antioquia y el 1.4% de Colombia (Departamento Administrativo Nacional de Estadística -en adelante DANE-, censo 2018 – actualizado 2019).

En el Oriente Antioqueño la Corporación Autónoma Regional de los Rios Negro y Nare CORNARE, es la encargada de realizar el monitoreo de las concentraciones de contaminantes y reportarlas al Subsistema de Información sobre Calidad del Aire – SISAIRE, el crecimiento exponencial de industria en esta zona del país ha sido un precursor del incremento de concentraciones de contaminantes criterio por fuentes antrópicas, lo que puede llegar a alterar de manera irreversible la calidad del aire en la región, para controlar eficazmente la contaminación atmosférica, es aún más urgente revelar la relación entre la contaminación atmosférica y las condiciones meteorológicas a partir de datos diarios a largo plazo o en tiempo real (Liu, Y., Zhou, Y. & Lu, J. , 2020).

# Propuesta de Solución y Alcance

Para la solución del negocio se propone utilizar lenguajes de programación como Python para desarrollar modelos estadísticos que analicen múltiples variables que influyen en la calidad del aire, como los contaminantes criterio: Material Particulado (PM10 y PM2.5), Dióxido de Azufre (SO2), Dióxido de Nitrógeno (NO2), Ozono troposférico (O3) y Monóxido de Carbono (CO), además variables meteorológicas como la humedad, temperatura, velocidad del viento, entre otras. El dataset será obtenido del reporte de Corporación Autónoma Regional de los Rios Negro y Nare CORNARE en la plataforma de Información sobre Calidad del Aire – SISAIRE.

Se aplicarán técnicas de machine learning para identificar patrones ocultos, analizar la relación entre las variables, como niveles de contaminación y condiciones meteorológicas, el análisis incluirá la correlación entre estas variables como la concentración de contaminantes, temperatura, velocidad del viento y otros factores que puedan influir en la calidad del aire mencionados anteriormente. Esta propuesta no solo busca proporcionar una solución técnica viable, sino también una herramienta de gestión que apoye en la toma de decisiones estratégicas para mejorar la calidad del aire en función de datos precisos y análisis científicos rigurosos

# Definición del Alcance

El proyecto abarca el desarrollo de herramientas de ciencias de datos, aplicándolas en datos atmosféricos y meteorológicos recolectados en el Oriente Antioqueño, específicamente en variables clave como Material Particulado (PM10 y PM2.5), Dióxido de Azufre (SO2), Dióxido de Nitrógeno (NO2), Ozono troposférico (O3) y Monóxido de Carbono (CO). Estas herramientas permitirán identificar patrones de contaminación del aire, evaluar su relación con las condiciones climáticas y determinar posibles escenarios críticos. Además, el proyecto proporcionará a las autoridades ambientales y demás gremios que les competa, herramientas de análisis que facilite la toma de decisiones oportunas, contribuyendo a la creación de mecanismos de prevención y control para mitigar los efectos negativos en la salud pública y el medio ambiente. Las soluciones propuestas se enfocarán en la automatización del análisis de datos y la generación de alertas ante niveles de contaminación elevados, con el objetivo de mejorar la gestión ambiental en la región.

# Riesgos e Impacto del Negocio

Al realizar este análisis de correlaciones, se podrán identificar patrones entre estos factores, como, por ejemplo, cómo la velocidad del viento puede influir en la dispersión de contaminantes o cómo las condiciones de alta temperatura y baja humedad pueden exacerbar la formación de ozono troposférico. El uso de técnicas estadísticas, como la regresión múltiple, permitirá cuantificar la magnitud de estas relaciones, proporcionando información valiosa para predecir la calidad del aire bajo diferentes condiciones. Un caso exitoso que podemos utilizar de referencia para nuestro negocio es en la Universidad Europea De Madrid, allí se realizó un artículo donde su objeto de estudio fue la Predicción De La Contaminación Del Aire En Madrid, en este se determinó la variable más influyente de su modelo de predicción, la velocidad del aire. Se concluyo que este dato meteorológico tiene una gran influencia en la contaminación del aire y esto se debe a que ciertos contaminantes, como es el caso del ozono, el cual se ve beneficiado (a aumentar) cuando existe una disminución de la velocidad del viento, ya que dificulta la dispersión del contaminante en la atmosfera (Iannelli, A. P., 2022).

Respecto a riesgos para desarrollar el proyecto de aprendizaje, se encuentra que:

* El conjunto de datos disponible puede que no suficiente para el entrenamiento de los algoritmos.
* El conjunto de datos disponible tenga muchos datos faltantes que impida la utilización de esto]