PROYECTO DE PROGRAMACIÓN

ACCIDENTES DE TRÁNSITO EN CHICAGO

[SAMUEL PATINO LUCUMI](mailto:samuel.patino@uao.edu.co)

INGENIERÍA DE DATOS E INTELIGENCIA ARTIFICIAL

[JOSE LUIS PANIAGUA JARAMILLO](mailto:jose_l.paniagua@uao.edu.co)

UNIVERSIDAD AUTÓNOMA DE OCCIDENTE.

2023

**Introducción y Contexto de información**

**Link del DataSet:** [**Chicago Car Crash Dataset (kaggle.com)**](https://www.kaggle.com/datasets/nathaniellybrand/chicago-car-crash-dataset)

Estos datos, publicados en Data.gov, indican que es el Hogar de los Datos Abiertos del Gobierno de los Estados Unidos. Aquí encontré datos, herramientas y recursos para llevar a cabo investigaciones, desarrollar aplicaciones web y móviles, diseñar visualizaciones de datos y mucho más.

Este Dataset contiene todos los datos de accidentes automovilísticos desde 2001 en adelante que ocurrieron en Chicago. El conjunto de datos incluye la fecha del accidente, el límite de velocidad, las condiciones climáticas, las lesiones, la latitud y longitud, y mucha más información sobre los accidentes en la ciudad.

Elegí este conjunto de datos porque pude ver que era un conjunto de datos muy grande que contenía mucha información, con más de 700,000 filas y 42 columnas. Estos datos que contiene este conjunto son desde 2001 hasta junio de 2023 de todos los accidentes reportados en la ciudad de Chicago, que es una de las ciudades más grandes de los Estados Unidos.

Por esta razón, también decidí trabajar este Dataset, ya que era un conjunto de datos muy grande en donde puedo extraer mucha información. Podré decidir cómo abordar el conjunto de datos porque contiene mucha información y puedo realizar diferentes tipos de transformaciones con él, además de que contiene algunos valores nulos y datos no entregados.

**Antecedentes**

Data.gov: Hogar de los Datos Abiertos

Data.gov es una plataforma del Gobierno de los Estados Unidos que alberga una amplia variedad de conjuntos de datos públicos. El conjunto de datos seleccionado para este proyecto es una muestra de la disponibilidad de información valiosa que se puede utilizar para investigaciones y análisis. En particular, estos datos sobre accidentes automovilísticos en Chicago son una fuente confiable y detallada que ofrece información relevante para abordar problemas relacionados con la seguridad vial.

**Objetivos**

Objetivo General

El objetivo general de este proyecto es realizar un análisis exhaustivo de los datos de accidentes automovilísticos en Chicago desde 2001 hasta junio de 2023, con el fin de identificar patrones, tendencias y factores de riesgo asociados a estos incidentes, con miras a mejorar la seguridad vial en la ciudad.

Objetivos Específicos

Realizar una limpieza y preparación de los datos, abordando valores nulos y datos no entregados para garantizar la calidad de la información.

Identificar las zonas de la ciudad con mayor incidencia de accidentes automovilísticos y analizar las posibles causas.

Analizar la relación entre las condiciones climáticas y la frecuencia de accidentes, así como su impacto en la seguridad vial.

Evaluar la influencia de la velocidad límite en la gravedad de los accidentes y proponer recomendaciones para mejorar la seguridad en las carreteras.

Diseñar visualizaciones interactivas de datos para difundir información sobre la seguridad vial en Chicago de manera accesible para la comunidad.

**Tipos de Análisis de Datos**

Con estos datos deseo principalmente poder analizar de diferentes maneras los accidentes que ocurren en Chicago Estados Unidos, en donde muchas variables influyen ya que un accidente puede ocurrir de distintas maneras y por diferentes factores, ya sea tiempo, clima, velocidad, hora del día, etc.

Para abordar los objetivos específicos mencionados anteriormente, se utilizarán diversas técnicas de analítica de datos, que incluyen:

-Análisis Descriptivo: Se realizará una exploración inicial de los datos para comprender su estructura y características principales.

-Análisis Espacial: Se utilizarán técnicas de análisis espacial para identificar áreas geográficas con alta concentración de accidentes.

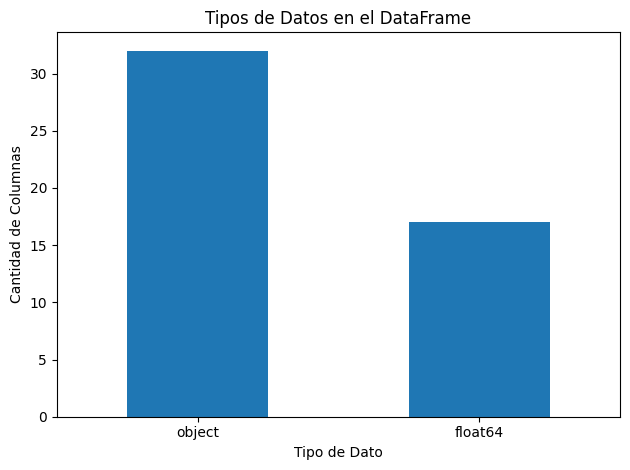
-Análisis de Series Temporales: Se analizarán las tendencias temporales de los accidentes a lo largo de los años.

-Modelos de Regresión: Se desarrollarán modelos de regresión para investigar las relaciones entre variables como la velocidad límite, las condiciones climáticas y la gravedad de los accidentes.

-Visualización de Datos: Se diseñarán visualizaciones interactivas para presentar los resultados de manera clara y accesible.

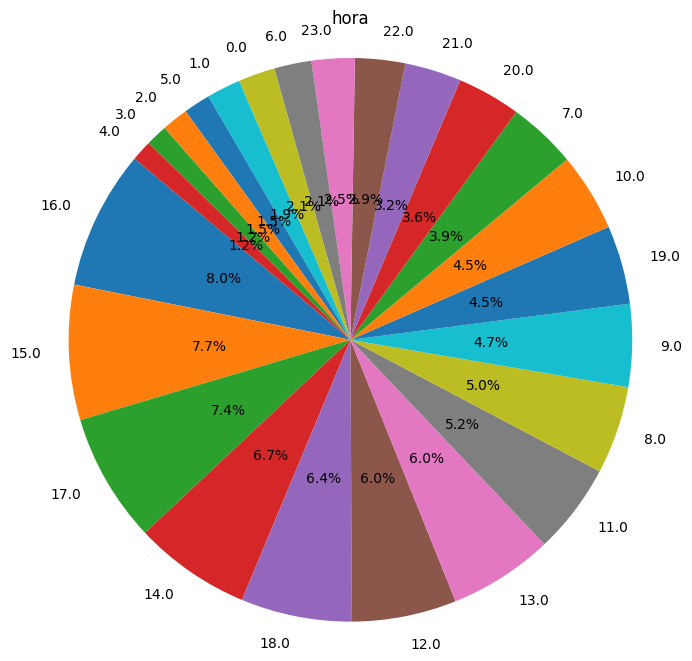
GRÁFICOS

-Este gráfico nos ayuda a entender de forma visual cuáles son los tipos de datos que se presentan con mayor frecuencia.

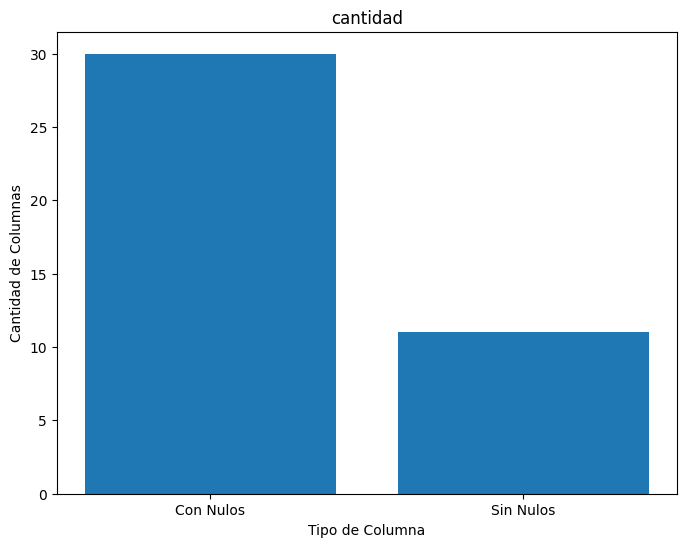


El gráfico muestra de forma visual el porcentaje de accidentes según el tiempo de

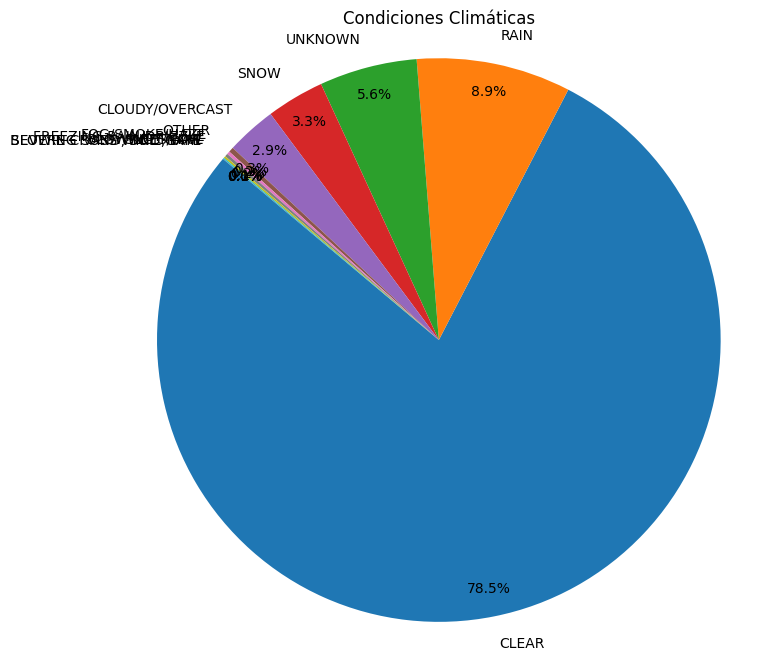
día



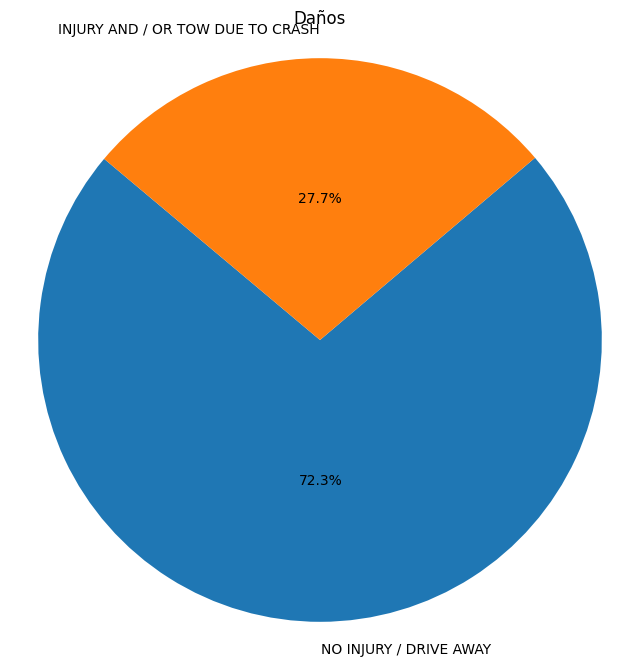
En esta gráfica podemos ver los valores nulos y los que no son nulos



Este gráfico fue tomado de la columna de condiciones climáticas para entender bajo qué condiciones ocurren los accidentes.

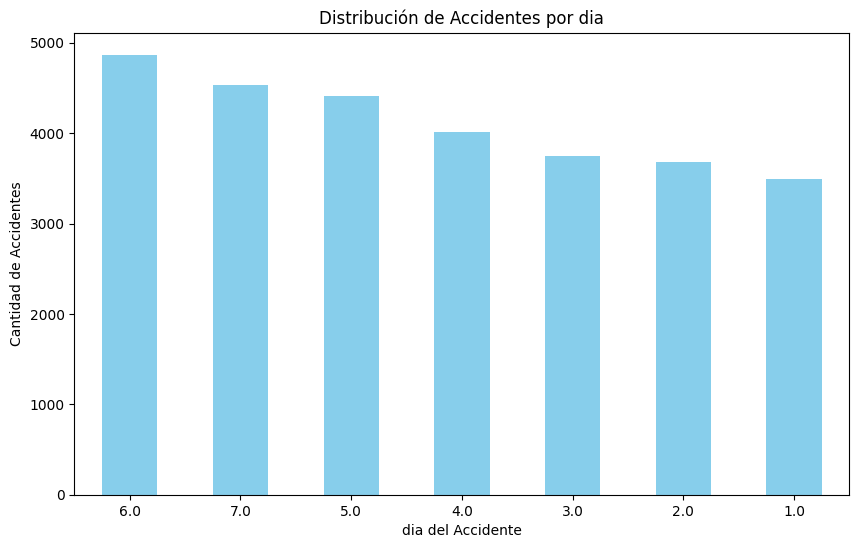


Esta parte nos muestra cuantos sufrieron lesiones y cuáles no, teniendo en cuenta que es un valor preliminar



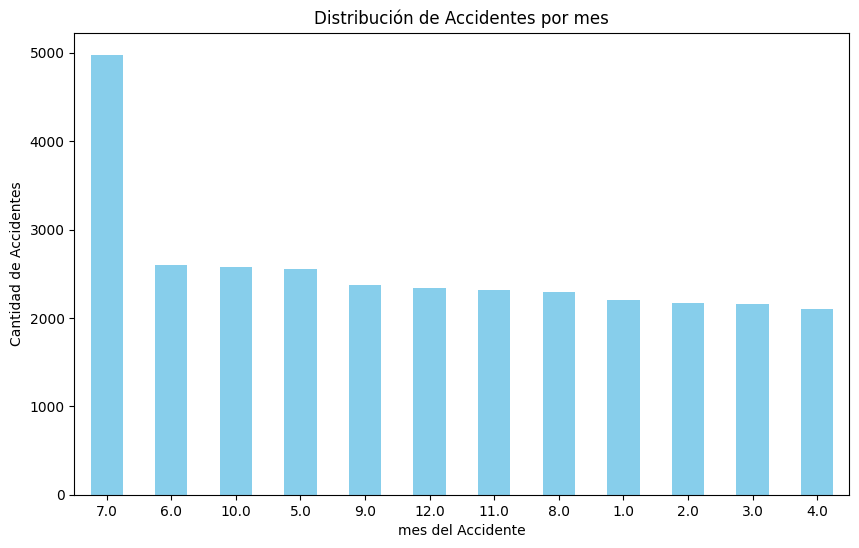
Este gráfico nos ayuda a entender de forma visual en qué día hay más

accidentes.



Este gráfico nos ayuda a entender de forma visual en qué mes hay más

accidentes



DASHBOARD

