**Resumen del Proyecto de Análisis y Predicción del Impacto del Consumo de Sustancias en Argentina**

1-LA PROBLEMÁTICA

Este proyecto busca aplicar técnicas avanzadas de machine learning para analizar los datos de la encuesta realizada por el INDEC sobre el consumo de sustancias como tabaco, alcohol, medicamentos, marihuana y cocaína en Argentina. El objetivo es detectar patrones ocultos, identificar posibles problemas y predecir efectos y soluciones futuras.

2- QUE INCOGNITAS NOS PRESENTA ESTA PROBLEMÁTICA ,QUE PODEMOS OBTENER COMO RESULTADOS

**¿Cuáles son los patrones de consumo?**

* + ¿Qué sustancias se consumen más y por qué grupos demográficos?
  + ¿Hay diferencias significativas en el consumo por edad, género, ubicación geográfica o nivel socioeconómico?

**¿Cómo ha evolucionado el consumo de sustancias a lo largo del tiempo?**

* + ¿Hay tendencias crecientes o decrecientes en el consumo de ciertas sustancias?
  + ¿Cómo han influido eventos recientes o políticas públicas en estos patrones?

**¿Cuáles son los factores de riesgo asociados con el consumo de sustancias?**

* + ¿Qué factores personales, sociales o económicos están correlacionados con un mayor consumo?
  + ¿Existen patrones de consumo múltiple o policonsumo?

**¿Qué impacto tiene el consumo de sustancias en la salud y la sociedad?**

* + ¿Qué problemas de salud están asociados con el consumo de diferentes sustancias?
  + ¿Cómo afecta el consumo de sustancias a la calidad de vida, la productividad laboral y las relaciones sociales?

**¿Qué intervenciones podrían ser efectivas para mitigar los problemas identificados?**

* + ¿Qué estrategias de prevención y tratamiento han sido efectivas en contextos similares?
  + ¿Cómo podrían los hallazgos informar políticas públicas o programas de salud?

**3 -Objetivos del Proyecto:**

**Identificar Tendencias y Patrones:**

* + Analizar los datos para descubrir tendencias clave y patrones de consumo de sustancias.

**Evaluar Factores de Riesgo y Causalidad:**

* + Determinar qué factores están más fuertemente asociados con el consumo de sustancias.

**Impacto en Salud y Sociedad:**

* + Evaluar cómo el consumo de sustancias afecta a diferentes aspectos de la salud y el bienestar social.

**Desarrollo de Modelos Predictivos:**

* + Utilizar técnicas de machine learning para predecir tendencias futuras y posibles puntos de intervención.

**Informar Políticas Públicas y Estrategias de Intervención:**

* + Proveer recomendaciones basadas en datos para políticas y programas que aborden el consumo de sustancias.

4- Objetivo Principal del Proyecto :

**"¿Cuáles son los principales factores predictivos del consumo problemático de sustancias en Argentina, y cómo se pueden utilizar estos factores para predecir las tendencias futuras y los posibles impactos en la salud pública y la sociedad?"**

Esta pregunta se centra en identificar variables específicas que puedan servir como indicadores predictivos del consumo problemático de sustancias. Estas variables podrían incluir, por ejemplo, la frecuencia y cantidad de consumo, patrones de uso (como el uso recreativo versus el uso compulsivo), factores demográficos (edad, género, ubicación geográfica), factores socioeconómicos (nivel de ingresos, educación), y otros factores de riesgo (como historial de salud mental o acceso a sustancias). El objetivo es utilizar estas variables para predecir tendencias y riesgos futuros, permitiendo una intervención más temprana y efectiva

5- Descripcion de DataSheet y Variables

Comentario de obtención de datos , y algunas de sus varables

6 – Librerias

Adjuntar las librerías usadas

7- Dataset

Mostrar con un dataframe como vienen los datos

8 – EDA análisis univariados,bivariados,multivariados

Ver los tipos de datos de las columnas y variables

Tratar datos nulos

Tratar líneas duplicadas

9 – tipos de variables Object- Descriptiva

10- Esatdisticas descriptivas

df.describe()

11- descripcion de las variables y sus tipos

12 – Analisys Univariado

Graficas y análisis de univariados

13 – Analisys Bivariados

Idem 12

14 analisys Multivariado

Analizar relaciones entre distintas variables y graficar las mas altas

15 – resumen de datos e informe hasta aquí

Resumen de todo lo analizado hasta acá

16 - analizar relaciones de variables y aplicar un filtro a los datos

Justificar

Medir varianza

17 – Análisis de componentes principales PCA

Realizar los gráficos de varianza

18 – REFORMULAR DATOS PARA VARIABLES Y PREDICCIONES

19 – APLICACIÓN DE MODELO ML

20-Metricas con distintos Modelos aboles,knn,etc

21- conclusiones y comparaciones de los distintos modelos aplicados