**Práctica 01 - ¿Cómo Hacer un Análisis Exploratorio?**

**Caso → Supervivencia del Hundimiento del Titanic**

**Primer Paso: Entendiendo el Contexto**

-Es de vital importancia entender el contexto de un problema al cual te enfrentas, generalmente se requiere un conocimiento general, pero ciertos detalles podrían ayudarte a comprender el problema de forma sistémica.

-Por ejemplo, el Titanic es uno de los naufragios más famosos de la historia, uno que se dio al principio del siglo XX de forma que tenemos que hacernos una mínima idea de cómo funcionaba la sociedad en ese momento.

-También es de utilidad saber que más de la mitad de formas perecieron en ese accidente, lo que en sí nos incentiva a pensar de qué dependía la tasa de supervivencia.

-Detalles extras como, el dinero que se gastó en construir el Titanic, o su causa de hundimiento por sí solas no podrían darnos mucha información, pero dependiendo los datos quizás una nueva perspectiva pueda surgir.

**Segundo Paso: Análisis Exploratorio de Datos [Netos]**

1. **Preparando el Entorno**

* Las librerías más importantes son, Numpy, pandas, Matplotlib.pyplot y seaborn.
* Podemos importar la librería warnings para ignorarlos
* %Matplotlib inline

1. **Una vez tengamos el DataFrame**
   * Lo importante será reconocer que tipos de variables tenemos, lo más importante es reconocer y entender qué significa cada una de ellas.
   * También es relevante reconocer que variables pueden encontrarse vacías, de modo que estemos en alerta cada que trabajemos en ellas.
   * Esto es sencillo con “tabla.isnull ().sum()”, en el caso de este proyecto, tenemos datos faltantes en las columnas: Edad, Cabina, Embarque.
2. **Elegir Hacia donde viraremos el Proyecto**

* Llegado a este punto tenemos que plantear la pregunta de cuál es el dato o datos que deseamos analizar con mucho más detenimiento.
* Por ejemplo, podríamos pensar que de un conjunto de datos de bienes raíces enfocarnos cuales son las variables que influyen o alterar el precio de la vivienda, entonces precio de la vivienda se convertiría en nuestra variable objetivo.
* Para resumir, debemos preguntarnos: ¿Qué variables están correlacionadas con nuestra variable objetivo?

1. **Análisis de Datos Nulos**
   * Es un error muy común querer eliminar o hacer nulas los campos de datos faltantes, sin embargo, esto puede ser peligroso ya que estamos arrebatando información en nuestra tabla que pudiera estar muy ligada a nuestra variable objetivo.
   * Lo primero que haremos será reconocer las variables nulas y evaluar si tienen un efecto muy significativo, para en este caso, la supervivencia del pasajero.
2. **Nuestras Primeras gráficas**
   * Teniendo nuestra variable objetivo definida, podemos en teoría graficar su comportamiento de forma básica, en este caso contamos con dos opciones: sobreviviente, y en estado de deceso.
   * Nos apoyaremos de Matplotlib, [Revisar el Notebook]

**SI REALICÉ MI CAMBIO EN GIT**

**3.- Tercer Paso: Aplicar Ingeniería de Características y Limpieza de Datos**