



# Tecnológico de Monterrey

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE  
MONTERREY CAMPUS PUEBLA

---

## REPORTE COMPARATIVO

**Materia:** Desarrollo de proyectos de análisis de datos

**Profesor:**

Mario Cardaña

Alfredo García

**Campus** Puebla

**Autor**

Pablo Antonio Gutierrez Piedras

A01736164

**Fecha de entrega:**

*27 de noviembre del 2022*

## Reporte

Durante el transcurso de este curso se han visto varios temas de programación los cuales nos han ayudado para resolver nuestro proyecto, sin embargo, aunque tienen similitudes bajo un punto de escritura, cada una de estas técnicas vistas tienen sus diferencias, por lo que en este reporte se pretende analizar cada una de estas diferencias.

Como se había dicho estas técnicas tienen algo de similitud en la escritura de código y otras no, pues al ser de diferente propósito cada una, esta difiere entre ellas:

**Valores nulos / Valores atípicos:** el valor nulo significa que es un valor desconocido, indefinido o no inicializado, mientras que los valores atípicos son observaciones cuyos valores son muy diferentes a las otras observaciones del mismo grupo de datos

**Extracción de datos / Extracción de características de Visualización:** Este apartado es donde más similitud hay ya que las dos hacen el mismo proceso, donde la diferencia se encuentra como su nombre lo indica, en un apartado más visual entre ambas

**Regresión Lineal múltiple / Regresión logística:** Por parte del primer tipo de regresión tenemos un modelo estadístico versátil para evaluar las relaciones entre un destino continuo y los predictores, y el del otro lado tenemos una técnica de análisis de datos que utiliza las matemáticas para encontrar las relaciones entre dos factores de datos. Luego, utiliza esta relación para predecir el valor de uno de esos factores basándose en el otro.

Como se puede ver todas estas técnicas vistas son diferentes entre ellas y con ello no se puede elegir una sobre las otras, pues cada una cumple una función diferente entre sí, y con al final cumplen un papel importante para el seguimiento del proyecto al servir de forma diferente y con un fin específico, por lo que al ser de diferente propósito ninguna es más importante o preferida que las otras.

En conclusión, la preferencia que se tendrá para cada una se definirá a partir del trabajo requerido, si se requiere identificar valores desconocidos se utilizarán valores nulos, si se requiere modelo estadístico versátil para evaluar las relaciones entre un destino continuo y los predictores se utilizará una regresión lineal múltiple, por lo que todo dependerá por lo que se requiere