



Herramientas computacionales: el arte de la analítica

TC1002S.222

Fecha de entrega: 11 de mayo del 2022

Equipo 4

Actividad : “Mapas de calor y boxplots ”

Profesor: Sergio Ruíz Loza

Nombres	Apellidos	Matrícula
Ana Lucía	Ahedo Reyes	A01661890
Andrea Samantha	Aguilar Ramírez	A01656200
Luis Alfonso	Pérez Mijangos	A01653848
Guillermo	González Jiménez	A01653930
Alfonso	Pineda Cedillo	A01660394

1. ¿Hay alguna variable que no aporte información?

Consideramos que las variables que no aportan información precisa son la, 4046, 4225 y 4770, ya que no se conoce el contexto al que hacen referencia, haciéndolas así difíciles de usar y comprender ante un usuario/lector que no tenga conocimiento de los tipos de aguacates.

2. Si tuvieras que eliminar variables, ¿cuáles quitarías y por qué?

Consideramos que si se tuviera que eliminar alguna variable sería la de año (Year), ya que se cuenta con la variable de fecha (Date) la cual aporta la misma información que la variable mencionada. También, se eliminarían las variables 4046, 4225 y 4770, ya que no aportan información útil al análisis de los datos.

3. ¿Existen variables que tengan datos extraños?

No, no existen variables con datos extraños ya que todos los datos que se muestran tienen una relación lógica con la propia variable.

4. Si comparas las variables, ¿todas están en rangos similares?

No, porque cada variable hace referencia a información completamente independiente, a pesar de que guarden una relación entre sí, no significa que los rangos tengan que ser similares.

5. ¿Crees que esto afecte?

No, porque son variables independientes entre sí, es decir, los valores de una columna no afectan el de otra; a excepción de "Total Bags", la cual es la suma total de "Small Bags", "Large Bags" y "XLarge Bags".

6. ¿Puedes encontrar grupos que se parezcan? ¿Qué grupos son estos?

Las columnas "Total Bags", "Small Bags", "Large Bags" y "XLarge Bags" guardan una relación entre sí, debido a que "Total Bags" es la suma total del resto de variables antes mencionadas. Otra correlación que

se puede suponer, es la que hay entre el volúmen total (Total Volume) con el total de bolsas (Total Bags), ya que a mayor volúmen habrá mayor número de bolsas.