# Respuestas Tarea 1

## Punto 1

1.1

GDP (USD Billion) description

mean: 8.750000000000002

median: 2.65

std: 19.579713140561246

Population (Millions) description

mean: 0.7310000000000001

median: 0.39

std: 1.330093605728559

Unemployment Rate (%) description

mean: 13.833333333333334

median: 13.45

std: 2.89555214462151

Average Age description

mean: 29.233333333333334

median: 29.0

std: 2.2012622641465405

Women (%) description

mean: 51.5

median: 51.0

std: 0.7637626158259734

Men (%) description

mean: 48.5

median: 49.0

std: 0.7637626158259734

Budget (USD Billion) description

mean: 1.65

median: 0.6

std: 3.3931794333142284

1.2

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente1.3

La covarianza entre GDP (USD Billion) y Population (Millions) es 25.795716666666667

1.4

La correlación entre GDP (USD Billion) y Population (Millions) es 0.9905104636501418

1.5

La correlación y la covarianza son dos medidas estadísticas que se utilizan para evaluar la relación entre dos variables.

La covarianza mide la variabilidad conjunta de dos variables, mientras que la correlación mide la relación lineal entre ellas y proporciona una medida estandarizada de esta relación.

1.6

Resultado en kaggle

1.7

Gráfico

Descripción generada automáticamente

## Punto 2

2.1

[[1.0344827586206893, 0.990510463650142],

[0.990510463650142, 1.03448275862069]]

2.2

(2.0249932222708313, 0.04397229497054733)

2.3

Varianza explicada eigenValue 1 97.87467240975688%

Varianza explicada eigenValue 2 2.125327590243122%

2.4

array([0.70710678, 0.70710678])

2.5

array([ 6.8502549 , 2.25429149, 1.28986071, 0.55599905, 0.22215487,

-0.13264057, -0.22552843, -0.1164 , -0.2495073 , -0.28371492,

-0.33878374, -0.31240449, -0.34981985, -0.33577783, -0.35363312,

-0.42826202, -0.4461173 , -0.46928881, -0.48353267, -0.50670418,

-0.53689669, -0.53348711, -0.58855592, -0.61874844, -0.58514634,

-0.64191995, -0.62958272, -0.65445902, -0.6740191 , -0.67763052])

2.6

Error de reconstrucción: 6.8882277654917115 %

2.7

Gráfico

Descripción generada automáticamente

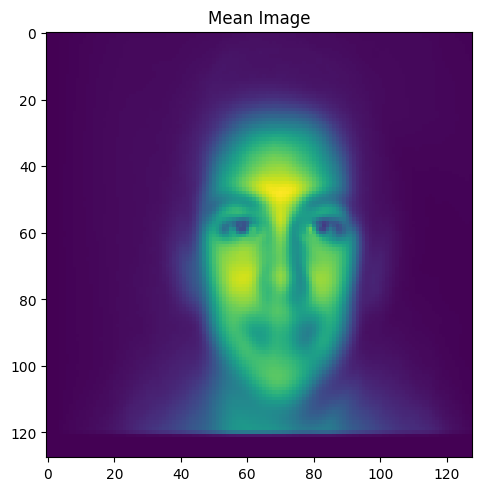
2.8

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

## Punto 3

3.1



3.2

Se deben utilizar **142** términos para asegurar el 95%

Interfaz de usuario gráfica

Descripción generada automáticamenteInterfaz de usuario gráfica

Descripción generada automáticamenteInterfaz de usuario gráfica

Descripción generada automáticamenteInterfaz de usuario gráfica

Descripción generada automáticamenteInterfaz de usuario gráfica

Descripción generada automáticamente con confianza media

Con el 95% de los componentes es suficiente para conservar una imagen de calidad, es decir, utilizando 142 componentes.

El modelo de test ha sido subido a kaggle

## Punto 4

### 4.1

4.1.1

Gráfico, Gráfico de barras

Descripción generada automáticamenteGráfico, Gráfico de barras

Descripción generada automáticamenteGráfico, Gráfico de barras

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamenteGráfico, Gráfico de barras

Descripción generada automáticamenteGráfico, Gráfico de barras

Descripción generada automáticamente

4.1.2

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std

Gráfico

Descripción generada automáticamente

Disney+ es un producto fuera de 5 std

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std y se observa falta de datos para esa variable

Gráfico, Histograma

Descripción generada automáticamente

No tiene productos lejos de 5 std y se observa falta de datos para esa variable

## 4.2

4.2.1

Gráfico

Descripción generada automáticamente Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamenteGráfico

Descripción generada automáticamente con confianza media

Gráfico

Descripción generada automáticamente Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente Gráfico, Gráfico de barras

Descripción generada automáticamente

La forma de interpretar estos graficos es que las cajas con un mayor tamaño tienen una valor que influye mas en el volume de ventas de los productos, podemos ver que las variables que mas influyen en el volumen de ventas es (is\_best\_seller y climate\_pledge\_friendly).

4.2.2

Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Gráfico

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Estos graficos se interpretan viendo el punto como la variable (el valor en el eje x) que influye en volumen de ventas (el valor en el eje y). Donde mas valores veamos es el valor que mas influye en esa variable sobre el volumen de ventas.

## 4.3

4.3.1

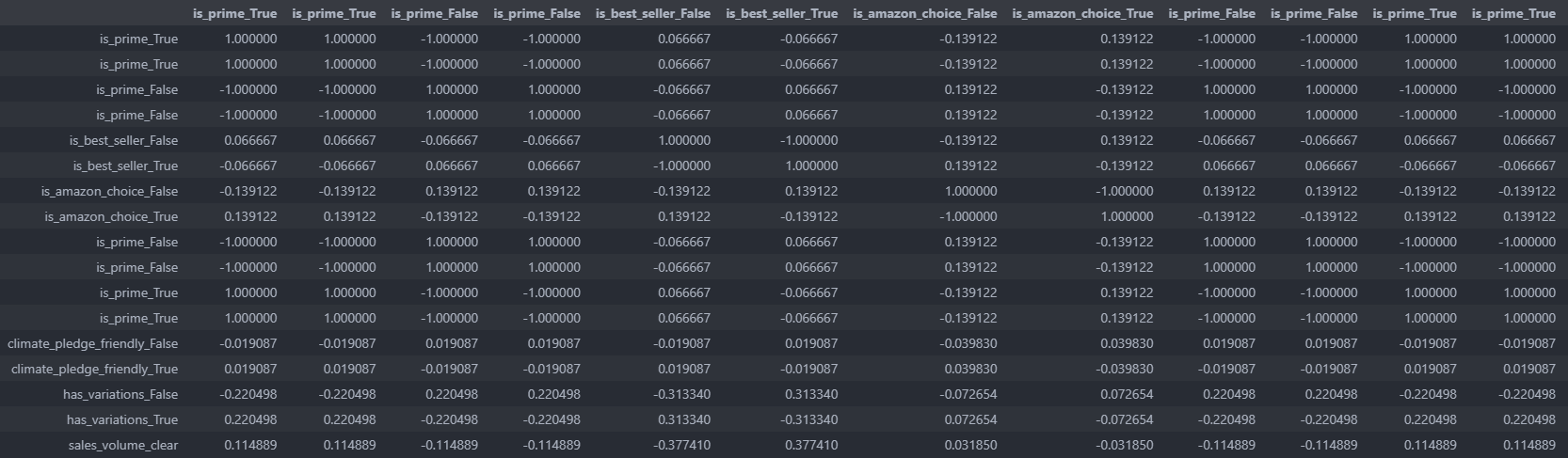
Captura de pantalla de computadora

Descripción generada automáticamente

La variables cuantitativas que más explican sales\_volume\_clear(volumen de ventas) es producto\_price(el precio del producto en dólares), con una relación negativa y product\_star\_rating(calificación en estrellas del producto) con una relación positiva. Lo que quiere decir que se puede explicar mucho del volumen de ventas de un producto en base a su precio y su calificación.

Podemos ver que hay coeficientes negativos y positivos, estar positivo es que tienen una relación directa(en cuanto la variable sube sales\_volume sube) y estar negativo es que tienen una relación inversa(en cuanto la variable sube sales\_volume baja).

4.3.2



Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente con confianza media

La variable con mayor relación es is\_best\_seller, en donde cuando es true(1) tiene una relación directa con sales\_volume, es decir, cuando un producto es best\_seller hace que tenga mayor volumen de ventas.

Y cuando es false(0) tiene una relación inversa, es decir, cuando un producto no es un best\_seller impacta en que este tenga menor volumen de ventas.