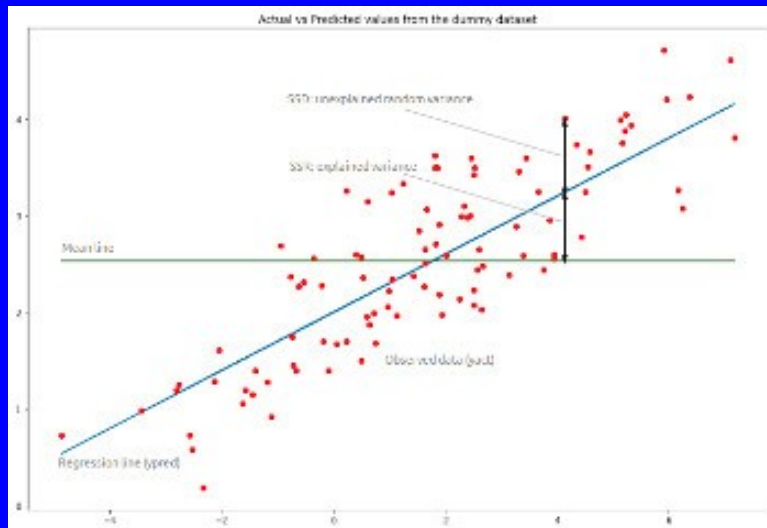


Python - Estadística

Regresión Lineal (Ejemplo)



Rogelio Ferreira Escutia

Profesor / Investigador
Tecnológico Nacional de México
Campus Morelia



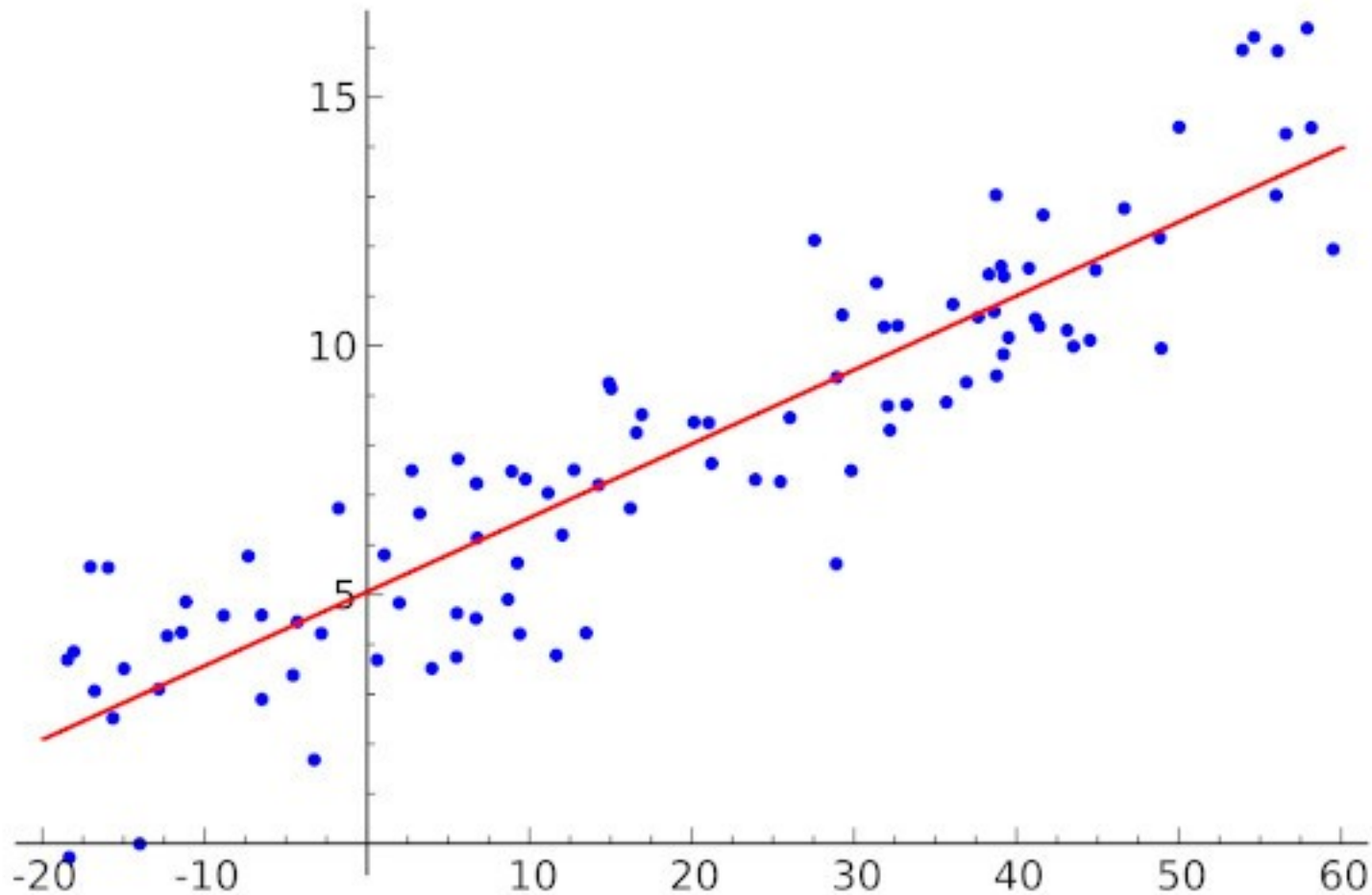
Estadística Inferencial

Regresión Lineal

- La “regresión lineal” es un modelo matemático para predecir una respuesta (variable dependiente) de acuerdo a una función de otra variable (variables independiente), es decir, dado un conjunto de datos podemos crear un modelo que nos permita predecir los valores futuros a que pueden llegar los datos.

Regresión Lineal

- Regresión lineal con una variable dependiente y una variable independiente.:



Ejemplo de una Regresión Lineal

Ejemplo de una Regresión Lineal

- **Juanito es un niño.**
- **Su Doctor particular lleva el siguiente registro de peso a través de los años con los siguientes datos:**
 - **Al año pesaba 6 kilogramos.**
 - **A los 5 años pesaba 15 kilogramos.**
 - **A los 10 años pesaba 40 kilogramos.**
- **Lo que se quiere saber es:**
 - **¿Cuánto pesará Juanito cuando tenga 20 años?**

Identificar datos y variables

- **Tebemos la edad, que será nuestra variable independiente (x).**
- **El peso depende de la edad, y se convierte en la variable dependiente (y).**
- **De lo anterior nos queda lo siguiente:**

x	1	5	10
y	6	15	40

Regresión Lineal

- **Código:**

```
1  # Importando matplotlib para graficar
2  import matplotlib.pyplot as plt
3  # Creamos una lista para indicar las edades
4  x_lista = [1, 5, 10]
5  # Creamos una lista para indicar los pesos
6  y_lista = [6, 15, 40]
7  # Hacemos un ciclo para imprimir todos los puntos
8  contador = 0
9  for ciclo in x_lista:
10     plt.plot(ciclo, y_lista[contador], marker="o", color="red")
11     contador = contador + 1
12 # Colocamos un título a nuestra gráfica y a nuestros ejes
13 plt.title("Gráfica de Edades y Peso de Juanito")
14 plt.xlabel("Edad")
15 plt.ylabel("Peso")
16 # Mostramos la gráfica en pantalla
17 plt.show()
```


Gráfica

- Calculamos su “diagrama de dispersión”, cargando los datos a una lista y los graficamos usando Matplotlib:

```
# Importando matplotlib para graficar
import matplotlib.pyplot as plt

# Creamos una lista para indicar los nombres
x_lista = [1, 5, 10]

# Creamos una lista para indicar las edades
y_lista = [6, 15, 40]

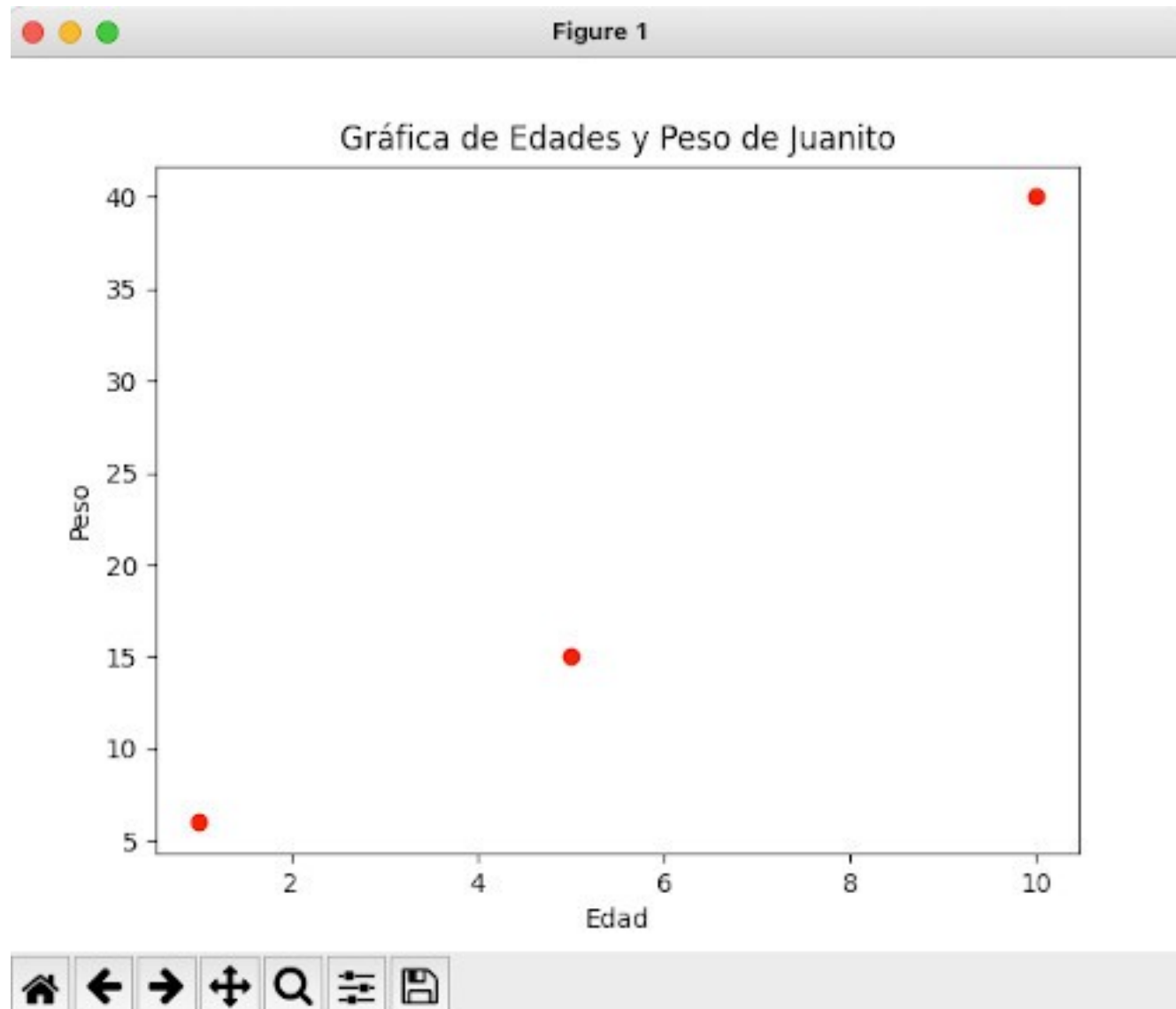
# Hacemos un ciclo para imprimir todos los puntos
contador = 0
for ciclo in x_lista:
    plt.plot(ciclo, y_lista[contador], marker="o", color="red")
    contador = contador + 1

# Colocamos un título a nuestra gráfica y a nuestros ejes
plt.title("Gráfica de Edades y Peso de Juanito")
plt.xlabel("Edad")
plt.ylabel("Peso")

# Mostramos la gráfica en pantalla
plt.show()
```

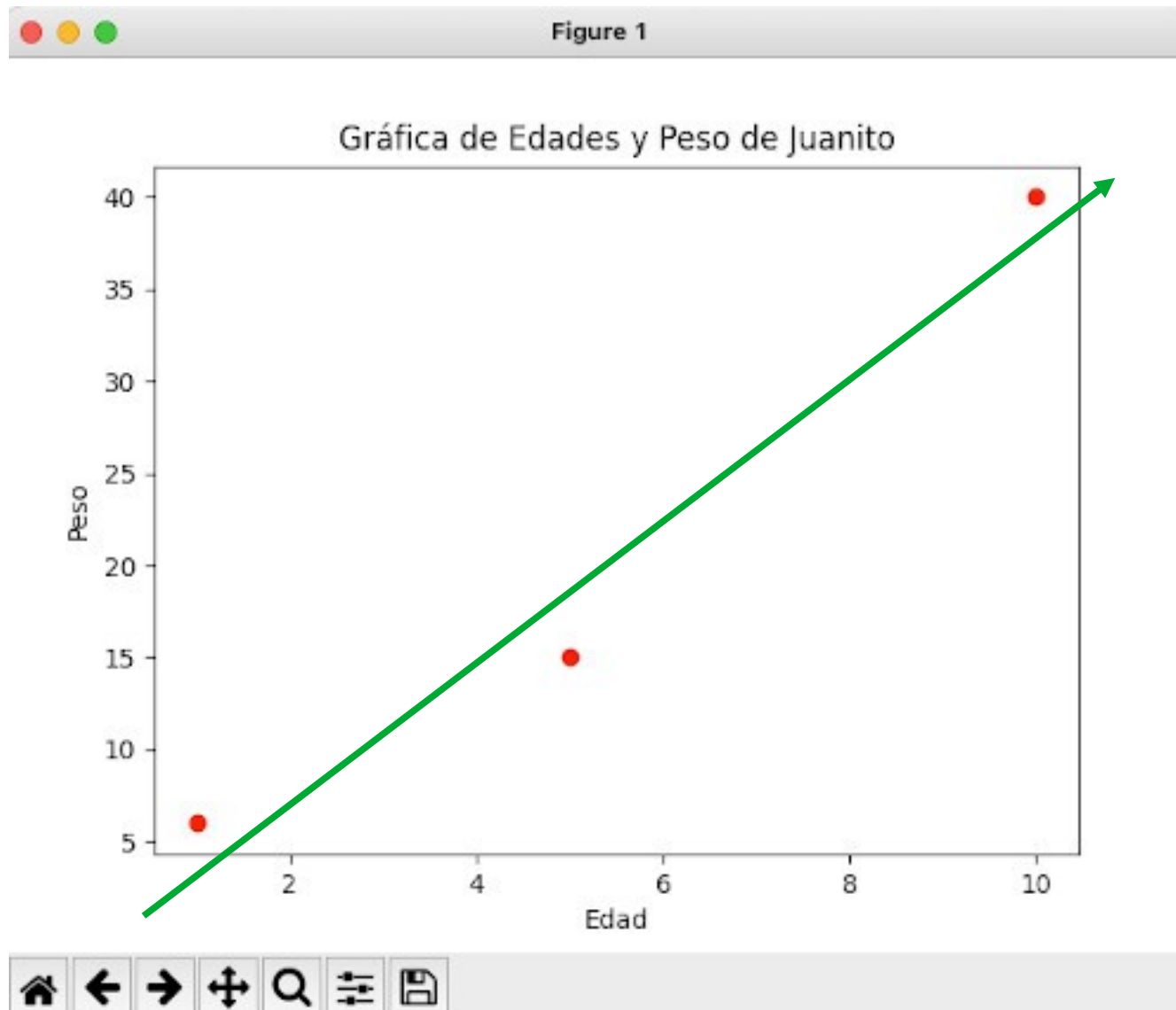
Diagrama de Dispersión

- Usando Matplotlib:



Análisis

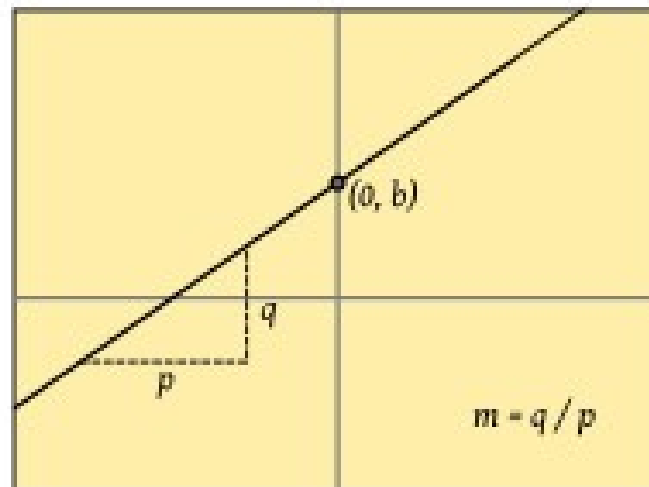
- En los datos se observa una tendencia lineal:



Modelo Propuesto

- Para modelar el comportamiento de los datos se propone una recta, donde conocemos “x”, “y”, y buscaremos el valor de “m” y “b” para poder tener nuestra función lineal que represente a nuestro modelo:

Función lineal



$$y = mx + b$$

Ecuaciones Normalizadas

- A partir del análisis de la ecuación de la recta, se obtienen las siguientes ecuaciones normalizadas:

$$\sum y = m \sum x + n b$$

$$\sum xy = m \sum x^2 + b \sum x$$

Datos a calcular

- Se requiere calcular los siguientes datos:

$$\sum x$$

$$\sum y$$

$$\sum x^2$$

$$\sum xy$$

Calculando

- Se hacen los siguientes cálculos:

$$\begin{aligned}\sum x &= x_0 + x_1 + \dots + x_n \\ &= 1 + 5 + 10 \\ &= 16\end{aligned}$$

$$\begin{aligned}\sum y &= y_0 + y_1 + \dots + y_n \\ &= 6 + 15 + 40 \\ &= 61\end{aligned}$$

Calculando

- Se hacen los siguientes cálculos:

$$\begin{aligned}\sum x^2 &= x_0^2 + x_1^2 + \dots + x_n^2 \\ &= (1)^2 + (5)^2 + (10)^2 \\ &= 1 + 25 + 100 \\ &= 126\end{aligned}$$

$$\begin{aligned}\sum xy &= x_0 y_0 + x_1 y_1 + \dots + x_n y_n \\ &= (1)(6) + (5)(15) + (10)(40) \\ &= 6 + 75 + 400 \\ &= 481\end{aligned}$$

Substituyendo

- Se hacen los siguientes cálculos:

$$\sum y = m \sum x + n b$$

$$61 = 16m + 3b$$

$$\sum xy = m \sum x^2 + b \sum x$$

$$481 = 126m + 16b$$

Substituyendo

- Se hacen los siguientes cálculos:

$$\textcircled{1} \quad 16m + 3b = 61$$

$$\textcircled{2} \quad 126m + 16b = 481$$

Despejando m en $\textcircled{1}$ Despejando m en $\textcircled{2}$

$$m = \frac{61 - 3b}{16}$$

$$m = \frac{481 - 16b}{126}$$

Igualando

$$\frac{61 - 3b}{16} = \frac{481 - 16b}{126}$$

Calculando b

- Se hacen los siguientes cálculos:

Realizando los cálculos

$$3.8125 - 0.1875b = 3.817460317460317 - 0.126984126984127b$$

Despejando para b

$$-0.060515873015873b = 0.004960317460317$$

$$\therefore b = -0.081967213114747$$

Calculando m

- Se hacen los siguientes cálculos:

Calculando m

sustituyendo b en ①

$$\textcircled{1} \quad 16m + 3b = 61$$

$$16m = 61 - 3b$$

$$m = \frac{61 - 3b}{16}$$

$$m = 3.8125 - 0.015368852459015$$

$$m = 3.797131147540985$$

Ecuación

- Finalmente encontramos nuestra ecuación de la recta:

Tenemos

$$y = mx + b$$

Sustituyendo m y b

$$y = 3.79x - 0.081$$

Predicción

- Para predecir:
 - ¿Cuánto pesará Juanito cuando tenga 20 años?
- Sustituimos $x = 20$ en nuestra ecuación de la recta

$$y = 3.79x - 0.081$$

$$y = (3.79)(20) - 0.081$$

$$y = 75.8 - 0.081$$

$$y = 75.719 \text{ Kilogramos}$$



Rogelio Ferreira Escutia

Profesor / Investigador
Tecnológico Nacional de México
Campus Morelia



rogelio.fe@morelia.tecnm.mx



rogeplus@gmail.com



xumarhu.net



[@rogeplus](https://twitter.com/rogeplus)



[https://www.youtube.com/
channel/UC0on88n3LwTKxJb8T09sGjg](https://www.youtube.com/channel/UC0on88n3LwTKxJb8T09sGjg)



[rogelioferreiraescutia](https://www.linkedin.com/in/rogelioferreiraescutia)

