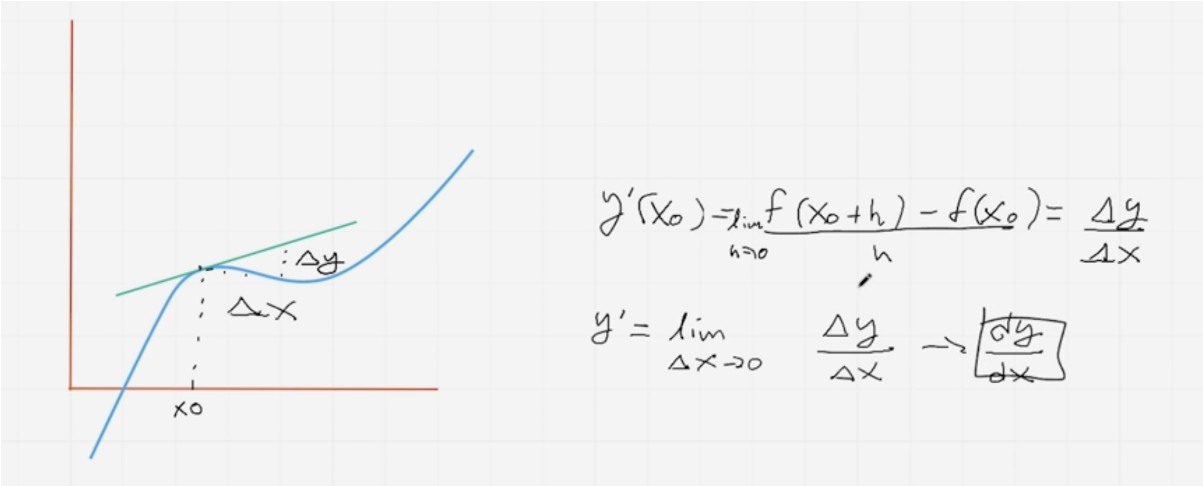
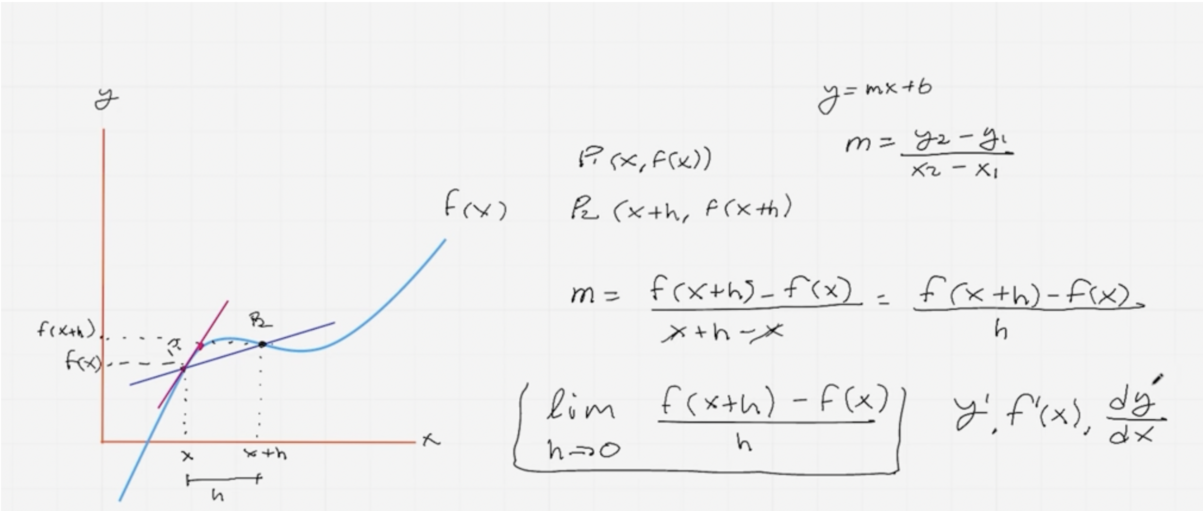
Derivadas

> ***La derivada*** *de una función* ***f(x)*** *es otra función* ***f '(x)*** *que representa las pendientes de las rectas tangentes para cualquier punto de f(x)*

Función simple	Derivada
$f(x) = k$	$f'(x) = 0$
$f(x) = x$	$f'(x) = 1$
$f(x) = ax + b$	$f'(x) = a$
$f(x) = u(x) + v(x)$	$f'(x) = u'(x) + v'(x)$
$f(x) = k \cdot u(x)$	$f'(x) = k \cdot u'(x)$
$f(x) = u(x) \cdot v(x)$	$f'(x) = u'(x) \cdot v(x) + u(x) \cdot v'(x)$
$f(x) = \frac{u(x)}{v(x)}$	$f'(x) = \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{v^2(x)}$
$f(x) = x^n$	$f'(x) = n \cdot x^{n-1}$
$f(x) = \ln x$	$f'(x) = \frac{1}{x}$
$f(x) = \log_a x$	$f'(x) = \frac{1}{x} \log_a e$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = a^x$	$f'(x) = a^x \cdot \ln a$
$f(x) = \operatorname{sen} x$	$f'(x) = \cos x$
$f(x) = \cos x$	$f'(x) = -\operatorname{sen} x$
$f(x) = \operatorname{tg} x$	$f'(x) = \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x$
$f(x) = \operatorname{cotg} x$	$f'(x) = \frac{-1}{\operatorname{sen}^2 x} = -\left(1 + \operatorname{cotg}^2 x\right)$
$f(x) = \operatorname{arcsen} x$	$f'(x) = \frac{1}{\sqrt{1-x^2}}$
$f(x) = \operatorname{arctg} x$	$f'(x) = \frac{-1}{\sqrt{1-x^2}}$

Ejemplos:





La derivada como razón de cambio

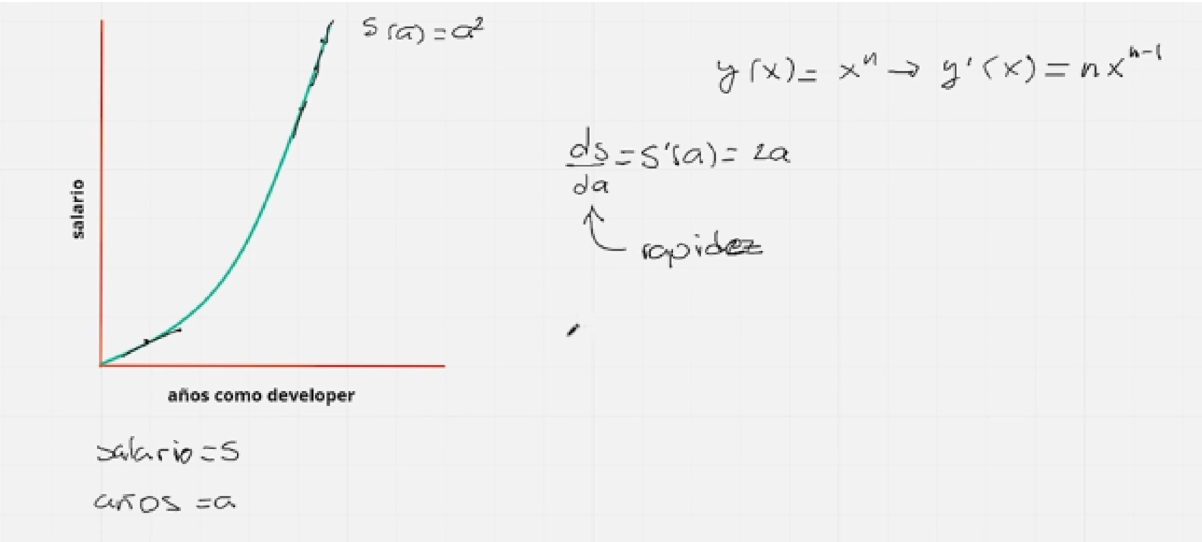
> La interpretación geométrica de la derivada es la pendiente de la recta tangente a la curva de la función en un punto dado. Si trazamos una recta tangente a la curva en un punto, la pendiente de esa recta representa la tasa de cambio instantánea de la función en ese punto específico.

```
Más formalmente, si tenemos una función f(x), la derivada de f(x) en un punto x = a, denotada como f'(a) o dy/dx |_(x=a), se define como el límite de la razón de cambio de f(x) respecto a x cuando x se acerca a a. Matemáticamente, se puede expresar como:
```

```
f'(a) = lim(x→a) [f(x) - f(a)] / [x - a]
```

> La derivada nos brinda información importante sobre la función. Por ejemplo:

- La derivada positiva indica que la función está creciendo en ese punto.
- La derivada negativa indica que la función está decreciendo en ese punto.
- La derivada igual a cero indica un posible máximo o mínimo local de la función.
- La derivada nos permite calcular la velocidad instantánea en problemas de movimiento.
- La derivada nos permite encontrar la tasa de cambio en problemas económicos y científicos.



Diferentes notaciones de la derivada

Existen diferentes formas de expresar la derivada si de notaciones hablamos. Cada una de ellas fue propuesta por un científico diferente al momento de desarrollar los principios del cálculo.

Si sabemos que la variable **x** es la variable independiente y **y** la variable dependiente a través de la relación **y=f(x)**. Algunas notaciones para la derivada son las siguientes:

$$f'(x) = y' = \frac{dy}{dx} = \frac{df}{dx} = \frac{d}{dx}\, f(x) = Df(x) = D_x f(x)$$

Notación de Leibniz

La notación de Leibniz surge del símbolo **dy/dx** que representa un operador de diferenciación y no debemos confundirlo como una división. Si quisiéramos expresar una segunda derivada usando la notación de Leibniz se puede mostrar como:

$$\frac{d^2 y}{dx^2}$$

Y para mostrar la n-ésima derivada se expresa de la forma:

$$\frac{d^n y}{dx^n}$$

Esta notación nos sirve para entender como la derivada puede ser expresada como los incrementos tanto de **x** como de **y** cuando el incremento de **x** tiende a cero.

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

Notación de Lagrange

La notación más sencilla de todas es la de Lagrange. Esta notación expresa que la función es una derivada usando una comilla simple antes del argumento, llamada *prima*.

$$f'(x)$$

Esta expresión se lee como "efe prima de equis". La cual representa la primera derivada de una función. Si deseamos expresar la segunda derivada sería:

$$f''(x)$$

Y para mostrar la n-ésima derivada se expresa de la forma:

$$f^{(n)}(x)$$

Notación de Newton

Por último tenemos la notación de Newton. Esta notación es muy usada en campos como la física y la ingeniería debido a su simplicidad para expresar la primera y segunda derivada. Se usa sobre todo en funciones relacionadas al tiempo en campos como la mecánica. Por ejemplo, como una función que representa el movimiento de una partícula.

Su representación de la primera y segunda derivada es la siguiente:

$\dot{x}$ $\ddot{x}$

Regla de la cadena

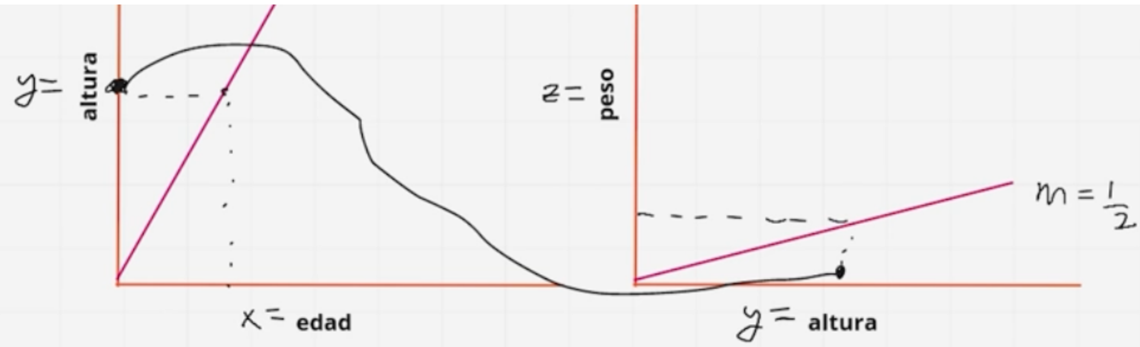
> **La regla de la cadena es una herramienta fundamental en cálculo diferencial que permite calcular la derivada de una función compuesta. Es esencial en muchas áreas de las matemáticas y la física, y tiene aplicaciones prácticas en diversos campos, como la física, la ingeniería, la economía y la ciencia de datos.**

$$\boxed{y = f(g(t)) \rightarrow y'(f(g(t))) = f'(g(t)) \cdot g'(t)}$$

$$\boxed{\frac{dy}{dt} = \frac{dy}{dg} \cdot \frac{dg}{dt}}$$

Para ver esto mejor, vamos a hacer un ejemplo:

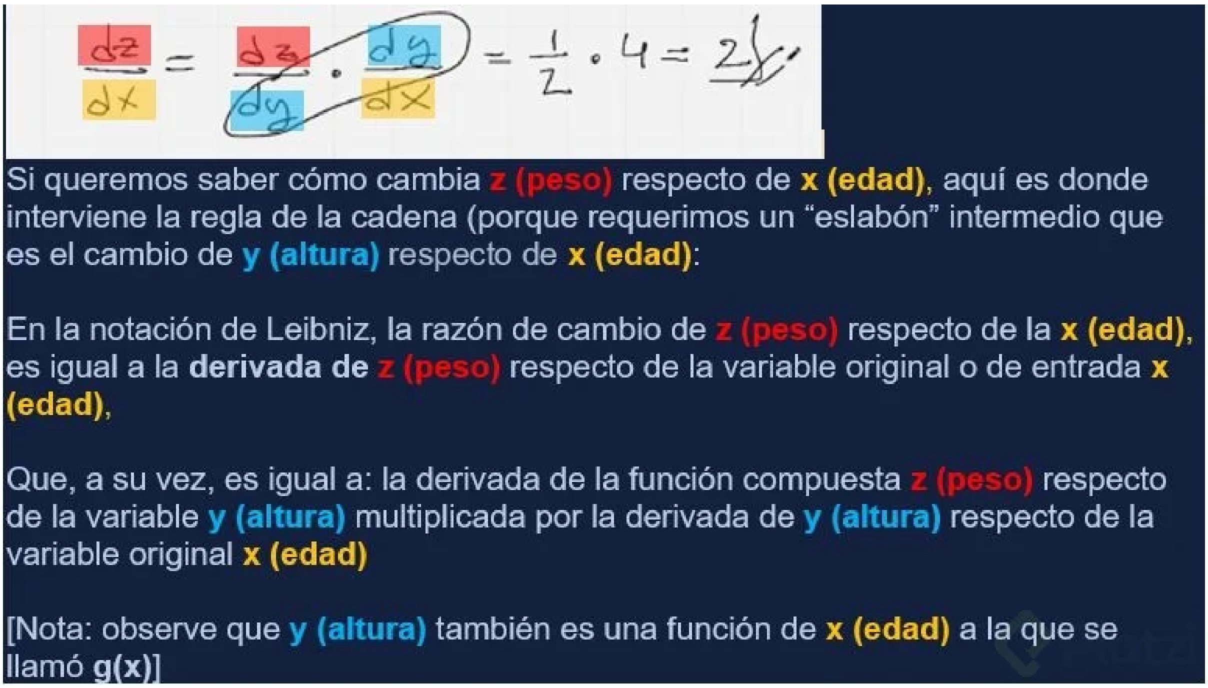
Para ver como varía la edad con respecto al peso, es decir, relacionar la edad vs peso, aplicaremos la regla de la cadena:



$$\frac{dy}{dx} = 4 \qquad \frac{dz}{dy} = \frac{1}{2}$$

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} = \frac{1}{2} \cdot 4 = \underline{2}$$

💡 La derivada nos da la pendiente de la recta tangente.



Máximo y mínimo

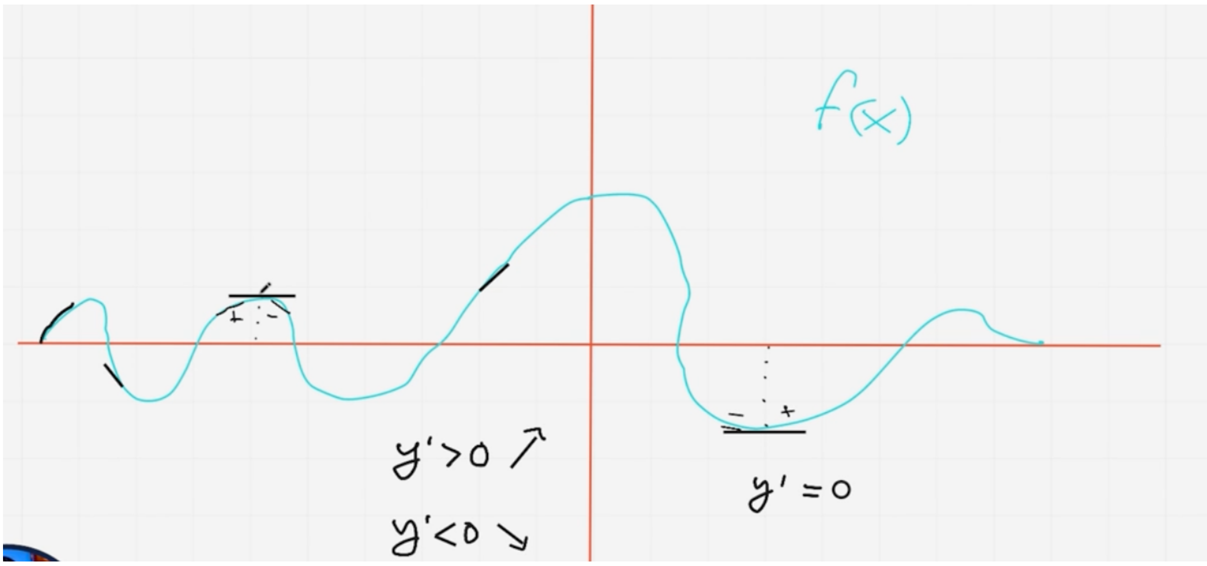
> **En matemáticas, un máximo y un mínimo son valores extremos de una función que representan los puntos más altos y más bajos, respectivamente. Estos extremos son importantes en el análisis de funciones, ya que brindan información sobre los puntos críticos y las características de la función. A continuación, se describen con más detalle cada uno de ellos:**

Máximo:

- Un máximo de una función es el valor más alto que la función alcanza en un determinado intervalo o en todo su dominio. Formalmente, sea f una función definida en un conjunto D, se dice que f tiene un máximo en el punto c si f(c) es mayor o igual que f(x) para todo x en D cercano a c. Un máximo absoluto es aquel que es el valor más alto de la función en todo su dominio, mientras que un máximo relativo es aquel que es el valor más alto en un intervalo específico.

Mínimo:

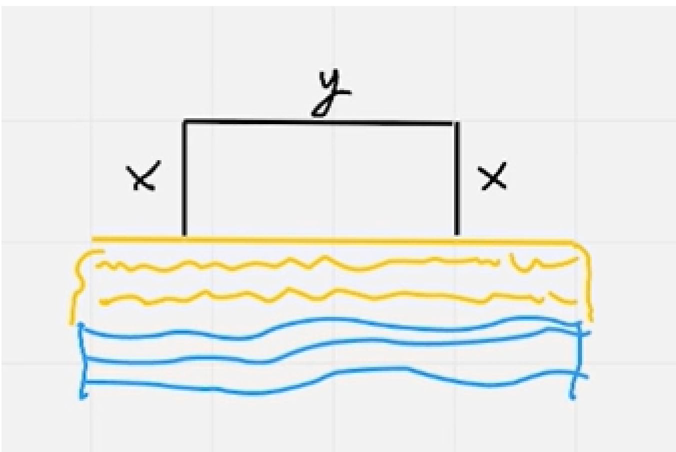
- Un mínimo de una función es el valor más bajo que la función alcanza en un determinado intervalo o en todo su dominio. Formalmente, sea f una función definida en un conjunto D, se dice que f tiene un mínimo en el punto c si f(c) es menor o igual que f(x) para todo x en D cercano a c. Un mínimo absoluto es aquel que es el valor más bajo de la función en todo su dominio, mientras que un mínimo relativo es aquel que es el valor más bajo en un intervalo específico.



Optimización

Los problemas de optimización en general requieren encontrar mínimos o máximos. Veamos esto con el siguiente problema: queremos construir una oficina con solo 50 mts de perímetro de paredes abarcando el área más grande posible (punto máximo). La oficina solo tiene 3 paredes y una vista al mar.

Resolviendo el problema



Primero debemos encontrar la función a optimizar. En este caso, queremos encontrar el área máxima. De acuerdo al dibujo de arriba, el área está dada por `A = x * y`, y el perímetro de las tres paredes está dado por `p=2x+y=50`. Teniendo estos datos, podríamos intentar resolverlo al tanteo. Sin embargo, la mejor forma de resolverlo parecido a un sistema de ecuaciones.

Si tenemos que `2x+y=50`, podemos reordenar y nos queda `y=50-2x`. Entonces sustituimos esta expresión en la fórmula del área para que nos quede solo en función de `x`. Nos queda: `A(x)=x(50-2x)=50x-2x*x`.

Ahora toca diferenciar la función del área e igualarla a cero. Quedando así: `A′(x) = -4x + 50 = 0`. Resolvemos para x, y nos da un punto crítico en `x=25/2`. Solo nos queda determinar si este es un máximo. Para ello evaluamos en la derivada un punto a la izquierda y a la derecha de 25/2. Es decir:

```
A'(12)=-4*12+50=2
A'(13)=-4*13+50=-2
```

Como a la izquierda de `x` la derivada es positiva, y a la derecha es negativa, podemos decir que `x=25/2` es un máximo. Solo queda sustituir en `y=50-2x`. Nos da que `y=25`. Por lo tanto, las medidas de las paredes de la oficina con la mayor área son `x=(25/2)m` y `y=25m` . El área nos da:

$$A = 12.5m * 25m = 312.5m^2$$

¿Cómo son las derivadas en las funciones de activación?

Las funciones de activación

> Son un componente clave de las redes neuronales artificiales. Su función es determinar la salida de una neurona dado un conjunto de entradas

Las funciones de activación más comunes son:

- Función sigmoide (buena para los procesos de aprendizaje, pero el costo computacional es demasiado alto)

- Función tangente hiperbólica (tanh) (buena para los procesos de aprendizaje, pero el costo computacional es demasiado alto)
- Función ReLU (Rectified Linear Unit) (buena para procesos de capas intermedias)
- Función Leaky ReLU

> 💡 La función sigmoide y la función tangente hiperbólica son buenas para la clasificación de probabilidades, pero cuando estemos optimizando, esta función va a complicar demasiado nuestro entrenamiento y tiene un problema cuando estamos tratando de optimizar, que satura el proceso de entrenamiento y que puede llevar a una “muerte neuronal” o más conocida como “vanish in gradient”