In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
```

Variables

x = altura de la persona
y = peso de la persona

In [2]:
```
y = [68.78,74.11,71.73,69.88,67.25,68.78,68.34,67.01,63.45,71.17,
    67.19,65.80,64.30,67.97,72.18,65.27,66.09,67.51,70.10,68.25,
    67.89,68.14,69.08,72.80,67.42,68.49,68.61,74.03,71.52,69.18]
x = [162,212,220,206,152,183,167,175,156,186,183,163,163,172,194,
    168,161,164,188,187,162,192,184,206,175,154,187,212,195,205]
```

In [3]:
```
n = len(y)

n
```

Out[3]:
```
30
```

In [4]:
```
x = np.array(x)
y = np.array(y)

x
```

Out[4]:
```
array([162, 212, 220, 206, 152, 183, 167, 175, 156, 186, 183, 163, 163,
       172, 194, 168, 161, 164, 188, 187, 162, 192, 184, 206, 175, 154,
       187, 212, 195, 205])
```

In [5]:
```
sumx = sum(x)
sumy = sum(y)
sumx2 = sum(x**2)
sumy2 = sum(y**2)
sumxy = sum(x*y)
promx = sumx/n
promy = sumy/n

sumx, sumx2, sumxy
```

Out[5]:
```
(5434, 994968, 374715.79)
```

In [6]:
```
m = (sumx*sumy - n*sumxy) / (sumx**2 - n*sumx2)
b = promy - m*promx

m, b
```

Out[6]:
```
(0.10860167641666033, 49.072616345062244)
```

Ecuación lineal

y = -218.68254697954592x + 5.816011292809254

In [7]:
```
sigmax = np.sqrt(sumx2/n - promx**2)
sigmay = np.sqrt(sumy2/n - promy**2)
sigmaxy = sumxy/n - promx*promy

R2 = (sigmaxy / (sigmax*sigmay))**2

R2
```

Out[7]:
```
0.6316285764573425
```

R2 = 63.1628 %

Gráfica

In [10]:
```
plt.plot(x, y, 'o', label = 'Datos')
plt.plot(x, m * x + b, label = 'Regresión')
plt.xlabel('\nAltura')
plt.ylabel('Peso\n')
plt.title('Gráfica de dispersión\n')
plt.grid()
plt.legend()
plt.show()
```