

stats2d

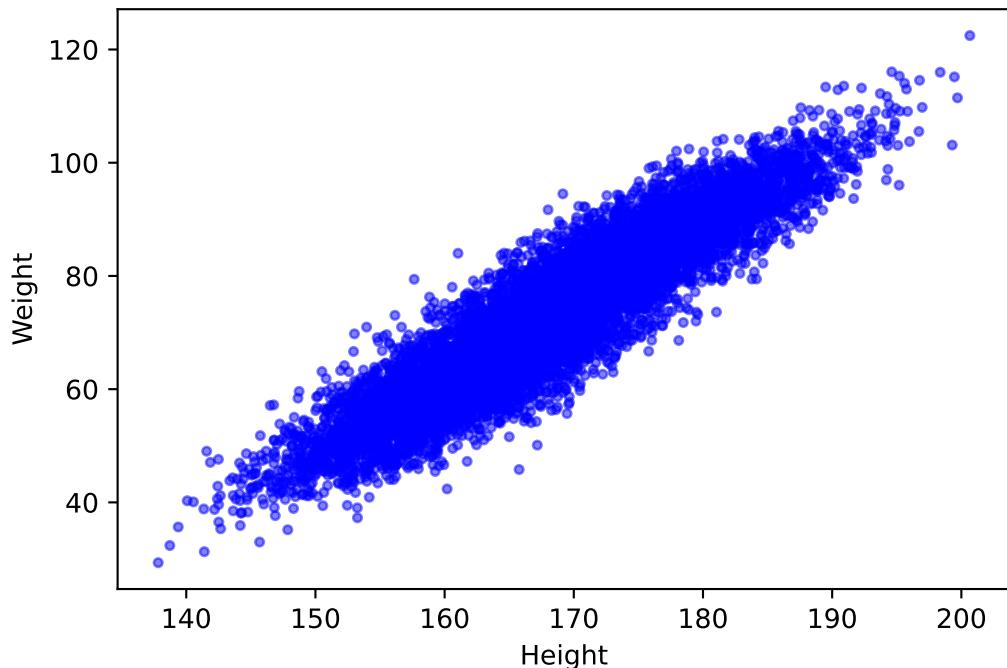
March 10, 2022

```
[ ]: import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
%config InlineBackend.figure_formats = ['svg']
```

```
[ ]: df = pd.read_csv('weight-height.csv')
df.head()
```

```
[ ]:   Gender      Height      Weight
0    Male    73.847017  241.893563
1    Male    68.781904  162.310473
2    Male    74.110105  212.740856
3    Male    71.730978  220.042470
4    Male    69.881796  206.349801
```

```
[ ]: X= df['Height'].values*2.54 #inch to cm
Y = df['Weight'].values*0.453592 # pound to kg
plt.scatter(X,Y, s=10, color="blue", alpha=0.5)
plt.xlabel('Height')
plt.ylabel('Weight')
plt.show()
```



1 Covariance

```
[ ]: N= X.shape[0]
Xm = X.mean()
Ym = Y.mean()
SumXY = X.dot(Y) #dot product ou produit scalaire
SEGxy= (SumXY/N)-Xm*Ym
print("N= ",N)
print('Xm= ', Xm)
print('Ym= ', Ym)
print("SumXY= ", SumXY)
print("COVxy= ",SEGxy)
```

```
N= 10000
Xm= 168.57360177724598
Ym= 73.22805433651739
SumXY= 124759253.78237706
COVxy= 131.60850759109417
```

2 Correlation Coefficient

```
[ ]: SEGx = X.std()
SEGy = Y.std()
r = SEGxy/(SEGx*SEGy)
print('SEGx= ', SEGx)
print('SEGy= ', SEGy)
print('r= ', r)
```

```
SEGx= 9.772232778476152
SEGy= 14.563402841085871
r= 0.924756298740881
```

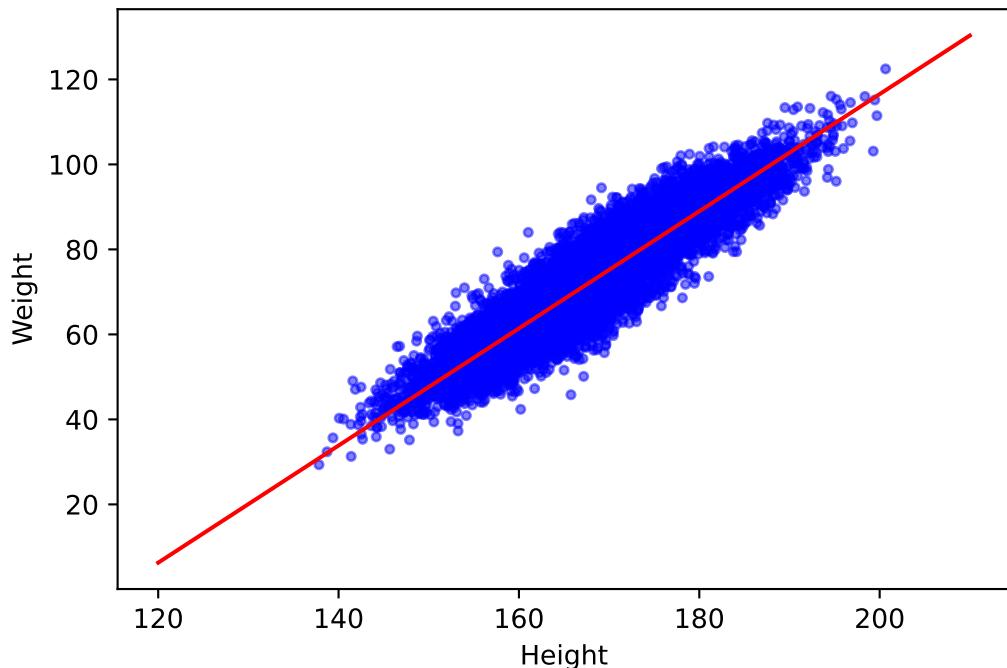
3 Linrear Regression

3.1 Regression line of Y on X: $y = ax + b$

```
[ ]: a = SEGxy/(SEGx**2)
b = -a*Xm + Ym
print('a= ', a)
print('b= ', b)
```

```
a= 1.3781495809287396
b= -159.0915843084424
```

```
[ ]: plt.figure()
plt.scatter(X,Y, s=10, color="blue", alpha=0.5)
x= np.linspace(120, 210,1000)
y=a*x+b
plt.xlabel('Height')
plt.ylabel('Weight')
plt.plot(x,y,color='red')
plt.show()
```



3.2 Regression line of X on Y: $x = ay + b$

```
[ ]: a2 = SEGxy/(SEGy**2)
b2 = -a2*Ym + Xm
print('a= ', a2)
print('b= ', b2)
```

```
a= 0.620523507676597
b= 123.13387264001777
```

```
[ ]: plt.figure()
plt.scatter(Y,X, s=10, color="blue", alpha=0.5)
y2= np.linspace(25, 130,1000)
x2=a2*y2+b2
plt.ylabel('Height')
plt.xlabel('Weight')
plt.plot(y2,x2,color='red')
plt.show()
```

