# Scaling Example

Ravinder Singh

# Who is more similar?

**A**

6 Ft

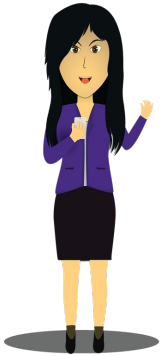
190 lbs

**B**

6 Ft

175 lbs

**C**



5 Ft

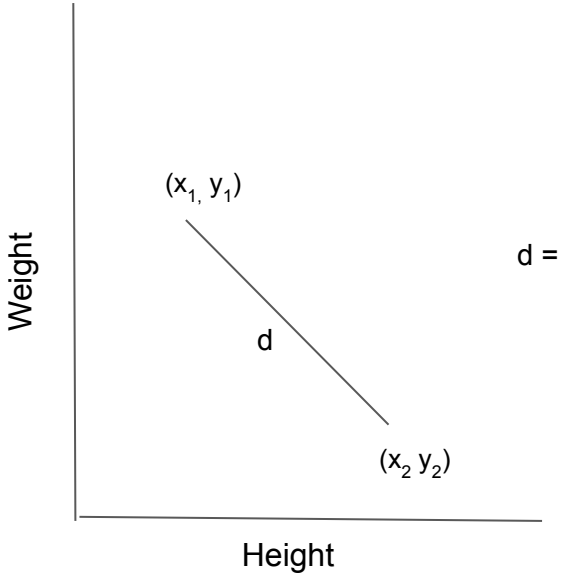
115 lbs

**D**

6 Ft

125 lbs

# Calculating Euclidean distance




$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

# Calculate distance

| | Height (ft) | Weight (lbs) |
|---|---|---|
| Male A | 6 | 190 |
| Male B | 6 | 175 |

$$d = \sqrt{(6-6)^2 + (190-175)^2} = 15$$

# Calculate distance

| | Height (ft) | Weight (lbs) |
|---|---|---|
| Female C | 5 | 115 |
| Female D | 6 | 125 |

$$d \simeq \sqrt{(6-5)^2 + (125-115)^2} \simeq 10.09$$

# Who is more similar?

**A**

**6 Ft**

**190 lbs**

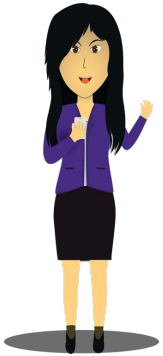
**B**

**6 Ft**

**175 lbs**

Unscaled distance: C & D are more similar than A & B

**C**



**5 Ft**

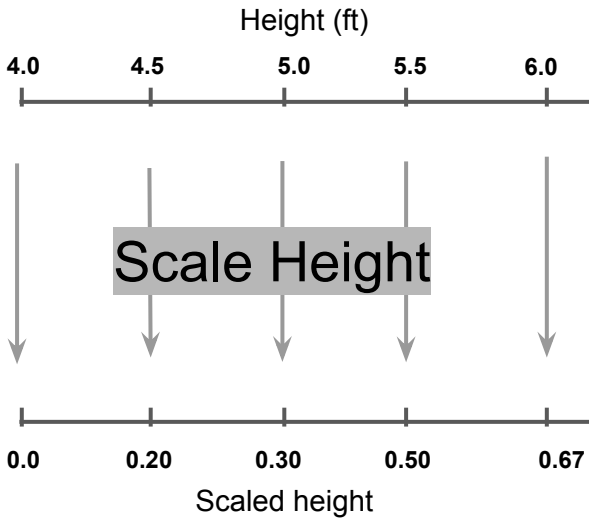
**115 lbs**

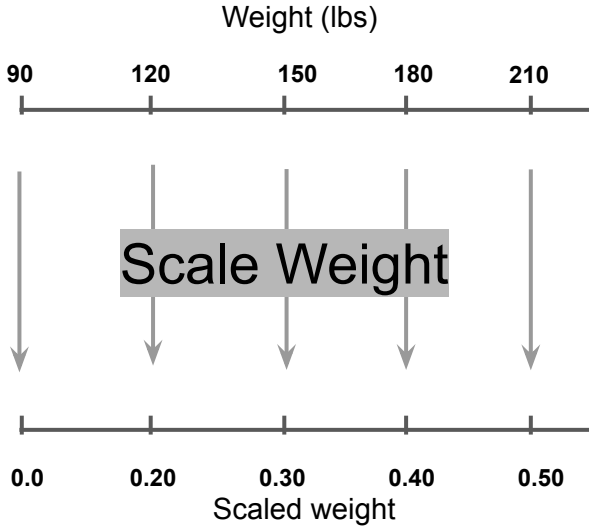
**D**

**6 Ft**

**125 lbs**

# Scale weight and height



# Calculate distance - scaled

| | Height (scaled) | Weight (scaled) |
|---|---|---|
| Male A | 0.67 | 0.47 |
| Male B | 0.67 | 0.40 |

$$d = \sqrt{(0.67 - 0.67)^2 + (0.47 - 0 \cdot 4)^2} = 0.07$$

# Calculate distance - scaled

| | Height (scaled) | Weight (scaled) |
|---|---|---|
| Female C | 0.30 | 0.02 |
| Female D | 0.67 | 0.05 |

$$d = \sqrt{(0 \cdot 67 - 0 \cdot 3)^2 + (0.05 - 0 \cdot 02)^2} = 0.37$$

# Who is more similar?

Scaled distance: A & B are more similar than C & D

**A**

**6 Ft**

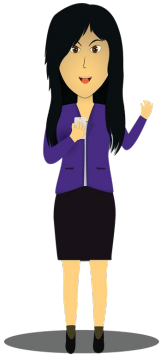
**190 lbs**

**B**

**6 Ft**

**175 lbs**

**C**



**5 Ft**

**115 lbs**

**D**

**6 Ft**

**125 lbs**

Available in Sklearn.preprocessing:

Basic
- **Min-Max Scaler**
- **Standard Scaler**
- MaxAbsScaler (sparse)
- Robust Scaler (outliers)

Advanced
- Quantile Transformer (Uniform)
- Quantile Transformer (Normal)
- Power Transformer (Box-Cox or Yeo-Johnson)
- Normalizer (Don’t use)