

# Analyse en composante principales

Point de vue géométrique

---

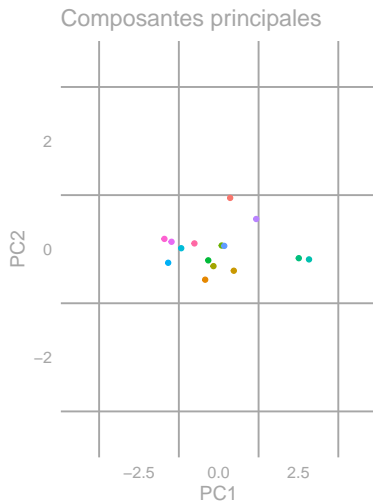
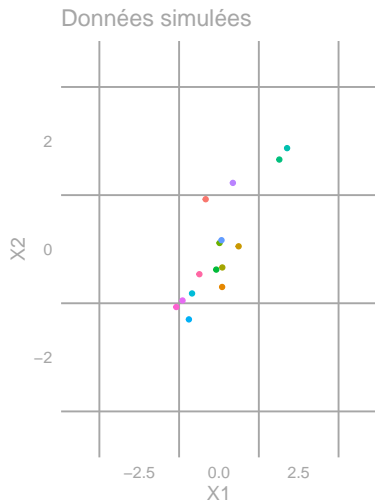
Véronique Tremblay

# Visualisation des observations

---

D'un point de vue géométrique, l'ACP projette les observations dans un sous-espace de dimensions inférieur.

Les composantes principales sont simplement les coordonnées des observations sur les nouveaux axes (axes factoriels).



## Préserve la distance entre les points

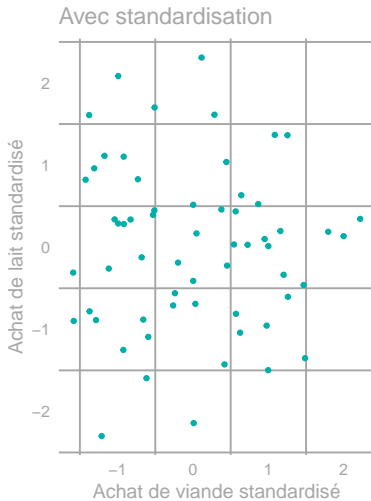
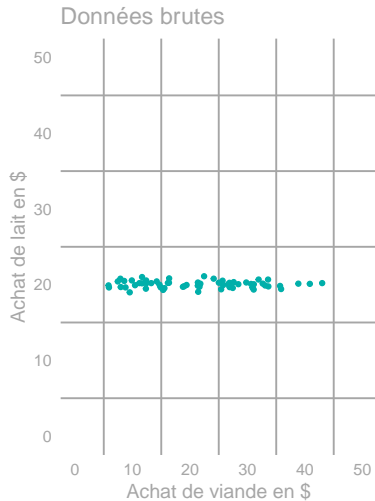
---

L'ACP préserve la distance entre les points.

$$\begin{aligned}\|Y_i - Y_j\|^2 &= (Y_i - Y_j)^\top (Y_i - Y_j) \\ &= \{A^\top (X_i - X_j)\}^\top A^\top (X_i - X_j) \\ &= (X_i - X_j)^\top A A^\top (X_i - X_j) \\ &= (X_i - X_j)^\top (X_i - X_j) \\ &= \|X_i - X_j\|^2,\end{aligned}$$

puisque  $A^\top = A^{-1}$ .

# Standardiser ou pas?



## Proportion de la variation expliquée

---

La **proportion de variation expliquée** la composante principale  $Y_i$  est

$$\frac{\lambda_i}{\lambda_1 + \dots + \lambda_p}.$$