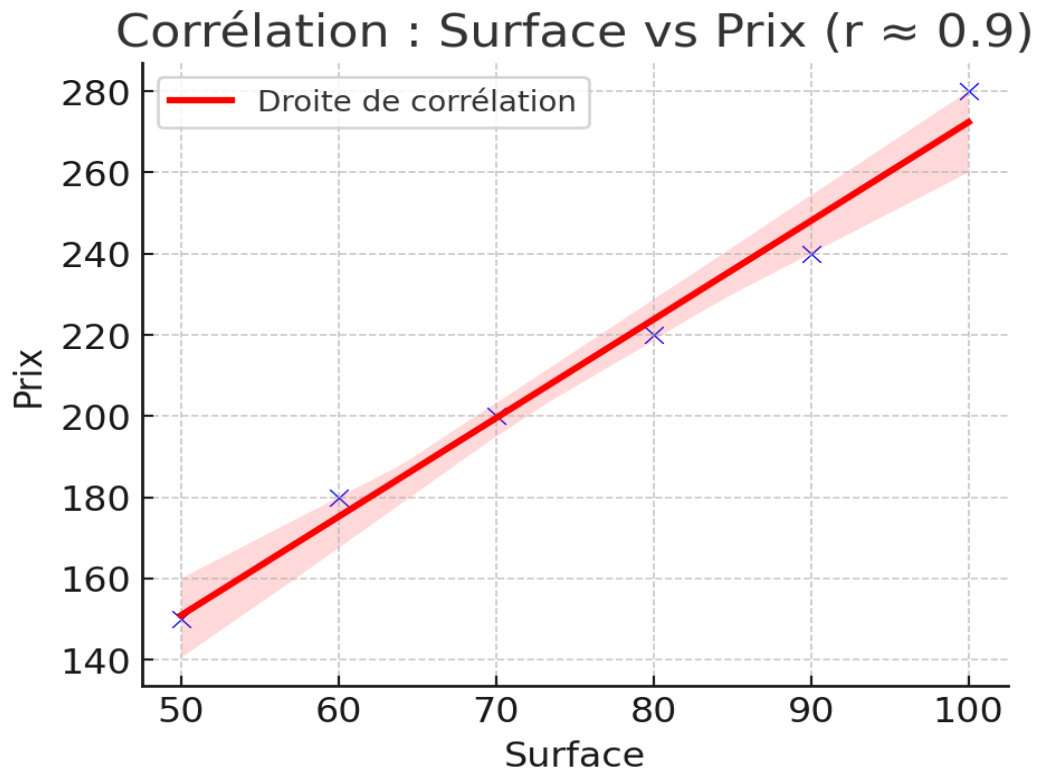


Corrélation, Régression et Feature Importance

1. Corrélation entre deux variables

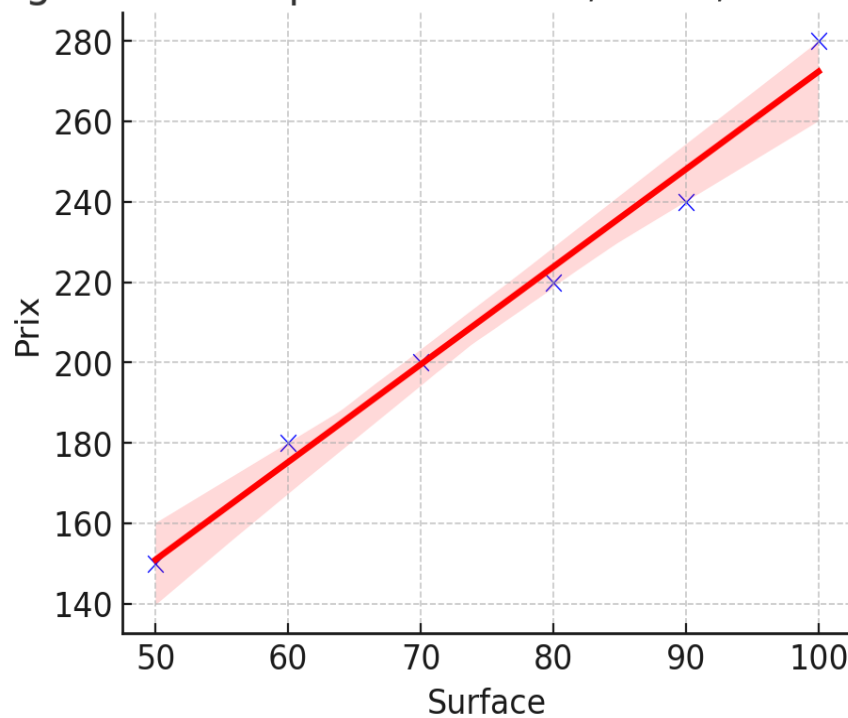
La corrélation de Pearson mesure le degré de relation linéaire entre deux variables. Ici, la surface et le prix ont une forte corrélation positive ($r \approx 0,9$).



2. Régression simple

La régression linéaire simple transforme cette corrélation en une équation prédictive : $\text{Prix} \approx 29,4 + 2,43 \times \text{Surface}$.

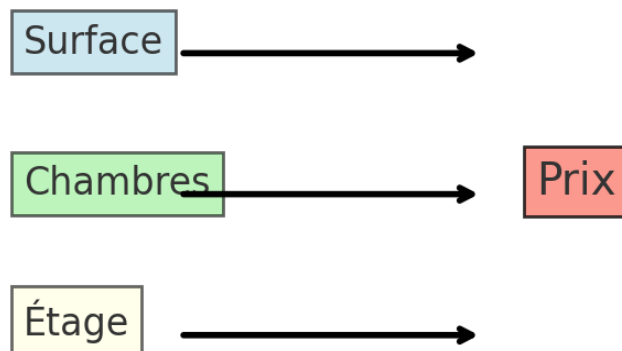
Régression simple : $\text{Prix} \approx 29,4 + 2,43 \times \text{Surface}$



3. Régression multiple

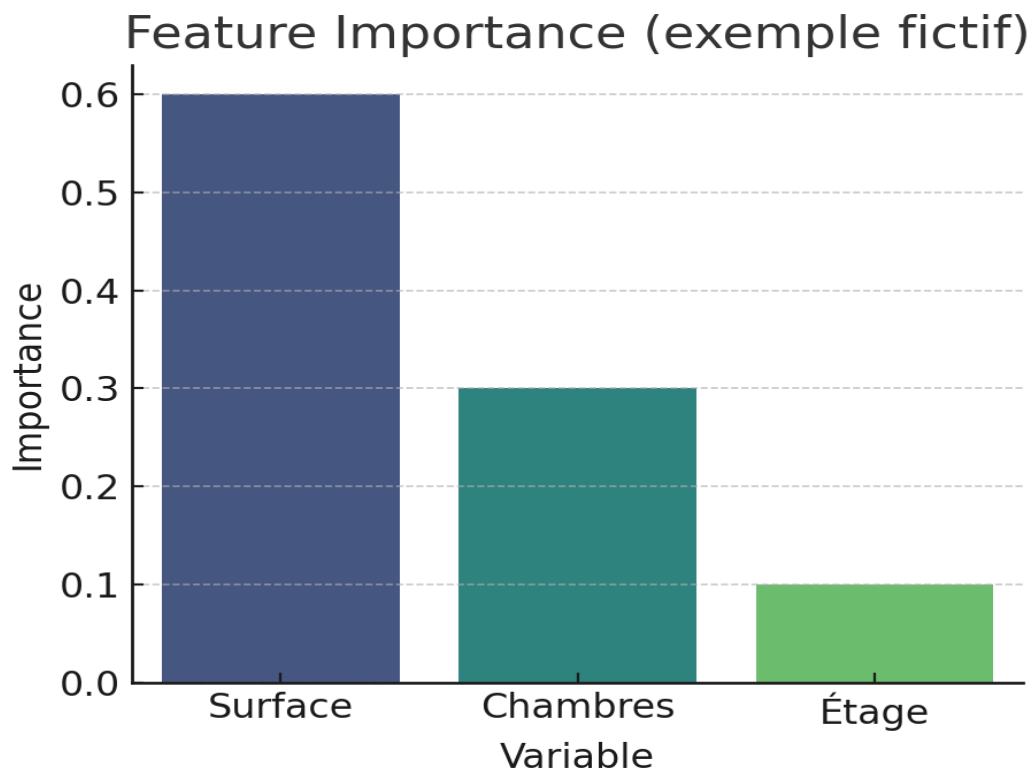
Dans la réalité, plusieurs variables peuvent expliquer Y. Ici par exemple : Surface, Nombre de chambres, Étage. La régression multiple prend en compte toutes ces relations simultanément.

Régression multiple : plusieurs variables expliquent Y



4. Feature importance

Les algorithmes de Machine Learning attribuent une importance relative à chaque variable, permettant d'identifier celles qui contribuent le plus à la prédiction.



Conclusion

- La corrélation est une mesure simple de relation linéaire.
- La régression linéaire transforme cette relation en équation prédictive.
- La régression multiple permet de prendre en compte plusieurs variables.
- La feature importance généralise la corrélation et mesure le poids relatif de chaque variable dans un modèle.