

- **(Multi-)Colinéarité**

- *Forte corrélation entre variables indépendantes*
  - *Faible contribution **spécifique** (unique) d'une ou plusieurs variables*

- **Diagnostic**

- **Examen de la matrice des corrélations bivariées (ou matrice de scatterplots)**

- Identification de corrélations bivariées fortes

- **Examen détaillé des corrélations partielles et semi-partielles**

- **Tolérance (1= absence totale de multicolinéarité)**

- Pour une variable indépendante particulière
  - Proportion de la variance non-partagée avec les autres variables indépendantes
    - $1-R^2$  de la régression de cette variables sur toutes les autres variables indépendantes)
    - $GD=0.963$ : Variation spécifique de la variable GD (seul 4% de la variance de la variable est commune avec les trois autres variable indépendantes)

	Collinearity Statistics	
	Tolerance	VIF
(Constant)		
Age du répondant	.967	1.034
Echelle Gauche-Droite	<b>.963</b>	1.038
Niveau d'éducation	.949	1.053
Taille de la commune	.965	1.036

- Tolérance très faible (*multicolinéarité forte*) → *La variable indépendante est redondante*
  - *Réexaminer le rôle de la variable dans le modèle*

- **VIF Facteur d'inflation de la variance**

- 1/tolérance
- $VIF=1$  (*Absence totale de multicolinéarité*); seuil de colinéarité? A vous de choisir!
  - Certains auteurs indiquent de valeurs de VIF de 4, 5, 10,... avant de s'inquiéter

- **Diagnostic basé sur une analyse en composantes principales (ACP)**
  - Colinéarité, si les premières valeurs propres (eigenvalues) sont proches de 0 (→ indices de condition élevés)
  - Analyse fine: Examen des proportions de variance
    - Colinéarité: Indices de condition élevés associées à des proportions de variance relativement importantes

Collinearity Diagnostics<sup>a</sup>

Dimension	Eigenvalue	Condition Index	Variance Proportions				
			(Constant)	Age	GD	Educ	Tcommune
1	4.479	1.000	.00	.00	.01	.01	.01
2	.248	4.251	.00	.01	.11	.65	.03
3	.131	5.837	.00	.03	.75	.04	.29
4	.114	6.267	.00	.54	.06	.01	.35
5	.027	12.763	.99	.41	.08	.28	.33

## • Solutions

- **Enlever les variables qui causent problème**
  - ... et si cela concerne une variable centrale de votre modèle?
- **Combiner les variables fortement corrélées (échelle, analyse en composantes principales)**
- **Revoir le modèle, si aucune solution est satisfaisante.**
- ...