

# Analyse Statistique : Régression Linéaire et Analyse de Sensibilité de la VAN.

Réalisé avec Python (Jupyter Notebook)

Mohamed FOFANA

Juillet 2025

## 1 Introduction générale

Ce rapport présente une étude statistique structurée en deux volets :

- La modélisation de la consommation énergétique via régression linéaire et ANOVA ;
- L'analyse de sensibilité de la Valeur Actuelle Nette (VAN) selon le taux d'actualisation (2025–2044).

## Partie I – Analyse Statistique Inférentielle

### 1. Régressions Linéaires Simples

**Objectif :** Étudier la relation entre la consommation totale d'énergie et les différentes sources : Électricité, Gaz naturel, Fioul domestique, Pétrole, Bois.

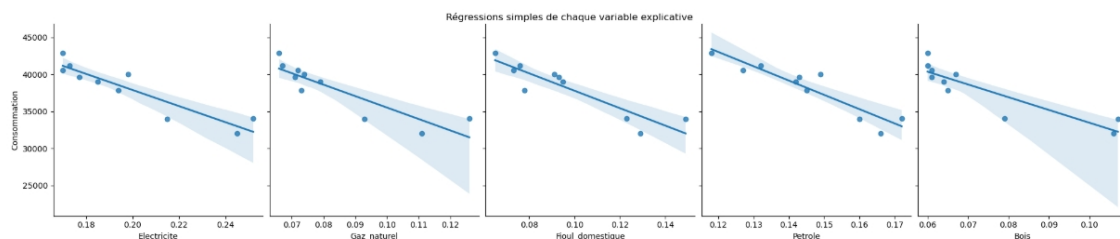


FIGURE 1 – Régressions simples : Consommation  $\sim$  Énergie (par source)

**Interprétation :** Les régressions simples indiquent que :

- La consommation est fortement corrélée à l'électricité ( $R^2$  élevé).
- Le gaz naturel et le fioul montrent également une corrélation modérée.
- Le pétrole et le bois ont un impact visiblement plus faible, voire marginal.

## 2. Test t de student et F de Fisher (F-statistique)

OLS Regression Results						
Dep. Variable:	Consommation	R-squared:	0.985			
Model:	OLS	Adj. R-squared:	0.966			
Method:	Least Squares	F-statistic:	52.57			
Date:	Sat, 26 Jul 2025	Prob (F-statistic):	0.000968			
Time:	20:43:51	Log-Likelihood:	-74.505			
No. Observations:	10	AIC:	161.0			
Df Residuals:	4	BIC:	162.8			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	6.541e+04	3446.078	18.980	0.000	5.58e+04	7.5e+04
Electricite	1.834e+05	7.01e+04	2.615	0.059	-1.13e+04	3.78e+05
Gaz_naturel	-1.876e+05	6.3e+04	-2.979	0.041	-3.62e+05	-1.28e+04
Fioul_domestique	1.385e+05	5.04e+04	2.746	0.052	-1540.541	2.79e+05
Petrole	-3.036e+05	8.25e+04	-3.679	0.021	-5.33e+05	-7.44e+04
Bois	-2.371e+05	5.76e+04	-4.118	0.015	-3.97e+05	-7.72e+04
Omnibus:	0.321	Durbin-Watson:	2.597			
Prob(Omnibus):	0.852	Jarque-Bera (JB):	0.178			
Skew:	-0.246	Prob(JB):	0.915			
Kurtosis:	2.570	Cond. No.	688.			

FIGURE 2 – Test t de student et F de Fisher

### Interprétation :

Les résultats de la régression multiple permettent de tester l'influence combinée et individuelle des différentes sources d'énergie sur la consommation totale.

- Le coefficient de détermination  $R^2 = 0.985$  indique que 98,5% de la variabilité de la consommation est expliquée par le modèle. Ce taux est très élevé, ce qui témoigne d'un bon ajustement.
- Le test de Fisher ( $F = 52.571$ ,  $p$ -value = 0.0010) est très significatif. Ainsi, au moins une des variables explicatives apporte une information significative au modèle. Cela valide l'intérêt du modèle multiple dans son ensemble.
- Les tests t permettent d'évaluer l'effet significatif individuel de chaque variable :
  - **Électricité** :  $p = 0.059$  — proche du seuil de 5%, effet potentiellement significatif.
  - **Gaz naturel** :  $p = 0.041$  — effet significatif, corrélation négative inattendue.
  - **Fioul domestique** :  $p = 0.052$  — proche du seuil, effet modérément significatif.
  - **Pétrole** :  $p = 0.021$  — effet significatif et négatif.
  - **Bois** :  $p = 0.015$  — effet fortement significatif et également négatif.
- Le modèle contient donc des variables avec des effets positifs (électricité, fioul) et négatifs (pétrole, bois, gaz naturel), ce qui peut refléter des réalités économiques ou des substitutions entre énergies.
- Enfin, les statistiques de normalité des résidus (test d'Omnibus  $p = 0.852$ , Jarque-Bera  $p = 0.915$ ) valident l'hypothèse de normalité, renforçant la fiabilité des tests t et F.

## 2 ANOVA de la régression

```
[6]: from statsmodels.stats.anova import anova_lm

anova_results = anova_lm(model_multiple)
print(anova_results)
```

	df	sum_sq	mean_sq	F	PR(>F)
Electricite	1.0	9.733364e+07	9.733364e+07	224.565531	0.000116
Gaz_naturel	1.0	5.969688e+03	5.969688e+03	0.013773	0.912233
Fioul_domestique	1.0	9.150997e+06	9.150997e+06	21.112933	0.010070
Petrole	1.0	9.038739e+04	9.038739e+04	0.208539	0.671613
Bois	1.0	7.348453e+06	7.348453e+06	16.954151	0.014642
Residual	4.0	1.733724e+06	4.334309e+05	NaN	NaN

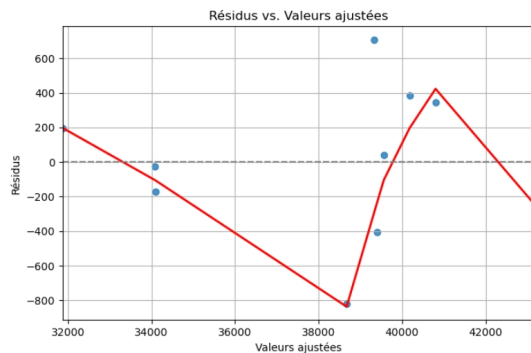
FIGURE 3 – Tableau ANOVA de la régression

**Interprétation :** L'analyse de variance (ANOVA) permet de tester si chacune des variables explicatives prises séparément (régression simple) a un effet significatif sur la consommation.

- **Électricité :**
  - $F = 224.57$ ,  $p\text{-value} = 0.000116 < 0.05$
  - Effet hautement significatif sur la consommation.
- **Gaz naturel :**
  - $F = 0.014$ ,  $p\text{-value} = 0.912$
  - Aucune influence significative détectée.
- **Fioul domestique :**
  - $F = 21.11$ ,  $p\text{-value} = 0.010 < 0.05$
  - Influence significative sur la consommation.
- **Pétrole :**
  - $F = 0.21$ ,  $p\text{-value} = 0.672$
  - Aucune relation significative.
- **Bois :**
  - $F = 16.95$ ,  $p\text{-value} = 0.015 < 0.05$
  - Influence significative détectée.

En résumé, seules l'électricité, le fioul domestique et le bois ont un effet significatif pris séparément. Le gaz naturel et le pétrole ne présentent pas d'effet significatif individuel selon l'ANOVA.

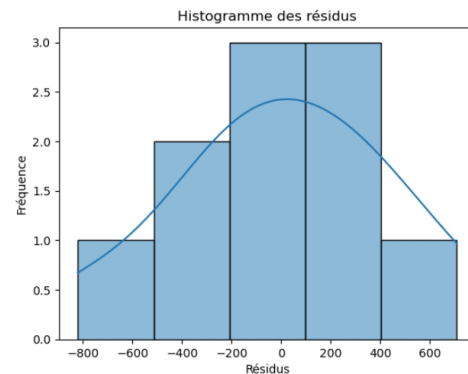
### 3. Analyse des Résidus



**Graphique 1 : Résidus vs valeurs ajustées**

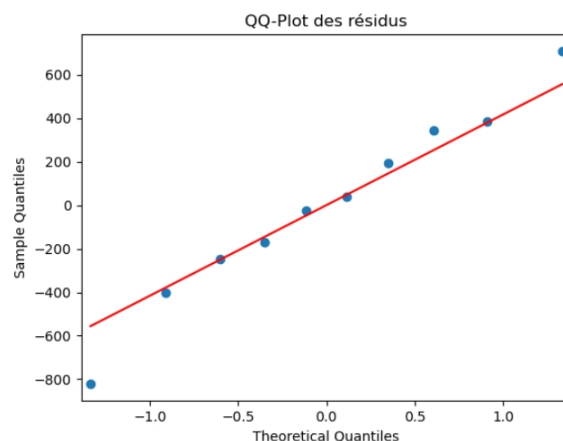
**Interprétation :** Le nuage semble centré autour de zéro, sans motif clair, ce qui suggère une variance constante (homoscédasticité). Toutefois, une légère forme pourrait indiquer une possible non-linéarité.

```
[12]: sns.histplot(residus, kde=True)
plt.title("Histogramme des résidus")
plt.xlabel("Résidus")
plt.ylabel("Fréquence")
plt.show()
```



**Graphique 2 : Histogramme des résidus**

**Interprétation :** Les résidus suivent approximativement une distribution normale, avec un pic centré autour de zéro. Cela soutient l'hypothèse de normalité.



**Graphique 3 : Q-Q plot des résidus**

**Interprétation :** La majorité des points suit la ligne théorique, ce qui indique que les résidus suivent une loi normale. On valide ainsi l'hypothèse de normalité pour les tests paramétriques précédents.

FIGURE 4 – Analyse graphique des résidus du modèle de régression multiple

## Partie II – Analyse de Sensibilité de la VAN

### Objectif

Évaluer la sensibilité de la Valeur Actuelle Nette (VAN) des flux énergétiques au taux d'actualisation pour la période 2025–2044. La VAN est un indicateur de rentabilité des investissements.

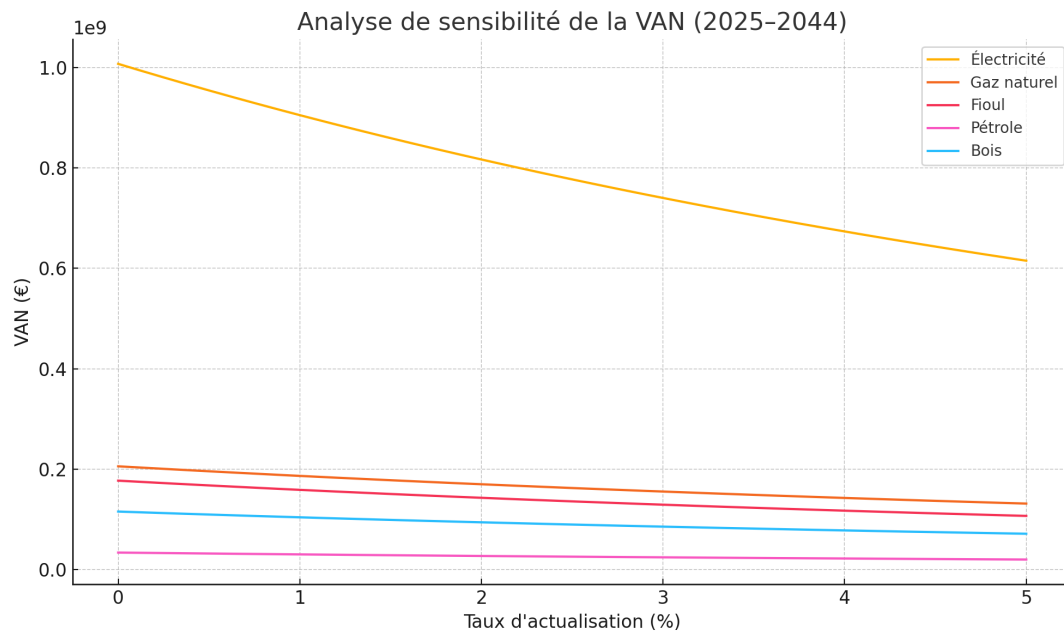


FIGURE 5 – Analyse de sensibilité de la VAN pour les 5 sources d'énergie

### Interprétation :

- La VAN diminue avec l'augmentation du taux d'actualisation, conformément aux principes financiers.
- L'électricité a la VAN la plus élevée, reflétant un apport massif en valeur.
- Le bois et le pétrole sont les moins sensibles, mais aussi les moins rentables à long terme.
- Le fioul et le gaz ont des positions intermédiaires, avec une sensibilité modérée.

*Conclusion :* Cette analyse met en lumière l'importance stratégique de l'électricité dans la consommation énergétique, ainsi que la nécessité de bien calibrer le taux d'actualisation dans les analyses économiques de transition énergétique.